Mateja Jamnik · Yuri Uesaka
Stephanie Elzer Schwartz (Eds.)

LNAI 9781

Diagrammatic Representation and Inference

9th International Conference, Diagrams 2016
Philadelphia, PA, USA, August 7–10, 2016
Proceedings

Springer

Lecture Notes in Artificial Intelligence 9781

Subseries of Lecture Notes in Computer Science

More information about this series at http://www.springer.com/series/1244

Mateja Jamnik · Yuri Uesaka
Stephanie Elzer Schwartz (Eds.)

Diagrammatic Representation and Inference

9th International Conference, Diagrams 2016
Philadelphia, PA, USA, August 7–10, 2016
Proceedings

Springer

Editors
Mateja Jamnik
The University of Cambridge
Cambridge
UK

Stephanie Elzer Schwartz
Millersville University
Millersville, PA
USA

Yuri Uesaka
The University of Tokyo
Tokyo
Japan

ISSN 0302-9743 ISSN 1611-3349 (electronic)
Lecture Notes in Artificial Intelligence
ISBN 978-3-319-42332-6 ISBN 978-3-319-42333-3 (eBook)
DOI 10.1007/978-3-319-42333-3

Library of Congress Control Number: 2016944340

LNCS Sublibrary: SL7 – Artificial Intelligence

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

The 9th International Conference on the Theory and Applications of Diagrams (Diagrams 2016) was held in Lafayette Hill just outside Philadelphia, USA, in August 2016. For the second time, Diagrams was co-located with the Annual Conference of the Cognitive Science Society (CogSci 2016), which encouraged the participation and stimulating interaction between researchers of both communities.

Diagrams is the only conference series that provides a united forum for all areas that are concerned with the study of diagrams. The authors continued to interpret this broadly, as intended by this interdisciplinary conference series. The result is a wide spectrum of contributions from diverse research cultures including cognitive science, computer science, psychology, philosophy, mathematics, design, history of science, logic, linguistics, and artificial intelligence.

Submissions to Diagrams 2016 were solicited in the form of long papers, short papers, and poster papers. The peer-review process entailed each paper being reviewed by at least three members of the Program Committee or a nominated external reviewer. The rebuttal phase followed next, where the authors had a chance to respond to the initial feedback. Finally, the reviewers discussed the papers, the feedback, and the responses from the authors, and amended their reviews as appropriate. This peer review was organized using the EasyChair system. The quality and substance of the reviewers' contributions allowed the program chairs to make decisions with confidence.

We are grateful to the 36 members of the Program Committee and the nine additional reviewers whose commitment, efforts, and diligence enabled us to select high-quality papers for the conference and this proceedings volume. A total of 48 submissions were received, of which 12 were accepted as long papers. A further 11 papers were accepted as short papers.

In addition to the main paper presentations, Diagrams 2016 included a graduate symposium, a workshop, and two tutorials. We are grateful to our two keynote speakers, Mary Hegarty of the University of California Santa Barbara, USA, and Isabel Meirelles of OCAD University, Toronto, Canada: They brought inspiration from their diverse fields of psychology and design. The abstracts of all of these events can be found in this proceedings volume.

The University of Millersville supported the conference's registration process. We also acknowledge the National Science Foundation (NSF) for funding the graduate symposium. Finally, we thank the Organizing Committee for managing their responsibilities so effectively: Richard Burns as local chair, Luana Micallef as graduate symposium chair, and Aidan Delaney as workshop and tutorial chair.

August 2016

Mateja Jamnik
Yuri Uesaka
Stephanie Elzer Schwartz

Organization

Executive Committee

General Chair

Stephanie Elzer Schwartz	Millersville University, USA

Program Chairs

Mateja Jamnik	University of Cambridge, UK
Yuri Uesaka	The University of Tokyo, Japan

Local Chair

Richard Burns	West Chester University, USA

Graduate Symposium Chair

Luana Micallef	Aalto University, Finland

Workshops and Tutorials Chair

Aidan Delaney	University of Brighton, UK

Program Committee

Dave Barker-Plummer	Stanford University, USA
Lisa Best	University of New Brunswick, Canada
Alan Blackwell	University of Cambridge, UK
Paolo Bottoni	Sapienza University of Rome, Italy
Richard Burns	West Chester University, USA
B. Chandrasekaran	The Ohio State University, USA
Peter C-H. Cheng	University of Sussex, UK
Richard Cox	Monash University, Australia
Aidan Delaney	University of Brighton, UK
Tim Dwyer	Monash University, Australia
Jacques Fleuriot	University of Edinburgh, UK
Jean Flower	Autodesk, UK
John Gero	George Mason University, USA
Ashok Goel	Georgia Institute of Technology, USA
John Howse	University of Brighton, UK
Yasuhiro Katagiri	Future University Hakodate, Japan
Maithilee Kunda	Georgia Institute of Technology, USA
John Lee	University of Edinburgh, UK

Emmanuel Manalo	Waseda University, Japan
Kim Marriott	Monash University, Australia
Luana Micallef	Aalto University, Finland
Nathaniel Miller	University of Northern Colorado, USA
Mark Minas	Universität der Bundeswehr München, Germany
Amirouche Mokfeti	Tallinn University of Technology, Estonia
Hari Narayanan	Auburn University, USA
Ian Oliver	Nokia, Finland
Beryl Plimmer	University of Auckland, New Zealand
Helen Purchase	University of Glasgow, UK
Peter Rodgers	University of Kent, UK
Frank Ruskey	University of Victoria, Canada
Atsushi Shimojima	Doshisha University, Japan
Sun-Joo Shin	Yale University, USA
Gem Stapleton	University of Brighton, UK
Ben Steichen	Santa Clara University, USA
Nik Swoboda	Universidad Politécnica de Madrid, Spain
Ryo Takemura	Nikon University, Japan

Additional Reviewers

Jim Burton
Chunlei Chang
Peter Chapman
Veronika Irvine
Steven Kieffer
Areti Manataki
Petros Papapanagiotou
Sonja Schimmler
Yalong Yang

Abstracts

The Power of Diagrams in Science and Challenges of Diagrams in Science Education

Mary Hegarty

Department of Psychological and Brain Sciences, University of California, Santa Barbara, USA
mary.hegarty@psych.ucsb.edu

Diagrams are used extensively by scientists to represent and reason about spatial structures and processes. Diagrams in science often represent things we cannot see in reality (e.g., molecules), show views of things that we cannot see in reality (e.g., cross sections of internal anatomy) and abstract from reality (e.g., various kinds of schematics). Scientists often use different diagrams of the same entities in different reasoning contexts. While experts are typically facile in using the various diagrammatic representations of their discipline, mastering these representations can be challenging for novices.

In this talk I will examine some of the cognitive processes that are involved in understanding diagrams in domains such as mechanics, organic chemistry, and anatomy. These processes include inferring 3-D structure from different 2-D representations, and inferring motion from static diagrams. They include a range of strategies including visualization and more analytic strategies, which are often automatic for an expert but a significant challenge for students. Examining how novices struggle with mastering the diagrams of different scientific domains can reveal these mental processes and provide guidance on how to teach diagrammatic comprehension and reasoning and how to develop spatial thinking more generally.

The Visualizing Spirit in the Twenty-First Century

Isabel Meirelles

OCAD University, Toronto, Canada
imeirelles@faculty.ocadu.ca

Our current obsession with collecting, quantifying and analyzing all types of data is not new. The eighteenth century saw most disciplines in the sciences and the humanities share a "quantifying spirit" characterized by the systematization of knowledge as well as a preoccupation with measuring all types of phenomena [1]. It also saw the creation of novel graphic methods, such as timelines and statistical graphics, to mention two. Jacques Barbeu-Dubourg devised the first known timeline in 1753, somewhat in parallel with Joseph Priestly's 1765 *Chart of Biography* [2]. Between 1786 and 1801, William Playfair created the pie chart, bar graph, and line and area graphs to depict economic data [3]. It is not uncommon that abundance of data is followed by the invention and application of novel graphic methods [4].

Humanistic and scientific knowledge production and dissemination have used graphic forms throughout history, though its adoption has been slow and in many cases non-existent [3, 5]. However, in recent years, information visualizations have gained unprecedented prominence. In all corners of academia and industry the use of visualizations has risen exponentially, fuelled in part by the need to extract meaning from huge amounts of data and our inabilities to make sense of them without the aid of external cognitive devices.

I would argue that a "visualizing spirit" better describes the present passion and widespread use of visual-spatial techniques in the already quantified sciences, humanities and the arts. Furthermore, I would suggest that the centrality of visualizations in many fields reflects their intrinsic value as teaching and learning tools.

This talk will examine the antecedents and significance of our present "visualizing spirit". I will focus on recent visualization trends, their roles, affordances and limitations in helping us explore, extract and interpret information.

References

1. Frangsmyr, T., Heilbron, J.L., Rider, R.E. (eds.): The Quantifying Spirit in the Eighteenth Century. University of California Press, Berkeley (1990)
2. Meirelles, I.: Design for Information. Rockport, Beverly (2013)
3. Playfair, W.: The Commercial and Political Atlas and Statistical Breviary (Edited and introduced by H. Wainer, I. Spence). Cambridge University Press, New York (2005)

4. Friendly, M.: A brief history of data visualization. In: Chen, C., Härdle, W., Unwin, A. (eds.) Handbook of Data Visualization, pp. 15–56. Springer, Heidelberg (2008)
5. Galison, P.: Images scatter into data, data gather into images In: Latour, B., Weibel, P. (eds.) Iconoclash: Beyond the Image Wars in Science, Religion and Art, pp. 300–323. MIT Press, Cambridge (2002)

Graduate Symposium

Luana Micallef

Helsinki Institute for Information Technology HIIT,
Aalto University, Espoo, Finland
luana.micallef@hiit.fi

The Diagrams Graduate Symposium (GS) provides Master and Doctoral students with an opportunity to present their work and get feedback from established researchers in the field. It is also a supportive environment for students to network and make contact with potential future colleagues or employers. The GS was an integral part of the Diagrams 2016 programme. As in previous years, lively discussions led to suggestions about the students' on-going research, and allowed experienced participants to hear fresh ideas and view some of the new trends in the field.

Students participating in the GS submitted a short paper describing their research. Each paper was reviewed by two distinguished scholars, and based on the reviews, six students were selected. Each of these students gave a presentation at the GS and also showcased a poster at the Diagrams poster session. One student with a poster paper in the main conference programme also gave a presentation at the GS, while five other students who had a Diagrams short or long paper (and who got financial support from the conference) also attended the GS. At the GS, a panel of experts gave feedback to the students about their presentations in an informal and constructive environment. The background of the GS students was widely diverse consisting of 30 % females, 20 % self-funded, 20 % part-time students, from nine different universities in six different countries (USA, Canada, UK, The Netherlands, India, Australia). Their research topics were also different, including: diagram drawing algorithms, evaluation of visualization designs and methods, diagrams in education and everyday life, and diagrams in connection to art, problem solving and reasoning.

The success of this year's GS is owed to, first and foremost, the National Science Foundation (NSF) who generously granted us a bursary of $20,000 USD to organize the GS and financially support all the students who sought funding to attend the conference and present their work. We are also grateful to the distinguished scholars for their insightful reviews of all the GS submissions, and the panel of experts who provided invaluable feedback to the students about their GS presentations. As part of the Organizing Committee of Diagrams 2016, we are indebted to the General Chair, Stephanie Elzer Schwartz, the Local Chair, Richard Burns, and the Program Chairs, Mateja Jamnik and Yuri Uesaka, for making the process of organizing the GS as smooth as possible. Finally, we would like to thank all those students who submitted their work to the GS.

The GS is an excellent opportunity for graduate students to improve their research, and an insightful experience for scholars to learn about the future of our field.

Workshop and Tutorials

Aidan Delaney

University of Brighton, Eastbourne, UK
aidan@ontologyengineering.org

The Tutorial and Workshop program at Diagrams 2016 has built upon the successes of previous conferences. The workshop at Diagrams 2016 on *Set Visualisation and Reasoning* expands the scope of the previous two instances of the Euler Diagrams workshop at Diagrams 2014 and 2012. Moreover, the two tutorials on *Programming Your Pictures* and *Visualizing "Information"* were chosen to both challenge and delight conference attendees.

The updated title for *Set Visualisation and Reasoning* describes the expanded scope of the workshop. Having previously focused on interpretation, reasoning and automated drawing of Euler Diagrams the workshop now describes its scope as including "established set representations like Euler/Venn diagrams and matrices as well as new formalisms, such as linear diagrams and line-sets. The workshop scope covers methods for laying out, reasoning with and evaluating set visualisations." We are indebted to Sven Linker and Peter Rodgers for continuing and expanding this workshop.

One tutorial drew from the praxis of incorporating diagrams into programming languages. Brent Yorgey and Jeffrey Rosenbluth presented a domain specific programming language for the representation of diagrams in Haskell. The Haskell diagrams framework is in the vanguard of approaches that allow exploration of data through programmed visualisations. Brent has a well considered teaching philosophy where he strives to encourage a welcoming community of learning. He liberally applied this philosophy in this tutorial.

Jenna Hartel, Rebecca Noone and Eden Rusnell presented their intriguing iSquared research programme. This programme is an interdisciplinary approach to exploring a space through diagrams. As well as having collected over 1500 hand-drawn diagrams, an iSquared protocol has been developed. The protocol supports the exposition of relationships between people and information in challenging environments such as school classrooms. As such, the tutorial encouraged participants to think outside of their main research focus and to engage with alternative approaches to understanding the intuitive power of diagrams.

The organisers of Diagrams 2016 extend their thanks to all the workshop and tutorial organisers. The dedication of the researchers to their respective areas was obvious from the quality of the workshop and the tutorials.

Contents

Cognitive Aspects of Diagrams

Diagrams Affect Choice of Strategy in Probability Problem Solving 3
Chenmu Xing, James E. Corter, and Doris Zahner

What Constitutes an Effective Representation? 17
Peter C-H. Cheng

Measuring User Comprehension of Inference Rules in Euler Diagrams 32
Sven Linker, Jim Burton, and Andrew Blake

Exploring Representations of Student Time-Use 40
Amy Rae Fox, Erica de Vries, Laurent Lima, and Savannah Loker

Logic and Diagrams

Evaluating Diagrammatic Patterns for Ontology Engineering 51
Eisa Alharbi, John Howse, Gem Stapleton, and Ali Hamie

The Interaction Between Logic and Geometry in Aristotelian Diagrams 67
Lorenz Demey and Hans Smessaert

A Generic Approach to Diagrammatic Representation: The Case of Single Feature Indicator Systems 83
Atsushi Shimojima and Dave Barker-Plummer

A Logical Investigation of Heterogeneous Reasoning with Graphs in Elementary Economics. 98
Ryo Takemura

Euler and Venn Diagrams

Minimizing Clutter Using Absence in Venn-i^e 107
Jim Burton, Mihir Chakraborty, Lopamudra Choudhury, and Gem Stapleton

Human Reasoning with Proportional Quantifiers and Its Support by Diagrams. 123
Yuri Sato and Koji Mineshima

Extensions of Euler Diagrams in Peirce's Four Manuscripts on Logical Graphs. 139
Ahti-Veikko Pietarinen

Diagrams and Education

Hint, Instruction, and Practice: The Necessary Components for Promoting Spontaneous Diagram Use in Students' Written Work? 157
Emmanuel Manalo and Yuri Uesaka

Promoting Multi-perspective Integration as a 21st Century Skill: The Effects of Instructional Methods Encouraging Students' Spontaneous Use of Tables for Organizing Information 172
Yuri Uesaka, Mika Igarashi, and Rei Suetsugu

Using Diagrammatic Drawings to Understand Fictional Spaces: Exploring the Buendía House in Gabriel García Márquez's *One Hundred Years of Solitude*. 187
Sarah Pérez-Kriz and Ricardo F. Vivancos-Pérez

Design Principles for Diagrams

An Investigation into OWL for Concrete Syntax Specification Using UML Notations 197
Anitta Thomas, Aurona J. Gerber, and Alta van der Merwe

Using One-Dimensional Compaction for Smaller Graph Drawings 212
Ulf Rüegg, Christoph Daniel Schulze, Daniel Grevismühl, and Reinhard von Hanxleden

On Comments in Visual Languages. 219
Christoph Daniel Schulze, Christina Plöger, and Reinhard von Hanxleden

Bistable Perception and Fractal Reasoning 226
Keith McGreggor and Ashok Goel

Diagrams Layout

Who, Where, When and with Whom? Evaluation of Group Meeting Visualizations. 235
Simone Kriglstein, Johanna Haider, Günter Wallner, and Margit Pohl

The Perception of Clutter in Linear Diagrams. 250
Mohanad Alqadah, Gem Stapleton, John Howse, and Peter Chapman

People, Place, and Time: Inferences from Diagrams 258
Barbara Tversky, Jie Gao, James E. Corter, Yuko Tanaka, and Jeffrey V. Nickerson

Exploring the Types of Messages that Pie Charts Convey in Popular Media . . . 265
Richard Burns, Eric Balawejder, Wiktoria Domanowska, Stephanie Elzer Schwartz, and Sandra Carberry

Comparison of Diagrams in Producing and Understanding Hierarchies in Three Different Application Domains. 272
Leonie Bosveld-de Smet and Robert-Jan Verheggen

Posters

Evaluating the Effects of Colour in LineSets 283
Dominique Tranquille, Gem Stapleton, Jim Burton, and Peter Rodgers

Negative Terms in Euler Diagrams: Peirce's Solution 286
Amirouche Moktefi and Ahti-Veikko Pietarinen

Visualising the Boolean Algebra $\mathbb{B}_4$ in 3D 289
Hans Smessaert and Lorenz Demey

Single Feature Indicator Systems in the Openbox 293
Dave Barker-Plummer and Michael Murray

Effective Representation of Information: Generalizing Free Rides 296
Gem Stapleton, Mateja Jamnik, and Atsushi Shimojima

Author Index 301

Cognitive Aspects of Diagrams

Diagrams Affect Choice of Strategy in Probability Problem Solving

Chenmu Xing[1(✉)], James E. Corter[1], and Doris Zahner[2]

[1] Teachers College, Columbia University, New York, USA
{cx2116, jec34}@tc.columbia.edu
[2] Council for Aid to Education, New York, USA
dzahner@cae.org

Abstract. We investigated whether diagrams influence strategy choice and success in solving elementary combinatorics problems. Generic diagrams (trees or two-way tables) were provided to solvers as aids. Participants' coded solution strategies revealed that problem solvers tended to utilize mathematical structures and solutions that easily mapped to the diagrams' visuospatial relations. For example, when provided with an unlabeled N-by-N table, solvers tended to proceed by defining an equally-likely outcome space (an "outcome-search" solution); when provided with a binary tree, solvers tended to adopt a "sequential" solution defining stage-wise simple or conditional probabilities; when provided with an N-ary tree cuing equally-likely outcomes, choices were split between the two solution types. Furthermore, the tree and the table showed different patterns of characteristic errors, and perhaps for this reason, the tree led to higher accuracy for one problem that involved sequential sampling without replacement, while the table was best for the other problem, involving independent events. The results support arguments that the content and structure of diagrams must be congruent to that of the problem at hand and be easily and accurately perceived to be effective, and demonstrate that diagrams can influence strategy choice in problem solving.

Keywords: Probability problem solving · Diagram congruence · Diagram design

1 Introduction

1.1 Background

Diagrams are essential tools for representation, communication, and reasoning. In education, diagrams have been used widely and play an important role in STEM learning and problem solving [e.g., 1–5], learning and comprehension of complex systems [2], judgment [6, 7], reasoning [8], analogical transfer [9, 10], planning [11], and data representation and interpretation [12, 13].

But like any tool, a diagram must be well chosen for the task at hand, and its use affects both the process and the product of the activity. First, as an external representation of a cognitive or educational problem, salient aspects of the diagram must map to relevant aspects of the problem [14, 15]. Second, structural, visuospatial, and

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 3–16, 2016.
DOI: 10.1007/978-3-319-42333-3_1

implicit aspects of the chosen diagram can influence and alter people's inferences, judgments, and the perceived relations and structures of the represented information [e.g., 6, 8, 13]. How a diagram steers people to making certain inferences and judgments concerning the represented concepts and relations is not arbitrary. Rather, it stems from cognitively natural ways of mapping visuospatial elements and relations to conceptual content and relations, externally or internally, based on shared metaphorical (or analogous) similarity of abstract relational structures [e.g., 6, 15–17].

A rich body of research has explored how specific types of diagrams affect inferences in reasoning and judgment tasks. When asked to describe the relation of individual data points shown in statistical graphs, people given a bar graph tended to describe the relation as comparisons of discrete entities, but as trends of continuous change when given the same information depicted as a line graph [13, 18]. To describe complex mechanical systems depicted by diagrams, people reading mechanical diagrams with arrows described the functions of the systems, whereas those reading the same diagrams without arrows gave structural descriptions [2]. Other evidence [8] has demonstrated that when people use diagrams to keep track of individuals' locations over different time points, presentation of different graph formats led to different inferences by participants.

In general, people's inferences using diagrams are systematically related to the schemas that different diagram formats convey. For example, lines connect and associate entities, indicating paths, relations, and movement [8, 13, 19]; bars and boxes suggest enclosures and separate categories [8, 13]; and arrows show asymmetric directions and sequences from actions to goals and causes to effects [2, 19].

In this paper, we investigate the role of diagrams in probability problem solving. One general question we address is: can diagrams affect the interpretation or processing of probability word problems? Another is: do diagrams facilitate solution of probability problems? These questions can be answered in different ways. First, authors of probability textbooks seem to believe that diagrams aid in the understanding of probability principles, since such textbooks sometimes feature outcome trees and tables in chapters on probability [20]. Second, there is evidence that spontaneous use of diagrams by probability problems solvers is associated with higher solution success, but only when the diagram type used is appropriate to the problem type [4, 21]. Finally, there is suggestive evidence that providing a generic diagram with a problem text may in some cases facilitate solution of the problem, especially when the problem is somewhat difficult for the would-be solver [22].

Providing a generic diagram might be beneficial to students solving a probability word problem simply because it offers another representation of the problem, and multiple representations have been argued to lead to a deeper understanding of the problem [23]. But if the structure of the diagram is indeed cognitively mapped to the problem structure and the diagram is then used to reason about the problem and its solution, using different types of diagrams to represent a given problem might steer the problem solver in different directions in terms of solution strategies or processes. Such process evidence would provide more diagnostic evidence than mere solution success for how diagrams steer, and ideally facilitate, problem solving processes.

1.2 The Present Study

In the present study, we explore how using different types of diagrams can affect both process (strategy choice) and product (solution success) in probability problem solving. Specifically, we focus on two types of schematic diagrams that are commonly used as visual aids in probability and statistics domains, trees and tables.

Trees and tables differ in diagrammatic structures, and thus may lead to inductions of different schematic information and applications. Novick and Hurley [15] analyzed the basic structures and schematic components of these two general types of diagrams. As they observed, the global structure of a table features the cross-classification of two variables or sets. The rows and columns each represent a variable or set, and the intersection cells represent the combinations of items from the two sets. The global structure of a tree diagram (or hierarchy) features hierarchical levels of events, generated by a branching process. Different levels of events are often dependent so that the identities of one level depend on the identities of the preceding level. Items listed at the same level are mutually exclusive and identical in status, whereas items listed at different levels differ in status or sequence.

Research suggests that these two types of diagrams are chosen for representing different situations and schemas [e.g., 24]. In probability problem solving, tables are used to represent factorial combinations [e.g., 10], and trees may be used to keep track of sequential selections [e.g., 25]. Zahner and Corter [4] found that using tree diagrams was particularly useful for solving conditional probability problems, a type of probability problems that involve sequential and dependent occurrence of events. These empirical findings suggest that trees and tables may evoke different schemas of mathematical relations and choices of different solutions.

Thus, we sought here to investigate whether and how trees and tables lead to different results in interpreting mathematical structures and selecting solutions. To do this, we used problems representing two topics in combinatorics: combinations and independent events. These specific types of problems were chosen because they admit of multiple types of solution strategies (described and illustrated in the Method section), and each can be represented by both trees and tables. We compared the effects of providing three different types of generic diagrams for these combinatorics problems: N-by-N tables, binary trees, and N-ary trees. The diagrams were designed to manipulate two factors: the type of diagram structure (tree versus table) and the abstraction level of the represented outcome space (either a large space of equally-likely outcomes, or a smaller space of unequally-likely outcomes).

We hypothesize that the choice of solution strategies can be influenced by the diagrammatic representation of a problem. The idea is that different types of diagrams should bias people to formulate different mathematical solutions, due to the similarity correspondences between a diagram's visuospatial structures and the selected solution's mathematical structure. That is, the problem solver tends to choose a solution with procedural or mathematical structure that can be easily aligned with the visuospatial relational structure of the diagram.

2 Method

2.1 Participants

The participants were 48 students (39 or 81.3 % female) recruited from a university in New York City. Most were paid $8 for their participation; the others participated for course credit. Their average age was 25.6 years. To be qualified for the experiment, a participant had to have taken at least one undergraduate- or graduate-level statistics course prior to participation. On average, participants reported to have taken 2.33 statistics courses. However, their level of training varied: 19 (or 38.6 %) participants reported having taken one statistics course, 11 (or 22.9 %) participants reported two, 10 (or 20.8 %) reported three, and 8 (or 16.7 %) participants reported four or more such courses.

2.2 Materials

Each participant solved three elementary probability word problems, representing three different probability topics: combinations, independent events, and conditional probability. The first two problems/topics were the target materials for this study, because they admit of two distinct salient solution strategies, which we refer to as outcome-search and sequential-sampling strategies (described below). The third problem, involving conditional probabilities (and referred to below as the Weather problem), does not admit of salient alternative solution strategies, and was treated merely as a filler problem for purposes of this investigation.

The problem text for the independent events problem (also called the Spinner problem) was:

> *Two spinners are constructed. Each spinner has 3 color sections of equal size: red, white, and blue. The two spinners are spun at the same time, and the result of each spinner is recorded. What is the probability of getting the same color on both spinners?*

In the diagram conditions, either a tree (Fig. 1a) or a table (Fig. 2a) was also provided to the problem solver. These diagrams were unlabeled, but the number of branches (or rows and columns) was appropriate to the problem.

The problem text for the combinations problem (also called the Work-Group problem) was as follows. The generic diagrams for this problem are presented in Figs. 3 and 4.

> *Five students are in a work group. The teacher randomly selects two of them to present the group work. If there are 2 boys and 3 girls in this group, what is the probability that the teacher selects 2 girls?*

Although these two problems represent two different probability topics, the same two broad strategies can be used to solve each one. We refer to these two approaches as "outcome-search" and "sequential-sampling". The outcome-search strategy finds the probability by first counting the total number of equally-likely outcomes in the outcome space. For the combinations (Work-Group) problem, this involves computing the

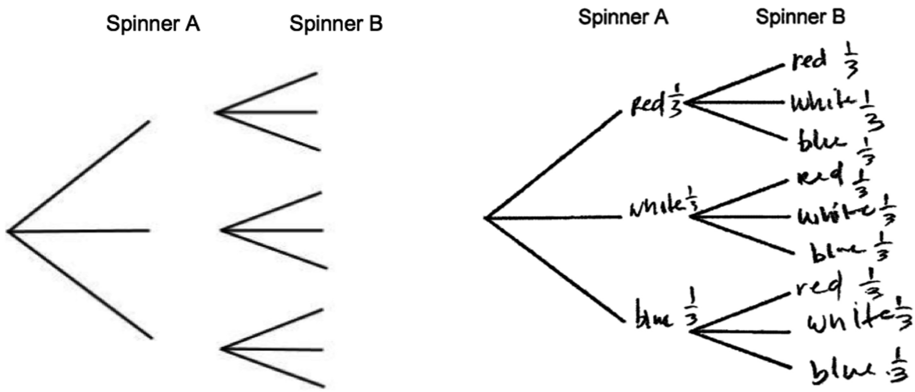


Fig. 1. Panel 1a (on left) shows the tree diagram provided with the Spinner problem for some participants. Panel 1b (on right) shows how the tree diagram was annotated by one participant.

Spinner A	Spinner B		
	____	____	____
____			
____			
____			

Spinner A	Spinner B		
	r	w	b
r	✓	X	X
w	X	✓	X
b	X	X	✓

Fig. 2. Panel 2a (on left) shows the table diagram provided with the Spinner problem for some participants. Panel 2b (on right) shows how the table diagram was annotated by one participant.

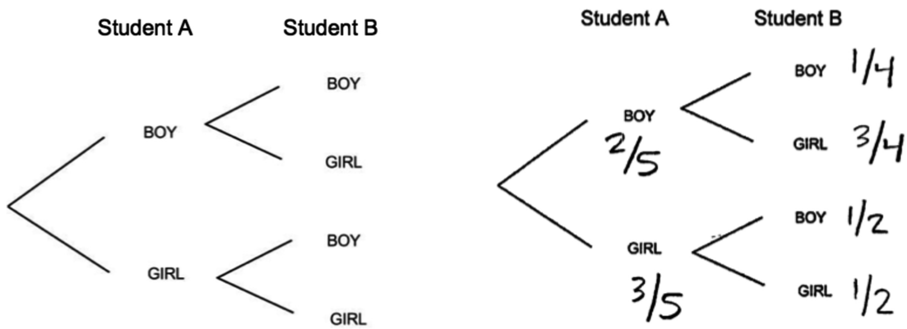


Fig. 3. Panel 3a (on left) shows the tree diagram provided to some participants for the Work-Group problem. Panel 3b (on right) shows how the tree diagram was annotated by one participant.

number of all possible combinations of 2 girls out of 3 girls (=3), and dividing into this the number of all possible combinations of 2 students out of 5 students (=10). Thus, the probability of selecting two girls is 3/10. The sequential-sampling strategy views the problem as sequential sampling without replacement, so the solver first defines the simple or conditional probability at each stage in a sequential order. For example, for

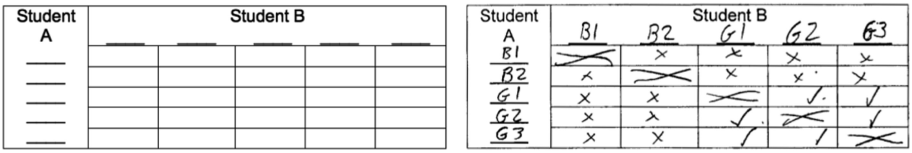


Fig. 4. Panel 4a (on left) shows the table diagram provided to some participants for the Work-Group problem. Panel 4b (on right) shows how the table diagram was annotated by one participant.

the combinations (Work-Group) problem, this strategy first defines $P(G_1) = 3/5$, the probability of selecting a girl on the first draw, and then $P(G_2|G_1) = 2/4$, the probability of selecting a girl again on the second draw, conditional on $P(G_1)$. Finally, $P(G_1 \cap G_2)$, the probability of selecting two girls, is equal to $P(G_1)P(G_2|G_1) = (3/5)(2/4) = 3/10$.

2.3 Design and Procedure

Three test forms were used, each presenting the problems in a different order (Table 1). The first problem for each test form was given in text only, with no provided diagram. For the second and the third problems, one was given with a generic tree diagram, and the other one was given with a generic table diagram. Problems were presented in different orders, counterbalanced to equate possible carry-over effects. Thus, each problem was attempted by three independent groups of participants, with one third of them solving it with no provided diagram, one-third with a tree diagram, and one-third with a table diagram. As Table 1 shows, three types of provided diagram were used: an N-by-N table and a binary tree for the Work-Group (combinations) problem; and an N-by-N table and an N-ary tree for the Spinner (independent events) problem.

Participants (N = 48) were randomly assigned to the three test forms, with 16 participants in each test form. In the task, each participant was given a booklet in which each problem was presented on a separate page. No explicit training was provided; they were asked to solve all three problems, showing their step-by-step solution procedure on the worksheet. To prevent participants from seeing more than one diagram type at a time, they were instructed not to look at any other problems when they were working on a problem. Participants were explicitly instructed to use the provided diagram for problem solving when one was provided. A probability formula sheet was also provided, although participants were told that the formula sheet was optional for them to use. Participants also filled out a brief survey about their basic demographic information and statistics training experiences.

Table 1. The three test forms and the design for the study.

	Problem 1	Problem 2	Problem 3
Form A	Spinner (no diagram)	Weather (tree: binary)	Work-Group (table: N-by-N)
Form B	Weather (no diagram)	Work-Group (tree: binary)	Spinner (table: N-by-N)
Form C	Work-Group (no diagram)	Weather (table: 2-by-2)	Spinner (tree: N-ary)

3 Results

3.1 Coding

Two outcome variables were of focal interest: choice of a solution strategy, and problem solving correctness. Solution types were coded based on whether an outcome-search strategy or a sequential-sampling strategy was used for problem solving. Use of each strategy was coded independently and dichotomously, with a value of 1 if the strategy was used, and 0 otherwise. This strategy determination was made without reference to correctness, which was coded independently. Examples of participants' work using each of the two solution strategies for the two problems are shown in Figs. 5 (the Work-Group problem) and 6 (the Spinner problem). To check reliability, a second coder independently coded 24 of the problem solutions. For each strategy, percent agreement was .96, with r = .92. The sole discrepancy occurred for an incorrect, somewhat disorganized response.

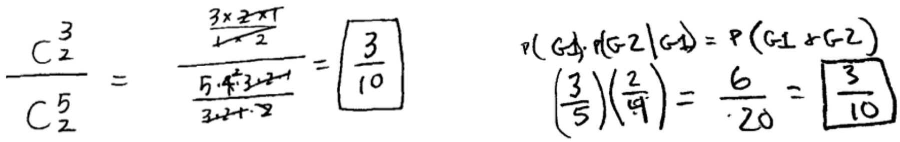


Fig. 5. Panel 5a (on left) shows an example of a solver using the outcome-search strategy for the Work-Group problem. Panel 5b (on right) shows an example of a solver using the sequential-sampling strategy for the Work-Group problem.

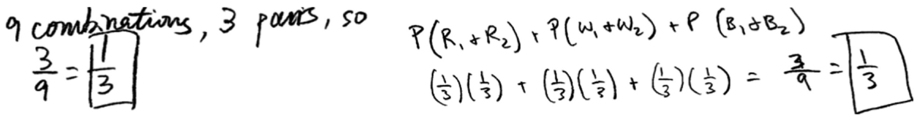


Fig. 6. Panel 6a (on left) shows an example of a solver using the outcome-search strategy for the Spinner problem. Panel 6b (on right) shows an example of a solver using the sequential-sampling strategy for the Spinner problem.

Solution success was assessed with two measures. Answer correctness was coded dichotomously, based on whether the final answer has the correct value, regardless of whether all the solution steps were correct. Procedural correctness was coded to indicate if the solution steps a solver followed were appropriate, regardless of final answer correctness. For example, if a solver's solution procedure was correct, but a computational error led to an incorrect final answer, procedure would be coded as correct (=1), but answer correctness would be coded 0. On the other hand, one participant obtained a correct final answer by an incorrect procedure. For these reasons, procedural correctness was analyzed as our main measure of solution success [cf. 26]. Inter-rater reliability for procedural correctness was perfect (=1.0) for the 24-solution set.

Table 2. Raw and proportional frequencies of solvers' solution choices for the Work-Group problem as a function of diagram type.

Diagram condition	Strategy		Correctness	
	Outcome search	Sequential sampling	Procedural correctness	Answer correctness
No diagram (N = 16)	10 (62.5 %)	7 (43.8 %)	13 (81.3 %)	12 (75.0 %)
Binary tree (N = 16)	4 (25.0 %)	13 (81.3 %)	13 (81.3 %)	11 (68.8 %)
N-by-N table (N = 16)	15 (93.8 %)	1 (6.3 %)	5 (31.3 %)	5 (31.3 %)
Overall (N = 48)	29 (60.4 %)	21 (43.8 %)	31 (64.6 %)	28 (58.3 %)

3.2 Analysis of the Work-Group (Combinations) Problem: Effects of the N-by-N Table Versus the Binary Tree

Strategy Choice. As hypothesized, solvers' choices of the solution strategies were strongly biased by diagram types (Table 2). Because use of each strategy was coded independently, neither or both might both be employed on a given problem; however use of both strategies was relatively rare. Participants given the N-by-N table more frequently used the outcome-search strategy (93.8 %) compared to participants given no diagram (62.5 %), $\chi^2(1,\ 32) = 4.571$, $p = 0.083$ (by Fisher's Exact Test); and less frequently used the sequential-sampling strategy compared to no diagram (6.3 % vs. 43.8 %), $\chi^2(1,\ 32) = 6.000$, $p = 0.037$ (by Fisher's Exact Test). On the other hand, participants given the binary tree showed significantly more use of the sequential-sampling strategy than those in the no-diagram condition (81.3 % vs. 43.8 %), $\chi^2(1,\ 32) = 4.800$, $p = 0.028$; but significantly less use of the outcome-search strategy than those in the no-diagram condition (25 % vs. 62.5 %), $\chi^2(1, 32) = 4.571$, $p = 0.033$.

Solution Success. Table 2 also shows participants' rates of procedural correctness and answer correctness. The procedural correctness rates for the no-diagram condition (81.3 %) and the binary tree condition (81.3 %) were high and identical. However, procedural correctness for the N-by-N table condition was significantly lower (31.3 %) than for the no-diagram condition (81.3 %), $\chi^2(1, 32) = 8.127$, $p = 0.004$.

Error Analysis. We analyzed unsuccessful solvers' error patterns to better understand why the table decreased problem solving success for this problem. When participants were given the N-by-N table for the Work-Group problem, 6 of the 11 erroneous solutions resulted from incorrectly defining the total number of equally likely outcomes, and 5 of these cases resulted directly from counting repeated selection of a single student (represented by the diagonal cells) as valid combinations. Thus, it was a common error to count all 25 cells as valid outcomes by using the 5-by-5 table (Fig. 7). This corresponds to treating the selection of two (distinct) students as sampling with replacement. Another 3 errors involved incorrect use of the combinations formula, or failure to convert the count of outcomes to probability.

Thus, more than 80 % of the procedural errors made with the N-by-N table involved the solver attempting to use the outcome-search strategy but being led astray by the

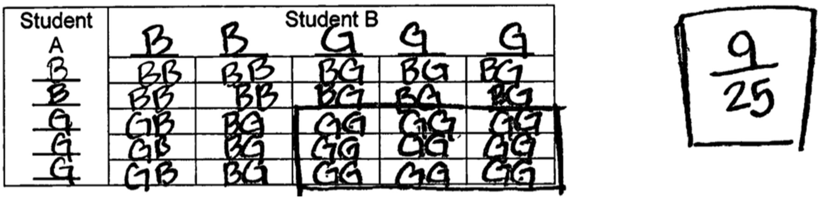


Fig. 7. Counting diagonal cells (=self-repeated combinations) as valid outcomes, a common misuse of the N-by-N table that led to erroneous solutions.

structure of the N-by-N table. Specifically, in order to use the table correctly, problem solvers must recognize that all the diagonal cells should not be used, because the self-repeated combinations that these cells represent are impossible outcomes when sampling without replacement. Put another way, the structure of the table does not map in a one-to-one way with the structure of the problem.

The binary tree led to distinctively different error types, so that three out of five incorrect answers involved incorrectly defining stage-wise probabilities. For example, the correct probabilities for the two sequential selections should be $P(G_1) = 3/5$ and $P(G_2|G_1) = 2/4$. However, the erroneous solutions involved incorrect probability values for these two events.

Discussion. For the Work-Group problem, the tree and the table altered the frequency of using particular strategies, with the N-by-N table biasing solvers to select an outcome-based strategy (the "outcome-search" strategy), and the binary tree leading them to select a sequential strategy based on an event-level representation (the "sequential-sampling" strategy). These differences in turn led to differential error rates for the provided-diagram conditions, and also to characteristic error patterns that are highly distinctive. In particular, the N-by-N table led to more use of an outcome-based strategy, and more errors in identifying the correct equally-likely outcome space, while the binary tree led to more use of a sequential strategy, with typical errors in identifying the correct stage-wise probabilities for unequally-likely outcomes.

We explain these solution-strategy effects in terms of the compatibility between the diagram and the relevant problem characteristics. Combinations problems (e.g., "How many ways can N objects be selected k at a time?") are typically interpreted as involving the simultaneous sampling of k entities from a larger set of N entities (e.g., by application of the formula for the number of combinations of n objects selected k at a time), but can also be formulated and solved as involving sequential sampling (k draws, a single object at a time, without replacement). However, in the latter case, order of selection is implied to be relevant, so answers may require appropriate adjustment.

The N-by-N table displays the outcomes in an outcome space simultaneously, by integrating all outcome cells into a single table. Therefore, it cues solvers to search for all possible outcomes in the whole outcome space and for the target event as a subset embedded in the whole outcome space on the table. Note specifically that an unlabeled N-by-N table implicitly cues solvers to consider all outcome cells as relevant to the problem, including those diagonal cells that represent self-repeated selections, although these are impossible outcomes for the combinations problem.

Thus, the table diagram led to more erroneous identifications of the correct outcome space, because its representational format violates the Principle of Congruence articulated by Tversky, Morrison and Betrancourt [27], in that there is poor fit between the structure of the diagram and the structure of the problem.

In contrast, the binary tree steers problem solvers towards use of an outcome space with only four *unequally-likely* outcomes: S = {BB, BG, GB, GG}. Further, we believe that the tree diagram's left-to-right hierarchical structure cues viewing the problem as involving sequential sampling without replacement. Therefore, the binary tree steers problem solvers towards a sequential-sampling strategy. Correct execution of this strategy requires correctly specifying conditional probabilities for the second student selected, as in Fig. 5b. Here, errors made by the binary tree condition tended to involve incorrect specification of these probabilities.

3.3 Analysis of the Spinner (Independent Events) Problem: Effects of the N-by-N Table Versus the N-ary Tree

The two generic diagrams offered as aids for the Work-Group problem actually vary two aspects of the diagram at once: the diagram structure (trees versus tables), and the abstraction level of outcome space (an equally-likely outcome space based on the specific students selected versus an unequally-likely outcome space based only on sex of the two selected students). For the Spinner problem, exemplifying the use of the fundamental principle of combinatorics, we chose to control for the abstraction level of the outcome space and vary only the diagram type: trees versus tables. We also refer to this problem as the "independent events" problem, because the outcome of the first spinner is independent of the outcome of the second spinner.

Strategy Choice. As seen in Table 3, for the Spinner problem, the N-by-N table diagram led to more frequent use of the outcome-search strategy (68.8 %), compared to 31.3 % when no diagram was provided, $\chi^2(1, 32) = 4.500$, $p = 0.034$. However, the N-ary tree, which cues both use of the equally-likely outcome space and a sequential strategy involving the definition of stage-wise probabilities, led to mixed choices of solutions: 50 % of participants in the N-ary tree condition used the outcome-search strategy, while 50 % of them used the sequential-sampling approach.

Solution Success. Table 3 also shows the frequency and percentage of participants in each condition who successfully solved the Spinner problem. Compared to the

Table 3. Raw and proportional frequencies of solvers' solution choices for the Spinner problem as a function of diagram type.

	Strategy		Correctness	
Diagram condition	Outcome search	Sequential sampling	Procedural correctness	Answer correctness
No diagram (N = 16)	5 (31.3 %)	9 (56.3 %)	8 (50 %)	8 (50 %)
N-ary tree (N = 16)	8 (50 %)	8 (50 %)	10 (62.5 %)	11 (68.8 %)
N-by-N table (N = 16)	11 (68.8 %)	5 (31.3 %)	14 (87.5 %)	14 (87.5 %)
Overall (N = 48)	24 (50 %)	22 (45.8 %)	32 (66.7 %)	33 (68.8 %)

no-diagram condition (50 % procedural correctness rate), both provided diagrams increased the percentage of procedurally correct solutions. The increase for the N-ary tree, to 62.5 %, was not significant, $\chi^2(1, 32) = 0.508$, $p = 0.476$. However, the increase to 87.5 % for the N-by-N table was significant, $\chi^2(1, 32) = 5.236$, $p = 0.022$.

Error Analysis. For the Spinner problem, the most common error was to calculate the probability of obtaining one specific color twice, instead of *any* color twice. Specifically, this error involved incomplete solution of the problem: finding the probability for the two spinners to both land on a particular color to be (1/3)·(1/3) = 1/9, and stopping there. However, the correct solution should be three times this probability (=1/3), because there are three colors on each spinner, and hence three same-color outcomes. In other words, solvers making this error failed to solve the problem completely due to failure to integrate intermediate results and all possible outcomes. This type of error and other types of errors were most likely to occur with no diagram, or with the tree. On the other hand, the N-by-N table, by displaying all possible outcomes for the independent events problem in a visually efficient way, facilitated coordinating the multiple substages of the problem solution, and thus improved the success rates.

Discussion. For the Spinner problem, the N-by-N table was the most effective representation, improving procedural correctness over the no-diagram condition. This is not surprising, since the table represents the N·N = 3·3 = 9 equally likely outcomes in a simple and direct way, even allowing space for labeling the 9 outcomes. Furthermore, the table representation naturally suggests the semantic aspects of the problem: that there are two different spinners that are of equal status or priority (corresponding to the rows and columns of the table) [cf. 15]. These findings support an explanation in terms of diagram congruency, that the more compatible the content and the structure of a diagram to that of the represented problem, the more likely it is to be facilitative for solving the problem.

The N-ary tree would also seem to offer advantages for this Spinner problem: it too displays the nine equally-likely outcomes with roughly equal salience, allowing space for convenient labeling. However, the tree's hierarchical structure suggests a sequential process, and here it is not explicitly stated whether the spinners are spun simultaneously or sequentially. Also, the tree does not emphasize the multiple target same-color outcomes for this problem to the same degree as the table; the table places these same-color outcomes on the main diagonal, where they are particularly prominent, and grouped after a fashion. In this way, the N-by-N table was able to provide extra external visual support for solvers to manage all the possible outcomes by chunking the essential information for computational efficiency [6, 28].

4 General Discussion

The present study explored whether and how different types of provided diagrams (trees or tables) influence problem solvers' choice of mathematical structures and solution methods for probability word problems. It was hypothesized that a diagram will cue use of a particular solution strategy to the extent that structural and connotative properties of the diagram match to structural properties of the mathematical solution procedure or the formal representation of the solution space [6, 13, 14, 29, 30].

Based on the structural differences of trees and tables, we specifically predicted that N-by-N tables would lead to more use of an outcome-search strategy, and trees to more use of a sequential-sampling strategy. This prediction was confirmed for both the problems studied: the tree diagram tended to elicit a sequential-sampling strategy, while the table increased use of the outcome-search strategy.

However, for the Work-Group problem there were confounds in our diagram manipulation: the diagram conditions not only differed in basic diagram structure, trees versus tables, they also differed in the abstraction levels that the diagram suggests for the outcome space. The diagrams contrasted for the Spinner problem (the N-ary tree and the N-by-N table) removed this confound. The results showed that the N-by-N table strongly cues the use of an outcome-based strategy. Note the N-ary tree has features that can cue either solution. The N-ary tree conveys a sequential schema, which maps to the mathematical relation of sequential definitions of simple or conditional probabilities; and it also offers the cue to use low-level and visualizable equally-likely outcomes. Indeed, solutions were evenly split between outcome-based and sequential solution strategies for this condition.

The present results support and extend previous findings on the cognitive effects that visual representations have on reasoning, inferences, and problem solving [e.g., 7, 13, 31]. In particular, this study builds on and extends previous research on the structural properties of tree and table diagrams [e.g., 15, 24], by assessing the effects of using trees and tables on probability problem solving behavior. In particular, our results suggest that performance is affected not only by diagram type (tree versus table), but also by the abstraction levels of the outcome space that they suggest. Most prior diagram studies have focused on problem solving success as the criterion variable. Importantly, the present study demonstrates that diagrams can also affect strategy choice and the types of characteristic errors made in mathematics problem solving. This finding has instructional implications, because it implies that while diagrams can be an important and powerful form of scaffolding, they should be used in a carefully targeted fashion, taking into account their specific effects on the solution process. Also, it suggests that student-generated diagrams might be useful diagnostic evidence for inferring student problem-solving strategies and solutions.

The present results also show that the degree of the structural compatibility between a diagram and the particular type of problem to be represented affects problem solving success. For a combinations problem, which can be thought of as sampling without replacement, the use of an N-by-N table actually decreased the proportion of correct solutions. Because the diagonal cells are depicted with equal status as all the other cells, the table structure can mislead solvers to perceive all types of combination outcomes as valid, including those self-repeated combinations that do not arise in sampling *without* replacement. On the other hand, for an independent events problem, which involves sampling with replacement, the N-by-N table significantly improved participants' solution correctness. For combinations of independent events, the table structure better conforms to the probability problem structure, cueing an appropriate solution.

Thus, the findings of the present research offer support for the Congruence Principle in diagram design, the idea that external visual representations must have analogous relational structures with the conceptual schemas and relations they are designed to represent in order to facilitate construction of appropriate mental representations and correct inferences [e.g., 6, 11, 27, 32–34]. Such congruence of structure also ensures that inferences or insights spurred by the diagrammatic representation can be accurately transferred back to the problem domain.

References

1. Hegarty, M., Kozhevnikov, M.: Types of visual–spatial representations and mathematical problem solving. J. Educ. Psychol. **91**(4), 684–689 (1999)
2. Heiser, J., Tversky, B.: Arrows in comprehending and producing mechanical diagrams. Cogn. Sci. **30**, 581–592 (2006)
3. Manalo, E., Uesaka, Y.: Quantity and quality of diagrams used in math word problem solving: a comparison between New Zealand and Japanese students. In: Paper Presented at the New Zealand Association for Research in Education (NZARE) National Conference, Rotorua, New Zealand, December 2006
4. Zahner, D., Corter, J.E.: The process of probability problem solving: use of external visual representations. Math. Thinking Learn. **12**(2), 177–204 (2010)
5. Novick, L.R., Catley, K.M.: Reasoning about evolution's grand patterns: college students' understanding of the tree of life. Am. Educ. Res. J. **50**, 138–177 (2013)
6. Gattis, M., Holyoak, K.J.: Mapping conceptual to spatial relations in visual reasoning. J. Exp. Psychol. Learn. Mem. Cogn. **22**(1), 231–239 (1996)
7. Simkin, D., Hastie, R.: An information-processing analysis of graph perception. J. Am. Stat. Assoc. **82**(398), 454–465 (1987)
8. Tversky, B., Corter, J.E., Gao, J., Tanaka, Y., Nickerson, J.: People, place, and time: inferences from diagrams. In: Proceedings of the 35th Annual Conference of the Cognitive Science Society, pp. 3593–3597. Cognitive Science Society, Austin (2013)
9. Gick, M.L., Holyoak, K.J.: Schema induction and analogical transfer. Cogn. Psychol. **15**, 1–38 (1983)
10. Novick, L.R.: Representational transfer in problem solving. Am. Psychol. Soc. **1**(2), 128–132 (1990)
11. Mason, D.L., Corter, J.E., Tversky, B., Nickerson, J.V.: Structure, space and time: some ways that diagrams affect inferences in a planning task. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 277–290. Springer, Heidelberg (2012)
12. Braithwaite, D.W., Goldstone, R.L.: Flexibility in data interpretation: effects of representational format. Front. Psychol. **4**, 1–16 (2013)
13. Zacks, J., Tversky, B.: Bars and lines: a study of graphic communication. Mem. Cogn. **27**(6), 1073–1079 (1999)
14. Markman, A.B.: Knowledge Representation. Lawrence Erlbaum Associates, Mahwah (1999)
15. Novick, L.R., Hurley, S.M.: To matrix, network, or hierarchy: that is the question. Cogn. Psychol. **42**(2), 158–216 (2001)
16. Gattis, M.: Mapping relational structure in spatial reasoning. Cogn. Sci. **28**, 589–610 (2004)
17. Tversky, B., Kugelmass, S., Winter, A.: Cross-cultural and developmental trends in graphic productions. Cogn. Psychol. **23**(4), 515–557 (1991)

18. Shah, P., Mayer, R.E., Hegarty, M.: Graphs as aids to knowledge construction: Signaling techniques for guiding the process of graph comprehension. J. Educ. Psychol. **91**(4), 680–702 (1999)
19. Tversky, B., Zacks, J., Lee, P., Heiser, J.: Lines, blobs, crosses and arrows: diagrammatic communication with schematic figures. In: Anderson, M., Cheng, P., Haarslev, V. (eds.) Diagrams 2000. LNCS (LNAI), vol. 1889, pp. 221–230. Springer, Heidelberg (2000)
20. Russell, W.E.: The use of visual devices in probability problem solving. Doctoral dissertation, Columbia University, Dissertation Abstracts International, 61, 1333 (2000)
21. Corter, J.E., Zahner, D.C.: Use of external visual representations in probability problem solving. Stat. Educ. Res. J. **6**(1), 22–50 (2007)
22. Bobek, E.J., Corter, J.E.: Effects of problem difficulty and student expertise on the utility of provided diagrams in probability problem solving. In: Ohlsson, S., Catrambone, R. (eds.) Proceedings of the 32nd Annual Conference of the Cognitive Science Society, pp. 2650–2655. Cognitive Science Society, Austin (2010)
23. Ainsworth, S.: The functions of multiple representations. Comput. Educ. **33**(2), 131–152 (1999)
24. Novick, L.R., Hurley, S.M., Francis, M.: Evidence for abstract, schematic knowledge of three spatial diagram representations. Mem. Cogn. **27**(2), 288–308 (1999)
25. Novick, L.R., Hmelo, C.E.: Transferring symbolic representations across nonisomorphic problems. J. Exp. Psychol. Learn. Mem. Cogn. **20**(6), 1296–1321 (1994)
26. Gugga, S.S., Corter, J.E.: Effects of temporal and causal schemas on probability problem solving. In: Bello, P., Guarini, M., Scassellati, B. (eds.) Proceedings of the 36th Annual Conference of the Cognitive Science Society, pp. 2650–2655. Cognitive Science Society, Austin (2014)
27. Tversky, B., Morrison, J.B., Betrancourt, M.: Animation: can it facilitate? Int. J. Hum. Comput. Stud. **57**(4), 247–262 (2002)
28. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. Cogn. Sci. **11**, 65–99 (1987)
29. Gentner, D., Markman, A.B.: Structure mapping in analogy and similarity. Am. Psychol. **52** (1), 45–56 (1997)
30. Tversky, B.: Visualizations of thought. Top. Cogn. Sci. **3**, 499–535 (2011)
31. Tversky, B., Corter, J.E., Yu, L., Mason, D.L., Nickerson, J.V.: Representing category and continuum: visualizing thought. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 23–34. Springer, Heidelberg (2012)
32. Bauer, M.I., Johnson-Laird, P.N.: How diagrams can improve reasoning. Psychol. Sci. **6**, 372–378 (1993)
33. Glenberg, A.M., Langston, W.E.: Comprehension of illustrated text: pictures help to build mental models. J. Mem. Lang. **31**, 129–151 (1992)
34. Pinker, S.: A theory of graph comprehension. In: Freedle, R. (ed.) Artificial Intelligence and the Future of Testing, pp. 73–126. Erlbaum, Hillsdale (1990)

What Constitutes an Effective Representation?

Peter C-H. Cheng(✉)

Department of Informatics, University of Sussex, Brighton, UK
p.c.h.cheng@sussex.ac.uk

Abstract. This paper presents a taxonomy of 19 cognitive criteria for judging what constitutes effective representational systems, particularly for knowledge rich topics. Two classes of cognitive criteria are discussed. The first concerns access to concepts by reading and making inferences from external representations. The second class addresses the generation and manipulation of external representations to fulfill reasoning or problem solving goals. Suggestions for the use of the classification are made. Examples of conventional representations and Law Encoding Diagrams for the conceptual challenging topic of particle collisions are provided throughout.

Keywords: External representation · Mental representation · Cognition · Knowledge domains · Law encoding diagrams · Algebraic expressions · Tables · Particle collisions · Problem solving

1 Introduction

What constitutes an effective representation? Here *representations* include abstract (non-figurative) encodings and presentations of information, such as tables, formal notations, maps, diagrams and interfaces to computers. The title question is important because the design of representations may have a dramatic impact on cognitive processes at different times scales – from perception on the order seconds, to problem solving over minutes, learning lasting hours and days, and discovery taking years. For example, isomorphic representations of the Tower of Hanoi problem can increase problem solution times by up to 16 times [20]. An empirical study [7] on the mechanics problem from Larkin and Simon's [21] seminal paper showed a six-fold benefit for diagrams over sentential representations. A computational study [12] on the topic of particle collisions showed how diagrams (such as that in Fig. 1A, below) might have been instrumental to the discovery of certain conservation laws in physics. So, this paper addresses the title question from a cognitive perspective, with a particular focus on representations for knowledge rich topics.

The question is challenging in cognitive terms. A cognitive answer must integrate: (a) considerations of the nature of external representations (ER); (b) considerations of the nature of the internal mental representations (IR); (c) investigate the rich and complex relations between the two – how ERs and IRs work together to encode knowledge. ERs may in themselves be complex [15, 17]. IRs are also complex [22] and must be examined in relation to the information processing capabilities of the human cognitive architecture [23], including visual perception, mental imagery, propositional

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 17–31, 2016.
DOI: 10.1007/978-3-319-42333-3_2

(verbal/logical) reasoning and spatial reasoning, which involve memory encoding and retrieval processes at many levels [28].

To be clear on terminology: an ER is a particular physically rendered instance of a representation in the external environment; an IR comprises the information associated with the representation in the internal mental environment. Here, *representation* will refer to the combination of the ER and IR, and the term *representational system*, RS, will be used when this needs to be explicit. An RS is a *representing world* that encodes knowledge about the *represented world* of the target concepts and ideas with which the user of a RS is engaged.

Many answers to the title question have been obtained from specific perspectives using a myriad of approaches, including: task analysis (e.g., [6, 8]); computational models (e.g., [12, 24]); empirical studies (e.g., [6, 8, 14, 29]); eye-movement studies (e.g., [24]); theoretical analyses (e.g., [26, 27, 16]). These studies span all levels of cognition from perception, reasoning, problem solving (e.g., [21, 24, 30]), learning (e.g., [5, 6]), and discovery (e.g., [12]). Thus, a single coherent answer to the title question does not appear feasible or even seem appropriate. Green's *Cognitive Dimensions* [16] is a particularly extensive set of heuristic "tools" for identifying poor notations and interfaces. So, this paper aims to collate the previously identified characteristics, or criteria, to propose additional criteria, and to present them in a cognitively motivated classification. The classification emphasizes (a) general classes of representationally related cognition and (b) many levels of cognitive processing. Regarding the first of these, the classification identifies two major classes of criteria. (1) How readily a RS provides *access* to concepts – what in the relation between an ER and IR makes reading and interpretation easy? (2) The *generativity* of a RS concerns the ease of producing and manipulating an ER to achieve task goals – what about the nature of RSs can make the transformation of ERs easier? Each class is present in a section below, but first sample RSs for a knowledge rich topic will be introduced as a source of illustrative examples.

2 Sample Topic and Representations: Particle Collisions

The classification is motivated by, and draws upon, the author's experience in the design and evaluation of *Law Encoding Diagrams*, LEDs, for educational domains [4-60] and to serve as graphical computer interfaces for complex problem solving [10, 120]. A LED is a special RS because it directly encodes the conceptual structure of a topic in the graphical format of its ERs using geometric, spatial and topological rules, such that each instantiation of an ER represents one state-of-affairs in the topic. Thus, LEDs provide useful theoretical leverage to study representational issues, because they combine characteristics of abstract general notations (c.f., formulas) and concrete particular displays of data (c.f., line graphs). The topic of particle collisions will provide a thoroughgoing set of examples. This will include, tables and algebraic equations, which are the conventional representations for this topic in physics texts, which will be compared with a LED.

A basic 1D head-on elastic collision between two bodies, body-1 and body-2, which have masses m_1 and m_2, may be characterized by the velocities before impact,

Table 1. Data and derived quantities for particle collisions (2F is not a valid case).

Case	m_1	m_2	U_1	U_2	V_1	V_2	M_{pre}	M_{post}	E_{pre}	E_{post}
1	5	3	2	−2	−1	3	4	4	16	16
2A	1	1	1	−1	−1	1	0	0	1	1
2B	1	1	1	0	0	1	1	1	0.5	0.5
2C	1	1	3	1	1	3	4	4	5	5
2D	2	0.1	0.1	−2	−0.1	2	0	0	0.21	0.21
2E	1.9	0.1	1	−1	0.8	2.8	1.8	1.8	1	1
2F	5	3	2	2	−1	3	16	4	16	16
5C	5	3	1	−3	−2	2	−4	−4	16	16
6A	5	3	2	−2	−0.5	2.167	4	4	16	7.667
6B	5	3	2	−2	0.5	0.5	4	4	16	1

U_1 and U_2, and the velocities after impact, V_1 and V_2. (Units of kg and m/s may, respectively, be assumed.) Table 1 shows a selection of cases; in each row it assigns values across the variables. Further, it displays derived quantities of momentum, M, and energy, E, that were computed elsewhere. In valid cases momentum is conserved and so it is equal before and after impact: $M_{pre} = M_{post}$. For elastic collisions (1-2E, 5C), energy, E, is also conserved, $E_{pre} = E_{post}$, but for inelastic cases some is lost in the collision: $E_{pre} > E_{post}$ (6A/B).

Expressing a case in a purely algebraic representation requires six equations: e.g.,

$$m_1 = 5, \quad m_2 = 3, \quad U_1 = 2, \quad U_2 = -2, \quad V_1 = -1, \quad V_2 = -3 \tag{1}$$

Physics texts typically analyze 1D elastic collisions in terms of the momentum and energy conservation laws, respectively:

$$m_1U_1 + m_2U_2 = m_1V_1 + m_2V_2 \tag{2}$$

$$\frac{1}{2}m_1U_1^2 + \frac{1}{2}m_1U_2^2 = \frac{1}{2}m_1V_1^2 + \frac{1}{2}m_1V_2^2. \tag{3}$$

With some algebraic manipulation of Eqs. 2 and 3 it is possible to eliminate the mass terms, to obtain the "velocity difference" equation:

$$U_1 - U_2 = V_2 - V_1. \tag{4}$$

To model elastic collisions an energy loss parameter or a coefficient of restitution are introduced as multiplicative factors to one side of Eq. 2 or 4, respectively.

A typical textbook problem is to compute values of V_1 and V_2 given the other variables. This requires many algebraic manipulation steps, the simultaneous solution of Eqs. (2) and (3), the application of the standard formula for quadratic equations, and the substitution of values from (1) into the resulting solution formulas.

Figure 1A shows one example of the diagrams that Huygens and Wren presented to the Royal Society of London in 1669 as models of 1D elastic collusions. It is a LED.

The diagram has been redrawn in Fig. 1B, with arrows for the velocities and line (segments) for the masses. In previous work, LEDs like these have been shown to enhance the learning [2, 3] and have been deployed in a computer-based discovery learning environment [4]. Figure 1C shows how the LEDs will be drawn here: they will be called *H&W diagrams*. This particular format allows the LEDs to be extended to cover sequences of collisions, inelastic impacts and 2D collisions (see below). Figure 1 shows the same collision as case 1 in Table 1 and Fig. 2 shows other examples that correspond to the cases in Table 1 with the same numbers.

Simple syntactic rules based on the relation of the arrows and lines to the central vertical *origin* and the parallelogram determine the structure of H&W diagrams. Most of the semantic rules for interpreting the diagrams are obvious but one should note that the mass lines are on the opposite side to their respective velocity arrows. Also, the slope of the parallelogram represents the overall momentum of the system. In Fig. 2A the overall momentum is zero, but increasing the mass of body-1 or decreasing the speed of body-2 will positively increase the overall momentum, Fig. 2E and B, respectively.

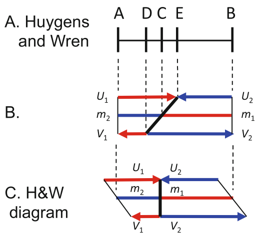


Fig. 1. Huygens, Wren and H&W Diagrams.

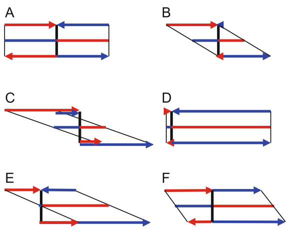


Fig. 2. H&W diagrams (F is invalid).

H&W diagrams can be derived from Eqs. 2 and 3 (and vice versa). For instance, Eq. 4 encodes the idea that the parallelogram has a constant width.

Figure 3 shows how the typical textbook problem mentioned above is solved. To find the final velocities, arrows for the initial velocities are first drawn to some chosen scale (Fig. 3A). Lines for the masses are drawn end to end to an arbitrary scale (B) and this line is rescaled to match the initial velocities (C). They are then aligned at the origin (D) and the parallelogram is completed with the final velocities (E), allowing their values to be read-off to scale. Other combinations of given variables may require some iteration of the diagrams. For example, given one initial and one final velocity (Fig. 3F) one must produce a parallelogram (I) by finding the correct length of the mass lines that is not too small (G) nor too large (H).

H&W diagrams may be composed to model sequences of collisions, such as two pairs of balls approaching at different speeds in a Newton's cradle, Fig. 4 – the middle balls rebound (row 1), collide with outer balls and head back to the centre (row 2), where they again rebound (row 3). Moving frames of reference is a core notion in physics that H&W diagrams usefully visualize. Figure 5A is a given collision, Fig. 5B gives the relative motion of an observer (say, on a train), and Fig. 5C shows what the observe sees (through the window). Although the same constant velocity (green arrow) has been deducted from all the velocities, it is clear that the H&W diagram is valid and

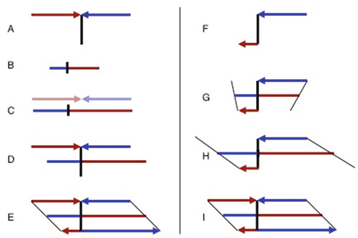


Fig. 3. Quantitative problem solution.

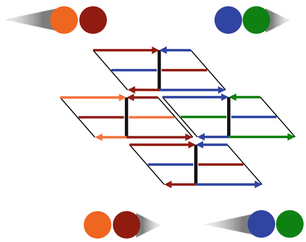


Fig. 4. Modelling sequence of impacts

would be so for any observer's velocity. Thus, the conservation laws are the same for all observers in uniform motion.

Figure 6 shows how H&W diagrams can be extended to model in-elastic collisions using the fact that the diagonal of the parallelogram represents the overall momentum of the system. The thick diagonal line bisecting the origin, o, runs parallel to the sides of the parallelogram. In the extreme case when the maximum amount of energy is lost, the bodies coalesce and c gives the velocities of the bodies after collision, p" and q" at B in Fig. 6. Between that extreme, c, and the fully elastic case, p and q, the overall momentum remains constant, thus any change to the momentum of one body must be compensated by the other, so the position of the arrow heads p' and q' from c must be in the same proportion as p and q are from c: i.e., p'c:q'c::pc:qc – e.g., A in Fig. 6. At c these ratios are both zero. Can all the energy of the system be lost? The general form of H&W diagrams reveals this can only occur when the overall momentum is zero, when the diagram is a rectangle and both final velocities tend to zero.

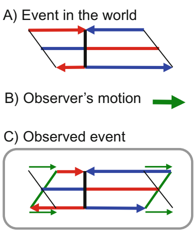


Fig. 5. Moving frame of reference.

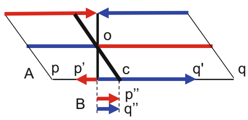


Fig. 6. Inelastic collision

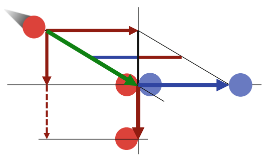


Fig. 7. 2D collisions

The modelling of two-dimensional impacts is illustrated in Fig. 7. When a moving ball, left (red), hits a stationary ball, right (blue), the departure angle between the balls is always 90°. Why? The diagram's orientation has been chosen so that the head-on and sliding components of the impact are horizontal and vertical, respectively. The sliding contact means the vertical motion is unchanged. The horizontal motion is simply modelled by Fig. 2B with all the motion of the first ball being transferred to the second, so after impact each ball has a motion just associated with one component of the initial motion, which are perpendicular by definition. To model 2D cases where both bodies are moving, the H&W diagram for a moving frame of reference, Fig. 5, can be used to decompose the situation into one similar to Fig. 7 and some uniform motion for the whole system.

The range of examples reveals H&W diagram's ability to model many types of collision situation and to support reasoning about important concepts of the topic. The contrast between H&W diagrams and the conventional representations will illustrate the effectiveness criteria in the following sections.

3 Access to Concepts: From ER to IR

The first class of effectiveness criteria concerns how readily concepts can be accessed in the IR from a given ER by the reading and interpreting the ER, without changes to its written or drawn form. Ready access to concepts is critical to the comprehensibility, memorability and learnability of a topic's content. Access is good when the cognitive demands, or work needed, to read and interpret information encoded in an ER is low. Further, with easy access related information will be readily retrieved from memory as the ER may provide rich cues for recall. An effective RS enables recognition of concepts and facilitates their interpretation. Poor access has negative consequences spanning all cognitive levels. It may reduce the rate at which meaningful cases that can be considered, increase the likelihood of interpretation errors and may hamper the spotting of errors when made. In learning contexts, poor access will increase the signal to noise ratio of positive learning episodes to negative ones [6]. Access will be considered in three sub-classes.

Elementary Encoding. This group of criteria considers how particular ways to encode concepts in ERs may affect the access of the concept in the IR.

A1.1. One Token for Each Type. Consider the elementary concepts of a topic, including properties, variables, entities and values. Access will be better when there is a one-to-one match between an elementary concept, or type, and a single symbol, or token, in the ER. Such mappings make the least cognitive demands because they avoid the work associated with managing complex associations between symbols and concepts, such as the need to exhaustively search for all occurrences of a symbol in the ER for a given variable [21]. In H&W diagrams, one graphical property encodes each type of elementary concept, but Eqs. 2 and 3 include two occurrences of letters for each velocity, two letters for each mass, and eight subscript numbers to denote the bodies. The original Huygens and Wren diagrams are poor, because many variables are mapped to different sections of one line (Fig. 1A).

A1.2. Reflects Structure of Concepts. Beyond elementary concepts, similar reasoning applies to the claim that the structure of expressions should reflect the topic's conceptual structure [29]; hierarchically related concepts should be encoded by hierarchically organized representations [1] and more generally they should be isomorphic [17]. Equations 1 and 2 clearly show how momentum and energy terms are sums of products of the variables. In general, however, equations tend to hide conceptual structures [6, 80]. (See [9] for an alternative RS for algebra that has one ER symbol for each variable and that shows the hierarchal relations among variables graphically.) Finding the conservation laws from data in Table 1 is a challenging inductive discovery problem [12]. The shape of H&W diagrams supports reasoning about momentum, but inferences about energy requires inferences about relative lengths of the mass lines and velocity arrows, without explicit support in the ER.

A1.3. Directly Depicts Structure of Cases. In additional to a topic's conceptual structure, it is desirable that ERs for individual cases reflect the concrete organization or physical structure of each case. H&W diagrams clearly do this, but the algebraic representation tends to hide such structure; for example, they do not explicitly encode the fact that the spatial ordering of the bodies is fixed.

In terms of Green's Cognitive Dimensions framework [16], criteria A1.2 and A1.3 are aspects of *closeness of mapping* and *consistency*.

A1.4. Exploits Spatial Indexing. Spatial indexing of information, rather than symbolic encoding, can make accessing concepts easier, by facilitating searches for information and the recognition of operators [21]. In H&W diagrams conceptually related information is often spatially co-located, and tables exploits spatial coordination in their columns and rows, but equations' linear concatenation of symbols tends to separate pieces of information that need to be related.

A1.5. Iconic Expressions. Expressions for important concepts should be iconic; they should consist of distinctive shapes or patterns that are clearly recognizable and particularly memorable. H&W diagrams are iconic at several levels: each pair of arrows forms a distinctive pattern, that are combined as unique parallelogram configurations (e.g., Figs. 1C, 2A–E), and in turn assemblies of H&W diagrams may themselves be iconic (e.g., Fig. 4). Whether a pattern is iconic depends on the user's level of experience with the RS and in some domains expertise is based on the acquisition of large number of perceptual chunks [28]. Scientists and engineers can instantly recognize that terms in Eq. 2 represents quantities of kinetic energy, but novices may perceive the expressions as arbitrary concatenations of symbols. Iconic expressions is one aspect of *visibility* in the Cognitive Dimensions framework [16].

Reading and Inference Operations. This group of criteria considers transfer of information in the ER to the IR and mentally transforming expressions of the IR.

A2.1. Prefer Low Cost Forms of Processing. Simply, ERs that invoke IRs and processes that have low cognitive demands will facilitate access to concepts. The Cognitive Dimensions framework [16] recommends avoiding *hard mental operators*, in general. More specifically, perceptual operators are easier than using visual imagery, and visual imagery is less demanding than verbal logical reasoning, in gross terms. For example, many important concepts can be accessed rapidly by visual inspection of H&W diagrams, but it is harder to imagine changing the shape of a H&W diagram in one's mind's eye (e.g., given Fig. 2A imagine Fig. 2C). It is harder still to use Eqs. 2 and 3 to mentally reason propositionally about the impact of changing some variables with others held constant. *Computational off-loading* [25] may be interpreted as the potential of some ERs to allow perceptual inferences to be substituted for purely mental forms of reasoning.

A2.2. Prefer Low Cost Operators. For a particular form of processing (whether perceptual, imagistic or verbal/propositional) some types of operator will be less demanding than others. For example, Cleveland and McGill [14] empirically established an order of effectiveness for simple perceptual operators used to judge quantities. In mental imagery, translation and rotation operations are likely to be easier than composing irregular shapes [18]. Verbal reasoning about chance situations is superior when probabilities are interpreted as frequencies rather than as decimal numbers [13].

A2.3. Invoke More Structured IRs. Cognitive scientists explain human ability on complex information processing tasks using a variety of types of IR, including associative semantic networks, condition-action production rules, semantic networks with inheritance, and schemas (or frames) [22, 28]. Cognitive benefits naturally arise from the use of IRs that are more systematic, arguably in the order just given, because more precise and rich indexing of information will aid contextualization and access to concepts. Thus, RSs that recruit IRs with good structure appear preferable. For example, schemas are IRs that possess particular *slots* which may be *filled* by certain types of information. This establishes specific relations among the pieces of information. Interpreted tables may be comprised of generic schemas that coordinate values in the columns and rows, but provide less in the way of topic specific relations. An equation may invoke a schema with slots for the left and right sides of an equation and that encodes the concept that they are equal. H&W diagrams may encourage users to develop a particularly effective schema – see next criteria.

A2.4. Support for Diagrammatic Configuration Schemas. Experts' proficiency in certain forms of problem solving may be attributable to their organization of information as a special form of schemas, *diagrammatic configuration schemas*, DCS, [19]. A DCS uses a diagram of a specific situation to coordinate what inferences can be made about the situation under given sets of constraints. This rich encoding allows experts to efficiently solve problems by rapidly planning effective sequences of operations, by decomposing the ER into characteristic patterns associated with DCSs and using the constraints to identify feasible inferences. Users familiar with H&W diagrams may possess DCS as IRs, because the rules governing the diagrams can be encoded as inference and applicability conditions. Such an encoding is unlikely for the algebraic representation, because the algebraic inference rules are diverse and generic and therefore not tightly and specifically associated with Eqs. 2 and 3.

Conceptual Transparency. This class of accessibility criteria considers the design of RSs when a full characterization of the conceptual structure of the topic is available. It differs from those above (esp. A1.2) by embracing the complexity and depth of ideas in knowledge rich topics. The *conceptual transparency* criteria (elsewhere called *semantic transparency* criteria [6, 8, 10]) consider how to make the full richness and range of important and distinctive concepts of a topic directly accessible as simple patterns in ERs. Such concepts include: the primary symmetries, invariants, laws and major regularities of the topic; alternative conceptions or ontological perspectives, such as taxonomic, causal processes and formal constraints; types of cases, including prototypical, special, extreme and limiting cases; valid versus invalid relations and cases [10]. Importantly, when diverse concepts are readily accessible simultaneously, they can provide mutual supportive contexts for each other's interpretations [10]. So, the challenging demand of conceptual transparency is to use what is known about the nature of a topic's content to encode it in a manner that allows it to be easily accessed and interrelated. The following criteria promote such encodings.

A3.1. A Format for Each Class of Primary Concepts. For conceptual transparency, a RS should provide a distinctive *representational format* for each of the *primary classes of concepts* of the topic [10]. A representational format is a particular type of graphical or notational scheme for encoding information, such as a spatial coordinate system, a set of visual properties, or formal rules applied to concatenations of symbols. Important types of concepts include: (a) properties and their values; taxonomic relations; (b) structural concepts; (c) temporal concepts; (d) behavioural concepts; (e) functional concepts; (f) formal relations (logical, mathematical). A topic might not include all of these classes. H&W diagrams has largely separate representational formats for each primary class of concepts: (a) velocity and mass and their values are represented by types and sizes of lines; (b) the structure of collisions is represented by the topology of the arrows (their relative left-right placement); (c) time is represented by relative vertical position; (d) collision behaviour is represented by the configuration of the arrows; (f) formal relations are encoded by the geometric rules of the diagram. The algebraic representation does not satisfy this criterion well, as types of alphanumeric symbols and algebraic relations span several classes of concepts.

A3.2. Coherent Encoding of Primary Concepts in a Format. For conceptual transparency, the format for each primary class of concepts should simultaneously differentiate and integrate the concepts in the class, so that the concepts can be readily distinguished but also to provide mutual contexts for each other's interpretation [10]. For example, all properties/quantities in H&W diagrams are line segments, but scalars are plain lines and vector are arrows, with the orientation of the arrows giving directions of motion. Equivalent information in equations is distributed across conventions on alphanumeric symbols and the assignment of numerical values to variables.

A3.3. Provide an Overarching Interpretive Scheme. For conceptual transparency a RS should have an overarching interpretive scheme to coherently combine the formats of the primary class of concepts [10]. The arrangement of the origin and parallelogram in H&W diagrams constitutes such an overarching interpretive scheme, whereby different properties, quantities, structure, times, behaviors, functions and formal relations are well integrated. Concatenation of symbols under algebraic rules provides little in the way of a topic-specific overarching interpretive scheme.

H&W diagrams largely satisfy A3.1–3 so they possess greater conceptual transparency than the equations. Both show the spatial and temporal symmetry of the collisions. The form of the equations is invariant across the identity of variables (body-1/body-2) and order of terms (pre/post collision). Valid H&W diagrams are produced when the whole diagram is reflected about the origin, or reflected vertically with the directions of the arrows reversed. However, the H&W diagrams simply encodes the notion of moving frames of reference (Fig. 5), but it is demanding problem to show that adding a constant to U_1 and U_2 in Eqs. 2 and 3 necessarily changes both V_1 and V_2 by the same amount. Prototypical (Fig. 1A/C), special (Fig. 1B/D) and limiting (Fig D/E) cases are distinctive H&W diagrams, but such cases are less apparent in Table 1. Invalid H&W diagrams often standout and what is wrong is often obvious (e.g., Figure 1F), but without the momentum values this case is not obvious in Table 1 nor is the source of the problem (a missing minus sign).

Representations with conceptual transparency may suffer less from the problem *diffuseness* as identified in the Cognitive Dimensions framework [16]. Shimojima [26] identifies three semantic properties of diagrams that appear to promote their access to information in the ER, specifically the potential for *free rides*, *consistency-checking* and *derivative meanings*. These three potentially beneficial properties may be interpreted in terms of conceptual transparency. So, criteria A3.1–3 may provide a means by which to design representations possessing the properties, for topics that are more knowledge rich than the cases examined by Shimojima [26].

4 Generating ERs

This part of the classification concerns the production of ERs through their modification or generation from scratch in order to revise or obtain new concepts. Given a new set of data one might add a row to Table 1 or draw a H&W Diagram; or to solve a problem one might write a new equation derived from Eqs. 1 and 2, or sketch a sequence of H&W diagrams. The ease of producing ERs will substantially impact the effectiveness of RS at multiple cognitive levels. Reasoning, problem solving, learning and discovery may be enhanced when generating ERs requires little effort and can be done so reliably. A RS in which it is complicated to do things, and in which great care is needed to avoid errors, is undesirable. Obviously, when a new ER has been generated the concepts contained within it are accessed, so the processes of generating and accessing ERs are symbiotic. The term *operation* is used for elementary manipulations of an ER and *procedure* denotes sequences of operations to achieve goals of ER transformation tasks. Unfortunately, it appears there is little prior work on the effective generation of ERs, per se. Two classes of criteria are considered.

Syntactic Plasticity. A RS is a medium for modelling ideas in a topic, much like materials are use to make physical models. A plastic material (e.g., clay) is good for creating a sculpture as it can be readily molded: it is not too brittle like chalk nor so fluid it flows in an unconstraint fashion like syrup. By analogy a RS should be syntactically plastic, with desired ERs being easy to produce, guided by the syntactic rules of the RS [10]. Producing ERs can, more formally, be treated as a form of problem solving and Newell and Simon's classical theory of problem solving applied [23]. The target ER is the *goal state* that is to be reached from an *initial state* of some given ER (or none) by the *search* of the *space* of possible partial ERs that can be generated using the RS's syntactic *operators*. *Tests* are applied at each production step to see if the goal has been reached. The search process may be conceptualized as a tree, the trunk being the initial state and leaves at the end of branches being completed ERs, one (or more) of which might be the desired goal. *Search heuristics* [23] guide navigation through the tree (*problem state space*). The following criteria consider the effectiveness of RSs in terms of the character of their problem states spaces, the demands of searching the tree.

G1.1. Simple Operations. A RS should possess simple operators that are easy to execute and that involve small amounts of cognitive effort. Drawing most components of a H&W diagram simply involves producing lines to scale, but in some cases a succession of sketches is needed to find the right line proportions (Fig. 3). Moving a whole term from one side of an equation to another is a simple operation (e.g., m_2u_2 to the right of Eq. 2), but many algebraic operations are more demanding, such as factoring a quadratic equation.

G1.2. Limited Types of Operators. A RS that has a small *set* of operators will tend to have a problem state space with a lower branching factor at each node: the tree will be narrower overall and thus tend to be simpler to search in general. Few operator types means fewer options to be consider at each inference step, which reduces the likelihood of selecting unproductive operators. The possible drawing operations for H&W diagrams is highly constrained, whereas a myriad of algebraic manipulations may be applied to a formula.

G1.1 and G1.2 and can be applied individually when all else is equal. Typically, however, comparisons between RSs will likely consider the trade-offs between them.

G1.3. Short Procedures. A RS with short procedures requires fewer executions of operators to complete each procedure. The problem state space will be shallower overall, so potential solution states are reached more quickly. Short procedures present less opportunity for error and are obviously less effortful to execute. For example, checking whether case 1 in Table 1 is valid given the masses and velocities requires few steps using H&W diagram (Fig. 3) but requires the substitution of all the values in Eq. 1 into Eq. 2 and a series of computations, and the same with Eq. 3. The full solution procedure for Eqs. 1 and 2 was outlined above. Modelling a series of collisions with H&W diagrams involves the simple composition of whole diagrams (Fig. 4) but may demand processing multiple simultaneous equations under the algebraic approach.

G1.4. Uniform Procedures. A representation should have similar procedures to handle most problems, so the chances of picking unfruitful strategies are lessened and so that few strategies and heuristics need to be learnt. If the shape of problem trees are limited to a small number of forms, the cost of choosing one, and the chances of picking an inappropriate one, are reduced. Solving problems with H&W diagrams involves variations on the construction of the diagram and complex situations may be resolved using moving frames of reference (Fig. 5) to decompose cases into iconic diagrams like those in Fig. 2.

A RS lacking syntactic plasticity may be considered to be *viscous* in the terms of the Cognitive Dimensions framework [16] and is likely to be *error prone*, to have *hidden dependencies* and involve *premature commitments*.

Conceptual-Syntactic Compatibility. Conceptual transparency and syntactic plasticity may complement each other to increase the effectiveness of a RS.

G2.1. Meaningful Syntactic Constraints. Generating ERs may be more effective in RSs that have conceptual transparency, because valid manipulations of the ER will likely correspond to meaningful variations of states of affairs in the topic. Any syntactically valid change to a H&W diagram yields a real collision, whereas valid algebraic manipulations often produce expressions whose interpretation are obscure relative to the topic content. When the syntax and encoding of concepts coincide in this way the actual problem context may directly support the selection of appropriate procedures to achieve task sub-goals. It may also highlight incorrect syntactic operations or invalid expressions, because the actual problem situation can be used as a test that a partial solution is sensible. This is like *progressive evaluation* in the Cognitive Dimensions framework [16].

G2.2. ER Construction Parallels Topic Processes. Extending the previous criteria, a RS will be more effective when processes to construct ERs coincide with the natural processes of the topic, and not just meaningful states of affair. For example, the assembly of H&W diagrams into larger diagrams for sequences of collisions mirrors the occurrence of impacts in such situations (e.g., Fig. 4). Writing equations to assign values to variables and writing multiple versions of Eqs. 1 and 2 for those variables does not directly reflect the impact sequence. Again, benefits accrue in relation to the easier selection and application of procedures.

G2.3. Separation of Modeling, Interpretation and Calculation. A RS should permit the separation of *situation modelling*, *interpretation* and *calculation* into distinct phases of problem solving [6]. Situation modeling involves the construction of an ER that interrelates the given information about the problem, including relevant laws: a H&W diagram is such a model. Interpretation identifies the target configurations in the ER associated with problem goal; for instance identifying a certain pattern of lines. Calculation finds the desired relation or computes the required quantity from the target configuration; for instance, the ratios of lines. The separation of these phases is has the benefit of disentangling considerations of what is known about the problem situation (modelling) from inferences needed to find a solution (interpretation). In contrast, the algebraic approaches typically involves a single phase of analysis prior to calculation, which depends on the selection of an appropriate solution strategy based on an abstract understanding of the nature of the problem prior to any solution activity. With no modelling phase, important information about the structure of the problem situation is not systematically examined so clues about appropriate strategies may be missed. For example, in one study, graduate physic students were asked to solve the textbook problem mentioned above [2]. All struggled to pick an appropriate solution strategy on the first attempt. When they eventually followed the typical strategy, some correctly derived the quadratic expression for velocities and applied the standard quadratic solution formula. However, from the two values obtained some picked a value that was one of the initial velocities without realizing so, which reveals they had little overall sense of the overall nature of the situation and problem they were attempting to solve.

5 Discussion

What constitutes an effective representation? Table 2 summarizes the 19 identified criteria, classified into two main classes and five sub-classes. The classification is more comprehensive than previous analyses taking a cognitive orientation. Whereas previous accounts have tended to focus largely on either (a) access to concepts in ERs [26] or (b) generation of ERs [16], the present classification covers both and also begins to

Table 2. The classification of characteristics of effective representations

A) Access to concepts: from ER to IR	A3.2. Coherent encoding of primary concepts
Elementary encoding	A3.3. Overarching interpretive scheme
A1.1. One token for each type	
A1.2. Reflects structure of concepts	**G) Generating ERs**
A1.3. Directly depicts structure of cases	Syntactic plasticity
A1.4. Exploits spatial indexing	G1.1. Simple operations
A1.5. Iconic expressions	G1.2. Limited types of operators
Reading and inference operations	G1.3. Short procedures
A2.1. Prefer low cost forms of processing	G1.4. Uniform procedures
A2.2. Prefer low cost operators.	Conceptual-syntactic compatibility
A2.3. Invoke more structured IRs	G2.1. Meaningful syntactic constraints
A2.4. Support for diagram config. schemas	G2.2. ER construction parallels topic process
Conceptual transparency	G2.3. Separation of modeling, interpretation and calculation
A3.1. A format for each class of 1° concepts	

consider how effective RSs obtain benefits when they are combined. The classification also reveals that factors that may positively impact the efficacy of a RS occur at many cognitive levels. No claim is made that the classification is complete. Nor is it claimed that the criteria are mutually exclusive in all respects, because some address related themes at different levels. At minimum Table 2 may serve as a checklist for those investigating or designing RSs. One may gain an overall sense of whether one RS is better than another, or identify the particular areas of strength and weakness of a RS. In this respect, this work follows Green's [16] approach with the Cognitive Dimensions framework.

How should one use the classification for RSs design? From the author's experience of designing RSs for knowledge rich topics (e.g., [6, 8, 9]) and graphical interfaces for complex problems (e.g., [10, 11]), the conceptual transparency criteria (A3.1–3) should be given priority, because the elementary encoding criteria (A1.1–5) and the reading and inference criteria (A2.1–4) tend also to be satisfied when one aims for conceptual transparency. Conceptual transparency focuses on the coherent encoding of conceptual structures in systematic representational formats at the level of individual classes of concepts and of the topic as a whole, which appears to yield representations that are simple and rational (e.g., [8, 9, 11]). Further, conceptually transparent RSs tend to satisfy the conceptual-syntactic compatibility criteria (G2.1–3) and thereby naturally meet many of the syntactic plasticity criteria (G1.1–4).

What constitutes a fair basis for applying the criteria to compare RSs? This is a fundamental issue that Larkin and Simon [21] recognized in their foundational paper on RSs. They asserted that two RSs must be *informationally equivalent* before one can compare their respective computational demands. Information inferable in one representation must also be inferable in the other. So, one basis for making comparisons between RSs across a range of tasks in *knowledge rich topics* is to ensure that they are *conceptually equivalent* [8]. This notion asserts that all the ideas that are required for a full range of tasks must be expressible in both RSs. Of course, when the coverage of concepts are not equivalent one could limit the comparison just to tasks that are within the scope of the RSs under consideration, but this a rather arbitrary approach that may introduce biases. Therefore, above the cognitive level considered here, the *conceptual coverage* of RSs with respect to target topics may be taken as another perspective for identifying a wider class of efficacy criteria. It is at this level that the greater generality of algebraic equations compared to H&W diagrams would be addressed.

What about relations between RSs? In real world circumstances RSs are not used in isolation. The table and equations are often found together and instruction with H&W diagrams is likely to refer to equations. This suggests that a further set of criteria is needed to address the effectiveness of coordinating information between RSs and the transformation of ERs in one RS into associated ERs in other RS.

To conclude, consider H&W diagrams one final time. Although Huygens and Wren's diagrams (Fig. 1A) have been considered elsewhere [2–4], their extension via H&W diagrams to model series of collisions, moving frames of references, inelastic collisions and 2D impact (Figs. 4, 5, 6 and 7) is novel. The contrast between H&W diagrams and the conventional representations is clear both in terms of access to concepts and also the generation of ERs. The previous research on RSs has largely focused on the former class of criteria but the contrast between the RSs in the examples here emphasizes the need to consider the processes of manipulating ERs and also the variety of diagrammatic operators that may be used to transform ERs. Further work is needed to classify and understand the full potential of such diagrammatic operators.

References

1. Barwise, J., Etchemendy, J.: Heterogeneous logic. In: Glasgow, J., Narayanan, N.H., Chandrasekaran, B. (eds.) Diagrammatic Reasoning: Cognitive and Computational Perspectives, pp. 211–234. AAAI Press, Cambridge (1995)
2. Cheng, P.C.-H.: Problem solving and learning with diagrammatic representations. In: Peterson, D. (ed.) Forms of Representation, pp. 47–66. Intellect Books, Bristol (1996)
3. Cheng, P.C.-H.: Learning qualitative relations in physics with Law Encoding Diagrams. In: Cottrell, G.W. (ed.) Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society, pp. 512–517. Lawrence Erlbaum (1996)
4. Cheng, P.C.-H.: Interactive law encoding diagrams for learning and instruction. Learn. Instr. **9**(4), 309–326 (1999)
5. Cheng, P.C.-H.: Unlocking conceptual learning in mathematics and science with effective representational systems. Comput. Educ. **33**(2–3), 109–130 (1999)
6. Cheng, P.C.-H.: Electrifying diagrams for learning: principles for effective representational systems. Cogn. Sci. **26**(6), 685–736 (2002)
7. Cheng, P.C.-H.: Why diagrams are (sometimes) six times easier than words: benefits beyond locational indexing. In: Blackwell, A.F., Marriott, K., Shimojima, A. (eds.) Diagrams 2004. LNCS (LNAI), vol. 2980. Springer, Heidelberg (2004)
8. Cheng, P.C.-H.: Probably good diagrams for learning: representational epistemic re-codification of probability theory. Topics Cogn. Sci. **3**(3), 475–498 (2011)
9. Cheng, P.C.-H.: Algebra diagrams: a HANDi introduction. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 178–192. Springer, Heidelberg (2012)

10. Cheng, P.C.-H., Barone, R.: Representing complex problems: a representational epistemic approach. In: Jonassen, D.H. (ed.) Learning to Solve Complex Scientific Problems, pp. 97–130. Lawrence Erlbaum Associates, London (2007)
11. Cheng, P.C.-H., Barone, R., Cowling, P.I., Ahmadi, S.: Opening the information bottleneck in complex scheduling problems with a novel representation: STARK diagrams. In: Hegarty, M., Meyer, B., Narayanan, N. (eds.) Diagrams 2002. LNCS (LNAI), vol. 2317, pp. 264–278. Springer, Heidelberg (2002)
12. Cheng, P.C.-H., Simon, H.A.: Scientific discovery and creative reasoning with diagrams. In: Smith, S., Ward, T., Finke, R. (eds.) The Creative Cognition Approach, pp. 205–228. MIT Press, Cambridge (1995)
13. Cosmides, L., Tooby, J.: Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition **58**, 1–73 (1996)
14. Cleveland, W.S., McGill, R.: Graphical perception and graphical methods for analysing scientific data. Science **229**, 828–833 (1985)
15. Engelhardt, J.: The language of graphics. ILLC, University of Amsterdam (2002)
16. Green, T., Blackwell, A.: Cognitive Dimensions of Information Artefacts: a tutorial. https://www.cl.cam.ac.uk/~afb21/CognitiveDimensions/CDtutorial.pdf
17. Gurr, C.A.: On the isomorphism, or lack of it, of representations. In: Marriott, K., Meyer, B. (eds.) Visual Language Theory, pp. 293–306. Springer, Heidelberg (1998)
18. Hoffman, D.D.: Visual Intelligence: How We Create What We See. WW Norton & Company, New York (2000)
19. Koedinger, K.R., Anderson, J.R.: Abstract planning and perceptual chunks: elements of expertise in geometry. Cogn. Sci. **14**, 511–550 (1990)
20. Kotovsky, K., Hayes, J.R., Simon, H.A.: Why are some problems hard? Cogn. Psychol. **17**, 248–294 (1985)
21. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. Cogn. Sci. **11**, 65–99 (1987)
22. Markman, A.B.: Knowledge Representation. Lawrence Erlbaum, London (1999)
23. Newell, A., Simon, H.A.: Human Problem Solving. Prentice-Hall, Englewood Cliffs (1972)
24. Peebles, D.J., Cheng, P.C.-H.: Modelling the effect of task and graphical representations on response latencies in a graph-reading task. Hum. Factors **45**(1), 28–45 (2003)
25. Scaife, M., Rogers, Y.: External cognition: how do graphical representations work? Int. J. Hum. Comput. Stud. **45**, 185–213 (1996)
26. Shimojima, A.: Semantic Properties of Diagrams and Their Cognitive Potentials. CSLI Press, Stanford (2015)
27. Stenning, K., Oberlander, J.: A cognitive theory of graphical and linguistic reasoning: logic and implementation. Cogn. Sci. **19**(1), 97–140 (1995)
28. Stillings, N.A., Weisler, S.E., Chase, C.H., Feinstein, M.H., Garfield, J.L., Rissland, E.L.: Cognitive Science: An Introduction, 2nd edn. MIT Press, Cambridge (1995)
29. Tversky, B.: Visualizing thought. Top. Cogn. Sci. **3**, 499–535 (2011)
30. Zhang, J., Norman, D.A.: Representations in distributed cognition tasks. Cogn. Sci. **18**(1), 87–122 (1994)

Measuring User Comprehension of Inference Rules in Euler Diagrams

Sven Linker(✉), Jim Burton, and Andrew Blake

Visual Modeling Group, University of Brighton, Brighton, UK
{s.linker,j.burton,a.l.blake}@brighton.ac.uk

Abstract. Proofs created by diagrammatic theorem provers are not designed with human readers in mind. We say that one proof, P_1, is more "readable" than another, P_2, if users make fewer errors in understanding which inference rules were applied in P_1 than in P_2, and do so in a shorter time. We analysed the readability of individual rules in an empirical study which required users to identify the rules used in inferences. We found that increased clutter (redundant syntax) in the premiss diagrams affects readability, and that rule applications which require the user to combine information from several diagrams are sometimes less readable than those which focus on a single diagram. We provide an explanation based on mental models.

1 Introduction

Interactive and automated theorem proving with diagrams has been explored in systems such as Speedith [6]. However, existing tools do not take into account the growing body of research on what the specific cognitive advantages of reasoning diagrammatically might be, and on where the source of these advantages, if they exist, might be located. This research includes neurological studies that examine brain activity of users reasoning with and without diagrams [5], and empirically-derived guidelines for producing diagrams that make good use of Gestalt principles relating to colour and form [2]. At the broadest level our research question asks *is it possible to develop a systematic understanding of readability in diagrammatic proofs?* We use the term "readable" to mean relatively easy to understand, and will use error rates and response times of users who read the proof as measures of relative readability.

Euler diagrams have been used as a formal logic since the 1990s. Figure 1 shows a theorem expressed using Euler diagrams, equivalent to the expression $B \subseteq A \wedge C - B = \emptyset \Rightarrow C \subseteq A$. In order to prove that this theorem is true, we need to apply inference rules which add and remove elements from the diagrams labelled 1 and 2 until we produce diagram 3.

In this paper we describe an empirical study in which we analyse and measure the factors that affect comprehension of individual inference steps when reasoning with Euler diagrams. The study measures the number of errors and the time taken to answer a series of questions about Euler inference rule applications.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 32–39, 2016.
DOI: 10.1007/978-3-319-42333-3_3

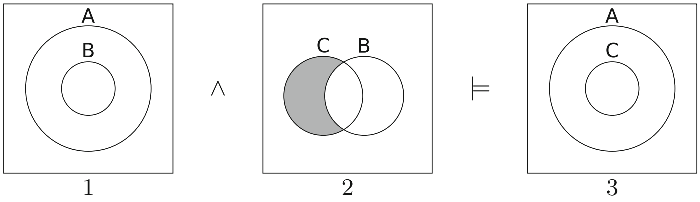


Fig. 1. An Euler diagram theorem

The present work is intended as a first step towards a notion of readability for whole proofs with Euler diagrams. To the best of our knowledge, there have been relatively few empirical studies of reasoning with Euler diagrams, e.g. the work of Sato and Mineshima such as [5]. These studies are concerned with the activity of proving itself, i.e., the creation of a proof, while we are interested in the task of understanding of a proof which already exists.

In the following section we give a brief explanation of Euler diagrams and their use in reasoning. In Sect. 3 we describe our experimental design, then in Sects. 4 and 5 we present our analysis and interpretation of the results. The training materials, questions and the (anonymised) data which was collected are available from our website, http://readableproofs.org/readability-study.

2 Euler Diagrams

Unitary Euler diagrams are drawn within a bounding rectangle representing the universe of discourse. Sets are represented by labelled *contours* (or *curves*) drawn within the rectangle. Topological relations between the curves specify the relations between sets. The curves divide the space within the diagram into *zones*. Zones may be shaded or non-shaded. The set represented by a shaded zone must be empty.

Unitary Euler diagrams can be composed to create *compound diagrams*. Within this paper, we will only use conjunction to compose diagrams.

The inference rules we consider are the following: 1. Erase Contour (EC), 2. Erase Shading (ES), 3. Combine (CO), 4. Copy Contour (CC), and 5. Copy Shading (CS). The effects of the rules are as follows. *Erase Contour* removes a contour from a unitary diagram. If this contour was separating a shaded zone from a non-shaded zone, the unified zone in the result will be non-shaded. *Erase Shading* removes the shading of a single zone from a unitary diagram. Note that rules 1 and 2 make changes to a single unitary diagram. We call these the *simple rules*. When using rule 3 to *combine* two unitary diagrams, both diagrams must contain the same set of zones. In the result, these two diagrams are replaced by a single diagram with the same set of zones and in which a zone is shaded if and only if it is shaded in one of the origin diagrams. For the copy rules, we have to identify which zones in different diagrams *correspond* [3] to each other. *Copy Contour* can be used to copy a contour c_1 from a unitary diagram d_1 to a second unitary diagram, d_2, respecting the topological information within d_1 about c_1

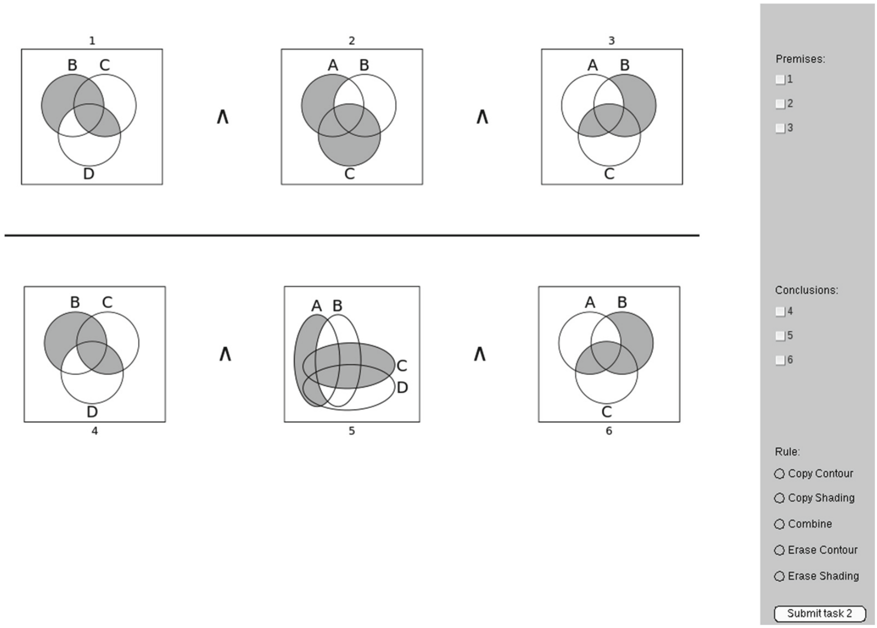


Fig. 2. Rule application: Copy Contour (cluttered version)

and the contours contained in both diagrams. Our last rule is *Copy Shading*: if a zone z_1 is not shaded in d_1 and corresponds to a shaded zone z_2 in d_2, then *Copy Shading* can be used to shade z_1. Rules 3, 4 and 5 depend on information from two unitary diagrams in the premiss. We call these rules *complex*.

An important notion for our work is the *clutter* of a diagram [1], which we measure by the *contour score* of a diagram. First, the contour score of a single zone is the number of contours it is enclosed in. The contour score of a unitary diagram is the sum of all contour scores of the zones present in the diagram. For example, within Fig. 2, all diagrams in the premiss have a contour score of 12, while diagram 5 has a score of 30.

3 Experimental Design

Our research questions are as follows: first, does the amount of clutter in the diagrams have an effect on the identification of rule applications? Secondly, are applications of complex rules less readable than applications of simple rules?

We designed our experimental tasks by first creating two semantical situations, being the relationships between four sets named A, B, C and D, chosen so as to ensure that all rules can be applied to such a situation. We constructed two compound diagrams representing each situation, with *high* and *low* clutter respectively. Each compound diagram consists of a conjunction of three unitary diagrams. From each of the four premises we constructed an instance of an application of each of the five rules, resulting in 20 such tasks for a within-group study. Given an application, the participants' task is to identify which rule has been applied, which unitary diagrams comprise its input, and which unitary diagram in the conclusion is the result of applying the rule.

The cluttered versions of the premises were created by using Venn-form diagrams. Diagrams were drawn according to well-formedness principles and using consistent font, font weight, line width and so on. All parts of the study took place in a specialised usability lab using the same equipment.

The study consisted of a paper-based introduction to Euler diagrams and the inference rules, a training phase and the main study. Within the training phase, the instructor actively worked with the participants to address misunderstandings and misconceptions directly as they arose. The participants had access to a "cheatsheet" containing an exemplary application of each rule with the answers the participants had to provide.

After the training phase, participants were presented with the 20 tasks that form the main study in a randomized order (see Fig. 2). For each question, the participants had to identify the rule that had been applied, the diagram(s) that rule had been applied to, and the diagram that was changed in the conclusion. The program recorded the answers of the participants as well as the time taken to finish the task. For the main study we recruited 30 undergraduate participants, 23 male and 7 female, from the ages of 18 to 34.

4 Analysis

Error Analysis. We excluded nine data points where the participant did not come to an answer within the time limit of 120 s.

Table 1. Errors: clutter

Clutter	Errors	Correct
High	69	226
Low	56	240

Table 2. Errors: type of rule

Rule	Errors	Correct
Combine	18	102
Copy Contour	28	89
Copy Shading	53	62
Erase Contour	6	114
Erase Shading	20	99

Table 1 shows errors aggregated by clutter. A Chi-square test reveals no significant differences for clutter. Table 2 shows the number of errors according to the rules used in the tasks. A Chi-square test shows that there is a significant difference between some pairs in the set ($\chi^2(df = 4, N = 591) = 66.26, p < 0.05$). We used the Chi-square test for all pairs in this set to find the significantly different entries with confidence of $p < 0.001$.

The results of these tests are shown in Table 3. Participants performed significantly worse for *Copy Shading*. The difference between *Copy Contour* and *Erase Contour* is also significant.

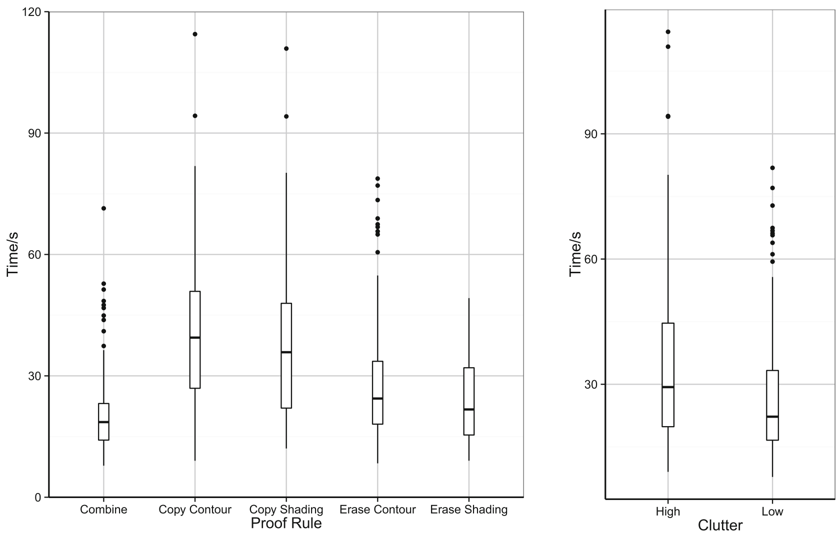
Table 3. Pairwise comparisons of rules (errors)

	CC	CS	EC	ES
CO	$\chi^2(1,237) = 2.48/$ $p = 0.116$	$\chi^2(1,235) = 25.46/$ $p < 0.001$	$\chi^2(1,240) = 5.60/$ $p = 0.018$	$\chi^2(1,239) = 0.04/$ $p = 0.838$
CC	-	$\chi^2(1,232) = 11.57/$ $p < 0.001$	$\chi^2(1,237) = 15.77/$ $p < 0.001$	$\chi^2(1,236) = 1.43/$ $p = 0.231$
CS		-	$\chi^2(1,235) = 50.56/$ $p < 0.001$	$\chi^2(1,234) = 22.02/$ $p < 0.001$
EC			-	$\chi^2(1,239) = 7.42/$ $p = 0.006$

Time Analysis. In this analysis we removed all errors, since we are interested in the time the participants needed to perform the tasks correctly, reducing our dataset to 466 data points. We distinguish the means according to the amount of clutter and according to the type of rule. Figures 3a and b show the interquartile ranges of the performance times. Participants took longest to identify *Copy Contour* and *Copy Shading*, while *Combine* has the lowest median.

Due to the removal of the errors and timeouts, our data is unbalanced. Hence, we analysed the set by using a RM-ANOVA test, comparing the performance time for clutter and the type of rules. Clutter levels had a significant impact on performance time ($F(1, 19) = 37.83, p < 0.05$). The effect size is $d = 0.45$, which corresponds to a percentile of 66 %–69 %. That is, approximately 2/3 of participants were, on average, faster at completing the tasks with low clutter than the average person completing the high clutter tasks.

Furthermore, the RM-ANOVA showed that a significant difference exists between at least one pair of rules ($F(4, 19) = 36.97, p < 0.05$). The results of a



(a) Rules (b) Clutter

Fig. 3. Performance times

Table 4. RM-ANOVA for pairwise comparisons of rules (time)

	CC	CS	EC	ES
CO	$F(1,19) = 107.46$/ $p < 0.001$/ $d = 1.32$/88–92%	$F(1,19) = 61.43$/ $p < 0.001$/ $d = 1.07$/84–88%	$F(1,19) = 16.89$/ $p < 0.001$/ $d = 0.53$/69–73%	$F(1,19) = 3.66$/ $p = 0.057$
CC	-	$F(1,19) = 1.54$/ $p = 0.217$	$F(1,19) = 34.97$/ $p < 0.001$/ $d = 0.75$/76–79%	$F(1,19) = 69.37$/ $p < 0.001$/ $d = 1.17$/84–88%
CS		-	$F(1,19) = 15.51$/ $p < 0.001$/ $d = 0.55$/69–73%	$F(1,19) = 36.79$/ $p < 0.001$/ $d = 0.92$/82–84%
EC			-	$F(1,19) = 4.88$/ $p = 0.028$

pairwise analysis, testing for an increased confidence level < 0.001, are shown in Table 4. This table allows us to group the rules into different (not necessarily disjoint) subsets. *Copy Contour* and *Copy Shading* are significantly different from all other rules. While *Erase Contour* and *Combine* are significantly different from each other, the difference to *Erase Shading* is not significant. Hence *Erase Contour* and *Erase Shading* constitute one subset, and *Combine* and *Erase Shading* constitute the last one.

The RM-ANOVA shows that there is significant interaction between the type of rule and the amount of clutter ($F(4,19) = 2.97, p < 0.05$). However, applying the test to pairs of rules and distinguishing the clutter level does not show a significant difference with a confidence < 0.01.

5 Interpretation

While higher amounts of clutter had no significant impact on the error rate, it did in fact increase the time the participants needed to solve the tasks. Our interpretation of this is as follows. Identifying the rules was a difficult task for our non-expert participants, requiring a meticulous analysis of the diagrams. A higher amount of clutter increases the number of single parts within the diagrams they had to look at. This explains the increased time the participants needed to solve the tasks. However, the high levels of concentration they had to maintain prevented them from being distracted by these additional elements, i.e., from making additional errors. Furthermore, in comparison to the diagrams used in the preceeding studies on clutter (e.g. [1]), our diagrams can still be considered to have a low amount of clutter.

The type of the rule had a much stronger impact. Our interpretation adopts the perspective of mental models, assuming that readers create and manipulate internal representations of diagrams [4]. We assume that the experimental tasks required participants to manipulate mental models in ways that correspond with syntactical manipulations of diagrams. By asking the participants to identify the premises and conclusion of a rule, we require them to consider each unitary diagram separately. That is, we expect that the participants create a mental model of each unitary diagram to analyse. As described by Johnson-Laird, numerous studies have corroborated the prediction that the "the more models we need to take into account to make an inference, the harder the inference should be" [4].

With regard to the performance time, we can group our rules into subsets. The first is the *complex* rules, where the participants needed to identify that information contained within one diagram in the premises was added to another diagram to yield the conclusion. Doing this would require them to inspect each diagram, p, from the premiss to decide whether their mental model of p can be manipulated in ways consistent with their mental model of the conclusion.

Our second subset comprises applications of *simple* rules. To identify these, the participants needed to find the right conclusion, create and manipulate a mental model of the diagram above it (by forgetting a part of the information within) and compare them. This difference in cognitive effort is reflected in the time the participants needed to perform the tasks.

Combine stands out from the other four rules, however, since it alone results in a conclusion containing only two unitary diagrams. Even though the participants would need to compare two diagrams from the premiss to see whether they yield the conclusion, the strong visual difference between the premiss and conclusion makes it obvious that this rule was applied.

6 Conclusion

We presented the results of a study which examined the impact of clutter and differences in inference rules on the ability of participants to identify applications of these rules. The amount of clutter did not significantly influence the number of errors made but did impact significantly on performance time. We found significant differences in performance time between the rules based on the number of unitary diagrams that need to be considered as input to the rule, modulo *Combine*. We attributed these differences to the cognitive effort the participants needed to make while identifying and validating rules applications.

Acknowledgements. This work is supported by EPSRC grant EP/M011763/1.

References

1. Alqadah, M., Stapleton, G., Howse, J., Chapman, P.: Evaluating the impact of clutter in Euler diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 108–122. Springer, Heidelberg (2014)
2. Blake, A., Stapleton, G., Rodgers, P., Cheek, L., Howse, J.: Improving user comprehension of Euler diagrams. In: VL/HCC, pp. 189–190. IEEE (2013)
3. Howse, J., Stapleton, G., Flower, J.: Corresponding regions in Euler diagrams. In: Hegarty, M., Meyer, B., Narayanan, N.H. (eds.) Diagrams 2002. LNCS (LNAI), vol. 2317, pp. 76–90. Springer, Heidelberg (2002)
4. Johnson-Laird, P.N.: Mental models and human reasoning. Proc. Nat. Acad. Sci. U.S.A **107**(43), 18243–18250 (2010)
5. Sato, Y., Masuda, S., Someya, Y., Tsujii, T., Watanabe, S.: An fMRI analysis of the efficacy of Euler diagrams in logical reasoning. In: VL/HCC, pp. 143–151. IEEE (2015)
6. Urbas, M., Jamnik, M., Stapleton, G.: Speedith: a reasoner for spider diagrams. J. Log. Lang. Inf. **24**, 1–54 (2015)

Exploring Representations of Student Time-Use

Amy Rae Fox[1(✉)], Erica de Vries[2], Laurent Lima[2], and Savannah Loker[1]

[1] California State University - Chico, Chico, CA 95929, USA
amyraefox@gmail.com
[2] Laboratory of Educational Sciences, University of Grenoble Alpes, Grenoble, France
Erica.deVries@univ-grenoble-alpes.fr

Abstract. How does one visually represent the use of time? We explored students' use of graphical metaphors by asking undergraduates at a public French university to generate representations of their personal time-use including: activities, sequence, duration, timing, and frequency. The resulting use of space and form was analyzed by way of an iteratively developed coding scheme. We discuss how the analyses of the spontaneous productions support previous research on spatial representations of time, and the implications for the design of time management tools for students.

Keywords: Spontaneous representation · Diagrams · Visualization · Student time-use · Coding scheme · Visual content analysis

1 Introduction

How students utilize time has important implications for their academic and professional success [5]. Students need to effectively allocate time between competing priorities such as homework, sleep, and extracurriculars, with such decisions constituting important developmental steps toward adulthood.

There are five components that can be derived from time-use data: activities, duration (quantity of time on activities), sequence (order of activities), timing (start/end of activities), and frequency (number of occurences in an interval). The visual representation of these components offer numerous applications, from resource allocation, to event planning, to detecting patterns in human behavior. In education, the applications are particularly compelling, as pedagogical activities challenging students to think critically about time-use are a first step toward aligning time-use with strategic goals. They also present a substantial challenge, as temporal phenomena are not strictly visual and cannot be visually compared to their representations.

When representing time in language, we routinely employ metaphors [7], such as, "The deadline is sneaking up on me, but my manuscript is ahead of schedule!" We use metaphor and analogy to create correspondence between abstract and familiar concepts. Similarly, these mappings can be applied in visual forms: *graphical metaphors*, constructed to convey meaning. In this work, we explore students' use of graphical metaphors when representing time-use.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 40–47, 2016.
DOI: 10.1007/978-3-319-42333-3_4

1.1 Space and Time as Representational Media

The advantages of space for representation have been extensively documented [1,11] and the observation that space is best represented in a spatial medium is common place [8]. A variety of cultural artifacts, such as calendars and clocks, represent time in the medium of space. These are extrinsic representations, where the properties of the domain must be enforced on the medium in an arbitrary manner [11]. To represent sequence in a two-dimensional plane, some artificial device is required to enforce the property of linearity. This can be accomplished by a graphical form (e.g. an arrow) by a domain-specific representational format (e.g. a calendar), or by relying on reading direction as a convention for imposing linearity. A fundamental asymmetry is evident when using space to represent anything other than space and *a fortiori* time. This asymmetry is consistent with research in cognitive linguistics [2] which suggests that individuals use spatial information when thinking about time, but rarely use temporal information when thinking about space.

Just as we must choose between myriad linguistic metaphors, when constructing graphic metaphors the designer must make choices for how to use media to represent. Tversky [12] examines how space and form are used to convey meaning in diagrams. She first identifies the use of space for depicting order, positioning forms (marks on the page) along horizontal, vertical and central-peripheral planes. An examination of productions by children revealed a consistent relationship between writing direction and depiction of temporal sequence [13]. While space exists in multiple dimensions, form is constructed. Complex forms are constructed by combining simple forms in purposeful configurations. The simplest forms are points and lines, which can be extended, contoured and combined to generate realistic depictions of worldly objects, or more abstract flights of imaginative fancy.

1.2 Flexibility in Representation

In response to an instruction to produce representations, participants can do a number of things depending on the affordances of the medium and their "catalogue" of available responses. Reuchlin (as cited in [9]) coined the term *vicariance* for the ways in which an individual may rely on a number of redundant mechanisms for performing a cognitive task. We view representation as a vicariant process, where the potential representational formats are determined by the individual's repertoire of available behaviors. In the case of students (with paper and pencils) we might expect to see:

- Letters and numbers: labels, digits, etc. (descriptive)
- Figurative drawings: attempted realism (depictive)
- Standard representations: histograms, pie charts, maps, etc.
- Domain-specific representations: particular to a specific domain [4].
- Ad-hoc representations: inventing a new context-specific representational format rather than using an existing one.

If like Tversky [12] we consider the process of representation an indicator of underlying thought, then examining the graphic productions of student time-use may help us understand how students conceptualize this important factor of their academic success. Do students think of their time as linear or cyclical? Do they emphasize order or timing of activities? As a first step, we explore the variability of representational behavior in a student population. Our investigation is guided by two questions: (1) How do students use space and form to represent time-use? (2) Which mechanisms are used to represent each component: sequence, timing, duration and frequency?

2 Method

Twenty-five (22 female, 3 male) undergraduate Education majors (median age = 23) at a public French university participated as a course requirement.

2.1 Materials and Procedure

After a demographic survey, students were given one sheet of paper, a variety of pens and colored pencils, and one hour to complete the exercise. Instructions prompted students to represent their time-use for a typical week during the academic year. They were explicitly directed to represent activity (what), sequence (order), duration (quantity), timing (chronology), and frequency (number of occurrences), in as many representations as necessary, using any graphic conventions. Only the term representation was used, avoiding biasing formats with terms such as: chart, graph, sketch, text, etc.

2.2 Coding Scheme

The coding scheme was developed using a directed approach [6]. Starting with the categories of 'use of space' and 'use of form', coding variables were chosen in alignment with Tversky's [12] discussion of space and form. Operational definitions were developed for each variable with values that were exhaustive and mutually exclusive. The resulting scheme[1] was applied by three psychology graduate students who coded the entire sample. Coding results for the entire sample were evaluated for inter-rater reliability, with positive outcomes: use of space $\alpha = 0.87$, use of form $\alpha = 1.00$, and primary mechanisms $\alpha = 0.97$.

3 Results

3.1 Use of Space

The use of space in the diagrams (n = 25) was consistent with our expectations based on reading direction in French. Twenty-one diagrams were characterized

[1] A full description of the coding scheme can be found at https://osf.io/ms9kq/.

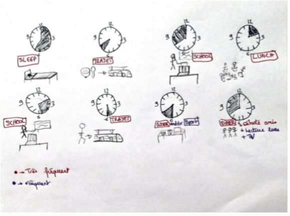


Fig. 1. Linear

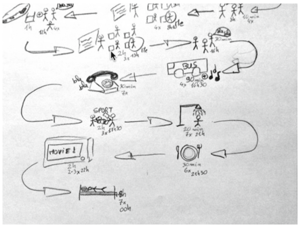


Fig. 2. Snake

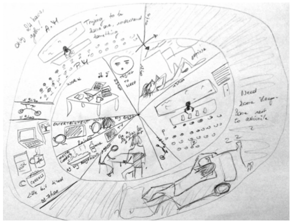


Fig. 3. Circular

as linear and four circular. Nearly all of the students (22) depicted the start of the day in the upper left corner of the page. Of the remaining three, two were circular representations placing the start of the day at the 12:00 position of a clock, and one was linear, starting in the lower left corner.

Nineteen of the 21 linear diagrams utilized both horizontal and vertical axes, while the remaining two used only the horizontal. Sixteen diagrams adopted a left-to-right orientation, while four alternated directions in a snake-like pattern. Figure 1 is a prototypical example of linear representation, with the origin in the top left corner. In these cases, the reader scans the diagram from left-to-right, jumping at the end of horizontal space, consistent with reading behavior. Four diagrams avoided the end of line scanning effect by alternating direction at the end of each line (Fig. 2). We dubbed these “snakes”, as the information appeared to slither across the page. In each case a form, either line or arrow, accompanied the transition to indicate the change in direction. We contrast this with Fig. 1 in which the student assumes the viewer will skip to the next line and continue reading left-to-right, without the need to provide a formal indicator of direction. Of the four circular diagrams, three presented information in the clockwise direction (Fig. 3). There was minimal use of the radial orientation, with only two diagrams depicting flow from the periphery of the circles toward the center in a spiral fashion.

3.2 Use of Form

The use of form in the sample was varied, suggesting the scenario was effective in motivating students without biasing their choice of form. Raters found few instances of meaningful glyphs [12] such as dots, lines, and boxes. Arrows were the noticeable exception, found in 21 diagrams, employed to orient the viewer from earlier to later activities. Number was the most prevalent form, found in 23 of the 25 diagrams, followed by text (21). Nineteen included depictive drawings, while only 13 utilized more than one color. Figures 4, 5 and 6 exemplify the range of forms in the drawings from highly depictive (employing analogs) to highly descriptive (employing symbols). As evident in these examples, the use of descriptive vs. depictive drawings to describe activities fell on a continuum.

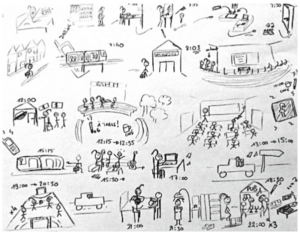


Fig. 4. Depictive

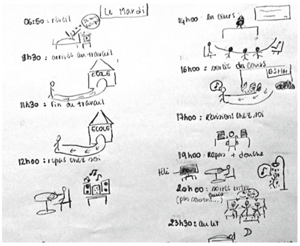


Fig. 5. Balanced

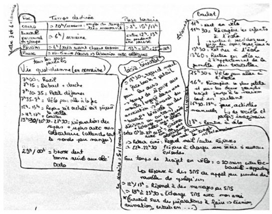


Fig. 6. Descriptive

There were no examples of prevalent but poorly executed drawings, suggesting that students utilized forms in accordance with their drawing ability.

3.3 Primary Mechanisms

Table 1 describes the percentage of diagrams that represented each component of time-use, as well as the number of diagrams utilizing each mechanism in the coding scheme. Only four individuals effectively represented all five components. Frequency was most commonly neglected, followed by duration, then timing. Sequence was always indicated by position, in many cases with the addition of arrows, while timing was almost exclusively represented by number. Two imaginative illustrations also utilized position to indicate timing by placing drawings around the corresponding positions of a clock face to indicate the time of day they began. Duration was often absent from the drawings, but when it was present, it was indicated by number. Two novel illustrations also utilized color to differentiate categories of activities.

Table 1. Frequency and methods of representing time-use components

Mechanism	Sequence	Timing	Duration	Frequency
% inclusion (n = 25)	100 %	96 %	54 %	46 %
Position	25	2	1	2
Size			4	
Text				
Number	2	21	9	6
Drawing		2	1	1
Arrow	15			
Color		2		2

4 Discussion

4.1 Student Productions

We found that in their use of space, students preferred linear patterns orientated with reading direction. Few students used circular patterns to indicate cyclical phenomena. Does this mean that students conceptualize their time as linear, rather than cylical? Tversky [12] made a similar observation, noting that students were reluctant to produce circular diagrams even when asked to model cycles and processes. She suggests that linear thinking is easier than cyclical, and students may prefer to consider a simple forward progression of time. Another explanation is the influence of cultural artifacts on the choice of representational format. The linear flow of information was evocative (though not strictly reflective) of calendars and agendas. Graphics conventionally used for planning may influence students' representational choices. In selecting form, students used both text and figurative drawings to represent activities, while number was used to describe timing, frequency and duration. Arrows were used exclusively to enforce linearity, directing the viewer's attention to the forward flow of time.

Of greatest interest were representational choices for the components of time-use. Although the instructions explicitly allowed for multiple representations, all students attempted to create a single integrated diagram. Alternatively, students could have created a series of representations for each component. Common formats such as charts and graphs were not exploited, despite their efficiency in communicating quantities (such as duration). It is possible that these formats were not familiar to the homogeneous sample of Education majors. In the future, we plan to sample students in science and engineering to explore variance in formats as a function of domain knowledge. The prevalance of depictions stands in contrast to the findings of Manalo and Useka [10], which suggest that students are reluctant to spontaneously produce diagrams given a communicative task. It is possible that the nature of the experimental task (representing activities with high personal involvement) as well as the imaginative nature of the instigating scenario may account for this discrepancy. It is possible that students are more comfortable constructing depictive representations of information that is personally relevant, as opposed to scientific phenomena. While no student constructed a complete domain-specific representation (e.g. diary, calendar), several utilized space in a fashion consistent with those conventions. The remaining productions demonstrated a preference for complex, integrated diagrams, reflecting an attempt at simultaneously inventing a representational format and expressing new content (ad-hoc context-specific representation). Alternatively, students might have placed a high value on informational efficiency. To examine this further, we suggest refining the stated goal from one of informing to differentiated tasks for planning, problem solving and exposition. While revealing sources of variation, a more strictly defined purpose might allow for more robust inferences about the underlying conceptual structure suggested by students graphic productions.

4.2 Implications and Future Work

Our analysis demonstrates conventional behavior in representations of time-use which can be applied to the production of organizational aids for students. Analysis of the pre-exercise survey showed, surprisingly, that few students consistently used paper or computer-based tools for planning and tracking time. Improving students' time allocation could help to improve performance and decrease failure rates in university [5]. Inspired by the preference for linear productions, it may be beneficial to design representations that draw attention to the cyclical nature of schedules, revealing the frequency of activities and supporting inferences about patterns of behavior. Rather than provide students with tools for planning and tracking, we suggest such tools be embedded in pedagogical activities on goal-setting. To maximize efficacy, the representations included should be consistent with students' conceptualizations of time-use (e.g. primarily linear, left-to-right), while simultaneously drawing attention to aspects of time-use neglected by students (e.g. the importance of sequence and frequency). As noted by Cox [3], individuals often perform better when utilizing self-generated external representations than those in formats invented by others. The exercise also prompted student reflection and discussion on time allocation as it pertains to prioritization and goal-setting. Preliminary analysis of the production content reveals diversity in the categorization of activities, which presents an additional question for research. We suggest that future work evaluate the content of productions by activity categories (school, homework, leisure, sleep, etc.) alongside pedagogical activities on time-use planning and evaluation. In the next phase of this research, we plan to repeat the diagramming exercise with students from different disciplines at an American University, utilizing digital pen systems to facilitate the content analysis. In addition to improving students meta-representational competency [4] we propose that constructing and analyzing representations of time-use may help students better understand how they allocate their time, and thus empower them to take control of this important factor of their academic success.

References

1. Bertin, J.: Semiology of Graphics: Diagrams, Networks, Maps. University of Wisconsin Press, Madison (1983)
2. Boroditsky, L.: Metaphoric structuring: understanding time through spatial metaphors. Cognition **75**(1), 1–28 (2000)
3. Cox, R.: Representation construction, externalised cognition and individual differences. Learn. Instr. **9**(4), 343–363 (1999)
4. DiSessa, A.A.: Metarepresentation: native competence and targets for instruction. Cogn. Instr. **22**(3), 293–331 (2004)
5. Fernex, A., Lima, L., de Vries, E.: Exploring time allocation for academic activities by university students in France. High. Educ., 399–420 (2014)
6. Hsieh, H.-F., Shannon, S.E.: Three approaches to qualitative content analysis. Qual. Health Res. **15**(9), 1277–1288 (2005)

7. Lakoff, G., Johnson, M.: Metaphors We Live By. University of Chicago Press, Chicago (1980)
8. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. Cogn. Sci. **99**, 65–99 (1987)
9. Lautrey, J.: A pluralistic approach to cognitive differentiation and development. In: Sternberg, R.J., Lautrey, J., Lubard, T.I. (eds.) Models of Intelligence: International Perspectives. APA Press, Washington (2003)
10. Manalo, E., Uesaka, Y.: Students' spontaneous use of diagrams in written communication: understanding variations according to purpose and cognitive cost entailed. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 78–92. Springer, Heidelberg (2014)
11. Palmer, S.E.: Fundamental aspects of cognitive representation. In: Rosch, E., Lloyd, B.B. (eds.) Cognition and Categorization, pp. 259–302. Erlbaum, Hillsdale (1978)
12. Tversky, B.: Visualizing thought. Top. Cogn. Sci. **3**, 499–535 (2011)
13. Tversky, B., Kugelmass, S., Winter, A.: Cross-cultural and developmental trends in graphic productions. Cogn. Psychol. **23**, 515–557 (1991)

Logic and Diagrams

Evaluating Diagrammatic Patterns for Ontology Engineering

Eisa Alharbi, John Howse, Gem Stapleton(✉), and Ali Hamie

University of Brighton, Brighton, UK
{e.alharbi2,John.Howse,g.e.stapleton,a.a.hamie}@brighton.ac.uk

Abstract. Diagrammatic logics have been widely studied since Shin's seminal work on Venn diagrams in the 1990s. There have been significant theoretical advances alongside empirical work investigating their efficacy with respect to symbolic notations. However, we have little understanding about how to choose between syntactically different diagrams when formulating logical axioms. This paper sets out to provide insight into such choices. By appealing to ontology engineering, we identify commonly required semantic properties that require axiomatization. We systematically identify three different ways of axiomatizing these properties using diagrammatic patterns. One way does not use explicit quantification. The other ways both use explicit quantification but employ different diagrammatic devices to capture the required semantics. We evaluated these competing patterns by conducting an empirical study, collecting performance data. We conclude that avoiding explicit quantification, and representing the information purely diagrammatically, best supports task performance. As a result, users and designers of diagrammatic logics are guided towards avoiding explicit quantification where possible.

Keywords: Ontologies · Axioms · Diagrammatic patterns · Visualization

1 Introduction

Understanding how to effectively represent information using diagrams is a major research goal of the Diagrams community. The focus of this paper is diagrams designed for making logical statements, in a precise and unambiguous way. Whilst a lot of research has been done into the design and theoretical development of diagrammatic logics, stemming from Shin's seminal work [12], little attention has been given to how to choose between semantically equivalent, yet syntactically different, diagrams. This paper begins to address this knowledge gap, by empirically evaluating competing choices of diagrams for axiomatizing semantic properties. To ensure practical relevance, and therefore wider significance of our research, we identified ontology engineering as a major endeavour where axioms are routinely defined.

Ontologies help us to structure and reason about information and data; formally, an ontology is a collection of axioms. With the abundance of data available

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 51–66, 2016.
DOI: 10.1007/978-3-319-42333-3_5

in this information age, ontology engineering is becoming increasingly important. Many different specialists are involved in the development of ontologies including domain experts, software engineers, data analysts and lawyers. Some of these stakeholders are not adept at using the existing approaches to ontology engineering, which involve the use of formal languages such as description logic [2] and OWL, the Web Ontology Language [1]. This implies that diagrammatic approaches to ontology engineering have the potential to appeal to ontology engineers without formal training in logic. With this in mind, concept diagrams [6] were designed to be used as an accessible ontology engineering language, usable by more stakeholders. Therefore, concept diagrams provide an ideal notation with which to provide an understanding into the relative effectiveness of different ways of axiomatizing semantic properties. Although specifically designed for ontology engineering, concept diagrams can be used in any logical context for which they are suitably expressive. As with any logic, it is frequently the case that any axiomatizable property can be represented by a variety of syntactically different concept diagrams but how to choose between the different representations is not obvious. This paper sets out to provide guidance on how to choose between such competing representations.

We briefly introduce the syntax and semantics of concept diagrams in Sect. 2. We identify commonly occurring ontology axioms in Sect. 3 where we also define different styles of diagrammatic patterns for representing them. The design and execution of an empirical study to determine which pattern styles are most accurately and most quickly interpreted by participants is described in Sect. 4. The analysis and results are presented in Sect. 5. We discuss the results in Sect. 6 and conclude in Sect. 7. Details of the questions and training material used in the study, along with the raw data collected, can be found at https://sites.google.com/site/eisamalharbi/DiagramsPatternsStudy.

2 Concept Diagrams

We present a brief overview of the syntax and semantics of concept diagrams, particularly with reference to the features occurring in this paper; a more detailed description of this fully formalized logic can be found in [14]. Closed curves represent sets which are called *concepts* in description logic and *classes* in OWL. Therefore concept diagrams are based on Euler diagrams. Binary relations, called *properties* or *roles* in ontology engineering, are represented by arrows. Individuals, or elements, are represented by dots or, more generally, trees.

Suppose that the individual Helen is a Person who is married to only the Person Poly (identified by the binary relation marriedTo) and that Helen owns exactly two pets (identified by the binary relation ownsPet), both of which are Dogs. These two pets include a Terrier called Lily. The left-hand diagram in Fig. 1 expresses this information, requiring three closed curves to represent the concepts Person, Dog, and Terrier. Person and Dog are disjoint and Terrier is subsumed by Dog. The location of the dots identifies the concepts of which they are instances; for example, Lily is located inside the curve labelled Terrier.

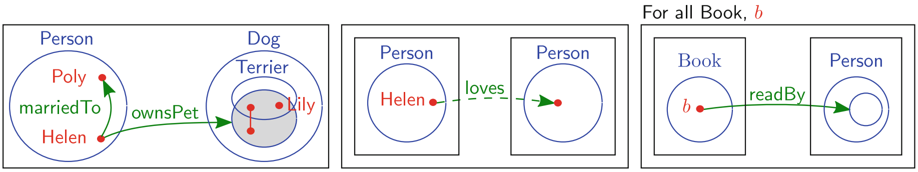


Fig. 1. Concept diagrams

The fact that Helen owns a set of Pets is expressed by the arrow labelled ownsPet, which hits an unlabelled curve. This curve is drawn inside Dog, to assert that the image of the relation ownsPet, with its domain restricted to Helen, is subsumed by Dog. The two trees inside this unlabelled curve tell us that Helen owns two Dogs. Helen's dog that is not Lily *might* be a Terrier. This uncertainty is captured by the use of the unlabelled tree with two nodes, one inside both the Dog and Terrier curves and the other inside the Dog curve but outside the Terrier curve. Shading is used to express that the only dogs owned by Helen are represented by the trees.

Concept diagrams use dashed arrows to represent partial information, such as Helen loves some Person and that Person *could* be Helen herself. A concept diagram expressing this is in the middle diagram of Fig. 1. The arrow connects diagrammatic syntax placed in different boxes to ensure that we have not asserted that the Person Helen loves is different from Helen. The right-hand diagram of Fig. 1 expresses that every Book is readBy only a subset of Person. The quantification expression written outside of the rectangles tells us that the diagram is making an assertion about all books. Lastly, we note that concept diagrams can also make assertions involving inverse relations, by annotating arrow labels using the symbol $^-$, and negation by labelling a bounding box with 'Not'. These will be discussed in more detail in Sect. 3.

3 Ontology Patterns

Concept diagrams are able to express commonly occurring ontology axioms in different ways. In this section we develop diagrammatic patterns for some types of axioms that commonly occur in ontology engineering: *subsumption*, *disjointness*, *All Values From*, *Some Values From*, *Domain* and *Range* [5].

3.1 Patterns Involving only Classes

The *Subsumption* axiom type is one of the simplest and widely used. Class C_1 *subsumes* Class C_2 if all members of C_2 are also members of C_1. Diagrammatically there is a natural way of representing subsumption, shown in the left-hand diagram of Fig. 2.

The *Disjointness* axiom type is also widely used. Classes C_1 and C_2 are *disjoint* if no element of C_1 is also an element C_2. Again, there is a natural way of expressing disjointness shown in the second diagram in Fig. 2.

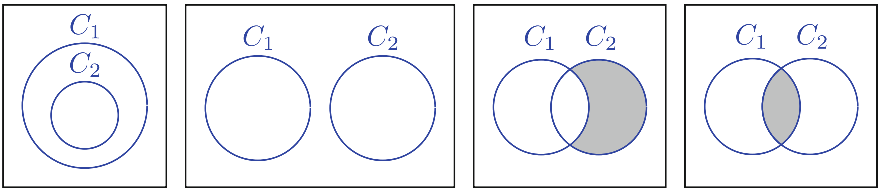


Fig. 2. *Subsumption* and *disjointness* patterns

There are other ways of representing subsumption and disjointness using concept diagrams. For example, we could use shading to indicate that a region is empty; this is the way that Venn diagrams represent such properties. The two right-hand diagrams of Fig. 2 show alternative patterns for subsumption and disjointness involving the use of shading. It is well established that a salient feature of diagrams is *well-matchedness* [4]. A notation is *well-matched to meaning* when its syntactic relationships reflect the semantic relationships being represented. In the left-hand diagram of Fig. 2, the curve labelled C_2 is enclosed by the curve labelled C_1 matching the semantic interpretation that C_2 is a subset of C_1. Similarly, in the adjacent diagram, the curves labelled C_1 and C_2 are disjoint, reflecting the interpretation that C_1 and C_2 are disjoint sets. However, the right-hand diagrams are not well-matched. The closed curves intersect giving no indication of the relationship between the sets they represent. Moreover, the shading is purely symbolic [8,13] and we have to learn that shaded regions represent the empty set. To confirm these theoretical insights, empirical studies have shown that users perform tasks more effectively when using well-matched Euler diagrams [3]. For these reasons, we recommend the well-matched subsumption and disjointness patterns for practical use by ontology engineers, and do not include them in our empirical study.

3.2 Patterns Involving Classes and Properties

In ontology engineering, a property can be considered as a mathematical (binary) relation between two classes. When we consider axioms involving properties, it is not clear what is the best diagrammatic way to represent these constructs. We consider four constructs involving properties: *All Values From*, *Some Values From*, *Domain* and *Range*. For each of these constructs we have systematically identified three different styles of diagrammatic patterns:

1. *Unquantified*
2. *Quantified with Solid Arrow*
3. *Quantified with Dashed Arrow*

The *Unquantified* patterns were first developed in [15].

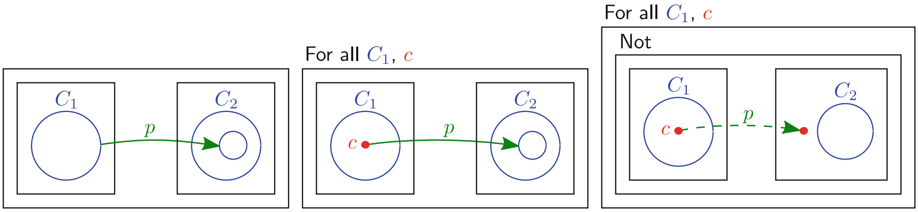


Fig. 3. *All Values From* patterns

3.3 All Values From Patterns

The *All Values From* axiom type represents a constraint involving two classes and a property: if each element of class C_1 is related, under property p, only to elements of class C_2 (if it is related to anything), then C_1 is said to have *All Values From* C_2 under p.

Unquantified Pattern. The left-hand diagram of Fig. 3 expresses that the image of property p, when its domain is restricted to C_1, is a subset of C_2. This axiomatizes the *All Values From* constraint. The closed curves representing classes C_1 and C_2 are each presented within a bounding rectangle because we do not want to express any relationship between C_1 and C_2.

Quantified with Solid Arrow Pattern. The middle diagram of Fig. 3 expresses that for each c in C_1, the image of property p, when its domain is restricted to c, is a subset of C_2. Thus C_1 has *All Values From* C_2 under p.

Quantified with Dashed Arrow Pattern. The right-hand diagram of Fig. 3 expresses that for each c in C_1, it is not the case that c is related, under p, to an element not in C_2. Thus no element of C_1 is related, under p, to an element not in C_2. Hence, each element of class C_1 is related, under property p, only to elements of class C_2.

3.4 Some Values From Patterns

The *Some Values From* axiom type also represents a constraint involving two classes and a property: if each element of class C_1 is related, under property p, to some element of class C_2, then C_1 has *Some Values From* C_2 under p.

Unquantified Pattern. The left-hand diagram of Fig. 4 expresses that the image of property p^-, when its domain is restricted to C_2, includes C_1. Therefore, for each a in C_1, there exists b in C_2 such that b is related to a under p^-. Hence, for each a in C_1, there is some b in C_2 such that a is related to b under p.

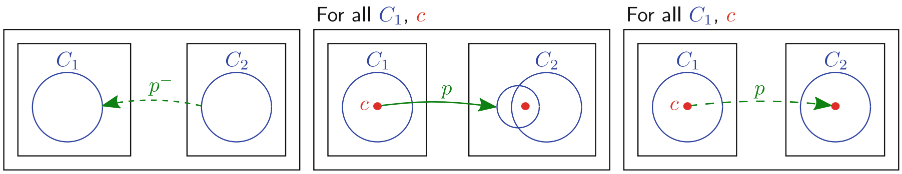


Fig. 4. *Some Values From* patterns

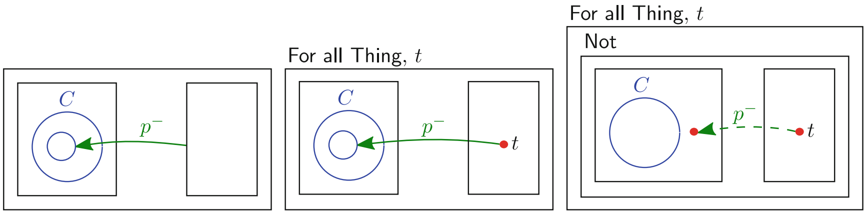


Fig. 5. *Domain* patterns

Quantified with Solid Arrow Pattern. The middle diagram of Fig. 4 expresses that for each c in C_1, the image of property p, when its domain is restricted to c, includes some element in C_2.

Quantified with Dashed Arrow Pattern. The right-hand diagram of Fig. 4 expresses that each c in C_1 is related, under p, to some element in C_2.

3.5 Domain Patterns

The *Domain* axiom type represents a constraint involving a class and a property: Class C is the *Domain* of property p if only elements from C are related to *something* under p. Each pattern for *Domain* will use the inverse of property p.

Unquantified Pattern. Noting that innermost rectangles represent the universal set, the left-hand diagram of Fig. 5 expresses that the image of property p^- is a subset of C. Hence, only elements in C are related to *something* by p.

Quantified with Solid Arrow Pattern. The middle diagram of Fig. 5 expresses that for each *Thing* t, the image of property p^-, when its domain is restricted to t, is a subset of C. Hence, only elements in C are related to *something* under p.

Quantified with Dashed Arrow Pattern. The right-hand diagram of Fig. 5 expresses that for each *Thing* t, it is not the case that t is related, by p^-, to an element not in C. Hence, only elements in C are related to *something* under p.

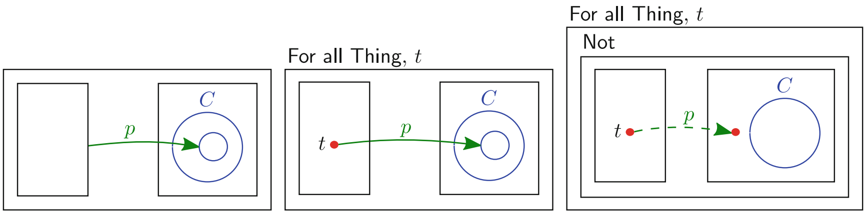


Fig. 6. *Range* patterns

3.6 Range Patterns

The *Range* axiom type also represents a constraint involving a class and a property: Class C is the *Range* of property p if *things* are related, under p, only to elements in C.

Unquantified Pattern. The left-hand diagram of Fig. 6 expresses that the image of property p is a subset of C. Hence, C is the *Range* of p.

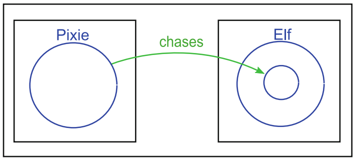
Quantified with Solid Arrow Pattern. The middle diagram of Fig. 6 expresses that for each *Thing* t, the image of property p, when its domain is restricted to t, is a subset of C. Hence, C is the *Range* of p.

Quantified with Dashed Arrow Pattern. The right-hand diagram of Fig. 6 expresses that for each *Thing* t, it is not the case that t is related, under p, to an element not in C. Hence, *things* are related, under p, only to elements in C.

4 Empirical Study

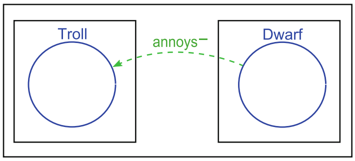
An empirical study was designed to determine which pattern style was more effective overall as well as for each of the four constructs, *All Values From*, *Some Values From*, *Domain* and *Range*. A pattern style was considered more effective than another if, on average, participants interpreted it with significantly fewer errors. If the pattern styles could not be distinguished on error rate then the pattern style that could be interpreted, on average, significantly more quickly was considered the most effective. In order to give some context to the questions used in the empirical study, a case study based on mythical creatures was developed. This context was chosen so that participants would be unable to guess the answers based on prior domain knowledge.

As described above, three different patterns for each of the four constructs were developed, giving a total of twelve different diagram patterns. Participants were shown 24 diagrams in total, with each different diagram pattern being shown twice, representing different information, in order to generate sufficient data points for statistical analysis. Each diagram was associated with a single



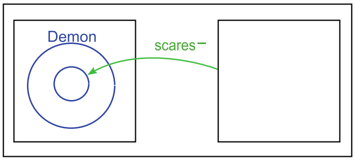
	Options
1	Pixies chase only Elves.
2	Pixies chase at least one Elf.
3	At least one Pixie chases only Elves.
4	At least one Pixie chases Elves.

Fig. 7. Unquantified *All Values From* pattern with multiple-choice answers



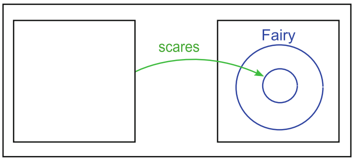
	Options
1	At least one Troll annoys Dwarves.
2	Trolls annoy only Dwarves.
3	At least one Troll annoys only Dwarves.
4	Trolls annoy at least one Dwarf.

Fig. 8. Unquantified *Some Values From* pattern with multiple-choice answers



	Options
1	Only Demons scare things.
2	Not only Demons scare things.
3	Demons scare at least one thing.
4	At least one Demon scare things.

Fig. 9. Unquantified *Domain* pattern with multiple-choice answers



	Options
1	Things scare at least one Fairy.
2	Things scare only Fairies.
3	At least one thing scares Fairies.
4	At least one thing scares only Fairies.

Fig. 10. Unquantified *Range* pattern with multiple-choice answers

question: "What does the diagram tell you?". The participants were provided with four multiple-choice options, presented in random order, exactly one of which was correct; the random order was the same for each participant. Figures 7, 8, 9 and 10 show example questions for each construct from the study, in each case the *Unquantified* pattern is used. The same multiple-choice options were used for each of the questions for each particular construct, but with the names changed to those given in the diagram (each diagram represented different information). The questions were presented in random order, generated uniquely for each participant. We set a time limit of two minutes to answer each question; attempts at the questions in the design phase by members of authors' research group indicated that the time taken to answer each question was usually much less than this. A time limit was deemed important so that the study did not continue indefinitely. We adopted a within-group design because there was unlikely to be any learning effect which could bias the results; each of the patterns has a different appearance and each diagram represented different information.

4.1 Experiment Execution

The experiment was performed within the university's usability laboratory, providing a quiet environment without interruption. Each participant was treated equally with the same environment, furniture, equipment, materials and procedures. Participants performed the experiment individually, and were provided with full details about the purpose of their role by an experiment facilitator who was present throughout.

At the beginning of the experiment, the facilitator introduced the participants to concept diagrams using paper-based training material. Participants were then given software training. They were shown three questions with a similar design to those in the main study in order to help familiarize them with the software's user interface. Finally, the facilitator allowed the participants to work on the study questions. Participants were able to refer to a hard copy single side of A4 paper detailing the elements of concept diagrams used in the study, which formed part of the training material. Upon completion of the experiment, each participant was provided with a debrief summary. Participants were offered a £6 canteen voucher for their time spent in the study, which was approximately 30 min.

A pilot study was conducted to test the experiment design, research software used to display the diagrams and questions, and the data collection process. Five participants (1F, 4M, ages 18–29) took part in the pilot study. As a result of the pilot, a minor change was made to the training material. Forty participants (12F, 28M, ages 18–38) participated in the main experiment, all students from the University of Brighton studying computing, mathematics or engineering. They reported no previous knowledge of concept diagrams, OWL or DL, but were familiar with Venn/Euler diagrams, first order logic and set theory.

5 Results

To determine whether there are differences between the interpretability of the three pattern styles, we analysed both errors and the time taken to answer each question. We performed this analysis on the pattern styles overall and separately for each of the four axiomatized constructs, *All Values From*, *Some Values From*, *Domain* and *Range*. For the errors, we performed chi-square tests. For the time analysis, we performed ANOVAs. However, as the time data were not normally distributed, we performed a log transformation to reduce the skewness to within tolerable levels for conducting robust ANOVAs. When the ANOVAs revealed significant differences, we proceeded to conduct Tukey tests to rank the pattern styles. The results are based on the data collected from 40 participants, with each participant answering 24 questions providing a total of 960 observations, 240 for each of the four axiom types and 80 for each diagrammatic pattern.

5.1 Overall Analysis

To determine which pattern style was most effective overall, we considered how the three pattern styles, *Unquantified*, *Quantified with Solid Arrow* and *Quantified with Dashed Arrow*, compared for the entire 24 questions. Firstly, we compared the error rates for each pattern style, which are summarised in Table 1; these data exclude six timeouts, and, thus, only include data from questions for which an answer was provided within the 2 min allowed. Conducting a chi-square test established that there was no significant difference in error rate between the *Unquantified* (Un) and *Quantified with Solid Arrow* (QwSA) pattern styles ($p = 0.205$). However, both of these pattern styles yielded significantly fewer errors than *Quantified with Dashed Arrow* (QwDA); in each case, $p = 0.000$. We can see that *Quantified with Dashed Arrow* yielded approximately 56 more errors for every 100 answers than *Unquantified*, which falls to 52 more errors as compared to *Quantified with Solid Arrow*.

To further distinguish the pattern styles, we analysed the time data. Consistent with Meulemans et al. [7], we only analyze the correct answers; it can be argued that it does not matter how long it takes to provide a wrong answer. The mean times and standard deviations are summarised in Table 1; these data are from questions for which a *correct* answer was provided within the 2 min allowed. Conducting an ANOVA test established that there were significant differences in the times taken between the three pattern styles ($p = 0.000$). To expose the nature of these differences, we proceeded to conduct a Tukey test in order to rank the pattern styles. This revealed that the *Unquantified* pattern style allowed participants to perform significantly faster than *Quantified with Solid Arrow* which, in turn, was significantly faster than *Quantified with Dashed Arrow*. In terms of time taken, we see that *Unquantified* is approximately 13.5 % faster than *Quantified with Solid Arrow* and approximately 57.9 % faster than *Quantified with Dashed Arrow*, on average.

Combining both our error analysis and time analysis, we conclude that using the *Unquantified* pattern style significantly improves overall task performance, as compared to the other two pattern styles.

Table 1. Overall summary

	Error analysis			Time analysis		
Pattern	N	Errors	Rate %	N	Mean	StDev
Un	320	75	23.44	245	18.40	14.66
QwSA	320	89	27.81	231	20.89	14.85
QwDA	314	250	79.62	64	29.05	21.62

Table 2. *All Values From* summary

	Error analysis			Time analysis		
Pattern	N	Errors	Rate %	N	Mean	StDev
Un	80	5	6.25	75	15.34	11.99
QwSA	80	7	8.75	73	18.57	11.53
QwDA	79	67	84.81	12	28.75	23.38

Table 3. *Some Values From* summary

	Error analysis			Time analysis		
Pattern	N	Errors	Rate %	N	Mean	StDev
Un	80	53	66.25	27	24.20	15.14
QwSA	80	67	83.75	13	38.75	25.83
QwDA	80	52	65.00	28	28.94	23.33

5.2 All Values from Analysis

Table 2 summarizes the error rates for each pattern; these data exclude a single timeout which was for *Quantified with Dashed Arrow*. Conducting a chi-square test showed that there was no significant difference between *Unquantified* and *Quantified with Solid Arrow*, with $p = 0.548$. However, *Unquantified* and *Quantified with Solid Arrow* both yielded significantly fewer errors than *Quantified with Dashed Arrow*, with $p = 0.000$ in each case. *Quantified with Dashed Arrow* yielded approximately 79 more errors for every 100 answers than *Unquantified*, which fell to 76 more errors as compared to *Quantified with Solid Arrow*.

The mean times and standard deviations for each pattern style are given in Table 2. An ANOVA test revealed that there were significant differences ($p = 0.005$) between the pattern styles. A Tukey test indicated that *Unquantified* was significantly faster than *Quantified with Solid Arrow* which, in turn, was significantly faster than *Quantified with Dashed Arrow*. We can see that *Unquantified* was approximately 87.4 % faster than *Quantified with Dashed Arrow* and approximately 21.1 % faster than *Quantified with Solid Arrow*, on average.

Combining the error and time analysis, we again conclude that the *Unquantified* pattern style significantly improves task performance, as compared to the other two pattern styles, in the case of *All Values From*.

5.3 Some Values from Analysis

Table 3 summarizes the error rates for each pattern; there were no timeouts. A chi-square test found no significant difference between *Unquantified* and *Quantified with Dashed Arrow*, with $p = 0.868$. However, *Unquantified* and *Quantified with Dashed Arrow* both yielded significantly fewer errors than *Quantified with Solid Arrow*, with $p = 0.011$ and $p = 0.007$ respectively. *Quantified with Solid Arrow* yielded approximately 18 or 19 more errors for every 100 answers than both *Unquantified* and *Quantified with Dashed Arrow*.

The mean times and standard deviations for each pattern style are given in Table 3. An ANOVA test revealed that there were no significant differences ($p = 0.167$) between the pattern styles. Therefore, we did not proceed to conduct a Tukey test. We conclude, on the basis of the error analysis, that using either the *Unquantified* or *Quantified with Dashed Arrow* best supports task performance for *Some Values From* axioms.

Table 4. *Domain* summary

	Error analysis			Time analysis		
Pattern	N	Errors	Rate %	N	Mean	StDev
Un	80	10	12.50	70	24.52	18.10
QwSA	80	7	8.75	73	24.42	16.41
QwDA	77	70	90.91	7	44.48	23.76

Table 5. *Domain* summary

	Error analysis			Time analysis		
Pattern	N	Errors	Rate %	N	Mean	StDev
Un	80	7	8.75	73	13.54	10.04
QwSA	80	8	10.00	72	16.44	9.84
QwDA	78	61	78.21	17	23.08	13.88

5.4 Domain Analysis

Table 4 summarizes the error rates for each pattern; there were three timeouts, all for *Quantified with Dashed Arrow*. A chi-square test found no significant difference between *Unquantified* and *Quantified with Solid Arrow*, with $p = 0.442$. However, *Unquantified* and *Quantified with Solid Arrow* both yielded significantly fewer errors than *Quantified with Dashed Arrow*, with $p = 0.000$ in each case. *Quantified with Dashed Arrow* yield approximately 82 more errors for every 100 answers than *Quantified with Solid Arrow* which slightly reduces to 78 more errors as compared to *Unquantified*.

The mean times and standard deviations for each pattern style are given in Table 4. An ANOVA test revealed that there were significant differences ($p = 0.041$) between the pattern styles. Therefore, we conducted a Tukey test, which ranked the pattern styles as follows: *Unquantified* and *Quantified with Solid Arrow* were not significantly different, but both were significantly faster than *Quantified with Dashed Arrow*. We can see that *Unquantified* and *Quantified with Solid Arrow* were approximately 81.4 % and 82.1 %, respectively, faster than *Quantified with Dashed Arrow*, on average.

Our error and time analysis consistently support the use of either *Unquantified* and *Quantified with Solid Arrow* over *Quantified with Dashed Arrow* for *Domain* axioms.

5.5 Range Analysis

Table 5 summarizes the error rates for each pattern; there were two timeouts, both for *Quantified with Dashed Arrow*. A chi-square test found no significant difference between *Unquantified* and *Quantified with Solid Arrow*, with $p = 0.786$. Again, we found that both *Unquantified* and *Quantified with Solid Arrow* yielded significantly fewer errors than *Quantified with Dashed Arrow*, with $p = 0.000$ in each case. *Quantified with Dashed Arrow* yield approximately 68 or 69 more errors for every 100 answers than *Unquantified* and *Quantified with Solid Arrow*.

The mean times and standard deviations for each pattern style are given in Table 5. An ANOVA test revealed that there were significant differences ($p = 0.001$) between the pattern styles. A Tukey test ranked the patterns as follows: *Unquantified* was significantly faster than *Quantified with Solid Arrow* which, in turn, was significantly faster than *Quantified with Dashed Arrow*. Participants' performance using the *Unquantified* pattern was approximately 70 % faster than

Quantified with Dashed Arrow and approximately 21.4 % faster than *Quantified with Solid Arrow*, on average.

Drawing on the error and time analysis, for *Range* using the *Unquantified* pattern most effectively supports user task performance and *Quantified with Solid Arrow* is placed second.

5.6 Summary of Analysis

As well as being ranked as the most effective overall, the *Unquantified* pattern style allows participants to perform at least as well, if not significantly better, than both the other pattern styles for each individual axiom type. Interestingly, in all but one case – namely *Some Values From* – *Quantified with Dashed Arrow* was ranked last by both errors and time taken. This indicates that using quantification with dashed arrows is particularly poor for cognition and an overall error rate of 79.62 % is not dissimilar to what is expected when randomly choosing one out of four options. Given this and that the overall error rates for the other two patterns styles are much lower, being 23.44 % for *Unquantified* and 27.81 % for *Quantified with Solid Arrow*, it is surprising that *lowest* error rate for *Quantified with Dashed Arrow* is for the axiom type *Some Values From* at 65 %. The other two styles have, by far, their *highest* error rates for this axiom type, namely 66.25 % and 83.75 %. These high error rates are, however, consistent with findings for symbolic approaches to ontology engineering, where it has been established that users have particular difficulty understanding *Some Values From* axioms [5,9–11,16]. We will further discuss this observation in Sect. 6.

6 Discussion

There are some interesting observations to be made from the results of the empirical study. In particular, the results for *Some Values From* are striking, with an overall error rate of 72.67 %. As just stated, it is well established that users have difficulties with this construct, so it may be the inherent conceptual difficulty of this axiom type that causes the high error rate. For example, Rector et al. [9,10] showed that new OWL students do not understand the exact meaning of *Some Values From*, and "are unsure if it means *all*, *any* or *nothing else*".

Delving deeper into the results of our study, we analysed the incorrect responses for the *Some Values From* construct. For the *Quantified with Solid Arrow* pattern 60 out of 80 (75 %) participants confused *Some Values From* with *All Values From*, for example, choosing "Trolls recruit only Goblins" rather than "Trolls recruit at least one Goblin". This is in agreement with other studies that report that users confuse *Some Values From* with *All Values From* [9,10]. In particular, one of the common logical errors made by ontology users is that C_1 has *All Values From* C_2 implies C_1 has *Some Values From* C_2. Furthermore, Schwitter and Tilbrook [11] showed that one of the common errors new OWL users make is to use the universal restriction *All Values From* as a default, when the existential restriction *Some Values From* actually applies. Interestingly, in our study, confusing *Some Values From* with *All Values From* was not the case for the *Unquantified* and *Quantified with Dashed Arrow* patterns: only three out of 80 (3.7 %) in both cases chose the *All Values From* option.

Other studies have also shown description logic users may interpret *Some Values From* incorrectly; for example, considering the pizza ontology, many users initially read, 'Pizza hasTopping MozzarellaTopping' to mean "some pizzas have toppings that are mozzarella topping", rather than the correct reading, "all pizzas have toppings that are some mozzarella topping" [5]. In our study, 41 out of 80 (51.25 %) for the *Unquantified* pattern and 36 out of 80 (45 %) for the *Quantified with Dashed Arrow* pattern made the same mistake as reported in the pizza example, choosing, for example, 'at least one Troll recruits Goblins' (equivalent to 'some Trolls recruit Goblins'). The reasons for this kind of misunderstanding are not clear, although it may have been that participants associated 'at least one' with the wrong class.

Moving on to consider the other constructs, it is surprising that there were differences in the results for *Domain* and *Range* in that they are diagrammatically 'mirror images' of each other. The difference, that there was a statistically significant best pattern for *Range* but not for *Domain*, could be explained by the use of the conceptually more difficult inverse property in *Domain*.

The results for the *Quantified with Dashed Arrow* pattern style were also striking, with an overall error rate of 79.62 %. This seems to imply that using a dashed arrow may be difficult to interpret. However, it may not be the dashed arrow but the other features of these patterns that cause problems. Three of the *Quantified with Dashed Arrow* patterns used explicit negation; no other pattern style used negation. All of the *Quantified with Dashed Arrow* patterns that involved negation performed badly, with error rates of 83.75 % for *All Values From*, 90.91 % for *Domain* and 78.21 % for *Range*. All six of the timeouts were for patterns involving negation, and each pattern involving negation had at least one timeout. By contrast, the *Quantified with Dashed Arrow* pattern that did not involve negation, for *Some Values From*, had the lowest absolute error rate at 65.00 % for this construct. In cognitive psychology, it is well-known that human reasoning with negation is harder than reasoning without [17]. We therefore conjecture that negation is a major contributor to the poor task performance observed for the *Quantified with Dashed Arrow* pattern styles.

Other factors may have influenced the results. The relative complexity of the diagrams could have an effect on performance. The two quantified styles are diagrammatically more complex than the unquantified style. They could also be considered as heterogeneous, in that they contain textual notation, rather than purely diagrammatic. This may be why the unquantified patterns are easier to deal with cognitively. Similarly, the unquantified styles may be better matched to meaning than the quantified styles. Little work has been carried out on well-matchedness in diagrammatic notations more expressive than Euler diagrams.

7 Conclusion

The aim of this paper was to provide insight into how to choose between syntactically different diagrams when formulating logical axioms, particularly from the perspective of ontology engineering. In the context of this empirical study, we conclude that avoiding explicit quantification and representing the information purely diagrammatically best supports task performance. As a result, we recommend that users and designers of diagrammatic logics, and in particular ontology engineers, avoid using explicit quantification where possible.

Having made this recommendation, it is important to determine whether there really is an advantage in using diagrammatic patterns over standard notations in ontology engineering. Further work is needed to empirically evaluate the recommended patterns from this paper, that is the *Unquantified* patterns for *Subsumption*, *Disjointness*, *All Values From*, *Some Values From*, *Domain* and *Range*, with equivalent axioms expressed in OWL and description logic. This will allow us to determine whether there is an advantage in performance when using these diagrammatic patterns over equivalent textual or symbolic representations. Further work is also required to determine whether it is negation that is causing poor task performance.

References

1. The OWL 2 Web Ontology Language (2016). http://www.w3.org/TR/owl2-overview/. Accessed April 2016
2. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P.F. (eds.): The Description Logic Handbook. Cambridge University Press, Cambridge (2003)
3. Chapman, P., Stapleton, G., Rodgers, P., Micallef, L., Blake, A.: Visualizing sets: an empirical comparison of diagram types. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 146–160. Springer, Heidelberg (2014)
4. Gurr, C.: Effective diagrammatic communication: syntactic, semantic and pragmatic issues. J. Vis. Lang. Comput. **10**(4), 317–342 (1999)
5. Horridge, M., Drummond, N., Goodwin, J., Rector, A.L., Stevens, R., Wang, H.: The Manchester OWL syntax. In: OWLed, vol. 216 (2006)
6. Howse, J., Stapleton, G., Taylor, K., Chapman, P.: Visualizing ontologies: a case study. In: Aroyo, L., Welty, C., Alani, H., Taylor, J., Bernstein, A., Kagal, L., Noy, N., Blomqvist, E. (eds.) ISWC 2011, Part I. LNCS, vol. 7031, pp. 257–272. Springer, Heidelberg (2011)
7. Meulemans, W., Henry Riche, N., Speckmann, B., Alper, B., Alper, B., Dwyer, T.: Kelpfusion: a hybrid set visualization technique. IEEE Trans. Vis. Comput. Graph. **19**(11), 1846–1858 (2013)
8. Peirce, C.: Collected Papers, vol. 4. Harvard University Press, Cambridge (1933)
9. Rector, A., Drummond, N., Horridge, M., Rogers, J., Knublauch, H., Stevens, R., Wang, H., Wroe, C.: OWL pizzas: practical experience of teaching OWL-DL: common errors & common patterns. In: Motta, E., Shadbolt, N.R., Stutt, A., Gibbins, N. (eds.) EKAW 2004. LNCS, vol. 3257, pp. 63–81. Springer, Berlin (2004)
10. Rector, A.L., Drummond, N., Horridge, M., Rogers, J., Knublauch, H., Stevens, R., Wang, H., Wroe, C.: Designing user interfaces to minimise common errors in ontology development: the CO-ODE and HyOntUse projects. In: Proceedings of the UK e-Science All Hands Meeting, vol. 2004, pp. 493–499 (2004)

11. Schwitter, R., Tilbrook, M.: Controlled natural language meets the semantic web. In: Proceedings of Australasian Language Technology Workshop, vol. 2004, pp. 55–62 (2004)
12. Shin, S.J.: The Logical Status of Diagrams. Cambridge UniversityPress, Cambridge (1994)
13. Shin, S.J.: The Iconic Logic of Peirce's Graphs. Bradford Book, Cambridge (2002)
14. Stapleton, G., Howse, J., Chapman, P., Delaney, A., Burton, J., Oliver, I.: Formalizing concept diagrams. In: Visual Languages and Computing, pp. 182–187. Knowledge Systems Institute (2013)
15. Stapleton, G., Howse, J., Taylor, K., Delaney, A., Burton, J., Chapman, P.: Towards diagrammatic ontology patterns. In: 4th Workshop on Ontology and Semantic Web Patterns, CEUR, Sydney, Australia, October 2013
16. Warren, P., Mulholland, P., Collins, T., Motta, E.: The usability of description logics. In: Presutti, V., d'Amato, C., Gandon, F., d'Aquin, M., Staab, S., Tordai, A. (eds.) ESWC 2014. LNCS, vol. 8465, pp. 550–564. Springer, Heidelberg (2014)
17. Wason, P., Johnson-Laird, P.: Psychology of Reasoning: Structure and Content. Harvard University Press, Cambridge (1972)

The Interaction Between Logic and Geometry in Aristotelian Diagrams

Lorenz Demey[1(✉)] and Hans Smessaert[2]

[1] Center for Logic and Analytic Philosophy, KU Leuven, Leuven, Belgium
Lorenz.Demey@hiw.kuleuven.be
[2] Department of Linguistics, KU Leuven, Leuven, Belgium
Hans.Smessaert@arts.kuleuven.be

Abstract. We develop a systematic approach for dealing with informationally equivalent Aristotelian diagrams, based on the interaction between the logical properties of the visualized information and the geometrical properties of the concrete polygon/polyhedron. To illustrate the account's fruitfulness, we apply it to all Aristotelian families of 4-formula fragments that are closed under negation (comparing square and rectangle) and to all Aristotelian families of 6-formula fragments that are closed under negation (comparing hexagon and octahedron).

Keywords: Aristotelian diagram · Logical geometry · Square of oppositions · Hexagon · Octahedron · Cross-polytope · Symmetry group

1 Introduction

Aristotelian diagrams are compact visual representations of the elements of some logical or conceptual field, and the logical relations holding between them. These diagrams have a long and rich history in philosophical logic [26]. Today, they are still widely used in logic [11,24], but also in fields such as cognitive science, linguistics, philosophy, neuroscience, law and computer science [17,19,21] (see [12, Sect. 1] for more examples). Aristotelian diagrams have thus come to serve "as a kind of *lingua franca*" [20, p. 81] for a highly interdisciplinary community of researchers concerned with logical reasoning. *Logical geometry*[1] systematically investigates Aristotelian diagrams as objects of independent interest (regardless of their role as lingua franca), exploring various abstract-logical topics [12,14, 15,33] as well as some more visual-geometrical issues [10,16,31,32].

One of the major visual-geometrical issues studied in logical geometry is the fact that a single logical structure often gives rise to a wide variety of different visualizations. In other words, even after all the strictly logical 'parameters' of a structure have been fixed, one is still confronted with several design choices when drawing the actual Aristotelian diagram for that structure. This phenomenon is widely manifested in the extant literature: there are numerous cases

[1] See www.logicalgeometry.org.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 67–82, 2016.
DOI: 10.1007/978-3-319-42333-3_6

of authors who use completely different Aristotelian diagrams to visualize one and the same underlying logical structure (concrete examples will be provided below). Furthermore, since authors typically use Aristotelian diagrams to help their readers gain a better insight into some underlying logical structure, the fact that a single structure can be visualized by means of different diagrams naturally raises the question whether some of these diagrams are perhaps more 'effective' (i.e. have a greater positive impact on readers' comprehension) than others.

This issue has already been partially addressed in other work. For example, [13] compares different Aristotelian diagrams (a hexagon and an octhaedron) for the Boolean algebra $\mathbb{B}_3$, and examines their geometrical connections with the Hasse diagram for $\mathbb{B}_3$. Similarly, [35] compares two Aristotelian diagrams (a rhombic dodecahedron and a nested tetrahedron) for the Boolean algebra $\mathbb{B}_4$. These existing studies have a number of limitations, however: on the logical side, they are restricted to structures that are Boolean closed (i.e. that constitute entire Boolean algebras), and on the geometrical side, they are restricted to comparing Aristotelian diagrams that are based on different geometric shapes.

The main aim of the present paper is therefore to propose and illustrate a new approach to systematically investigate different Aristotelian diagrams for a given underlying logical structure. We will show that this approach does not suffer from the limitations present in other work: logically speaking, it applies to structures that are Boolean closed as well as to structures that are not Boolean closed, and geometrically speaking, it applies to Aristotelian diagrams that are based on different geometric shapes as well as to Aristotelian diagrams that are based on the same geometric shape. The key idea of the new approach is that for any given set of logical formulas $\mathcal{F}$, one can calculate a number $\ell(\mathcal{F})$ based on strictly *logical* considerations; similarly, for any concrete Aristotelian diagram $\mathcal{P}_\mathcal{F}$ that visualizes $\mathcal{F}$, one can calculate a number $g(\mathcal{P}_\mathcal{F})$ based on strictly *geometrical* properties. The interaction between $\ell(\mathcal{F})$ and $g(\mathcal{P}_\mathcal{F})$ will turn out to be very informative about the quality of $\mathcal{P}_\mathcal{F}$ as a visualization of $\mathcal{F}$.

The paper is organized as follows. Section 2 introduces some basic notions from logical geometry, and explains the distinction between informational and computational equivalence of Aristotelian diagrams. Section 3 then discusses the interaction between logical and geometrical properties of Aristotelian diagrams on a wholly general level. Next, Sects. 4 and 5 investigate the concrete details of this logico-geometrical interaction in Aristotelian diagrams with 4 and 6 formulas, respectively. Finally, Sect. 6 summarizes the results obtained in this paper, and discusses the advantages as well as the limitations of the logico-geometrical perspective.

2 Informational and Computational Equivalence

Given a logical system S and a set $\mathcal{F}$ of formulas from that system, an Aristotelian diagram for $\mathcal{F}$ in S is a diagram in which the formulas of $\mathcal{F}$ and the Aristotelian relations holding between those formulas are visualized by means of

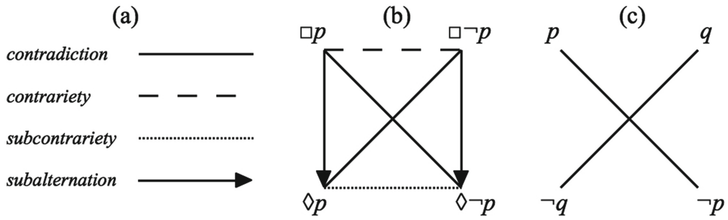


Fig. 1. (a) Visual code for the Aristotelian relations, (b) classical square for formulas from the modal logic S5 ($\Box$ and $\Diamond$ should be read as 'necessarily' and 'possibly', respectively), (c) degenerated square for formulas from propositional logic.

points and lines connecting those points, respectively. The Aristotelian relations are defined as follows: two formulas φ and ψ are said to be

S-contradictory	iff	$\mathsf{S} \models \neg(\varphi \wedge \psi)$	and	$\mathsf{S} \models \varphi \vee \psi$,
S-contrary	iff	$\mathsf{S} \models \neg(\varphi \wedge \psi)$	and	$\mathsf{S} \not\models \varphi \vee \psi$,
S-subcontrary	iff	$\mathsf{S} \not\models \neg(\varphi \wedge \psi)$	and	$\mathsf{S} \models \varphi \vee \psi$,
in S-subalternation	iff	$\mathsf{S} \models \varphi \rightarrow \psi$	and	$\mathsf{S} \not\models \psi \rightarrow \varphi$.

Informally, the first three relations are concerned with whether the formulas can be true/false together, whereas the fourth relation is concerned with truth propagation [33]. These relations will be visualized using the code shown in Fig. 1(a). Finally, two formulas are said to be *unconnected* iff they do not stand in any Aristotelian relation at all.

The contemporary literature on Aristotelian diagrams has mainly focused on the *logical* aspects of these diagrams. This is clearly manifested in the classification of Aristotelian diagrams into different families. For example, with respect to 4-formula-diagrams, we distinguish between the 'classical square of oppositions' and the 'degenerated square', as shown in Fig. 1(b) and (c), respectively. Similarly, with respect to 6-formula-diagrams, we distinguish between the 'Jacoby-Sesmat-Blanché (JSB) hexagon', the 'Sherwood-Czężowski (SC) hexagon' and the 'U4 hexagon' (among others), as shown in Fig. 2(a), (b) and (c), respectively.[2] The differences between these families of Aristotelian diagrams are all based on their logical properties. First of all, different families often have different Aristotelian relations; e.g. the classical square does not contain pairs of unconnected formulas, whereas the degenerated square contains 4 unconnected pairs. Secondly, different families may have different constellations of Aristotelian relations; e.g. the JSB hexagon and the SC hexagon both have 3 contrarieties, but in the former they constitute a closed triangle, whereas in the latter they do not [31]. Thirdly, different families often have a different Boolean structure; e.g. the Boolean closure of the classical square is (isomorphic to) $\mathbb{B}_3$, whereas the Boolean closure of the degenerated square is (isomorphic to) $\mathbb{B}_4$ [14].

Using terminology from Larkin and Simon [23], the Aristotelian diagrams in Figs. 1(b–c) and 2(a–c) are not *informationally equivalent*. They visualize

[2] See [15] for some historical background on this nomenclature.

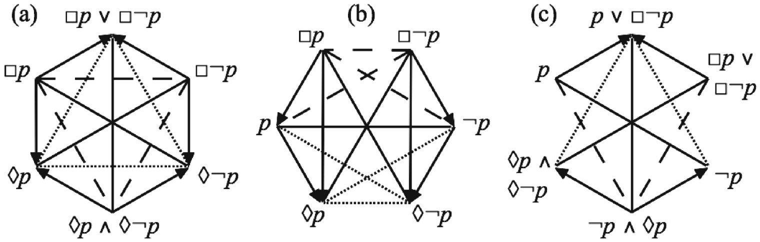


Fig. 2. (a) JSB hexagon, (b) SC hexagon and (c) U4 hexagon for formulas from S5.

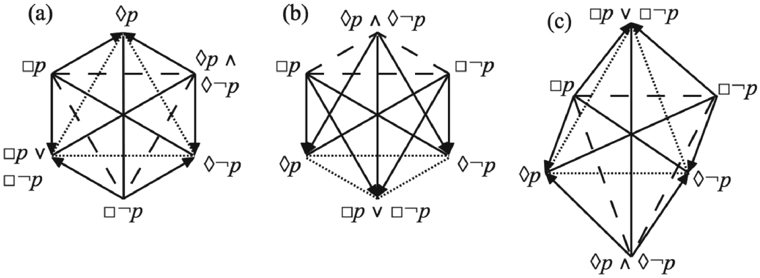


Fig. 3. Three visual alternatives to the JSB hexagon in Fig. 2(a).

different logical structures—i.e. different chunks of logical information—, and hence, the differences between these diagrams are entirely due to differences between their underlying logical structures.

A completely different type of question arises when we decide to focus on a single logical structure (i.e. one set of formulas in one logical system), and investigate the various Aristotelian diagrams that have been used to visualize this single structure. For example, given four formulas that constitute a classical square, this square is usually drawn as in Fig. 1(b), but it has also been drawn with the subalternations pointing upwards, from left to right, etc. [20]. Similarly, given six formulas that constitute a JSB hexagon, this hexagon is usually drawn as in Fig. 2(a), but it has also been drawn as shown in Fig. 3(a) [29] or Fig. 3(b) [5,25]. This structure has also been visualized by means of a different geometric shape altogether, viz. an octahedron, as shown in Fig. 3(c) [22,30].

Again using terminology from Larkin and Simon [23], the Aristotelian diagrams in Figs. 2(a) and 3(a–c) are informationally equivalent (they are different visualizations of one and the same logical structure), but they are not *computationally equivalent*. After all, even though the visual differences between these diagrams are irrelevant from a strictly logical perspective, they can significantly influence the diagrams' effectiveness in increasing user comprehension.

3 Logic Versus Geometry in Aristotelian Diagrams

We will now present a general approach to study the interaction between logical and geometrical properties of informationally equivalent Aristotelian diagrams.

We will focus exclusively on fragments $\mathcal{F}$ that are *closed under negation*, i.e. if $\varphi \in \mathcal{F}$, then also $\neg\varphi \in \mathcal{F}$ (up to logical equivalence). Such fragments always have an even number of formulas, and it will be fruitful to view them not only as consisting of $2n$ formulas, but also as n "pairs of contradictory formulas" (PCDs). Additionally, we will only deal with Aristotelian diagrams in which negation is visually represented by means of *central symmetry*, so that φ and $\neg\varphi$ correspond to diametrically opposed points in the diagram. It should be emphasized that both the logical condition (closed under negation) and the geometrical condition (central symmetry) are satisfied in nearly every Aristotelian diagram that has ever been produced,[3] and are thus very mild restrictions.

We know from basic combinatorics that a fragment of $2n$ formulas (i.e. n PCDs) can be ordered in exactly $(2n)!$ ways.[4] However, this number does not take into account the fragment's PCD-structure, in the sense that a formula and its negation are not treated any differently from any other pair of formulas. If we only consider orderings that respect the fragment's PCD-structure, we find the number $2^n \times n!$. On the one hand, there are n PCDs to be ordered, which yields the second factor ($n!$); on the other hand, each of these n PCDs has 2 'orientations', viz. $(\varphi, \neg\varphi)$ and $(\neg\varphi, \varphi)$, thus yielding the first factor (2^n). Note that this formula is strictly based on the *logical* properties of the fragment, viz. the facts that it contains $2n$ formulas and is closed under negation.

A fragment of n PCDs can be visualized by means of a polygon or polyhedron that has $2n$ vertices and is centrally symmetric (so that the diagram's vertices correspond to the fragment's formulas, and the diagram's central symmetry corresponds to the fragment's PCD-structure).[5] Each such polygon/polyhedron $\mathcal{P}$ has a number of reflectional and rotational symmetries, which constitute a group under the composition operation. This group is the *symmetry group* of $\mathcal{P}$ and will be denoted $\mathcal{S}_\mathcal{P}$ [28, p. 67]. Its cardinality $|\mathcal{S}_\mathcal{P}|$ measures how symmetric $\mathcal{P}$ is, and is thus strictly based on the *geometrical* properties of the polygon/polyhedron.

We now turn to the interaction between the numbers $2^n \times n!$ and $|\mathcal{S}_\mathcal{P}|$. First of all, it should be noted that the former is typically larger than the latter, since every symmetry of $\mathcal{P}$ can also be seen as the result of permuting and changing the orientation of the PCDs that are visualized by $\mathcal{P}$, but not vice versa. Note, for example, that the hexagon in Fig. 3(a) can be seen as the result of reflecting the hexagon in Fig. 2(a) around the axis defined by $\Box p$ and $\Diamond\neg p$, but it can equally validly be seen as the result of permuting the PCDs $(\Diamond p, \Box\neg p)$ and $(\Box p \vee \Box\neg p, \Diamond p \wedge \Diamond\neg p)$ in the latter hexagon. By contrast, note that the hexagon in Fig. 3(b) can be seen as the result of changing the orientation of the PCD

[3] One counterexample is Chow [7], who studies Aristotelian diagrams that satisfy the logical condition, but *not* the geometrical condition.

[4] In [20, p. 77] this formula is applied to a fragment of 4 formulas (so $n = 2$).

[5] In this paper, we will mainly focus on *regular* polygons and polyhedra (the only exception being the brief discussion of rectangles in Sect. 4). However, this restriction is only made for reasons of space; in principle, the account presented here can be applied to regular and non-regular shapes alike.

$(\Box p \vee \Box\neg p, \Diamond p \wedge \Diamond\neg p)$ in the hexagon in Fig. 2(a), but that it is *not* the result of applying any reflection or rotation to the latter hexagon.

The key idea is now that the $2^n \times n!$ different ways of ordering n PCDs can be partitioned based on whether they yield variants of $\mathcal{P}$ that can be obtained from each other via reflections or rotations.[6] This partition has $\frac{2^n \times n!}{|\mathcal{S}_\mathcal{P}|}$ cells, which will be called *fundamental forms*. It follows immediately that diagrams with different fundamental forms are not reflectional or rotational variants of each other; and each fundamental form yields exactly $|\mathcal{S}_\mathcal{P}|$ diagrams that are all reflectional or rotational variants of each other. For example, the hexagons in Figs. 2(a) and 3(a) have the same fundamental form, whereas the hexagons in Figs. 2(a) and 3(b) have different fundamental forms.

Suppose now that we have two distinct polygons/polyhedra $\mathcal{P}$ and $\mathcal{P}'$ that visualize the same $2n$-formula fragment $\mathcal{F}$. Suppose, furthermore, that $\mathcal{P}$ is less symmetric than $\mathcal{P}'$. This means that $|\mathcal{S}_\mathcal{P}| < |\mathcal{S}_{\mathcal{P}'}|$, and hence $\frac{2^n \times n!}{|\mathcal{S}_\mathcal{P}|} > \frac{2^n \times n!}{|\mathcal{S}_{\mathcal{P}'}|}$,[7] i.e. $\mathcal{P}$ has more fundamental forms than $\mathcal{P}'$. In other words, by having fewer symmetries, $\mathcal{P}$ makes a number of visual distinctions that are not made by $\mathcal{P}'$. The quality of $\mathcal{P}$ and $\mathcal{P}'$ as Aristotelian diagrams for the fragment $\mathcal{F}$ depends on whether these visual distinctions correspond to any logical distinctions in $\mathcal{F}$. On the one hand, if there are such logical distinctions present in $\mathcal{F}$, then $\mathcal{P}$ is to be preferred over $\mathcal{P}'$, since $\mathcal{P}$ allows us to visualize these logical distinctions by mapping them onto the visual distinctions of its fundamental forms, whereas $\mathcal{P}'$ would simply force us to collapse them. On the other hand, if there are no such logical distinctions present in $\mathcal{F}$, then $\mathcal{P}'$ is to be preferred over $\mathcal{P}$, since in this case, the visual distinctions between the fundamental forms of $\mathcal{P}$ do not correspond to any logical differences in $\mathcal{F}$, but are merely by-products of the lack of symmetry in $\mathcal{P}$.[8]

4 Aristotelian Diagrams with Two PCDs

We will now apply the logico-geometrical account presented in the previous section to Aristotelian diagrams for fragments consisting of 2 PCDs. On the logical side, one can show that the 2-PCD Aristotelian diagrams can be classified into

[6] The idea of working *up to symmetry* can already be found in [4, p. 315], where it is stated that Aristotelian squares that are symmetrical variants of each other should be "counted as being of the same type". The assumed irrelevance of symmetry considerations for diagram design is also in line with work on other types of diagrams, such as Euler diagrams [27]: several of their visual characteristics have been investigated [1,3], but it has been found that rotation has no significant influence on user comprehension of Euler diagrams [2].

[7] Note that both fractions have the same *numerator* (since the two Aristotelian diagrams have the same *logical* properties, viz. they both visualize the fragment $\mathcal{F}$), but different *denominators* (since the two diagrams have different *geometrical* properties, viz. $\mathcal{P}$ is less symmetric than $\mathcal{P}'$).

[8] These considerations can be viewed as an application of the congruity/isomorphism principle in diagram design [18,36]: the visual properties of the diagram should closely correspond to the logical properties of the visualized fragment.

exactly 2 Aristotelian families, viz. *classical* and *degenerated* [32,33]. On the geometrical side, visualizing such 2-PCD diagrams requires a polygon/polyhedron that has 4 vertices and is centrally symmetric. In this section, we will focus on two such polygons, viz. the *square* and the *rectangle* (also recall Footnote 5).

Logically speaking, a 2-PCD fragment can be ordered in $2^2 \times 2! = 8$ distinct ways. Geometrically speaking, the symmetry group $\mathcal{S}_{sq}$ of a square has order 8, whereas the symmetry group $\mathcal{S}_{rect}$ of a proper (i.e. non-square) rectangle has order 4. This difference reflects the fact that a square is a more symmetrical shape than a (proper) rectangle, since a rectangle distinguishes between its long and short edges, whereas the square collapses this distinction (by having 4 edges of the same length). Consequently, when a 2-PCD fragment is visualized by means of a square, this yields $\frac{2^2 \times 2!}{|\mathcal{S}_{sq}|} = \frac{8}{8} = 1$ fundamental form; by contrast, when it is visualized by means of a rectangle, this yields $\frac{2^2 \times 2!}{|\mathcal{S}_{rect}|} = \frac{8}{4} = 2$ fundamental forms.

Because it is less symmetrical, the rectangle makes more visual distinctions (long/short edges) than the square. In order to determine which shape is the most effective visualization of a 2-PCD fragment, we should investigate whether these visual distinctions correspond to any logical distinctions in the fragment. We will now do this for each of the two Aristotelian families of 2-PCD fragments.

Classical 2-PCD Fragments. Visualizing a classical 2-PCD fragment using a square yields 1 fundamental form; see e.g. Fig. 1(b). This means that all oriented permutations of the 2 PCDs yield diagrams that are rotational or reflectional variants of each other, regardless of where the (sub)contrarieties and subalternations are in the diagram. By contrast, visualizing this fragment by means of a (proper) rectangle yields 2 fundamental forms, as shown in Fig. 4(a–b). In the first fundamental form, the (sub)contrarieties occupy the rectangle's long edges and the subalternations occupy its short edges, whereas in the second fundamental form it is the other way around.

Some authors have claimed that there is an important logical difference between the Aristotelian relations of (sub)contrariety on the one hand, and subalternation on the other. They distinguish between two complementary perspectives on the classical square[9] of opposition: as a theory of negation and as a theory of logical consequence [33]. The former focuses on (sub)contrariety, while the latter focuses on subalternation. Furthermore, it has been argued that these two perspectives are linked to different scholarly traditions of Aristotle's logical works: the former is mainly found in commentaries on *De Interpretatione*, whereas the latter is central in commentaries on the *Prior Analytics* [8].

If a classical 2-PCD fragment is visualized by means of a *rectangle*, this logical distinction can directly be visualized, by putting the (sub)contrarieties and subalternations on edges of different lengths. For example, if one primarily focuses on the theory of negation, then one can put the (sub)contrarieties on the long edges, thus giving them more visual prominence [37, pp. 515–516],

[9] We are using the term 'square' in a strictly historical sense here, regardless of its concrete geometrical properties.

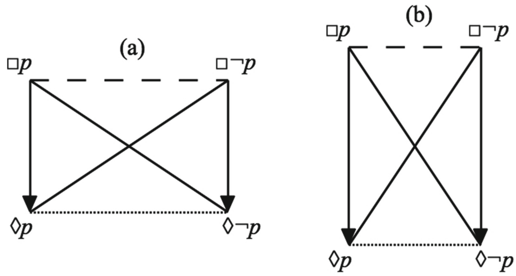


Fig. 4. The two fundamental forms of a (proper) rectangle for the classical 2-PCD fragment of S5-formulas that was already visualized by means of a square in Fig. 1(b).

while if one's focus is on the theory of consequence, one should rather put the subalternations on the long edges. By contrast, if the fragment is visualized using a *square*, then the distinction between (sub)contrariety and subalternation cannot be visualized in this way, since the square's edges are all of the same length.

Next to those who focus on the differences between (sub)contrariety and subalternation, there are also authors who rather emphasize the logical unity of these relations. They point out, for example, that every (sub)contrariety gives rise to two subalternations, and every subalternation gives rise to a contrariety and a subcontrariety [15,33].[10] This has important consequences for the optimal visualization of a classical 2-PCD fragment. If the fragment is visualized by means of a *square*, then the unity of (sub)contrariety and subalternation is visualized by putting them all on edges of the same length. By contrast, if the fragment is visualized by means of a *rectangle*, then one will be forced to put either the (sub)contrarieties or the subaltnerations on the rectangle's long edges; however, this visual difference is not motivated by any logical considerations, but is merely a by-product of the lack of symmetry in the rectangle.

In sum: whether a square or a rectangle is the most suitable diagram for visualizing a classical 2-PCD fragment depends on one's logical views. If one focuses on the *differences* between the Aristotelian relations of (sub)contrariety and subalternation, then the rectangle is the optimal diagram, but if one rather focuses on the *unity* between those relations, then the square seems most suitable.

Degenerated 2-PCD Fragments. The formulas in a degenerated 2-PCD fragment are all pairwise unconnected (except for the two pairs of contradictory formulas, of course). Because of the strictly negative characterization of unconnectedness (absence of all Aristotelian relations), there do not seem to be any

[10] In particular: (i) a contrariety between φ and ψ yields subalternations from φ to $\neg\psi$ and from ψ to $\neg\varphi$; (ii) a subcontrariety between φ and ψ yields subalternations from $\neg\varphi$ to ψ and from $\neg\psi$ to φ; (iii) a subalternation from φ to ψ yields a contrariety between φ and $\neg\psi$ and a subcontrariety between $\neg\varphi$ and ψ.

logical grounds for further differentiating between these pairs of unconnected formulas. Consequently, if the fragment is visualized by means of a *square*, then the equal logical status of the four pairs of unconnected formulas is visually represented by the fact that they are all on edges of the same length; see e.g. Fig. 1(c). By contrast, if one were to visualize the fragment using a *rectangle*, then one would be forced to choose two unconnected pairs to put on the rectangle's long edges, without having any logical motivation for doing so. The optimal diagram for visualizing a degenerated 2-PCD fragment thus seems to be a square, rather than a rectangle.

5 Aristotelian Diagrams with Three PCDs

In this section we continue our exploration of the logico-geometrical account presented in Sect. 3, by applying it to Aristotelian diagrams for fragments consisting of 3 PCDs. On the logical side, one can show that the 3-PCD Aristotelian diagrams can be classified into exactly 5 Aristotelian families, viz. *JSB*, *SC*, *U4*, *U8* and *U12* [31,33]. On the geometrical side, visualizing such 3-PCD diagrams requires a centrally symmetric polygon/polyhedron with 6 vertices. We will consider two such shapes, viz. the *hexagon* (2D) and the *octahedron* (3D).

Logically speaking, a 3-PCD fragment can be ordered in $2^3 \times 3! = 48$ distinct ways. Geometrically speaking, the symmetry group $\mathcal{S}_{hex}$ of a hexagon has order 12, whereas the symmetry group $\mathcal{S}_{oct}$ of an octahedron has order 48. This difference reflects the fact that an octahedron is higher-dimensional than a hexagon, which allows it to have more symmetries (viz. one additional rotation axis). Consequently, when a 3-PCD fragment is visualized by means of a hexagon, this yields $\frac{2^3 \times 3!}{|\mathcal{S}_{hex}|} = \frac{48}{12} = 4$ fundamental forms; by contrast, when it is visualized by means of an octahedron, this yields $\frac{2^3 \times 3!}{|\mathcal{S}_{oct}|} = \frac{48}{48} = 1$ fundamental form.

Because it is less symmetrical, the hexagon makes more visual distinctions than the octahedron. In order to determine which shape is the most effective visualization of a 3-PCD fragment, we should investigate whether these visual distinctions correspond to any logical distinctions in the fragment. This is exactly what we will do next, for each of the 5 Aristotelian families of 3-PCD fragments.[11]

JSB 3-PCD Fragments. Visualizing a JSB fragment by means of an octahedron yields 1 fundamental form; see e.g. Fig. 3(c). By contrast, visualizing it by means of a hexagon yields 4 fundamental forms; see e.g. Figs. 2(a) and 5(a–c). In the first fundamental form, the 3 lines connecting the contrary formulas are all equally long, and thus constitute an equilateral triangle. In the other three fundamental forms, one line of contrariety is longer than the other two, yielding a (proper) isosceles triangle.

If the JSB fragment being visualized is (isomorphic to) the Boolean algebra $\mathbb{B}_3$ (except for its $\top$- and $\bot$-elements), then its 3 pairwise contrary formulas

[11] For reasons of space, our discussion of the visualizations of these 5 families will be fairly brief; however, much more can (and should) be said about each of them.

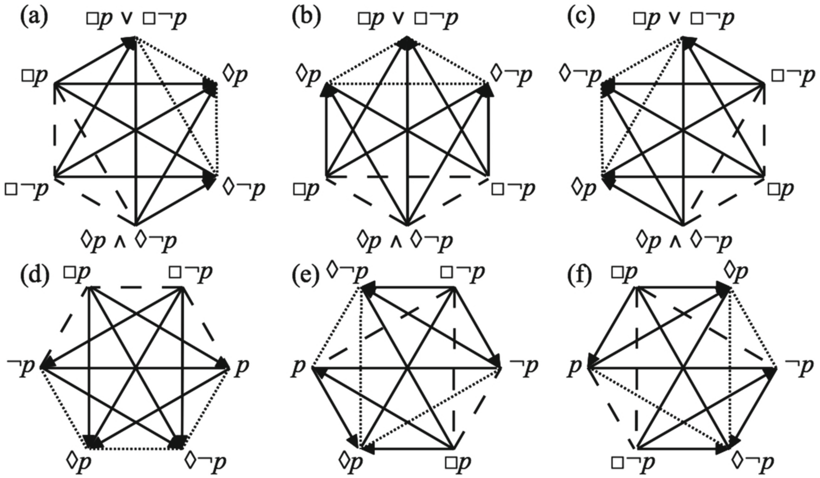


Fig. 5. (a–c) The three remaining fundamental forms of a hexagon for the JSB fragment of S5-formulas whose first fundamental form was already shown in Fig. 2(a); (d–f) the three remaining fundamental forms of a hexagon for the SC fragment of S5-formulas whose first fundamental form was already shown in Fig. 2(b).

(in our S5-example: $\Box p$, $\Diamond p \wedge \Diamond\neg p$ and $\Box\neg p$) are all of the same level: their canonical bitstring representations are 100, 010 and 001 [14]. Consequently, the 3 contrarieties holding between them are all equally 'strong', and are thus best visualized using 3 lines of equal length [18,36], as in the hexagon in Fig. 2(a).

However, for linguistic-cognitive reasons it is sometimes useful to view a JSB fragment as (isomorphic to) a fragment of a much *larger* Boolean algebra, e.g. $\mathbb{B}_5$, since this allows us to treat the 3 pairwise contrary formulas as belonging to different levels. For example, in our S5-example, it makes sense to treat $\Box p$ and $\Box\neg p$ as the level-1 bitstrings 10000 and 00001, resp., since these formulas represent the two 'extremes' of a 'modal scale', while $\Diamond p \wedge \Diamond\neg p$ is treated as the level-3 bitstring 01110, since it represents the entire 'interior' of that modal scale [34]. Consequently, the contrarieties holding between these formulas are of different 'strengths': the extremes $\Box p$ and $\Box\neg p$ are much more contrary to each other than they are to the intermediate $\Diamond p \wedge \Diamond\neg p$. It therefore makes sense to visualize the *strongest* contrariety (between $\Box p$ and $\Box\neg p$) by means of a line that is *longer* than the lines representing the two other contrarieties. This is exactly the case with the contrariety triangle in the hexagon in Fig. 5(b).

In sum: if the JSB fragment is visualized by means of a *hexagon*, then various logical distinctions can directly be visualized by using different fundamental forms for different cases. By contrast, if the fragment is visualized using an *octahedron*, then these distinctions are collapsed, since the octahedron has just a single fundamental form. The optimal diagram for visualizing a JSB fragment is thus a hexagon, rather than an octahedron.[12]

[12] This result is in line with earlier work on visualizations for JSB fragments [13,16].

SC 3-PCD Fragments. Visualizing an SC fragment by means of a hexagon yields 4 fundamental forms; see e.g. Figs. 2(b) and 5(d–f).[13] In the two hexagons in Fig. 5(e–f), the subalternations do not share one common direction; in other words, the idea that the formulas are *ordered* (according to the strict partial order of subalternation) is not at all visualized in these hexagons. Furthermore, the (sub)contrarieties are visually most prominent: most of them are in the *center* of the diagram (the subalternations are in the periphery) and they are also *longer* than most of the subalternations [37]. Turning to the hexagon in Fig. 5(d), we see that the subalternations do share a common direction (they all go downward), and hence, this diagram directly visualizes the ordering induced by the subalternations (lower in the diagram corresponds to further in the subalternation ordering). Furthermore, in this hexagon the subalternations are visually most prominent: they are all in the center of the diagram, and they are longer than all the (sub)contrarieties. Finally, the hexagon in Fig. 2(b) also directly visualizes the ordering induced by the subalternations (again: lower in the diagram corresponds to further in the subalternation ordering). In this hexagon, however, the (sub)contrarieties are visually most prominent: most of them are in the diagram's center, and they are longer than most of the subalternations.

Putting everything together, we thus find that the two hexagons in Fig. 5(e–f) primarily draw the user's attention to the (sub)contrarieties in the SC fragment, and in order to achieve this, they even distort the ordering induced by the subalternations [36, p. 37]. Next, the hexagon in Fig. 5(d) focuses on the fragment's subalternation structure, by making the subalternations visually most prominent and also respecting the ordering induced by these subalternations. Finally, the hexagon in Fig. 2(b) strikes an ideal balance between these two extremes: it primarily draws the user's attention to the (sub)contrarieties in the SC fragment, but does so while still respecting the ordering induced by the subalternations.

Hence, if the SC fragment is visualized by means of a *hexagon*, then different fundamental forms can be used to visually emphasize different logical aspects of the fragment. By contrast, if one were to visualize the fragment by means of an *octahedron*, then this would no longer be possible, since the octahedron has just a single fundamental form. Consquently, the best diagram for visualizing a SC fragment seems to be a hexagon, rather than an octahedron.

U4 3-PCD Fragments. Visualizing a U4 fragment by means of a hexagon yields 4 fundamental forms; see e.g. Figs. 2(c) and 6(a–c). In Fig. 2(c), the (sub)contrarieties are visually most prominent: they are in the center of the diagram and they are also longer than the subalternations. In Fig. 6(b), it is exactly the other way around: the subalternations are in the center of the diagram, and they are longer than the (sub)contrarieties.[14] The hexagons in Figs. 2(c) and 6(b) thus draw the user's attention to either the (sub)contrarieties or the

[13] Hexagons 1, 3 and 6 in [6, pp. 131–132] visualize an SC fragment using three distinct fundamental forms, viz. those shown in Figs. 2(b), 5(e) and (f), respectively.

[14] The hexagons in Fig. 6(a) and (c) strike a balance between the (sub)contrarieties and subalternations, by distributing visual prominence equally among them.

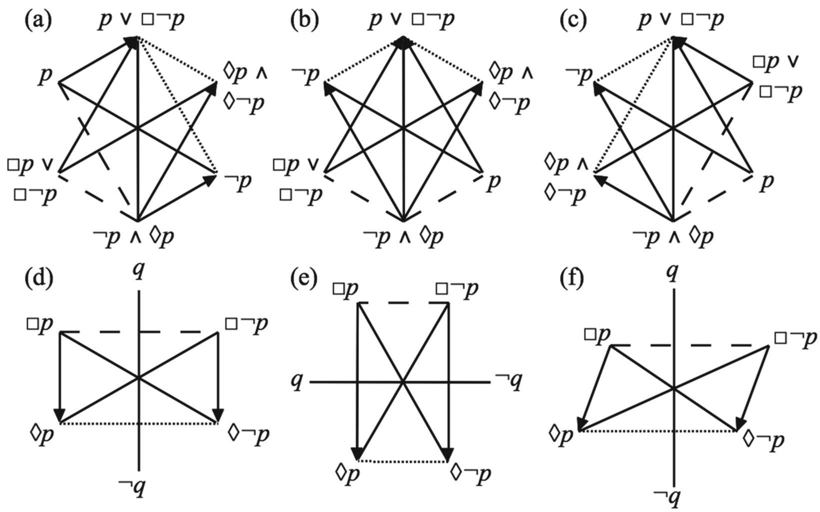


Fig. 6. (a–c) The three remaining fundamental forms of a hexagon for the U4 fragment of S5-formulas whose first fundamental form was already shown in Fig. 2(c); (d–e) two of the four fundamental forms of a hexagon for the U8 fragment $\{\Box p, \Box\neg p, \Diamond p, \Diamond\neg p, q, \neg q\}$; (f) (the unique fundamental form of) an octahedron for the same fragment.

subalternations. Recalling the logical importance of the distinction between these two types of Aristotelian relations (cf. Sect. 4), these two hexagons will thus be particularly useful, depending on the author's concrete purposes: does she want her audience to focus on the (sub)contrarieties or rather on the subalternations?

In sum: if the U4 fragment is visualized by means of a *hexagon*, then different fundamental forms can be used to visually emphasize different logical relations inside the fragment. By contrast, if one were to visualize the fragment by means of an *octahedron*, then this would no longer be possible, since the octahedron has just a single fundamental form. The optimal diagram for visualizing a U4 fragment is thus a hexagon, rather than an octahedron.

U8 3-PCD Fragments. A U8 fragment consists of four formulas that constitute a classical 2-PCD fragment, together with an additional pair of contradictory formulas that are unconnected to the first four. A typical example is the S5-fragment $\{\Box p, \Box\neg p, \Diamond p, \Diamond\neg p, q, \neg q\}$. Visualizing a U8 fragment by means of a *hexagon* yields 4 fundamental forms, viz. 2 fundamental forms in which the additional PCD is parallel to the subalternations, as in Fig. 6(d), and 2 fundamental forms in which the additional PCD is parallel to the (sub)contrarieties, as in Fig. 6(e). Since there does not seem to be any logical reason for preferring one option over the other, the differences between the 4 fundamental forms are thus mere side-products of the hexagon's lack of symmetry. By contrast, if one visualizes the fragment using an *octahedron*, then one can put the additional PCD perpendicular to the subalternations as well as the (sub)contrarieties (thereby avoiding any unmotivated design decisions), as in Fig. 6(f). In sum, then, the best diagram for visualizing a U8 fragment seems to be an octahedron.

U12 3-PCD Fragments. The U12 fragments are the perfect analogues of the degenerated 2-PCD fragments, in the sense that their formulas are all pairwise unconnected (except for the 3 pairs of contradictory formulas, of course). A typical example is the CPL-fragment $\{p, \neg p, q, \neg q, r, \neg r\}$. If such a fragment is visualized by means of an *octahedron*, then the equal logical status of its 12 pairs of unconnected formulas is visually represented by the fact that they are all on lines of equal length (viz. the 12 edges of the octahedron). By contrast, if one were to visualize the fragment using a *hexagon*, one would be forced to put these unconnected pairs on lines of different lengths, without any logical motivation. The optimal diagram for visualizing a U12 fragment is thus an octahedron, rather than a hexagon.

6 Conclusions and Future Work

In this paper we have presented a systematic approach for dealing with informationally equivalent Aristotelian diagrams. The account is based on the interaction between the logical properties of the visualized fragment and the geometrical properties of the concrete polygon/polyhedron. Applying this account to all Aristotelian families of 2-PCD and 3-PCD fragments has led to several new insights: as to the 2-PCD fragments, the classical ones are best visualized by means of a rectangle if one focuses on the distinction between (sub)contrariety and subalternation and by means of a square otherwise, and the degenerated ones by means of a square; as to the 3-PCD fragments, JSB, SC and U4 are best visualized using a hexagon, and U8 and U12 using an octahedron.

A natural next step involves applying the account to 4-PCD fragments. This is by no means trivial, since there exist 18 Aristotelian families of 4-PCD fragments, only a few of which are currently well-understood. As for the geometric shapes to be used, obvious candidates include the (regular) octagon and the cube, which have symmetry groups of order resp. 16 and 48, and thus yield resp. $\frac{2^4 \times 4!}{16} = \frac{384}{16} = 24$ and $\frac{2^4 \times 4!}{48} = \frac{384}{48} = 8$ fundamental forms. However, when dealing with Aristotelian families that do not have any relevant logical distinctions to be visualized, one might also want to consider shapes with a symmetry group of order 384, since these will yield exactly $\frac{384}{384} = 1$ fundamental form.

On a more general level, when visualizing an n-PCD fragment, one might want to consider a polytope[15] that is (i) centrally symmetric, (ii) has $2n$ vertices, and (iii) has a symmetry group of order $2^n \times n!$ (since such a polytope will yield exactly $\frac{2^n \times n!}{2^n \times n!} = 1$ fundamental form). There indeed exists a polytope satisfying these criteria, for all n, viz. the *cross-polytope of dimension* n, which is the dual polytope of the n-dimensional hypercube [9, pp. 121, 294]. In case $n = 2$, this is

[15] The term ‘polytope’ is a generalization of the terms ‘polygon’ and ‘polyhedron’ to arbitrary dimensions [9].

the dual of a square, which is itself also a *square* (cf. Sect. 4); in case $n = 3$, this is the dual of a cube, which is an *octahedron* (cf. Sect. 5).

The practical usefulness of these last observations is fairly limited, because they involve (cross-)polytopes of arbitrarily high dimensions, which are not very useful for concrete visual-diagrammatic purposes. For example, visualizing a 4-PCD fragment would require a so-called *16-cell* (i.e. the dual of the 4-dimensional hypercube) [9, p. 292], which can be studied abstractly, but goes beyond human visual cognition.[16] Nevertheless, the theoretical importance of these observations should not be underestimated, since they show that the diagrams that several logicians have come up with to visualize 2- and 3-PCD fragments 'up to symmetry' (i.e. having a unique fundamental form), viz. the square and the octahedron, are the first few instances of a well-defined, infinite series of polytopes.

Acknowledgements. We would like to thank Dany Jaspers and Margaux Smets for their valuable feedback on an earlier version of this paper. The first author holds a Postdoctoral Fellowship of the Research Foundation–Flanders (FWO).

References

1. Benoy, F., Rodgers, P.: Evaluating the comprehension of Euler diagrams. In: 11th International Conference on Information Visualization, pp. 771–778. IEEE Computer Society (2007)
2. Blake, A., Stapleton, G., Rodgers, P., Cheek, L., Howse, J.: Does the orientation of an Euler diagram affect user comprehension? In: 18th International Conference on Distributed Multimedia Systems, pp. 185–190. Knowledge Systems Institute (2012)
3. Blake, A., Stapleton, G., Rodgers, P., Cheek, L., Howse, J.: The impact of shape on the perception of Euler diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 123–137. Springer, Heidelberg (2014)
4. Brown, M.: Generalized quantifiers and the square of opposition. Notre Dame J. Formal Logic **25**, 303–322 (1984)
5. Cavaliere, F.: Fuzzy syllogisms, numerical square, triangles of contraries, interbivalence. In: Béziau, J.Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 241–260. Springer, Heidelberg (2012)
6. Chatti, S., Schang, F.: The cube, the square and the problem of existential import. Hist. Philos. Logic **32**, 101–132 (2013)
7. Chow, K.F.: General patterns of opposition squares and 2n-gons. In: Béziau, J.Y., Jacquette, D. (eds.) Around and Beyond the Square, pp. 263–275. Springer, Heidelberg (2012)

[16] One might object that higher-dimensional cross-polytopes *can* be visualized, by considering their lower-dimensional projections. This objection fails to take into account, however, that such projections will *break* some of the cross-polytopes' symmetries, and thus not be useful for our current purposes. This phenomenon can already be observed in very low dimensions; for example, the hexagon in Fig. 2(a) can be seen as the 2D projection of the octahedron in Fig. 3(c), but this projection drastically reduces the number of symmetries (from the octahedron's 48 to the hexagon's 12).

8. Correia, M.: Boethius on the square of opposition. In: Béziau, J.Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 41–52. Springer, Heidelberg (2012)
9. Coxeter, H.S.M.: Regular Polytopes. Dover Publications, Mineola (1973)
10. Demey, L.: Algebraic aspects of duality diagrams. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 300–302. Springer, Heidelberg (2012)
11. Demey, L.: Structures of oppositions for public announcement logic. In: Béziau, J.Y., Jacquette, D. (eds.) Around and Beyond the Square of Opposition, pp. 313–339. Springer, Heidelberg (2012)
12. Demey, L.: Interactively illustrating the context-sensitivity of Aristotelian diagrams. In: Christiansen, H., Stojanovic, I., Papadopoulos, G. (eds.) CONTEXT 2015. LNCS, vol. 9405, pp. 331–345. Springer, Heidelberg (2015)
13. Demey, L., Smessaert, H.: The relationship between Aristotelian and Hasse diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 213–227. Springer, Heidelberg (2014)
14. Demey, L., Smessaert, H.: Combinatorial bitstring semantics for arbitrary logical fragments (2015, submitted)
15. Demey, L., Smessaert, H.: Metalogical decorations of logical diagrams. Log. Univers. **10**, 233–292 (2016)
16. Demey, L., Smessaert, H.: Shape heuristics in Aristotelian diagrams. In: Kutz, O., Borgo, S., Bhatt, M. (eds.) Shapes 3.0 Proceedings. CEUR-WS (forthcoming)
17. Dubois, D., Prade, H.: From Blanché's hexagonal organization of concepts to formal concept analysis and possibility theory. Log. Univers. **6**, 149–169 (2012)
18. Gurr, C.: Effective diagrammatic communication: syntactic, semantic and pragmatic issues. J. Vis. Lang. Comput. **10**, 317–342 (1999)
19. Horn, L.R.: A Natural History of Negation. University of Chicago Press, Chicago (1989)
20. Jacquette, D.: Thinking outside the square of opposition box. In: Béziau, J.Y., Jacquette, D. (eds.) Around and Beyond the Square, pp. 73–92. Springer, Heidelberg (2012)
21. Joerden, J.: Logik im Recht. Springer, Heidelberg (2010)
22. Kraszewski, Z.: Logika stosunków zakresowych [Logic of extensional relations]. Stud. Logica. **4**, 63–116 (1956)
23. Larkin, J., Simon, H.: Why a diagram is (sometimes) worth ten thousand words. Cogn. Sci. **11**, 65–99 (1987)
24. McNamara, P.: Deontic logic. In: Zalta, E.N. (ed.) Stanford Encyclopedia of Philosophy. CSLI (2010)
25. Moretti, A.: The geometry of standard deontic logic. Log. Univers. **3**, 19–57 (2009)
26. Parsons, T.: The traditional square of opposition. In: Zalta, E.N. (ed.) Stanford Encyclopedia of Philosophy. CSLI (2006)
27. Rodgers, P.: A survey of Euler diagrams. J. Vis. Lang. Comput. **25**, 134–155 (2014)
28. Rotman, J.J.: An Introduction to the Theory of Groups, 4th edn. Springer, New York (1995)
29. Seuren, P., Jaspers, D.: Logico-cognitive structure in the lexicon. Language **90**, 607–643 (2014)
30. Smessaert, H.: On the 3D visualisation of logical relations. Log. Univers. **3**, 303–332 (2009)
31. Smessaert, H.: Boolean differences between two hexagonal extensions of the logical square of oppositions. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 193–199. Springer, Heidelberg (2012)

32. Smessaert, H., Demey, L.: Logical and geometrical complementarities between Aristotelian diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 246–260. Springer, Heidelberg (2014)
33. Smessaert, H., Demey, L.: Logical geometries and information in the square of opposition. J. Logic Lang. Inform. **23**, 527–565 (2014)
34. Smessaert, H., Demey, L.: The unreasonable effectiveness of bitstrings in logical geometry. In: Béziau, J.Y. (ed.) Proceedings of Square (2014, forthcoming)
35. Smessaert, H., Demey, L.: Visualising the Boolean algebra $\mathbb{B}_4$ in 3D. In: Jamnik, M., Uesaka, Y., Elzer Schwartz, S. (eds.) Diagrams 2016. LNCS, vol. 9781, pp. 289–292. Springer, Heidelberg (2016)
36. Tversky, B.: Prolegomenon to scientific visualizations. In: Gilbert, J.K. (ed.) Visualization in Science Education, pp. 29–42. Springer, Heidelberg (2005)
37. Tversky, B.: Visualizing thought. Top. Cogn. Sci. **3**, 499–535 (2011)

A Generic Approach to Diagrammatic Representation: The Case of Single Feature Indicator Systems

Atsushi Shimojima[1(✉)] and Dave Barker-Plummer[2]

[1] Faculty of Culture and Information Science, Doshisha University, 1-3 Tatara-Miyakodani, Kyotanabe 610-0394, Japan
ashimoji@me.com
[2] CSLI/Stanford University, Cordura Hall, 210 Panama Street, Stanford, CA 94305, USA
dbp@stanford.edu

Abstract. In this paper we take a *generic* approach to developing a theory of representation systems. Our approach involves giving an abstract formal characterization of a class of representation systems, and proving formal results based on this characterization.

We illustrate this approach by defining and investigating two closely related classes of representations that we call *Single Feature Indicator Systems* (SFIS), with and without *neutrality*. Many common representations including *tables*, such as timetables and work schedules; *connectivity graphs*, including route maps and circuit diagrams; and *statistical charts* such as bar graphs, either are SFIS or contain one as a component.

By describing SFIS abstractly, we are able to prove some properties of all of these representation systems by virtue of the fact that the properties can be proved on the basis of the abstract definition only. In particular we show that certain abstract inference rules are sound, and that each instance admits concrete inference rules obtained by instantiating the abstract counterparts.

1 Introduction

In this paper we adopt a *generic* approach to developing a theory of representation systems in general, with diagrammatic systems as a special case. Our approach involves giving an abstract formal characterization of a class of representation systems, and then proving results about the properties of all members of the class in the abstract setting. By adopting this approach, we are able to short-circuit investigation of individual representation systems, and also to assign the responsibility for the possession of various properties of an individual representation system to its membership in the class. Specifically, we do three things in this paper:

1. **Describe and formalize our view of a *representation system*.** Our formalization uses channel theory, a formal framework for modeling information

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 83–97, 2016.
DOI: 10.1007/978-3-319-42333-3_7

flow, of which representation is a special case [3]. This task occupies Sects. 2 and 3 of this paper.

2. **Show how to model particular types of representation systems within the channel theory framework**. We focus on two closely related classes of representation systems: *Single Feature Indicator Systems with and without neutrality* (SFIS). This is the content of Sects. 4 and 5. SFISs are among the simplest representation systems that we can think of, and are built into a number of important, familiar diagrammatic representation systems. Each of the diagrams presented in Fig. 1 illustrates a system that either is an SFIS itself or has an SFIS as its main component. We will refer to these examples throughout this paper.
3. **Demonstrate properties held in common between all instances of the class of SFIS**. Our formalization of SFIS allows us to prove that they all share important properties. One important goal of the diagrams research community has been the development of diagrammatic proof editors, with the ability to verify the application of inference rules to diagrammatic representations. Hyperproof [2], Diamond [4], and CDEG [5] are examples of such proof editor/checkers. MixR and Openbox are frameworks for constructing heterogeneous proof systems for arbitrary representations [1,7]. For the purpose of developing such systems, it is useful to have a generic view of a set of inference rules that are guaranteed to be valid in any member of a class of diagrammatic representation systems. We discuss this in Sects. 6 and 7.

Sections 2 and 4 describe the intuitions guiding this paper, while Sects. 3 and 5 describe the corresponding formalization of these ideas. We recommend reading the informal sections first, before delving into their formalization.

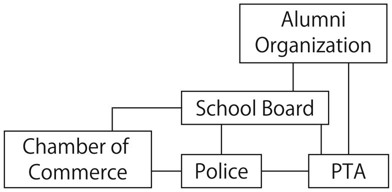
2 The General Picture

We begin by sketching the general picture of representation systems that forms the basis of the development of the theory that we present here. Our notion of representation system is designed to capture important semantic properties of a *representational practice* followed by a group of people. A representational practice is a recurrent pattern in which people express information by creating a (typically proximal) object and extract the information from it. In many cases, the information thus expressed is about a particular (distal) object or situation. We call a proximal object created on a particular occasion a *representation*. When a representation s is created to express information about a particular distal object or situation t, we say s *represents* t.

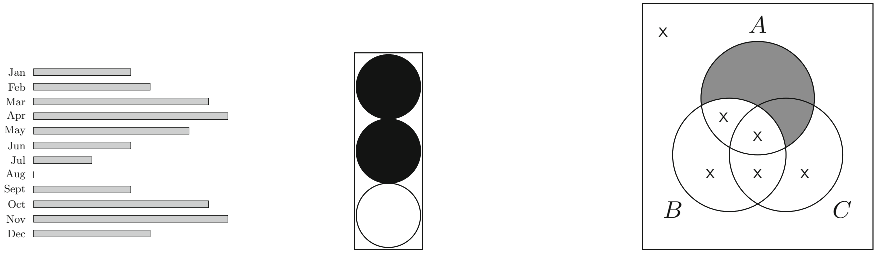
For example, a project leader may create the table in Fig. 1a to express information about the work schedules of four workers at a research project. Many people know how to extract the information expressed in this table and they do extract information from it. Here we see a representational practice followed by the project leader and these people. We will refer to this representational practice and its formalization as $\mathcal{R}_t$.

	Mon	Tues	Wed	Thurs	Fri
Atsushi	✓	✓	✓	✓	✓
Dave	✗	✗	✗	✗	✗
John	✓	✗	✓	✗	✗
Mike	✓	✗	✓	✓	✗

(a) Scheduling Table.



(b) Organization Chart.



(c) Bar Chart. (d) Traffic Light. (e) Venn Diagram.

Fig. 1. Diagrams illustrating representation systems that either are SFISs themselves or amplifications of SFISs

Typically, a representational practice is governed by various constraints of different origins, and the effectiveness of the practice deeply depends on these constraints. They consist of *source constraints* concerning what kinds of symbols appear in a representation and how they are arranged, *semantic constraints* concerning what arrangement of symbols indicate what information, and *target constraints* holding among the pieces of information expressible in the representational practice in question.

The source constraints in $\mathcal{R}_t$ include the fact that each cell of the table contains one, and only one, of the symbols ✓ and ✗. The target constraints include the facts that every worker either does, or does not, work on a particular day. The semantic constraints include the fact that a cell has the symbol ✓ only if the relevant worker works on that particular day. The project manager, his workers, and other users know these constraints and respect them to make their communication based on scheduling tables reliable and efficient.

The source constraints, semantic constraints, and target constraints governing a representational practice can be considered to make up a system, which we call a *representation system*. Thus, we can think of the system $\mathcal{R}_t$ of scheduling tables for this research project, as well as the systems underlying the other diagrams listed in Fig. 1a—the systems of connectivity maps between these organizations, of bar charts representing sales of a particular book, of diagrams of a particular traffic light, and of Venn diagrams concerning these particular sets.

In the next section we present a formalization of representation systems using channel theory [3]. In Sect. 4 we will define the notion of Single Feature Indicator System using this general theory.

3 Channel Theory and Representation Systems

Channel theory provides us with a formal framework for describing information flow, and it can be used to describe representation systems as a paradigmatic case [3]. This section is a free-standing presentation, simplified to fit our purposes, of Barwise and Seligman's discussion of representation systems within the channel theory framework (see Chap. 20 of [3]).

3.1 Types

As we discussed above, a representation system can be decomposed into three systems of constraints. Each system of constraints is modeled as a *theory*, while a theory is built on top of a set of *types*. The relevant set of types depends on the system of constraints that we are considering. When we consider source constraints of the system $\mathcal{R}_t$, for example, the set of types include the following types:

(σ_1) The intersection of a row labeled "Atsushi" and a column labeled "Tues" has a ✓.
(σ_2) The intersection of a row labeled "Atsushi" and a column labeled "Tues" has an ✗.
(σ_3) A column labeled "Tues" has at least one ✓.

while when we consider that system's target constraints, the set of types include:

(θ_1) Atsushi works on Tuesday.
(θ_2) Atsushi does not work on Tuesday.
(θ_3) At least one person is working on Tuesday.

3.2 Constraints

We represent constraints using Gentzen sequents, which are pairs of sets of types. When we write $\Gamma \vdash \Delta$, we refer to the pair of sets Γ and Δ, and indicate that this pair is a member of the set $\vdash$, or that the relation $\vdash$ holds between them.

We use lowercase Greek letters to refer to types, and uppercase Greek letters to refer to sets of types and adopt a common abuse of notation and use types and set of types interchangeably in sequents.

Types on the left hand side of the $\vdash$ are interpreted conjunctively, and on the right hand side, disjunctively. A sequent of the form $\alpha_1, \alpha_2, \alpha_3 \vdash \beta_1, \beta_2$ represents the constraint that any object which is of type α_1, α_2 and α_3, is also of one of the types β_1 or β_2 (or both). As a consequence:

1. $\alpha \vdash \beta$, represents the constraint that everything of type α is also of type β,
2. $\emptyset \vdash \alpha$, represents the constraint that everything is of type α,
3. $\alpha, \beta \vdash \emptyset$, represents the constraint that types α and β do not hold together.

For example, the Gentzen sequents $\sigma_1, \sigma_2 \vdash \emptyset$ and $\sigma_1 \vdash \sigma_3$ capture plausible source constraints in the system $\mathcal{R}_t$. Source constraints often originate in syntactic conventions combined with natural, spatial constraints on the arrangement of symbols. For example, neither of these example constraints would hold without syntactic conventions saying that there can be only one column labeled "Tues" and that a cross or a check appearing in a cell has a certain minimal size and may not overlap other marks. Both example constraints do hold in the presence of such syntactic conventions.

3.3 Theories

A theory captures a set of constraints holding in a domain by modeling them as a set of Gentzen sequents defined over a fixed set of types.

Definition 1 (Theory). *A* theory *is a pair $T = \langle \Upsilon, \vdash \rangle$, where $\vdash$ is a set of Gentzen sequents over Υ. A* constraint *of the theory T is a sequent $\langle \Gamma, \Delta \rangle$ in $\vdash$.*

When the set of constraints of a theory is logically closed, it is called a "regular theory".

Definition 2 (Regular Theory). *A theory $T = \langle \Upsilon, \vdash \rangle$ is* regular *if it satisfies the following closure conditions:*

- **Identity:** $\alpha \vdash \alpha$*, for all types* α
- **Weakening:** *If $\Gamma \vdash \Delta$, then $\Psi_1, \Gamma \vdash \Delta, \Psi_2$ for any sets of types Ψ_1, Ψ_2,*
- **Global Cut:** *If $\Psi_1, \Gamma \vdash \Delta, \Psi_2$ for any partition of any set Ψ into Ψ_1, Ψ_2, then $\Gamma \vdash \Delta$.*

The following proposition from [3] shows that any set $\vdash$ of Gentzen sequents has a unique regular theory that minimally extends it.

Proposition 1. *For every theory $T = \langle \Upsilon, \vdash \rangle$, there is a smallest regular theory on Υ containing the sequents in Σ as constraints. This is called* the regular closure *of T.*

Proof: See [3], Proposition 9.7.

3.4 Representation Systems

As we described in Sect. 2, a representation system consists of three parts, a system of source constraints (pertaining to representations), a system of target constraints (pertaining to the represented situations), and a system of semantic constraints linking representations to the represented situations. Each of these components is represented as its own theory.

Definition 3 (Representation System). *A* representation system *is a triple* $\langle T_s, T_c, T_t \rangle$*, where*

1. T_s *is a theory* $\langle \Upsilon_s, \vdash_s \rangle$*, this is the* source *theory,*
2. T_t *is a theory* $\langle \Upsilon_t, \vdash_t \rangle$*, this is the* target *theory,*
3. T_c *is a theory* $\langle \Upsilon_c, \vdash_c \rangle$ *where* $\Upsilon_c = \Upsilon_s \uplus \Upsilon_t$*, this is the* semantic *theory.*

Among the three theories posited in this definition, T_c requires further explanation.[1] As we have explained above, the semantic conventions observed in a representational practice can be considered as constraints on what arrangement of symbols indicate what information. Take the previous example of the semantic convention in $\mathcal{R}_t$, according to which a check mark in the intersection of a row labeled "Atsushi" and a column labeled "Tuesday" indicates that Atsushi works on Tuesday. Since the participants of this practice generally follow this convention, it gives rise to a constraint in their local environment, according to which σ_1 holds of the scheduling table only if θ_1 holds in the work place. The theory $T_c = \langle \Upsilon_c, \vdash_c \rangle$ captures constraints of this sort. Since the relevant constraints to capture are ones from subsets of Υ_s to subsets of Υ_t, we define the set Υ_c to be the disjoint union of these two sets: $\Upsilon_s \uplus \Upsilon_t$. The theory T_c then lists the relevant constraint from σ_1 to θ_1 as a sequent in $\vdash_c$ (i.e., $\sigma_1 \vdash_c \theta_1$). When a type σ in Υ_s and a type θ in Υ_t are connected in this way, we say that σ *indicates* θ in the system $\mathcal{R}_t$.

The above definition of representation systems significantly simplifies the one proposed by [3] (Definition 20.1) while preserving the idea that three systems of constraints make up a representation system with one system providing a semantic connection between the other two.

4 Observations Underlying the Concept of SFIS

In this section we make some observations about similarities among many familiar representation systems. We will abstract these observations into a definition of Single Feature Indicator Systems, the class of representation systems sharing these similarities. In Sect. 5 we present a formalization of this definition.

4.1 First Observation: Roles

Many diagrammatic representations consist of basic components playing certain common *roles*. These roles are *common* in the sense that they are played not only by components of particular diagrams, but by components of all diagrams used in the given representational practice. For example, each scheduling table in the system $\mathcal{R}_t$ has a basic component that plays the role of [the intersection of the row labeled "Atsushi" and the column labeled "Mon"]. The existence of such a component is mandated by syntactic stipulations and spatial constraints on scheduling tables in the system $\mathcal{R}_t$. In this way, we can think of $4 \times 5 = 20$

[1] We sometime call this the *core* theory, hence the subscript "c".

common roles for the source domain of $\mathcal{R}_t$. Such roles appear in every diagram in this representational practice, for example, if one schedule is made for each week for a year.

Similarly, consider a Venn diagram representation system, $\mathcal{R}_v$, with circles labelled A, B and C, such as depicted in Fig. 1e. Every diagram used in the system $\mathcal{R}_v$ has a component that plays the role of [the set of points inside all of the circles labeled "A," "B," and "C"]. Another role is that of [the set of points inside the circles labeled "A" and "B," but outside the circle labeled "C"]. In this way, we can think of $2^3 = 8$ roles in the source domain of $\mathcal{R}_v$. When a symbol or a place in a diagram plays a common role in this sense, we call it a *basic element* of that diagram.

4.2 Second Observation: Values

In many diagrammatic representation systems there is a fixed range of possible values that a basic element can take. Further, each basic element must take at least one value (value existence condition) but not more than one (value uniqueness condition). For example, a basic element of a scheduling table must have either a ✓ or ✗ (existence) but cannot have both (uniqueness). The basic elements of the bar chart in Fig. 1c are individual bars, and each has a certain height (existence) but not more than one height (uniqueness).

4.3 Third Observation: Features in the Source Domain

The combination of roles and values that satisfy the value existence condition and the value uniqueness condition give rise to a structure that we call a *feature*.

For example, the source domain of our scheduling tables involves a feature consisting of 20 common roles (played by cells) and 2 values (having a ✓ or ✗). The source domain of our bar charts involves a feature consisting of 12 roles (played by bars) and an infinite number of values (heights). The source domain of our organization charts involves a feature consisting of ${}_5C_2 = 10$ roles (played by pairs of organization names) and 2 values (directly connected or not).

Notice that in each of the example representation systems, the values taken by the various roles in the source domain are *independent*. That is, as far as the syntactic stipulations and spatial constraints are concerned, the basic element playing a role can take any value without consideration of the values of other basic elements. We call this the *independence condition*.

4.4 Fourth Observation: Feature in the Target Domain

Often the target domain in a diagrammatic system makes up a feature too. For example, in any given week represented by a scheduling table, Atsushi either works or does not work on Monday, but he cannot do both. We can restate this as the fact that the element of the described situation playing the role of $\langle \text{Atsushi}, \text{Monday} \rangle$ must have either the property [working_on] or the property [not_working_on] (value existence) but cannot have both (value uniqueness). Here, $\langle \text{Atsushi}, \text{Monday} \rangle$ is a role, and the properties [working_on] and [not_working_on] make up the value range. The other roles are $\langle \text{Atsushi}, \text{Tuesday} \rangle$, $\langle \text{Mike}, \text{Friday} \rangle$, and so on, counting up to $4 \times 5 = 20$ pairs.

Similarly, in any particular group of people represented by an $\mathcal{R}_v$-diagram, the set of $A \cap B \cap C$ must be either empty or non-empty (existence) but cannot be both (uniqueness). In this case $2^3 = 8$ sets, including $A \cap B \cap C$, $A \cap B \cap \overline{C}$ and $A \cap \overline{B} \cap C$, constitute the set of roles, and the value range is $\{\text{empty}, \text{non-empty}\}$.

4.5 Fifth Observation: Semantic Correspondence of Features

We have just seen that the source and the target domain of a diagrammatic system often make up features. Our final observation is that these features typically stand in a close correspondence through the system's semantic conventions.

Take the case of bar charts. The source role of [the bar labeled "Jan"] corresponds to the target role [January]. In this way, a natural one-one correspondence holds between the set of source roles and the set of target roles in this system. A natural correspondence holds between the sets of values too. Each possible height taken by a bar corresponds to a possible number of books sold in the corresponding month. In the case of scheduling tables, the source role of [the intersection of the row labeled "Atsushi" and the column labeled "Mon"] corresponds to the target role of (Atsushi, Monday) and similarly for the other roles. The two source values, having a ✓ or ✗, each corresponds to a unique target value, [working_on] or [not_working_on]. The reader can easily check a similar two-fold correspondence holds between the source feature and the target feature involved in each of the other systems illustrated in Fig. 1.

These correspondences underlie semantic conventions in these systems. In the system of scheduling tables $\mathcal{R}_t$, if the intersection labeled "John" and the column labeled "Mon" has a ✓ in a scheduling table, it indicates that John works on Monday. In the system of bar charts, that the bar labelled January having a height of 10 mm indicates that the number of book sales in the month of January being 100 units. In this way, many diagrammatic representation systems employ semantic conventions with the form that, if a basic element playing role r has the value v, it indicates that the element playing the role corresponding to r has the value corresponding to v. We call representation systems having this form of semantic convention "Single Feature Indicator Systems" or "SFISs" for short.

5 Single Feature Indicator Systems – Formalized

One of the contributions that we make in this paper is a demonstration of how channel theory can be used to formalize classes of representation systems. This section of the paper, where we formalize the notion of Single Feature Indicator System, is focussed on this task. Our approach is to specialize Barwise and Seligman's definition of representation system that we presented in Definition 3, so that the types and constraints in the component theories capture the observations about roles and values that we outlined in Sect. 4.

5.1 Features

In Sect. 4, we observed that the target and source domains of many diagrammatic representation systems can be characterized as consisting of roles which take on specific values. We formalize this idea by using an ordered pair $\langle r, v\rangle$ to model the type that holds on a representation d if and only if the element of d playing the role r has the value v. For example, a diagram in the traffic light representation system illustrated in Fig. 1d is of type $\langle \mathit{uppermost\_circle}, \mathit{white}\rangle$ if the uppermost circle of d is white. A *feature* is a specialized theory over these types:

Definition 4 (Feature). *A* feature *is a regular theory* $T = \langle \Upsilon, \vdash\rangle$ *for which there are sets* R *and* V *such that:*

1. $R \times V = \Upsilon$,
2. *For every* $r \in R$,
   (a) $\vdash \{\langle r, v\rangle : v \in V\}$,
   (b) $\langle r, v\rangle, \langle r, v'\rangle \vdash \emptyset$ *for all distinct* $v, v' \in V$.

When these conditions hold, R *and* V *are called the* set of roles *and the* set of values *of the feature* T*, respectively.*

Clause 1 declares that this theory is concerned with pairs of the form $\langle r, v\rangle$ with a role $r \in R$ and a value $v \in V$. Clauses 2a and 2b capture the value existence condition and the value uniqueness conditions respectively.

5.2 Single Feature Indicator Systems

A Single Feature Indicator System is a representation system whose source and target theories are features with appropriate connections provided by the semantic theory.

Definition 5 (Single Feature Indicator System (SFIS)). *A representation system* $\langle S, C, T\rangle$ *is a* Single Feature Indicator System (SFIS) *iff:*

1. S *is a feature with the set* R_s *of roles and the set* V_s *of values*
2. T *is a feature with the set* R_t *of roles and the set* V_t *of values*
3. *Every assignment* $f : R_s \to V_s$ *of values in* V_s *to roles in* R_s *is consistent, i.e.,* $f \not\vdash_S \emptyset$
4. C *is the theory* $\langle \Upsilon_c, \vdash_c\rangle$ *where there are bijections* p_r *from* R_s *to* R_t *and* p_v *from* V_s *to* V_t *such that* $\vdash_c$ *is the regular closure of the set of all sequents* $\langle\{\langle r, v\rangle\}, \{\langle p_r(r), p_v(v)\rangle\}\rangle$ *where* $\langle r, v\rangle \in \Sigma_S$.

Conditions 1 and 2 state that the source and target theories are features. The additional condition on the source feature expressed by Clause 3 is the *independence condition*, which implies that basic elements of a representation can take any value no matter what values are taken by other basic elements.

The target theory T does not necessarily satisfy such a condition, which is to say that the constraints on T may result in some assignments of values not being permitted. In our model of a representation system, the source theory captures only those constraints originating in spatial constraints and syntactic stipulations associated with a representational practice. We have seen that as far as these constraints are concerned, the assignments of values to roles are independent of one another. On the other hand, T is intended to capture any constraint that holds on the types Σ_T in the target domain. So for example, there are no spatial or syntactic constraints preventing the drawing of a traffic light diagram with both the uppermost and lowermost circles being white, but there are additional constraints in the target domain which should prevent such a combination (on the assumption that a white circle indicates that the corresponding lamp is illuminated).

Finally, condition 4 tells us about the connections between the source and target theories. The projection functions p_r and p_v respectively associate roles in the source with roles in the target, and values in the source with values in the target. This clause requires that the system's semantic theory respects the correspondence between source types and target types established by these projection functions.

5.3 Single Feature Indicator System with Neutrality

Before we discuss the logical properties of Single Feature Indicator Systems we will introduce a closely related, and more interesting, class of representation systems that we call Single Feature Indicator System with Neutrality.

Consider a situation where you are observing the author of the scheduling table in Fig. 1a as it is being constructed. Perhaps all of the row and column labels are present, but the author has not yet filled in all of the cells. Such a representation carries *partial* information about the target that it describes. We can see, perhaps, that Dave will not be working on Monday, but whether or not Atsushi is working that day is not represented in the diagram.

We can model this situation by introducing a third kind of source value, a blank, in addition to ✓ and ✗. But we do not want to assign a target value to this source value. The function p_v defined to map source properties to target properties must be partial with respect to this blank property value.

The definition of a Single Feature Indicator System with Neutrality is similar to that of an SFIS.

Definition 6 (SFIS with Neutrality (SFIS$^\perp$)). *A representation system* $\langle S, C, T\rangle$ *is a* Single Feature Indicator System with Neutrality (SFIS$^\perp$) *iff there are sets* R_s, V_s, R_t, V_t *such that*

1. *S is a feature with the set R_s of roles and the set V_s of values*
2. *V_s contains a distinguished value $\perp$*
3. *Every assignment $f : R_s \to V_s$ of roles in R_s to values in V_s is consistent, i.e., $f \not\vdash_S \emptyset$*
4. *T is a feature with the set R_t of roles and the set V_t of values*
5. *There is a bijection p_r from R_s to R_t and a bijection p_v from $V_s - \{\perp\}$ to V_t such that $f_s(\langle r, v\rangle) \vdash_C f_t(\langle p_r(r), p_v(v)\rangle)$ for every $\langle r, v\rangle \in (R_s \times V_s - \{\perp\})$.*

The critical difference between this definition and the definition of Single Feature Indicator System *simpliciter* is that the source domain contains a neutral value $\bot$, and that this value is not in the domain of the bijection p_v, and therefore carries no information about the target.

6 Specifications and Semantic Consequence

We now turn out attention to some inference rules supported by every SFIS.

Before we can define these inference rules, we need to have a clear notion of consequence between diagrams. That is, we must define what it means for one diagram to follow from another. But before we can do this we need a way to describe complete diagrams.

Let $\langle S, C, T\rangle$ be an SFIS (with or without neutrality). We call any function $\Sigma : R_s \to V_s$ a *complete specification of the source.* Σ is a set of types assigning a unique value to every role in R_s. Such a function completely describes a source representation by associating a value with each role. Indeed, we can think of the complete specification *as* the representation with the visual appearance of the roles and values abstracted away. As the range of diagrams in Fig. 1 attest, the values associated with roles can be represented in a variety of ways, but semantically, only the particular association of roles to values matters.

If Σ is a complete specification, we can define

$$Ind(\Sigma) = \{\langle p_r(r), p_v(v)\rangle : \langle r, v\rangle \in \Sigma \text{ and } v \neq \bot\}$$

$Ind(\Sigma)$ is the set of target types *indicated* by the source types in Σ.

The critical definition is of what it means for a specification be a consequence of other specifications. If S is a set of representations (or more precisely, their complete specifications) and Σ_0 another representation, then Σ_0 is a logical consequence of S if everything indicated by Σ_0 is entailed by the union of types indicated by the members of S, or formally:

Definition 7 (Semantic Consequence). *Given a collection S of complete specifications of source (the premises), and a complete specification of source Σ_0 (the conclusion), we say that Σ_0 is a* semantic consequence *of S and write $S \Longrightarrow \Sigma_0$ iff $\bigcup\{Ind(\Sigma) : \Sigma \in S\} \vdash_t \sigma$, for all $\sigma \in Ind(\Sigma_0)$.*

7 Inference Rules

We now have everything that we need in order to define inference rules for SFISs, and to demonstrate their soundness in our theory.

7.1 Contradiction

Definition 8 (Value Conflict). *Two complete specifications of source Σ_1 and Σ_2 have an* value conflict *iff there is some role $r \in R_s$ and values v_1 and v_2 in V_s such that $\langle r, v_1\rangle \in \Sigma_1$ and $\langle r, v_2\rangle \in \Sigma_2$ and $v_1 \neq v_2$, $v_1 \neq \bot$ and $v_2 \neq \bot$.*

Theorem 1 (Contradiction). *Given complete specifications of source Σ_1 and Σ_2, if Σ_1 and Σ_2 have a value conflict, then $\Sigma_1, \Sigma_2 \Longrightarrow \Sigma$ for any complete specification of source Σ.*
Proof: *$Ind(\Sigma_1) \cup Ind(\Sigma_2)$ has both $\langle p_r(r), p_v(v_1)\rangle$ and $\langle p_r(r), p_v(v_2)\rangle$ for some $r \in R_s$ and distinct v_1, v_2. $\langle p_r(r), p_v(v_1)\rangle, \langle p_r(r), p_v(v_2)\rangle \vdash_t \emptyset$ because T is a feature and $p_v(v_1) \neq p_v(v_2)$. By weakening, $\langle p_r(r), p_v(v_1)\rangle, \langle p_r(r), p_v(v_2)\rangle \vdash_t \sigma$ for any target type σ, so certainly for any $\sigma \in Ind(\Sigma)$.*

We therefore obtain a generic inference rule which we will call SFIS-Contradiction. This a *generic* rule since the rule may be specialized to a particular concrete version of the rule in any particular SFIS, whether it is in the system of scheduling tables, that of Venn diagrams, or that of connectivity maps. For example, if one organization chart shows a connection between [Police] and [PTA], and another shows no such connection, the rule lets us derive any organization chart from the two.

We now turn our attention to additional generic inference rules which are available only within SFIS$^\perp$ since they require the existence of a neutral value for their specification.

7.2 Erasure

In what follows, we need some definitions which will help us to describe manipulations of complete specifications of the source (manipulations of the diagrams that they describe).

Definition 9 (Point Substitution). *Suppose that Σ is complete specification of source. Let $\Sigma^{\langle r,v\rangle}$ be defined in the following way:*

$$\Sigma^{\langle r,v\rangle} = (\Sigma - \{\langle r, v'\rangle\}) \cup \{\langle r, v\rangle\}$$

where v' is the value taken by r in Σ. We call $\Sigma^{\langle r,v\rangle}$ the $\langle r, v\rangle$-substitution *of Σ.*

The $\langle r, v\rangle$-substitution of Σ is just like Σ, except that $\langle r, v\rangle$ is a member of $\Sigma^{\langle r,v\rangle}$, instead of $\langle r, v'\rangle$. Note that, since Σ is a complete specification of source, there is some v' such that $\langle r, v'\rangle \in \Sigma$, and that $\Sigma^{\langle r,v\rangle}$ is also a complete specification of source.

As special cases of point substitution, we call the $\langle r, \perp\rangle$-substitution of Σ, the *r-erasure* of Σ, and, if $\langle r, \perp\rangle \in \Sigma$ and $v \neq \perp$, we call $\Sigma^{\langle r,v\rangle}$ the *$\langle r, v\rangle$-extension* of Σ.

Theorem 2. *If Σ_1 is a complete specification of source in an SFIS$^\perp$ and Σ_2 is the r-erasure of Σ_1, then $\Sigma_1 \Longrightarrow \Sigma_2$.*
Proof: *The result follows from Identity (Definition 2), since $Ind(\Sigma_2) \subset Ind(\Sigma_1)$.*

In this case, we say that Σ_2 may be obtained from Σ_1 by SFIS$^\perp$-Erasure. As an example of its use: if an organization chart showing a connection between [Police] and [PTA] represents the world accurately, then a chart which is otherwise identical but is non-committal about the existence of a connection between these organizations is also accurate, though less informative.

If Σ_2 may be obtained from Σ_1 by repeated use of this rule, then we say that Σ_2 is *an erasure of* Σ_1.

Corollary 1. *If Σ_2 is an erasure of Σ_1, then $\Sigma_1 \Longrightarrow \Sigma_2$.*
Proof: *Trivial using Theorem 2.*

7.3 Proof by Cases

Any SFIS$^\perp$ supports an inference rule allowing proof by cases. We know that in the target theory, there is some property enjoyed by the target role, which corresponds to a source role having one or other of the definite properties. If some source role in a diagram token has the neutral property $\perp$ then there is a collection of new representations, which differ from the original, and from each other, only in their assignment of a source property to this role. One of these representations is a faithful representation of the target.

Suppose that $\langle r, \perp\rangle \in \Sigma$ for some role r. Then we define the set of r-extensions of Σ, denoted $Ext_r(\Sigma)$ to be the set containing the $\langle r, v\rangle$-extensions of Σ for $v \in V_s$. If the set V_s is finite, then so is $Ext_r(\Sigma)$.

Theorem 3. *If $\Sigma' \Longrightarrow \Sigma^*$ for every $\Sigma' \in Ext_r(\Sigma)$, then $\Sigma \Longrightarrow \Sigma^*$.*
Proof: *Assume the antecedent, and let σ be an arbitrary member of $Ind(\Sigma^*)$. It suffices to show $Ind(\Sigma) \vdash_t \sigma$. Let $A = \{\langle p_r(r), p_v(v)\rangle : v \in V_s\}$. To apply Global Cut, we show $A_1, Ind(\Sigma) \vdash_t \sigma, A_2$ for every partition $\langle A_1, A_2\rangle$ of A. Suppose $A_1 = \emptyset$. Then $A_2 = A$. By the definition of a feature, $\vdash_t A$. Thus, by Weakening, $A_1, Ind(\Sigma) \vdash_t \sigma, A_2$. Suppose on the other hand that $A_1 \neq \emptyset$. Then $\langle p_r(r), p_v(v')\rangle \in A_1$ for some $v' \in V_s$. Let Σ' be the r-extension of Σ for v'. Then $Ind(\Sigma') = Ind(\{\langle r, v'\rangle\} \cup \Sigma)$, and so $Ind(\Sigma') \subseteq A_1 \cup Ind(\Sigma)$. But $Ind(\Sigma') \vdash_t \sigma$ by assumption. So, by Weakening, $A_1, Ind(\Sigma) \vdash_t \sigma, A_2$.*

This permits the definition of a generic inference rule SFIS$^\perp$-Cases, which lets us derive any representation Σ' from Σ if Σ' is derivable from every r-extension of Σ for some role r.

7.4 Disjunction

Every SFIS$^\perp$ supports a rule which allows the weakening of information from a collection of representations into a single representation.

Definition 10 ($\Sigma_{\bigvee S}$). *Let S be a set of complete specifications of source, define $\Sigma_{\bigvee S}$ as follows:*

For each $r \in R_s$:

(a) $\Sigma_{\bigvee S}(r) = v$ *iff* $\Sigma(r) = v$ *for all* $\Sigma \in S$
(b) $\Sigma_{\bigvee S}(r) = \bot$ *otherwise*

We call $\Sigma_{\bigvee S}$ the disjunction of S.

Theorem 4. *Suppose that* $\langle r, \bot \rangle \in \Sigma$ *for some role* r. *If* S *is a set of complete specifications of source, and for each* $\Sigma' \in Ext_r(\Sigma)$ *there is* $\Sigma^* \in S$ *such that* $\Sigma' \Longrightarrow \Sigma^*$, *then* $\Sigma \Longrightarrow \Sigma_{\bigvee S}$.
Proof: *Observe* $\Sigma_{\bigvee S}$ *is a consequence of every* r*-extension of* Σ, *since* $\Sigma_{\bigvee S}$ *is a consequence of every member of* S *(Corollary 1) while every* r*-extension of* Σ *has some member of* S *as a consequence (assumption). Then apply Theorem 3.*

This theorem immediately permits the definition of an abstract inference rule, SFIS$^{\bot}$-Merge, which is a more useful version of proof by cases. Rather than insisting that each subproof of the proof by cases derive the same specification, possibly involving uses of erasure, we can allow the different cases to derive different specifications, but the disjunction of those specifications is exported into the main proof. An instance of SFIS$^{\bot}$-Merge is implemented in the Hyperproof program [2] under the name of MERGE.

7.5 Conjunction

Definition 11. **($\Sigma_{\bigwedge S}$).** *Let* S *be a set of complete specifications of source. Define* $\Sigma_{\bigwedge S}$ *as follows:*

1. $\Sigma_{\bigwedge S}$ *is undefined if* S *contains a value conflict,*
2. *for each* $r \in R_s$*:*
   (a) $\Sigma_{\bigwedge S}(r) = v$ *iff* $\Sigma(r) = v$ *for some* $\Sigma \in S$, *and* $v \neq \bot$
   (b) $\Sigma_{\bigwedge S}(r) = \bot$ *otherwise*

We call $\Sigma_{\bigwedge S}$ the conjunction of S.

Theorem 5. *If* S *is a set of complete specifications of source with no value conflict, then* $S \Longrightarrow \Sigma_{\bigwedge S}$.
Proof: $\bigcup\{Ind(\Sigma) : \Sigma \in S\} = Ind(\Sigma_{\bigwedge S})$.

8 Conclusion

In this paper, we have launched what may be called *generic approach* to the formalization of diagrammatic proof systems. The strategy is to characterize a class of diagrammatic systems using a formal property they commonly have. On the basis of this characterization we can investigate other properties that hold of this class of systems.

In this particular paper we define and analyze two related classes of representation systems, namely, SFIS and SFIS$^{\bot}$. Although these systems are rather straightforward to characterize, they include a surprisingly large category of diagrammatic representation systems. In this paper, we have developed a set of generic inference rules whose soundness is provable on the basis of membership in the class of these systems. SFIS and SFIS$^{\perp}$ support a substantial set of generic inference rules, including Contradiction, Proof by Cases, Conjunction and Disjunction, which can serve the foundation for more complex inference rules that are applicable in more sophisticated diagrammatic representation systems.

In future work, we propose to investigate other properties of SFIS, for example their ability to support free rides and diagrammatic consistency-check [6]. Again, we will investigate these ideas in the abstract and demonstrate the conditions under which SFIS have such properties.

In our view, many more sophisticated systems are derivatives of SFIS. We have already discussed one derivative of SFIS, namely, SFIS$^{\perp}$. We can also think of derivatives whose indication relation is amplified through the meaning derivation mechanism [6] or through the introduction of concrete symbols that indicate abstract information. Each of these different kinds of variants seem to define its own class of diagrammatic representation systems (just as SFIS does) and allow the generic approach to the set of inferences rules applicable in all representation systems in the category. This will open a way to a generic diagrammatic proof editor and checker that incorporates a wide range of diagrammatic systems in an incremental, but systematic manner.

References

1. Barker-Plummer, D., Etchemendy, J., Murray, M., Pease, E., Swoboda, N.: Learning to use the openbox: a framework for the implementation of heterogeneous reasoning. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 3–3. Springer, Heidelberg (2012)
2. Barwise, J., Etchemendy, J.: Hyperproof. CSLI Publications, Stanford (1994)
3. Barwise, J., Seligman, J.: Information Flow: The Logic of Distributed Systems. Cambridge University Press, Cambridge (1997)
4. Jamnik, M.: Mathematical Reasoning with Diagrams. CSLI Publications, Stanford (2001)
5. Miller, N.: Euclid and His Twentieth Century Rivals: Diagrams in the Logic of Euclidean Geometry. CSLI Publications, Stanford (2007)
6. Shimojima, A.: Semantic Properties of Diagrams and Their Cognitive Potentials. Studies in the Theory and Applications of Diagrams. CSLI Publications, Stanford (2015)
7. Urbas, M., Jamnik, M.: A framework for heterogeneous reasoning in formal and informal domains. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 277–292. Springer, Heidelberg (2014)

A Logical Investigation of Heterogeneous Reasoning with Graphs in Elementary Economics

Ryo Takemura(✉)

Nihon University, Tokyo, Japan
takemura.ryo@nihon-u.ac.jp

Abstract. Heterogeneous reasoning is a salient component of logic, mathematics, and computer science. Another remarkable field it applies to is economics. In this paper, we apply the proof-theoretic techniques developed in our previous studies [7,8] to heterogeneous reasoning with graphs in elementary economics. We apply the natural deduction-style formalization, which makes it possible to apply well-developed proof-theoretic techniques to the analysis of heterogeneous reasoning with graphs. We also apply the proof-theoretic analysis of free rides developed in [7], and analyze the efficiency of heterogeneous reasoning with graphs. We further discuss abductive reasoning in elementary economics. Abduction has been discussed by philosophers and logicians, and has been extensively studied in the literature on artificial intelligence (see, for example, [2]). In the context of heterogeneous reasoning, we are able to formalize abductive reasoning in elementary economics in the style we employ in our actual reasoning.

1 Reasoning with Graphs in Economics

Because of space limitations, we omit some details. For them, see an extended version of this paper: http://abelard.flet.keio.ac.jp/person/takemura/.

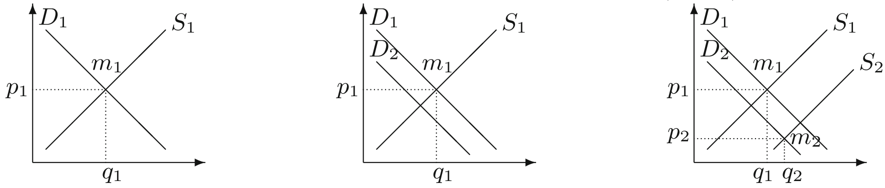
Let us examine the following example of reasoning with graphs in elementary economics, which is a slight modification of an example given in [5].

Example 1 ([5] p. 94 by Krugman and Wells). When a new, faster computer chip is introduced, (1) demand for computers using the older, slower chips decreases. (This graphically corresponds to a leftward shift of the demand curve from the original D_1 to D_2, which we express as $D_2 \leftarrow D_1$.) Simultaneously, (2) computer makers increase their production of computers containing the old chips in order to clear out their stocks of old chips. (Graphically, this corresponds to a rightward shift of the supply curve from the original S_1 to S_2; $S_1 \rightarrow S_2$.) Furthermore, (3) it is widely known that there is only a minor change in the new computer chip, and it does not make computers dramatically faster. That is, the decrease in demand is small relative to the increase in supply. What happens to the equilibrium price and quantity of computers?

In economics, graphs of demand and supply functions are conventionally drawn in a two-dimensional plane, where the vertical axis represents price and the

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 98–104, 2016.
DOI: 10.1007/978-3-319-42333-3_8

horizontal axis represents quantity demanded or supplied. In the above example, there are no concrete demand and supply functions. Hence, we draw demand and supply curves in the simplest manner, i.e., as straight lines with slopes of -1 and 1, respectively, as in the following G_1. We assume that G_1 represents the initial state of the given market, and its equilibrium is $m_1(q_1, p_1)$.



G_1 G_2 G_3

From premise (1), the original demand curve D_1 shifts to D_2, as in G_2. From (2), the supply curve S_1 shifts to S_2, as in G_3. Although we do not know how much D_1 (resp. S_1) shifts to D_2 (resp. S_2), we can infer from (3) that the horizontal shift of the supply curve is greater than that between D_1 and D_2, as expressed in G_3. These shifts lead to the new equilibrium $m_2(q_2, p_2)$. By comparing m_2 and the original m_1, we find $q_1 < q_2$ and $p_1 > p_2$.

Based on the above example, let us investigate the structure of reasoning with graphs in elementary economics. The reasoning in Example 1 goes as follows.

1. An appropriate graph is given, which describes the initial state of a market.
2. We shift a curve based on the given premise, which represents an increase or decrease in demand or supply. This step may be repeated several times. This shifting operation may be considered as the addition of a new curve, since it is convenient to keep the original curve to compare equilibriums at a later point.
3. With this shift in a curve, a new intersection (equilibrium) arises between the demand and supply curves.
4. We compare the new intersection and the original one, and read off the changes in price and quantity.

Let us compare the above graphical reasoning with algebraic reasoning, where we solve simultaneous equations describing given demand and supply functions.

1. Let the given demand function D_1 be $y = -x + \gamma$, and the supply function S_1 be $y = x + \delta$, where γ, δ are real numbers.
2. For some real numbers $\alpha > 0$ and $\beta > 0$ such that $\alpha < \beta$, D_2 can be expressed as $y = -x + \gamma - \alpha$, and S_2 as $y = x + \delta - \beta$.
3. By solving the simultaneous equations D_1 and S_1, we find $q_1 = \frac{\gamma - \delta}{2}$ and $p_1 = \frac{\gamma + \delta}{2}$, which represent the original equilibrium quantity and price.
4. Similarly, by solving D_2 and S_2, we find $q_2 = \frac{\gamma - \delta - \alpha + \beta}{2}$ and $p_2 = \frac{\gamma + \delta - \alpha - \beta}{2}$, which represent the new equilibrium quantity and price.
5. By comparing the equilibrium quantities, we find that $q_1 - q_2 = \frac{\alpha - \beta}{2} < 0$ (since $\alpha < \beta$), and hence, we have $q_1 < q_2$.
6. By comparing the equilibrium prices, we find $p_1 - p_2 = \frac{\alpha + \beta}{2} > 0$, i.e., $p_1 > p_2$.

Although the above calculation is not difficult, it is slightly cumbersome compared with our graphical reasoning. Furthermore, if we formalize it in the framework of mathematical logic, a considerable number of steps are required.

Economic reasoning similar to our example has been studied in the framework of qualitative reasoning, e.g., [3,4]. In qualitative reasoning studies, with the aim of implementation, economic reasoning "without graphs" is investigated. Such a formalization in the framework of qualitative reasoning is considered as another symbolic or linguistic counterpart of our graphical formalization. In some aspects, the economic reasoning we investigate here is an extension of previous research, where either the demand or supply curve is allowed to shift just once. Such an analysis has been extended to a more complicated, multivariable setting [3]. However, we concentrate on analyzing the basic demand and supply market, but allow *simultaneous shifts* of the demand and supply curves.

2 Heterogeneous Logic with Graphs in Economics HLGe

We assume the shift size is specified when we consider the shift in a curve. However, in our qualitative framework, the exact value of the shift is not as significant as the relation between the magnitude of the shifts. Thus, we do not express the shift size as a numeral, but as a constant a that represents some real number. A formula $C \xrightarrow{a} C'$ then means "C shifts rightward to C' with shift width a."

For our heterogeneous system HLGe, we use the following symbols: *Connectives* $\&, \vee, \Rightarrow, \Leftrightarrow, \neg, \forall, \exists$; *Constants for widths* a, b, c; *Constants for coordinates* p, q, r; *Variables for coordinates* x, y; *Curves* D, S, C, B. We also use typical mathematical function symbols such as $+$ and $-$ and predicates such as $=$ and $<$.

Among the usual mathematical formulas, we distinguish the following special formulas in HLGe. A **demand (resp. Supply) curve** is written as $D(x) = -x + r$ (resp. $S(x) = x + r$) for some r. When (q, p) is an **intersection point** of C and C', we write $C \cap C'(q) = p$. We define **shift formulas** as follows:

- $D \xrightarrow{a} D' := \forall x(D(x) = -x + r \Leftrightarrow D'(x) = -x + r + a)$
- $D' \xleftarrow{a} D := \forall x(D(x) = -x + r \Leftrightarrow D'(x) = -x + r - a)$

Similarly for $S \xrightarrow{a} S'$ and $S' \xleftarrow{a} S$.

Definition 1 (Graph). A **graph** in HLGe consists of the following items:

- The first quadrant of the xy-coordinate space.
- Straight lines of slope 1, called **supply curves** and named $S, S', S_1, \ldots$; and of slope -1, called **demand curves** and named $D, D', D_1, \ldots$. When we do not distinguish between them, we denote a curve by $C, C', C_1, \ldots$.
- Every point of intersection of straight lines is accompanied by its coordinates.

Definition 2 (Width). Let C_i and C_j be a pair of lines that are parallel in a graph. Let q_i (and q_j) be the intersection point of C_i (resp. C_j) and the vertical axis when C_i (resp. C_j) is extended as necessary. We define the **width** $w(C_i, C_j)$ between C_i and C_j as $|q_i - q_j|$.
When G is a graph, by $w(G)$, we denote the set of all widths in G.

In contrast to a graph drawn as a diagram, we consider the *type* of a graph, which is a symbolic specification. The type of a graph also defines what kind of information we can extract from it; cf. our inference rule Observe in Definition 6.

Definition 3 (Type). The **type of a graph** G is $(\mathcal{C}, l_w, \mathcal{E}, l_p, l_q)$, where:

- $\mathcal{C}$ is two sequences $D_i \to D_j \to \cdots \to D_n$; $S_k \to S_l \to \cdots \to S_m$ of demand curves and supply curves in G, respectively, which are ordered from left to right as they are in the drawn graph G.
- By allowing equality, i.e., some elements are equal, l_w is the linearly ordered set $w(C_i, C_j) < \cdots\cdots < w(C_k, C_l) < \cdots$ of all widths in G.
- $\mathcal{E}$ is the set of points of intersection in G of the form $D_i \cap S_j(q_k) = p_k$.
- l_p is the linearly ordered set $p_i < p_j < \cdots$ of all y-coordinates of intersections.
- l_q is the linearly ordered set $q_i < q_j < \cdots$ of all x-coordinates of intersections.

The translation of our graphs into first-order formulas is straightforward based on the type of graph.

Definition 4 (Translation of graphs). A graph G of $(\mathcal{C}, l_w, \mathcal{E}, l_p, l_q)$ is translated into a conjunctive formula $\bigwedge \mathcal{C} \& \bigwedge l_2 \& \bigwedge \mathcal{E} \& \bigwedge l_p \& \bigwedge l_q$, where $\bigwedge X$ denotes the conjunction of all corresponding formulas contained in the set X.

For the set-theoretical semantics of HLGe, it is sufficient to employ a domain of real numbers in which arithmetic operations such as $+$ and $-$ are defined. Hence, we provide the real closed field with the ordering relation $<$ as our model. Then, graphs in HLGe are interpreted as follows.

Definition 5 (Interpretation of graphs). Let M be a model. Let G be a graph of $(\mathcal{C}, l_w, \mathcal{E}, l_p, l_q)$, where $\mathcal{C} = D_1 \to D_2 \to \cdots D_n$; $S_1 \to S_2 \to \cdots S_m$, and $l_w = w(C_1, C_2) < w(C_3, C_4) < \cdots < w(C_k, C_l)$. Then, $M \models G$ if and only if

- $M \models D_1 \xrightarrow{w(D_1,D_2)} D_2 \& \cdots \& D_{n-1} \xrightarrow{w(D_{n-1},D_n)} D_n$; and
- $M \models S_1 \xrightarrow{w(S_1,S_2)} S_2 \& \cdots \& S_{m-1} \xrightarrow{w(S_{m-1},S_m)} S_m$; and
- $M \models w(C_1, C_2) < w(C_3, C_4) < \cdots < w(C_k, C_l)$; and
- $M \models \mathcal{E}$, that is, $M \models C \cap C'(q) = p$ for all $C \cap C'(q) = p \in \mathcal{E}$.

The inference rules for HLGe consist of the usual natural deduction rules for first-order formulas and rules for graphs. Our rules for graphs are the following Apply and Observe as in Hyperproof [1].

Definition 6 (Inference rules for graphs of HLGe).

Apply: Let G be a graph, that contains a curve C but does not contain C'. Let $C \xrightarrow{a} C'$ be a shift formula. Let l be an ordering condition that specifies a linear ordering of all widths in $w(G) + C' = w(G) \cup \{w(C', B) \mid B$ is a curve parallel to C (including C) in $G\}$:

$$\frac{G \quad C \xrightarrow{a} C' \quad l}{G'} \text{ Apply}$$

where G' is obtained from G by adding the curve C' so that (1) C' is parallel to C; (2) C' is orthogonal to every curve that is orthogonal to C; (3) the width between C' and C is a; (4) the widths including a satisfy l.
Similarly for $C' \overset{a}{\leftarrow} C$.

Observe: From a given graph G, we can extract, as a conclusion, any corresponding formula contained in the type of G.

When the given ordering condition l does not fully specify a linear ordering among $w(G) + C'$, we cannot apply Apply. In such a case, we enumerate all possible linear orderings of $w(G) + C'$ and apply the $\vee$-elimination rule ($\vee E$) of natural deduction: Let $\{l_1, \ldots, l_n\}$ be the enumeration of all possible linear orderings of $w(G) + C'$ that satisfies the given l. Since $l_1 \vee \cdots \vee l_n$ is provable from l, we divide the cases according to $l_1, \ldots, l_n$ by using $\vee E$, and then, apply Apply in every case as follows:

$$\dfrac{\dfrac{l}{l_1 \vee \cdots \vee l_n} \quad \begin{matrix} \dfrac{G \quad C \overset{a}{\rightarrow} C' \quad [l_1]^m}{G_1}\,\text{Apply} \\ \vdots \\ G'/\psi \end{matrix} \quad \cdots \quad \begin{matrix} \dfrac{G \quad C \overset{a}{\rightarrow} C' \quad [l_n]^m}{G_n}\,\text{Apply} \\ \vdots \\ G'/\psi \end{matrix}}{G'/\psi}\,\vee E, m$$

where G'/ψ denotes that either a graph G' or a first-order formula ψ is obtained, and $[l_i]^m$ denotes the assumption l_i is closed as usual in natural deduction. By regarding the above part of a proof as an inference rule, we call it the rule of Cases.

Example 2 (A proof in HLGe*).* Figure 1 is an example of a proof in HLGe.

It is shown that HLGe can handle simultaneous curve shifts even though Apply (and Cases) is applied in order during a proof.

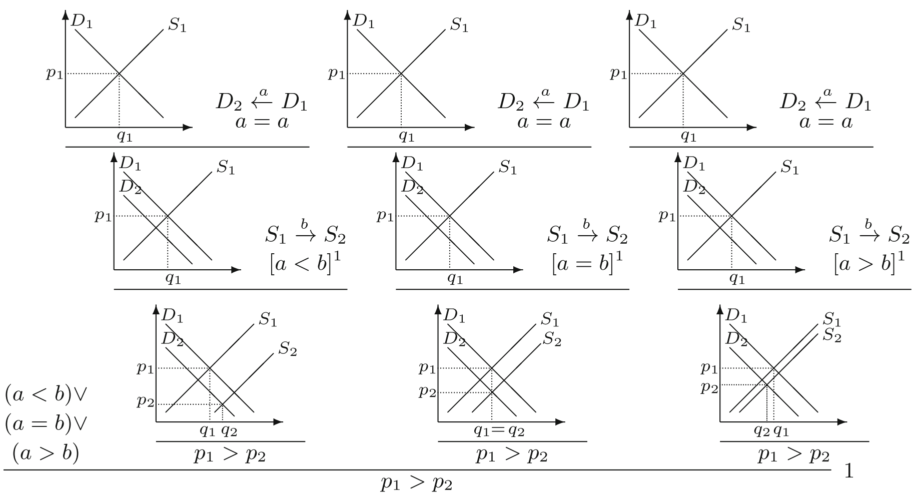


Fig. 1. A proof of $D_2 \overset{a}{\leftarrow} D_1$, $S_1 \overset{b}{\rightarrow} S_2$, $D_1 \cap S_1(p_1) = q_1$, $D_2 \cap S_2(p_2) = q_2 \vdash p_1 > p_2$, which describes the situation of Example 1 without condition (3).

The soundness theorem for HLGe is proved, after dividing several cases, in a similar way to that described in Sect. 1 by algebraic calculation.

Theorem 1 (Soundness). *Let $\mathcal{S}$ be a set of shift formulas; $\mathcal{E}$ be a set of intersections; $\mathcal{O}$ be a set of ordering conditions among widths; and $\mathcal{A}$ be a conjunction of formulas comparing x- and y-coordinates. If $\mathcal{S}, \mathcal{E}, \mathcal{O} \vdash \mathcal{A}$, then $\mathcal{S}, \mathcal{E}, \mathcal{O} \models \mathcal{A}$.*

By slightly extending the notion of *free ride* [6], we refer to diagrammatic objects, or the translated formulas thereof, as **free rides** if they do not appear in the given premise diagrams or sentences, but (automatically) appear in the conclusion after adding pieces of information to the given premise diagrams. The notion of free rides enables us to analyze the effectiveness of each inference rule. (Cf. [7,8].) Let us consider our Apply. We compare the types, or translated formulas, of graphs of premises and the conclusion. Let $G = (\mathcal{C}, l_w, \mathcal{E}, l_p, l_q)$. Then, G' is $(\mathcal{C}', l'_w, \mathcal{E}', l'_p, l'_q)$, where:

- $\mathcal{C}' = \mathcal{C} \cup \{C \to C'\}$, and $l'_w = l$,
- $\mathcal{E}' = \mathcal{E} \cup \{C' \cap B(q) = p \mid B \text{ is orthogonal to } C \text{ in } G\}$,
- l'_p is the linear ordering of $l_p \cup \{p \mid C' \cap B(q) = p \in \mathcal{E}'\}$,
- l'_q is the linear ordering of $l_q \cup \{q \mid C' \cap B(q) = p \in \mathcal{E}'\}$.

Observe that $\mathcal{C}'$ and l'_w are already given in the premises of Apply. On the other hand, the differences between $\mathcal{E}'$ and $\mathcal{E}$, l'_p and l_p, and l'_q and l_q, respectively are free rides of Apply, as they do not appear in the premises.

3 Abduction in Economic Reasoning

Example 3. When a new, faster computer chip is introduced, (1) demand for computers using the older, slower chips decreases (i.e., $D_2 \overset{a}{\leftarrow} D_1$). Simultaneously, (2) computer makers increase their production of computers containing the old chips in order to clear out their stocks of old chips (i.e., $S_1 \overset{b}{\rightarrow} S_2$). When the equilibrium quantity falls in response to these events, what possible explanations are there for this change?

Let $D_1 \cap S_1(q_1) = p_1$ and $D_2 \cap S_2(q_2) = p_2$. First, note that we cannot prove $q_1 > q_2$ under the given premises (1) and (2), as observed in Example 2. Thus, our task in this question is to find a possible explanation H such that $D_2 \overset{a}{\leftarrow} D_1$, $S_1 \overset{b}{\rightarrow} S_2, D_1 \cap S_1(q_1) = p_1, D_2 \cap S_2(q_2) = p_2, H \vdash q_1 > q_2$ holds. In Example 2, the two given premises (1) and (2) provide three graphs, according to whether $a < b, a = b$, or $a > b$, as shown in Fig. 1. Among these three graphs, we find a graph (the third one) in which $q_1 > q_2$ holds. Thus, we know that $q_1 > q_2$ holds when $a > b$ holds for the shift widths of the demand and supply curves. Hence, we can propose $a > b$ as a possible explanation H.

This type of reasoning is called *abduction*, and frequently appears in scientific reasoning. Abduction has been extensively studied in the literature on artificial intelligence AI. In the framework of AI, abduction is usually formalized as the task of finding a hypothesis H that explains a given observation O under a theory T such that O is a logical consequence of T and H, i.e., $T, H \vdash O$, and T, H are consistent. To solve abductive problems, the usual strategy such as resolution and proof-search to construct deductive proofs are applied. (See, for example, [2] for surveys of abduction in AI.) Our strategy in this paper can be considered as a kind of model enumeration. Our inference using graphs in HLGe essentially corresponds to model construction by regarding our graph as a certain kind of representative model. When there is insufficient information on the shift widths of curves, we enumerate all possible cases (i.e., models) by using Cases. We can then determine the required explanation from among these cases. To describe our abductive reasoning more formally, we modify Cases as follows:

$$\dfrac{\dfrac{l}{\overline{\overline{l_1 \vee \cdots \vee l_n}}} \qquad \dfrac{G \quad C \xrightarrow{a} C' \quad \underline{l_i}}{G_i}\ \mathsf{Apply}}{G_i}\ \mathsf{AbCases}$$

where l_i is one of the linear orderings of $l_1, \dots, l_n$, and the underline indicates a proposed explanation. Similarly for $C' \xleftarrow{a} C$.

We formalize our procedure as follows. Let $\mathcal{S}, \mathcal{E}, \mathcal{O}$ be given premises, and $\mathcal{A}$ be a given conclusion or observation. Our task is to find an explanation H such that $\mathcal{S}, \mathcal{E}, \mathcal{O}, H \vdash \mathcal{A}$ holds, where we restrict H to be an ordering condition. (1) We construct a proof of $\mathcal{S}, \mathcal{E}, \mathcal{O} \vdash \mathcal{A}$ by using AbCases as well as our Apply, Observe (and Cases) for HLGe. (2) Among the applications of AbCases, we choose a linear ordering $\underline{l_i}$ that has the maximal length, and set $H = l_i$.

In this paper, we concentrated on a competitive market described by supply and demand models. However, extending our HLGe would enable the investigation of economic reasoning with graphs employed in various other analyses, such as a consumer's optimal consumption analysis and IS-LM analysis.

References

1. Barwise, J., Etchemendy, J.: Hyperproof: For Macintosh. CSLI Publications, Stanford (1995)
2. Denecker, M., Kakas, A.C.: Abduction in logic programming. In: Kakas, A.C., Sadri, F. (eds.) Computational Logic: Logic Programming and Beyond. LNCS (LNAI), vol. 2407, pp. 402–436. Springer, Heidelberg (2002)
3. Farley, A.M., Lin, K.P.: Qualitative reasoning in economics. J. Econ. Dyn. Control **14**, 465–490 (1990)
4. Kalagnanam, J., Simon, H.A., Iwasaki, Y.: The mathematical bases for qualitative reasoning. IEEE Expert **6**(2), 11–19 (1991)
5. Krugman, P.R., Wells, R.: Economics, 3rd edn. Worth Publishers, New York (2012)
6. Shimojima, A.: On the efficacy of representation. Ph.D. thesis, Indiana University (1996)
7. Takemura, R.: Proof theory for reasoning with Euler diagrams: a logic translation and normalization. Studia Logica **101**(1), 157–191 (2013)
8. Takemura, R., Shimojima, A., Katagiri, Y.: Logical investigation of reasoning with tables. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 261–276. Springer, Heidelberg (2014)

Euler and Venn Diagrams

Minimizing Clutter Using Absence in Venn-i^{e}

Jim Burton[1], Mihir Chakraborty[2], Lopamudra Choudhury[2], and Gem Stapleton[1(✉)]

[1] Visual Modelling Group, University of Brighton, Brighton, UK
{j.burton,g.e.stapleton}@brighton.ac.uk
[2] Jadavpur University, Kolkata, India
mihirc4@gmail.com, choudhuryl@yahoo.com

Abstract. Over the last two decades substantial advances have been made in our understanding of diagrammatic logics. Many of these logics have the expressiveness of monadic first-order logic, sometimes with equality, and are equipped with sound and complete inference rules. A particular challenge is the representation of *negated* statements. This paper addresses the problem of how to represent negated statements involving constants, thus asserting the absence of specific individuals, in the context of Euler-diagram-based logics. Our first contribution is to explore the potential benefits of explicitly representing absence using constants, in terms of clutter reduction, and to highlight ontological issues that arise. We go on to define a measure of clutter arising from constants. By defining a set of semantics-preserving inference rules, we are able to algorithmically minimize diagram clutter, in part made possible by the inclusion of absence. Consequently, information about individuals can be represented in a minimally cluttered way.

1 Introduction

Negation, closely related to the notion of absence, plays a crucial role in all logics. Indeed, "The capacity to negate is the capacity to refuse, to contradict, to lie, to speak ironically, to distinguish truth from falsity – in short, the capacity to be human" [8]. It has long been recognized that diagrams are sometimes unable to explicitly represent negated statements. Indeed, many of the logics based on Euler diagrams do not permit statements such as $a \notin P$ to be made explicitly. Instead, one has to assert that $a \in P'$, where P' is the complement of P. There is, however, one exception to this: Choudhury and Chakraborty developed a classical logic called Venn-i that allows $a \notin P$ to be *directly* expressed [5].

Venn-i extends Shin's Venn-I system, which includes Peirce's $\otimes$-sequences to assert non-emptiness of sets [13], alongside i-sequences and $\bar{i}$-sequences to represent individuals and their absence. Since Choudhury and Chakraborty adopt a classical interpretation, the absence of an individual from one set implies its presence in the complement. In Fig. 1, D_1 uses an i-sequence to assert $a \in P \backslash Q$, using an a, and D_2 negates this statement, expressing $a \notin P \backslash Q$, using an $\bar{i}$-sequence. Moreover, D_2 is semantically equivalent to D_3, which expresses

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 107–122, 2016.
DOI: 10.1007/978-3-319-42333-3_9

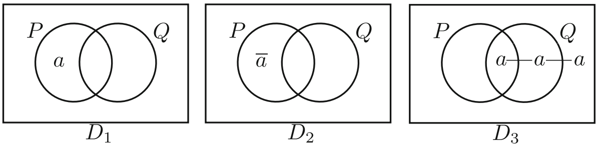


Fig. 1. Asserting presence and absence.

$a \in (P \cap Q) \cup (Q \backslash P) \cup ((U \backslash P) \cap (U \backslash Q))$ using an i-sequence, namely $a - a - a$. An inspiration for Choudhury's and Chakraborty's work came from the notion of *abhāva* (absence). Abhāva, an important feature of ancient Indian knowledge systems, allocates a first class status to the absence of individuals. A philosophical account of absence can be found in [4].

Speaking from the point of view of cognitive science, absence would indicate that though we do not directly perceive the object, we do perceive its absence; there is a mental imagery of the absent object. Thus, when considering a particular individual (of which we have a mental image) we check whether it is in a particular locus and directly perceive its absence. This is reflected by the treatment of Venn-i as a classical logic, where the law of excluded middle holds, as opposed to a sort of constructivist logic where the absence of an individual from one set need not imply its presence in the complement [3].

As we will demonstrate, explicitly representing the absence of individuals allows information to be presented in a less cluttered way. Clutter in Euler diagrams, which are closely related to Venn diagrams, was studied by John et al. [11]: they devised a theoretical measure of clutter. Alqadah et al. established that increased levels of clutter in Euler diagrams negatively impacts user task performance [1]. Hence, there is clearly a need to theoretically understand clutter in diagrams generally and its impact on end-user task performance.

This paper takes the first step towards understanding clutter arising from the sequences in an extended version of Venn-i, which we call Venn-i^{e}, by:

- Discussing the interplay between absence and presence, as well as highlighting their asymmetry (Sect. 2),
- Formalizing the syntax and semantics of Venn-i^{e}, which use Euler diagrams as a basis[1] (Sect. 3),
- Defining a measure of clutter arising from $\otimes$-sequences, i-sequences and $\bar{i}$-sequences (Sect. 4); we note here that i-sequences can comprise many nodes whereas $\bar{i}$-sequences always have a single node,
- Identifying necessary and sufficient conditions for Venn-i^{e} diagrams to be unsatisfiable (Sect. 5),

[1] Using Euler diagrams as a basis renders Swoboda's and Allwein's Euler/Venn logic, which does not include $\bar{i}$-sequences, a proper fragment of Venn-i^{e}. Indeed, many techniques that have been devised for visualizing sets extend Euler diagrams (or variations of them) by the inclusion of individuals, such as [6,12,14]. Whilst not viewed as logics, our work is relevant to these systems since absence provides an alternative way of asserting the set to which an individual belongs.

– Demonstrating how to minimize clutter in satisfiable diagrams by defining inference rules for altering sequences (Sect. 6), and
– Discussing the role of absence in clutter reduction and its potential implications on task performance (Sect. 7).

We conclude and discuss future work in Sect. 8.

2 Representing Absence Diagrammatically

Semantically equivalent statements can be made about the sets in which an individual lies using either *positive* or *negative* statements, such as $a \in P \cup Q$ versus $a \notin P' \cap Q'$ respectively. Whilst various diagrammatic logics include syntax to explicitly make positive statements like $a \in P \cup Q$, including [18,19], they have overlooked the possibility of making negative statements like $a \notin P' \cap Q'$. One benefit of allowing diagrams to make negative statements is that less cluttered diagrams can be formed: using $\overline{a}$ signs in diagrams can be more succinct, relative to diagrams using a signs; see Fig. 2. As previously noted, clutter can have a significant negative impact on diagram comprehension.

In Fig. 3, the three diagrams are semantically equivalent. We can reduce the clutter in D_1 by substituting $\overline{a}$ for the a-sequence, with the result shown in D_2. As well as swapping syntax that makes positive (resp. negative) statements for syntax that makes negative (resp. positive) statements, clutter can also be reduced by removing redundant syntax. The diagram D_3 has more syntax than D_1, such as two additional $\otimes$-sequences, and is more cluttered as a result.

There are fundamental ontological differences between pieces of syntax representing presence and absence. This is because, although syntactically similar, the semantic status of a and $\overline{a}$ signs is different. Firstly, there are differences relating to their locations within a diagram. If there are distinct i-sequences with the same label placed in disjoint regions then the diagram is inconsistent: disjoint regions represent disjoint sets and a given individual cannot be in two disjoint sets. By contrast, $\overline{a}$ can be placed in several disjoint regions without giving rise to inconsistency per se: it is entirely possible for an individual to be absent from two disjoint sets, for instance.

Secondly, we observe that the presence of a sequence, either of the form a or $\overline{a}$, in some region, r, carries existential import. However, this existential import behaves differently: we see that a drawn inside r implies the set, s, that r represents is not empty, whereas $\overline{a}$ drawn in r implies the complement of s is not

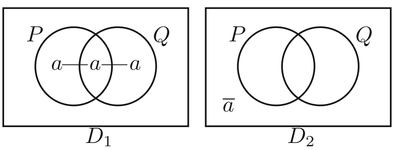


Fig. 2. Making positive (left) and negative (right) statements.

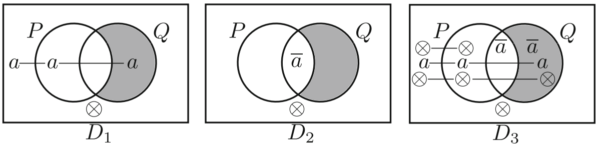


Fig. 3. Diagrams with different levels of clutter.

empty. Thus, the role of absence in terms of existential import is asymmetrical with presence. This may affect the way diagrams are understood by users.

Thirdly, the interaction of absence with subsumption may contradict intuition. In Fig. 4, D_1 tells us that $Q \subseteq P$ and $a \notin Q$. However, this does not imply $a \notin P$, so D_1 does not imply D_2; by contrast, $Q \subseteq P$ and $a \in Q$ implies $a \in P$. This behaviour runs counter to the iconicity [17] of Euler diagrams, which are known to support inference through mechanisms such as free rides [15]. Iconicity is exploited in Euler diagrams through the way that containment indicates subsumption: elements that belong to a set represented by a contained circle belong, "naturally", to the set represented by the containing circle. On the other hand, the absence of an individual from a set represented by a contained circle does not imply absence from the set represented by the containing circle. Thus, with regard to subsumption, $\overline{a}$ does not behave transitively, unlike a.

To summarize, explicitly representing the absence of individuals allows clutter to be reduced in diagrams. Moreover, we must be mindful of various ontological differences between a and $\overline{a}$ when reasoning.

3 Syntax and Semantics of Venn-i^e

Venn-i^e extends Venn-i introduced in [5], relaxing the restriction to Venn diagrams by allowing Euler diagrams to be used. In turn, Venn-i extends Shin's Venn-I system [16]. As is typical, the abstract syntax is given alongside an informal description of the concrete syntax.

Consider the Venn-i^e diagram in Fig. 5. There are two closed curves, labelled P and Q. We conflate the closed curves with their labels and simply say 'the curve P', or just 'P'. The curves give rise to three *zones*: a zone is a region inside some (possibly no) curves and outside the remaining curves. In Fig. 5, the only *shaded* zone is inside Q but outside P. The diagram also contains four graphs:

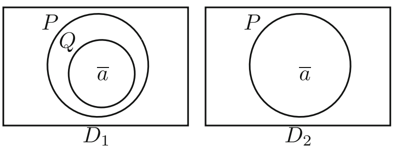


Fig. 4. The interplay between absence and subsumption.

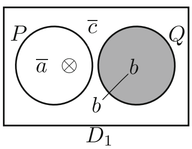


Fig. 5. Syntax: Venn-i^e.

1. One $\otimes$-sequence which comprises a single node,
2. One i-sequence (i for individual), namely b, comprising two nodes joined by one edge, and
3. Two $\bar{i}$-sequences, namely $\bar{a}$ and $\bar{c}$, both of which comprise a single node.

Typically, the abstract syntax for an Euler diagram, D, comprises a set of labels, a set of zones, and a set of shaded zones, written $D = (L, Z, ShZ)$. Zones are ordered pairs of finite, disjoint sets of labels, (in, out), where in (resp. out) denotes the (labels of) the curves that the zone is inside (resp. outside). The zone outside all of the curves, namely $(\emptyset, L)$, must be in D and any zone in D satisfies $in \cup out = L$. The set ShZ of shaded zones only contains zones in Z.

In Fig. 5, the underlying Euler diagram is (L, Z, ShZ), where $L = \{P, Q\}$, $Z = \{(\emptyset, \{P, Q\}), (\{P\}, \{Q\}), (\{Q\}, \{P\})\}$ and $ShZ = \{(\{Q\}, \{P\})\}$. The zone $(\emptyset, \{P, Q\})$ is that which is outside all of the curves, hence the first part of the ordered pair being $\emptyset$ and the second part containing both P and Q. As we shall see, this zone denotes the set $P' \cap Q'$.

It is helpful for us to have a set of labels from which all labels used in any diagram are drawn; we call this set $\mathcal{L}$. When making general statements, we take $\mathcal{L} = \{\lambda_1, \lambda_2, ...\}$ whereas in examples we use P, Q, R, and so forth. Given $\mathcal{L}$, the set of all zones is denoted $\mathcal{Z}$. We also have a set of constant symbols, denoted $\mathcal{C}$, which gives rise to i-sequences and $\bar{i}$-sequences. We take $\mathcal{C} = \{\iota_1, \iota_2, ...\}$; in examples, we use a, b, c, and so forth.

The *regions* (i.e. non-empty sets of zones) in a diagram need to be associated with the sequences drawn in them. In general, $\otimes$-sequences and i-sequences can have nodes placed in many zones, whereas $\bar{i}$-sequences always have a single node. This reflects the dual role of i-sequences and $\bar{i}$-sequences: an a-sequence in the region $\{z_1, ..., z_n\}$ asserts that $a \in z_1 \vee ... \vee a \in z_n$ which is equivalent to $a \notin z_{n+1} \wedge ... \wedge a \notin z_{n+m}$, where $z_{n+1}, ..., z_{n+m}$ are the zones not in r. This equivalent statement can be made by a set of $\bar{a}$-sequences, one in each zone not in r. To identify the sequences in each region, we use three binary relations $\rho_\otimes$, ρ_i and $\rho_{\bar{i}}$. In Fig. 5, $\rho_\otimes = \{(\{(\{P\}, \{Q\})\}, \otimes_1)\}$, $\rho_i = \{(\{(\{Q\}, \{P\}), (\emptyset, \{P, Q\})\}, b)\}$, and $\rho_{\bar{i}} = \{((\{P\}, \{Q\}), a), ((\emptyset, \{P, Q\}), c)\}$.

Definition 1. *A* ***Venn-i^e*** *diagram, D, is a tuple, $D = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\bar{i}})$ such that:*

1. *L is a finite set of labels chosen from $\mathcal{L}$.*
2. *Z is a set of zones where $(\emptyset, L) \in Z$ and for all (in, out) in Z, $in \cup out = L$.*
3. *ShZ is a subset of Z whose elements are called* ***shaded*** *zones.*
4. *$\rho_\otimes \subseteq (\mathbb{P}Z \backslash \{\emptyset\}) \times \{\otimes\}$ is a finite binary relation that associates non-empty regions with $\otimes$ symbols. The elements of $\rho_\otimes$ are called* ***$\otimes$-sequences****.*
5. *$\rho_i \subseteq (\mathbb{P}Z \backslash \{\emptyset\}) \times \mathcal{C}$ is a finite binary relation that associates non-empty regions with constant symbols. The elements of ρ_i are called* ***i-sequences****.*
6. *$\rho_{\bar{i}} \subseteq Z \times \mathcal{C}$ is a finite binary relation that associates zones with constant symbols. The elements of $\rho_{\bar{i}}$ are called* ***$\bar{i}$-sequences****.*

The ***missing zones*** *of D are elements of $MZ = \{(in, out) \in \mathcal{Z} : in \cup out = L\} \backslash Z$. Furthermore, given a constant, ι, the set of ι-sequences in D is denoted*

$I(\iota)$ where $I(\iota) = \{(r,\iota) : (r,\iota) \in \rho_i\}$. Similarly, the set of $\bar{\iota}$-sequences in D is denoted $I(\bar{\iota})$ where $I(\bar{\iota}) = \{(z,\iota) : (z,\iota) \in \rho_{\bar{i}}\}$.

The underlying Euler diagrams have the typical semantics: the closed curves represent sets and their spatial relationships correspond to set-theoretic relationships. Shading asserts emptiness, as seen in Shin's systems [16]. Sequences give information about the location of elements in sets. First, $\otimes$-sequences, introduced by Peirce [13], assert the non-emptiness of sets. Second, i-sequences assert that the denoted individuals are in the sets represented by the regions in which they are placed. Lastly, each $\bar{i}$-sequence asserts the absence of the denoted individual from the set represented by the zone in which it is placed. In Fig. 5, the b-sequence asserts that $b \in Q \cap P'$ *or* $b \in P' \cap Q'$, since b is in the two zone region $\{(\{Q\},\{P\}),(\emptyset,\{P,Q\})\}$. Likewise, the $\bar{c}$-sequence is in the zone $(\emptyset,\{P,Q\})$ which means that $c \notin P' \cap Q'$. To formalize the semantics, we adopt a standard model-theoretic approach.

Definition 2. *An **interpretation**, $\mathcal{I}$, is a triple, $\mathcal{I} = (U, \psi, \Psi)$, such that*

1. *U is a non-empty set, called the **universal set**,*
2. *$\psi\colon \mathcal{C} \to U$ maps constants to elements in U, and*
3. *$\Psi\colon \mathcal{L} \to \mathbb{P}U$ maps curve labels to subsets of U.*

The function Ψ is extended to interpret zones as follows: for each zone, (in, out),

$$\Psi(z) = \bigcap_{l \in in} \Psi(l) \cap \bigcap_{l \in out} (U \backslash \Psi(l)).$$

Definition 3. *Let $D = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\bar{i}})$ be a Venn-i^e diagram and let $\mathcal{I} = (U, \psi, \Psi)$ be an interpretation. Then $\mathcal{I}$ is a **model** for D provided the following conditions all hold.*

1. ***Missing Zones Condition:*** *for each $z \in MZ$, $\Psi(z) = \emptyset$.*
2. ***Shaded Zones Condition:*** *for each $z \in ShZ$, $\Psi(z) = \emptyset$.*
3. ***$\otimes$-Sequence Condition:*** *for each $(r, \otimes) \in \rho_\otimes$, $\Psi(z) \neq \emptyset$ for some $z \in r$.*
4. ***i-Sequence Condition:*** *for each $(r, \iota) \in \rho_i$, $\psi(\iota) \in \Psi(z)$ for some $z \in r$.*
5. **$\bar{i}$-Sequence Condition:** *for each $(z, \iota) \in \rho_{\bar{i}}$, $\psi(\iota) \notin \Psi(z)$.*

*If $\mathcal{I}$ models D then $\mathcal{I}$ **satisfies** D. Diagrams with no models are **unsatisfiable**.*

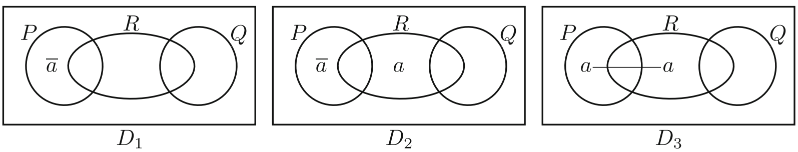


Fig. 6. Measuring clutter in Venn-i^e diagrams.

4 Measuring Clutter

We require a measure of clutter arising from the sequences. Figure 6 shows three simple examples, all with the same underlying Euler diagram. The lefthand diagram with just one node, namely $\overline{a}$, is less cluttered than the middle diagram. The righthand diagram is the most cluttered, since this has two nodes (both named a) *and* a connecting edge. Thus, to measure the clutter arising from the sequences, we count the number of nodes and the number of edges.

Definition 4. *Let* $D = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\overline{i}})$ *be a Venn-i*^{e} *diagram. The* ***sequence clutter score*** *for* D*, denoted* $\mathcal{SCS}(D)$*, is*

$$\mathcal{SCS}(D) = \left(\sum_{(r,\otimes)\in\rho_\otimes} (2|r| - 1) \right) + \left(\sum_{(r,\iota)\in\rho_i} (2|r| - 1) \right) + |\rho_{\overline{i}}|$$

The three diagrams in Fig. 6 have sequence clutter scores 1, 2, and 3 respectively. From this point forward, we simply say *clutter score*.

Definition 5. *Let* $D_1 = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\overline{i}})$ *be a Venn-i*^{e} *diagram. Then* D *is* ***minimally cluttered*** *if there does not exist a semantically equivalent diagram,* $D' = (L, Z, ShZ, \rho'_\otimes, \rho'_i, \rho'_{\overline{i}})$*, such that* $\mathcal{SCS}(D') < \mathcal{SCS}(D)$*.*

5 Minimizing Clutter in Inconsistent Diagrams

Figure 7 shows a minimally cluttered inconsistent diagram, namely D_1: it has a clutter score of 0; thus, any inconsistent diagram is semantically equivalent to D_1. To allow us to focus on consistent diagrams, when algorithmically reducing clutter, we need to identify syntactic conditions which capture inconsistency. There are various ways in which Venn-i^e diagrams can be inconsistent:

1. All interpretations have a non-empty universal set, so a diagram is inconsistent if it is entirely shaded. See D_1 in Fig. 7.
2. Shaded regions containing entire $\otimes$-sequences or i-sequences are inconsistent since the shading asserts set emptiness whereas the sequence implies set non-emptiness. See D_2 in Fig. 7, where each sequence gives rise to inconsistency.
3. There are i-sequences placed in regions that do not share a common non-shaded zone, z, where z does not contain an $\overline{a}$-sequence. Intuitively, for each ι, the individual represented must lie in the set denoted by a non-shaded zone that is shared by all ι-sequences in $I(\iota)$. If all such zones include $\overline{\iota}$ then the diagram also asserts that the individual is absent from the sets represented by those zones. See D_3 in Fig. 7.

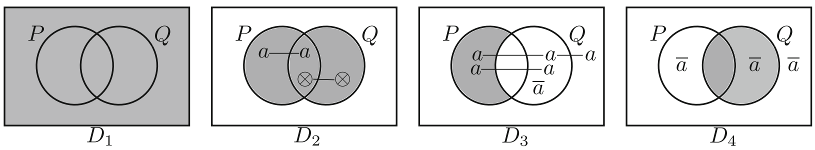


Fig. 7. Inconsistent Venn-i^e diagrams.

4. The set of $\bar{\imath}$-sequences for constant symbol ι, namely $I(\bar{\iota})$, cannot include all non-shaded zones. If all non-shaded zones were included then the law of excluded middle tells us that the represented individual must be an element of the empty set. See D_4 in Fig. 7.

Definition 6 (Inconsistency). *Let $D = (L, Z, ShZ, \rho_{\otimes}, \rho_i, \rho_{\bar{i}})$ be a Venn-i^e diagram. Whenever any one of the following conditions holds D is* ***inconsistent****.*

1. *All zones are shaded: $Z = ShZ$.*
2. *There is an $\otimes$-sequence, say $(r, \otimes)$, in D such that $r \subseteq ShZ$.*
3. *There is an i-sequence, say (r, ι), in D such that for all zones, z, in $\bigcap_{r' \in I(\iota)} r'$, either z is shaded or $(z, \iota) \in I(\bar{\iota})$.*
4. *There is an $\bar{\imath}$-sequence, say (z, ι), in D such that $Z \backslash \{z' : (z', \iota) \in I(\bar{\iota})\} \subseteq ShZ$.*

If D is not inconsistent then D is ***consistent****.*

Theorem 1 (Inconsistent). *D is inconsistent iff D is unsatisfiable.*

Using Theorem 1 we can therefore identify whether any given diagram is inconsistent. Given such a diagram $D = (L, Z, Z, \rho_{\otimes}, \rho_i, \rho_{\bar{i}})$ we can see that a minimally cluttered, semantically equivalent diagram is $D_{min} = (L, Z, Z, \emptyset, \emptyset, \emptyset)$.

6 Minimizing Clutter in Consistent Diagrams

The goal of this section is to produce minimally cluttered diagrams using inference rules that alter their sequences. To this end, we first define some useful transformations on diagrams.

Transformation 1 (Sequence Removal). *Let $D = (L, Z, ShZ, \rho_{\otimes}, \rho_i, \rho_{\bar{i}})$ be a Venn-i^e diagram. Let $(r, \bullet)$ be a sequence in D. We define three removal operations on D:*

1. *If $(r, \bullet) \in \rho_{\otimes}$ then $D - (r, \bullet) = (L, Z, ShZ, \rho_{\otimes} \backslash \{(r, \bullet)\}, \rho_i, \rho_{\bar{i}})$.*
2. *If $(r, \bullet) \in \rho_i$ then $D - (r, \bullet) = (L, Z, ShZ, \rho_{\otimes}, \rho_i \backslash \{(r, \bullet)\}, \rho_{\bar{i}})$.*
3. *If $(r, \bullet) \in \rho_{\bar{i}}$ then $D - \overline{(r, \bullet)} = (L, Z, ShZ, \rho_{\otimes}, \rho_i, \rho_{\bar{i}} \backslash \{(r, \bullet)\})$.*

Transformation 2 (Sequence Addition). *Let $D = (L, Z, ShZ, \rho_{\otimes}, \rho_i, \rho_{\bar{i}})$ be a Venn-i^e diagram. Let $(r, \bullet)$ be a sequence such that $r \subseteq Z$ or $r \in Z$. We define three addition operations on D:*

1. *If $r \subseteq Z$ and $\bullet = \otimes$ then $D + (r, \bullet) = (L, Z, ShZ, \rho_{\otimes} \cup \{(r, \bullet)\}, \rho_i, \rho_{\bar{i}})$.*
2. *If $r \subseteq Z$ and $\bullet \in \mathcal{C}$ then $D + (r, \bullet) = (L, Z, ShZ, \rho_{\otimes}, \rho_i \cup \{(r, \bullet)\}, \rho_{\bar{i}})$.*
3. *If $r \in Z$ and $\bullet \in \mathcal{C}$ then $D + \overline{(r, \bullet)} = (L, Z, ShZ, \rho_{\otimes}, \rho_i, \rho_{\bar{i}} \cup \{(z, \bullet)\})$.*

Before we present our inference rules, we work through an example showing how to minimize clutter. Consider Fig. 8. Here, the diagram D is consistent, but not minimally cluttered. To reduce clutter, we make various observations and adopt the following process:

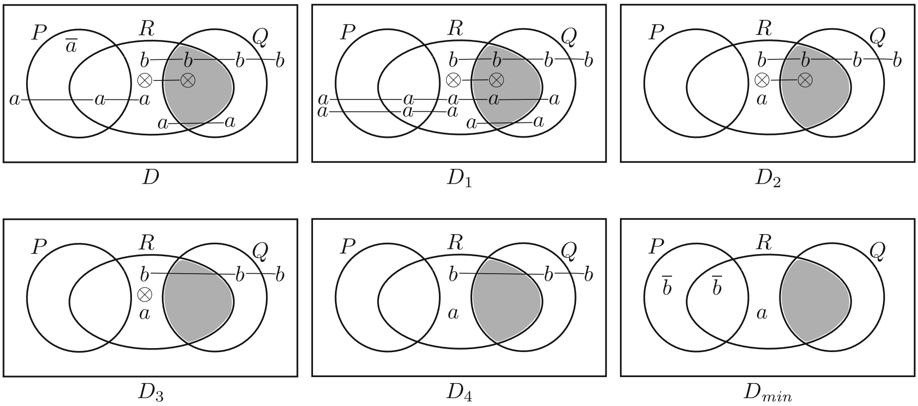


Fig. 8. Clutter reduction in consistent diagrams.

1. First we observe that whenever we express information using $\bar{i}$-sequences, we can instead use an i-sequence. Thus, in D we can swap $\overline{a}$ for an a-sequence, as shown in D_1. In general, this swap may result in the clutter score increasing, but it allows us to more easily identify, syntactically, the region in which a must represent an element.
2. Next, we observe that in D_1 (and, in any diagram), we only need one occurrence of each constant symbol to specify in which set it lies. So, we can reduce the three a-sequences in D_1 to a single a-sequence shown in D_2. This single a-sequence is placed in the zone common to all of the a-sequences in D_1, thus allowing us to see which region contains the individual a. In this step, the clutter score of D_2 is lower than that of D_1.
3. Reductions can also be made to sequences that are placed in regions which contain shaded zones, since shaded zones represent empty sets. The diagram D_2 contains two such sequences, (b, r_b) and $(\otimes, r_\otimes)$, and can be replaced by D_3.
4. Some sequences can be redundant from diagrams. In D_3, the $\otimes$-sequence is redundant since it tells us that $Q \backslash (P \cup R) \neq \emptyset$ which can be deduced from the a-sequence. So D_3 can be replaced by D_4.
5. Lastly, we examine each i-sequence in turn. If its contribution to the clutter score can be reduced by swapping it for $\bar{i}$-sequences then this swap is performed. Here, the b-sequence is swapped for two $\bar{b}$-sequences, resulting in D_{min}. This last step exploits the use of absence to reduce diagram clutter.

As we have just seen, it is possible to swap i-sequences for $\bar{i}$-sequences, and vice versa, reflecting their dual roles. For example, in Fig. 9, the a-sequence in D_1 tells us $a \in P \cap Q' \cap R'$ or $a \in P \cap Q \cap R'$. Given the shading and the spatial relationships between the curves, asserting $a \notin P' \cap Q' \cap R'$ is equivalent. This alternative representation is seen in D_2. We can *swap* the i-sequence $(\{(\{P\}, \{Q, R\}), (\{P, Q\}, \{R\})\}, a)$ for the $\bar{i}$-sequence $((\emptyset, \{P, Q, R\}), a)$.

Inference Rule 1 (Swap i-Sequence). *Let* $D = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\bar{i}})$ *be a Venn-*i^e *diagram. Let* (r, ι) *be an* i*-sequence in* D*. Then* (r, ι) *may be swapped for the set* $\{(z, \iota) : z \in Z \backslash (ShZ \cup r)\} = \{\overline{(z_1, \iota)}, ..., \overline{(z_n, \iota)}\}$ *of* $\bar{i}$*-sequences. That is,* D *may be replaced by* $D - (r, \iota) + \overline{(z_1, \iota)} + ... + \overline{(z_n, \iota)}$ *and vice versa.*

Inference Rule 2 (Swap $\bar{i}$-Sequences). *Let* $D = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\bar{i}})$ *be a Venn-*i^e *diagram. Let* ι *be a constant symbol such that* $I(\bar{\iota}) \neq \emptyset$ *and* $Z \backslash (ShZ \cup \{z_1, ..., z_n\}) \neq \emptyset$*, where* $I(\bar{\iota}) = \{(z_1, \iota), ..., (z_n, \iota)\}$*. Then* $I(\bar{\iota})$ *may be swapped for the* i*-sequence* $(Z \backslash (ShZ \cup \{z_1, ..., z_n\}), \iota)$*. That is,* D *may be replaced by*

$$D - \overline{(z_1, \iota)} - ... - \overline{(z_n, \iota)} + (Z \backslash (ShZ \cup \{z_1, ..., z_n\}), \iota)$$

and vice versa.

There are also occasions when we can remove parts of sequences: when the region in which a sequence is placed includes a shaded zone, the part in the shaded zone can be deleted, thus reducing the sequence. Moreover, we have also seen that sets of i-sequences can be reduced.

Inference Rule 3 (Reduce Sequence). *Let* $D = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\bar{i}})$ *be a Venn-*i^e *diagram. Let* $(r, \bullet)$ *be a sequence in* D *such that* r *contains at least two zones, one of which,* z *say, is shaded. Then* D *may be replaced by* $D - (r, \bullet) + (r \backslash \{z\}, \bullet)$ *and vice versa. Such a sequence is said to be* ***reducible*** *in* D*.*

Inference Rule 4 (Reduce a Set of Sequences). *Let* $D = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\bar{i}})$ *be a Venn-*i^e *diagram. Let* ι *be a constant symbol such that* $I(\iota) \neq \emptyset$ *and* $r \neq \emptyset$ *where*

$$r = \Big(\bigcap_{(r_i, \iota) \in I(\iota)} r_i \Big) \backslash \Big(ShZ \cup \bigcup_{(z, \iota) \in I(\bar{\iota})} \{z\} \Big).$$

Then D *may be replaced by*

$$D - (r_1, \iota) - ... - (r_n, \iota) + (r, \iota)$$

and vice versa, where $I(\iota) = \{(r_1, \iota), ..., (r_n, \iota)\}$*. The set* $I(\iota)$ *of sequences is said to be* ***reducible*** *in* D*.*

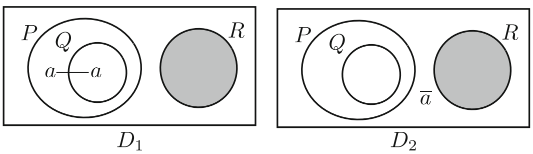


Fig. 9. Swapping sequences.

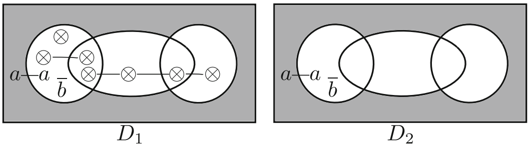


Fig. 10. Redundant $\otimes$-sequences.

There are various ways in which an $\otimes$-sequence can be redundant in a diagram, in the sense that its removal does not alter the semantics:

1. An $\otimes$-sequence, $(r, \otimes)$, that includes all of the non-shaded zones in r is redundant, since this amounts to asserting that $U \neq \emptyset$ which is necessarily true in all interpretations.
2. In D_1, Fig. 10, the single-node $\otimes$-sequence asserts $P \cap Q' \cap R' \neq \emptyset$. From this we can deduce $P \neq \emptyset$, asserted by the two-node $\otimes$-sequence which is, thus, redundant.
3. In D_1, the a-sequence tells us that $a \in P \cap Q' \cap R'$ or $a \in P' \cap Q' \cap R'$. The shading asserts $P' \cap Q' \cap R' = \emptyset$, so $a \in P \cap Q' \cap R'$. This implies that $P \cap Q' \cap R' \neq \emptyset$, so the single-node $\otimes$-sequence is also redundant. The *presence* of the individual a has permitted a reduction in diagram clutter.
4. Lastly, in D_1 the location of $\bar{b}$ tells us that $b \notin P \cap Q' \cap R'$, from which – together with the shading – it follows that $b \in Q \cup R$. Therefore, the four-node $\otimes$-sequence asserting $Q \cup R \neq \emptyset$ is redundant. The *absence* of the individual b has permitted a reduction in diagram clutter.

Removing the $\otimes$-sequences from D_1 in Fig. 10 to give D_2 reduces the clutter score from 15 to 4.

Inference Rule 5 (Remove $\otimes$-Sequence). *Let $D_1 = (L, Z, ShZ, \rho_\otimes, \rho_i, \rho_{\bar{i}})$ be a Venn-i^e diagram and let $(r, \otimes)$ be an $\otimes$-sequence in D such that either:*

1. *The region r includes all non-shaded zones: $Z \backslash ShZ \subseteq r$,*
2. *There is a distinct $\otimes$-sequence, say $(r', \otimes)$, in D where $r' \backslash ShZ \subseteq r$,*
3. *There is an i-sequence, say (r', ι), in D such that $r' \backslash ShZ \subseteq r$, or*
4. *Given $I(\bar{\iota}) = \{(z_1, \iota), ..., (z_n, \iota)\}$, it is the case that $(Z \backslash (ShZ \cup \{z_1, ..., z_n\})) \subseteq r$.*

Then D can be replaced by $D - (r, \otimes)$ and vice versa and we say $(r, \otimes)$ is ***redundant*** *in D.*

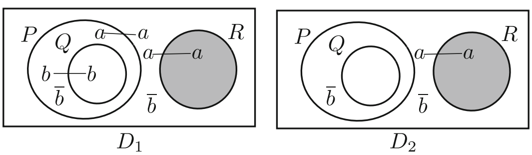


Fig. 11. Redundant i-sequences.

Considering i-sequences, in Fig. 11 two of them are redundant in D_1:

1. The a-sequence with a node in the shaded region tells us that $a \in P' \cap Q' \cap R'$. From this, we can deduce that $a \in P' \cap Q' \cap R'$ or $a \in P \cap Q' \cap R'$, asserted by the other a-sequence, so this second a-sequence is redundant.
2. The presence of the two $\bar{b}$-sequences, together with the shading, allows us to infer that $b \in P \cap Q' \cap R'$ or $b \in P \cap Q \cap R'$, expressed by the b-sequence. Thus, the b-sequence is redundant. Again, we see that the *absence* of the individual b has permitted a reduction in diagram clutter.

Removing the i-sequences from D_1 reduces the clutter score from 11 to 5 in D_2.

Inference Rule 6 (Remove i-Sequence). *Let $D_1 = (L, Z, ShZ, \rho_{\otimes}, \rho_i, \rho_{\bar{i}})$ be a Venn-i^e diagram and let (r, ι) be an i-sequence in D such that either:*

1. *the region r includes all non-shaded zones: $Z \setminus ShZ \subseteq r$,*
2. *there is a distinct i-sequence, (r', ι), in D such that $r' \setminus ShZ \subseteq r$, or*
3. *there is a set of $\bar{i}$-sequences, say $I = \{(z_1, \iota), ..., (z_n, \iota)\}$ such that $(r \cup \{z_1, ..., z_n\}) \setminus ShZ = Z \setminus ShZ$.*

Then D can be replaced by $D - (r, \iota)$ and vice versa and we say (r, ι) is ***redundant*** *in D.*

Importantly, all inference rules preserve semantics. In addition, other than the swap rules, applying them never increases diagram clutter. These two properties are captured in Theorem 2.

Theorem 2 (Soundness and Clutter Reduction). *Let D and D' be a Venn-i^e diagrams such that D' is obtained from D by applying one of the inference rules. Then D and D' are semantically equivalent and if the inference rule applied was not a swap rule then the clutter score of D' is at most that of D.*

We are now in a position to show how to minimize clutter in consistent diagrams. Algorithm 1 presents the steps in detail. Referring to Fig. 8, the input to Algorithm 1 is D. Step 1 iteratively removes $\bar{i}$-sequences using inference rule 2, of which D has just one (namely $\bar{a}$), to give D_1. Step 2 iteratively reduces sets of i-sequences using inference rule 4. In this case, the set of a-sequences is reducible and the result is shown in D_2. Taking D_2, step 3 reduces all reducible sequences using inference rule 3; here the result is D_3, where two sequences have altered due to the presence of shading. Step 4 proceeds to remove redundant sequences using inference rule 5, resulting in D_4. Lastly, step 5 inspects the i-sequences to see whether clutter is reduced by swapping them for $\bar{i}$-sequences. In this case, it is beneficial to swap b for two $\bar{b}$s: the b-sequence contributes 5 to the clutter score, whereas the (swapped) $\bar{b}$-sequences in D_{min} contribute just 2. However, the a-sequence contributes only 1 to the clutter score of D_4, so is retained, not swapped. D_{min} is the output from Algorithm 1. Lastly, we note that minimally cluttered diagrams are not, in general, unique. It should be clear from the last step of Algorithm 1 that it is sometimes possible to swap sequences without altering the clutter score.

Theorem 3 (Clutter Minimization). *Let D be a consistent Venn-i^e diagram and let D_{min} be the result of applying Algorithm 1 to D. Then D and D_{min} are semantically equivalent and D_{min} is minimally cluttered.*

The proof can be found online [2].

Algorithm 1. Clutter Minimization

Input: a consistent diagram $D = (L, Z, ShZ, \rho_{\otimes}, \rho_i, \rho_{\bar{i}})$.
Minimise the clutter in D using the following steps:

1. Iteratively swap all $\bar{i}$-sequences, (z, ι), in D for an i-sequence using rule 2. Call the resulting diagram D_1.
2. Iteratively reduce all reducible sets of sequences in D_1 using rule 4 until no reducible sets of sequences remain. Call the resulting diagram D_2.
3. Iteratively reduce all reducible sequences in D_2 using rule 3 until no reducible sequences remain. Call the resulting diagram D_3.
4. Iteratively remove redundant sequences from D_3 using rules 5 and 6 until no redundant sequences remain. Call the resulting diagram D_4.
5. Swap all i-sequences, (r, ι), in D_4, where

$$|Z \backslash (ShZ \cup r)| < 2|r| - 1,$$

   for $\bar{i}$-sequences using rule 1. Call the resulting diagram D_{min}.

Output: D_{min}.

---

7 Cognitive Implications

As we have seen, it is possible to reduce clutter in a diagram by removing sequences, reducing them and swapping between i-sequences and $\bar{i}$-sequences. Whilst earlier research into diagram clutter has established that increasing clutter levels correlates with decreased task performance, it is unclear whether and when this remains true for Venn-i^e diagrams. We conjecture that the impact of clutter on task performance will be task dependent.

For instance, consider the semantically equivalent diagrams in Fig. 8 and suppose that we are asked to determine the set in which the individual a lies. We conjecture that this task is easier to perform by studying D_{min} than by studying D. This is because a is more salient in D_{min}, due to the reduced amount of syntax present: this could make it quicker to identify the location of a. Thus, for this task, it could be that D_{min} promotes improved task performance.

Suppose now that our task is to determine whether b is not in P. D_{min} explicitly represents this information using absence (i.e. $\bar{b}$), whereas it must be deduced from D: identify the location of b and deduce that b is not in P. Here, we conjecture that the use of absence has *directly* aided performance. Indeed, there are other tasks for which neither D nor D_{min} are potentially 'optimal'. For example, suppose we wish to determine the set in which b lies. Perhaps the best representation of this information is D_4, which includes a three-node b-sequence (by contrast, D, D_1, D_2 and D_3 are more cluttered). From D_4, we can read off the fact that either b is in just R, b is in just Q, or b is in none of P, Q, and R.

In summary, these examples demonstrate that the diagram that best supports task performance need not be that which is minimally cluttered. There is likely to be trade-off between clutter and directly representing statements of interest, using either absence or presence information. There is clearly an interplay between diagram clutter, the use of syntax to represent presence versus absence and task performance. It is an interesting avenue of future work to explore, empirically, the relationship between diagram clutter and the directness of information representation with respect to task performance.

8 Discussion and Conclusion

In this paper we have explored the potential cognitive benefits of directly representing the absence of individuals in Euler diagram logics. Through identifying sound inference rules, and conditions under which diagrams are inconsistent, we have been able to algorithmically produce minimally cluttered Venn-i^e diagrams. As a consequence, it is possible to represent information about sets and their elements in a minimally cluttered way. The inspiration for this research was derived from related work on Euler diagrams which established that increasing levels of clutter diminished task performance. Our discussion above highlights that the case for reducing clutter in Venn-i^e diagrams, as a way of improving task performance, is less clear cut. Our results lay an essential foundation for empirically evaluating the impact of clutter from this perspective.

As well as empirical research, future work also includes considering clutter and absence in non-classical logics. In our interpretation of Venn-i^e, $\overline{a}$ is syntactic sugar of which we have made use for its practical ability to reduce clutter. There are two other (non-classical) interpretations of Venn-i^e, explored in Choudhury and Chakrabory's work [3]:

1. The absence of a in P does not necessarily imply a is in the complement of P, and
2. The universe is open, so the complement of P does not exist.

In our opinion, the two alternative interpretations are interesting from the point of view of the philosophy and logic of diagrams, and we plan to make them the subject of future work. In the first interpretation, we can represent recursively enumerable sets, which have many important applications in computer science and elsewhere. In the second interpretation since P' does not exist it is also the case that $\overline{a} \in P$ does not imply $a \in P'$. The implications of diagrammatic reasoning with an open universe is an interesting and open topic. Lastly, the use of absence could be incorporated into other Euler-diagram-based logics, such as spider diagrams [9], Euler/Venn diagrams [19], constraint diagrams [7] and concept diagrams [10].

Acknowledgement. This research was supported by EPSRC grant EP/M011763/1.

References

1. Alqadah, M., Stapleton, G., Howse, J., Chapman, P.: Evaluating the impact of clutter in Euler diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 108–122. Springer, Heidelberg (2014)
2. Burton, J., Chakraborty, M., Choudhury, L., Stapleton, G.: Technical report: minimizing clutter using absence in Venn-i^e. http://readableproofs.org/diagrams-2016-constants-and-absence-technical-report. Accessed March 2016
3. Choudhury, L., Chakraborty, M.: On representing open universe. Stud. Logic **5**(11), 96–112 (2012)
4. Choudhury, L., Chakraborty, M.: Singular propositions, negation and the square of opposition. Log. Univers. **10**(2), 215–231 (2016)
5. Choudhury, L., Chakraborty, M.K.: On Extending Venn Diagram by Augmenting Names of Individuals. In: Blackwell, Alan F., Marriott, Kim, Shimojima, Atsushi (eds.) Diagrams 2004. LNCS (LNAI), vol. 2980, pp. 142–146. Springer, Heidelberg (2004)
6. Collins, C., Penn, G., Sheelagh, M., Carpendale, T.: Bubble sets: revealing set relations with isocontours over existing visualizations. IEEE Trans. Vis. Comput. Graph. **15**(6), 1009–1016 (2009)
7. Fish, A., Flower, J., Howse, J.: The semantics of augmented constraint diagrams. J. Vis. Lang. Comput. **16**, 541–573 (2005)
8. Horn, L.: A Natural History of Negation, 1st edn. CSLI Lecture Notes. Center for the Study of Language and Information, Stanford (2001)
9. Howse, J., Stapleton, G., Taylor, J.: Spider diagrams. LMS J. Comput. Math. **8**, 145–194 (2005)
10. Howse, J., Stapleton, G., Taylor, K., Chapman, P.: Visualizing ontologies: a case study. In: Aroyo, L., Welty, C., Alani, H., Taylor, J., Bernstein, A., Kagal, L., Noy, N., Blomqvist, E. (eds.) ISWC 2011, Part I. LNCS, vol. 7031, pp. 257–272. Springer, Heidelberg (2011)
11. John, C., Fish, A., Howse, J., Taylor, J.: Exploring the notion of 'clutter' in Euler diagrams. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) Diagrams 2006. LNCS (LNAI), vol. 4045, pp. 267–282. Springer, Heidelberg (2006)
12. Meulemans, W., Riche, N.H., Speckmann, B., Alper, B., Dwyer, T.: Kelpfusion: a hybrid set visualization technique. IEEE Trans. Vis. Comput. Graph. **19**(11), 1846–1858 (2013)

13. Peirce, C.: Collected Papers, vol. 4. Harvard University Press, Massachusetts (1933)
14. Riche, N., Dwyer, T.: Untangling Euler diagrams. IEEE Trans. Vis. Comput. Graph. **16**(6), 1090–1099 (2010)
15. Shimojima, A.: Inferential and expressive capacities of graphical representations: survey and some generalizations. In: Blackwell, A.F., Marriott, K., Shimojima, A. (eds.) Diagrammatic Representation and Inference. LNCS, vol. 2980, pp. 18–21. Springer, Heidelberg (2004)
16. Shin, S.-J.: The Logical Status of Diagrams. Cambridge University Press, Cambridge (1994)
17. Shin, S.-J.: The Iconic Logic of Peirce's Graphs. Bradford Book, Bradford (2002)
18. Stapleton, G., Taylor, J., Howse, J., Thompson, S.: The expressiveness of spider diagrams augmented with constants. J. Vis. Lang. Comput. **20**, 30–49 (2009)
19. Swoboda, N., Allwein, G.: Heterogeneous reasoning with Euler/Venn diagrams containing namedconstants and FOL. In: Euler Diagrams 2004. ENTCS, vol. 134. Elsevier (2005)

Human Reasoning with Proportional Quantifiers and Its Support by Diagrams

Yuri Sato[1(✉)] and Koji Mineshima[2]

[1] Interfaculty Initiative in Information Studies, The University of Tokyo, Tokyo, Japan
satoyuri0@gmail.com

[2] Center for Simulation Sciences, Ochanomizu University, Tokyo, Japan
mineshima.koji@ocha.ac.jp

Abstract. In this paper, we study the cognitive effectiveness of diagrammatic reasoning with proportional quantifiers such as *most*. We first examine how Euler-style diagrams can represent syllogistic reasoning with proportional quantifiers, building on previous work on diagrams for the so-called plurative syllogism (Rescher and Gallagher, 1965). We then conduct an experiment to compare performances on syllogistic reasoning tasks of two groups: those who use only linguistic material (two sentential premises and one conclusion) and those who are also given Euler diagrams corresponding to the two premises. Our experiment showed that (a) in both groups, the speed and accuracy of syllogistic reasoning tasks with proportional quantifiers like *most* were worse than those with standard first-order quantifiers such as *all* and *no*, and (b) in both standard and non-standard (proportional) syllogisms, speed and accuracy for the group provided with diagrams were significantly better than the group provided only with sentential premises. These results suggest that syllogistic reasoning with proportional quantifiers like *most* is cognitively complex, yet can be effectively supported by Euler diagrams that represent the proportionality relationships between sets in a suitable way.

Keywords: Euler diagrams · Proportional quantifiers · Reasoning · Logic and cognition

1 Introduction

Euler diagrams have been used to represent various set-theoretical properties and relations. We can distinguish three types:

(i) The basic inclusion and exclusion relations between sets, as represented by sentences like *All A are B* and *No A are B*;
(ii) The proportionality relationship between sets, as represented by sentences like *Most A are B* and *More than half of A are B*;
(iii) The cardinality of sets, represented by sentences like *Three A are B*, *More than three A are B*, and *Less than three A are B*.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 123–138, 2016.
DOI: 10.1007/978-3-319-42333-3_10

In previous cognitive studies on Euler diagrams [16,17], empirical evidence has been found in support of the effectiveness of diagrams with respect to (i). However, whether and how diagrams can also be effective in representing and reasoning about (ii) and (iii) still remains to be explored. In this paper, we will focus on (ii), i.e., the proportionality relationship between sets, and examine the cognitive effectiveness of diagrams to represent proportional quantification as expressed by a sentence like *Most A are B*.

As we will review in Sect. 2, logical properties of proportional quantifiers like *most* have been the focus of recent research on quantification in logic and linguistics. In logical and cognitive studies on diagrams, however, although Euler and Venn diagrams are widely used to represent and reasoning with quantified statements, little has been discussed about how they can represent proportional quantification and how effective they can be in actual reasoning. This paper is a first step to fill this gap.

The structure of this paper is as follows. In Sect. 2, we present backgrounds on proportional quantifiers in logical, computational and cognitive studies. In Sect. 3, we analyze diagrammatic representations (Sect. 3.1) and diagrammatic inferences (Sect. 3.2) for proportional quantifiers in order to generate predictions (Sect. 3.3). In Sect. 4, we report the results of an experiment comparing participants' performance in solving proportional syllogism with and without diagrams. Section 5 concludes the paper with a summary and discusses some directions for future work.

2 Background on Proportional Quantifiers

We will start with a brief overview of the logical, computational and cognitive properties of proportional quantifiers. This provides the necessary background information about the main issue in this paper.

Logic. Natural languages use many expressions of quantification. Among them, quantifiers such as *all*, *some* and *no* can be represented within first-order logic, using the *unary* quantifiers $\forall$ and $\exists$ and other logical connectives. Thus, the sentence *All A are B* is represented as $\forall x(Ax \to Bx)$, the sentence *Some A are B* as $\exists x(Ax \wedge Bx)$, and the sentence *No A are B* as $\forall x(Ax \to \neg Bx)$.

In contrast, it is known that quantifiers that denote the proportionality relation between sets are not definable within first-order logic [3]. A typical example is the quantifier *most*, where *Most A are B* is usually analyzed to mean *More than half of A are B*, symbolized as $|\mathbf{A} \cap \mathbf{B}| > |\mathbf{A}|/2$, or equivalently, $|\mathbf{A} \cap \mathbf{B}| > |\mathbf{A} - \mathbf{B}|$. As these paraphrases show, the quantifier *most* essentially denotes the binary relation between sets, which is not reducible to a standard unary quantification (i.e., a property of a set). Throughout this paper, we call a non-first-order quantifier like *most* a *proportional quantifier*, in contrast to a standard first-order quantifier like *all* and *no*.

Computation. Quantifier interpretations in terms of generalized quantifier theory have also been analyzed from computational perspectives. In a seminal work, van Benthem [4] uses automata to model semantic computing of quantified sentences. For example, in the standard quantified sentence of the form *All A are B*, the machine reads the states of objects. If the object is *B*, the transition to an accepting state occurs; otherwise, the transition to a rejecting state occurs. Thus, the machine need not memorize data at each process, and the machine's system can be realized by a simple finite automaton. By contrast, a proportional quantified sentence of the form *Most A are B* can only be modeled by push-down automata. In the case of proportional quantifiers, if the result of reading the states of objects is equal to the information in the top stack, the result is stored in a pushdown stack; otherwise, information in the stack is removed. Thus, a memory device (pushdown stack) is needed to give a computational modeling of proportional quantifiers. The resulting system is realized not by a finite automaton but by a push-down automaton. In this sense, the difference between proportional and non-proportional quantification also appears in the semantic automata approach to quantifiers.

Cognition. In cognitive psychology, it has been discussed whether the logical and computational difference between proportional and non-proportional quantifiers is reflected in the difference in the actual processing of quantifier expressions (see [23] and references given there). In these studies, it has been widely observed that proportional quantifiers take longer time to process and are interpreted less accurately than standard quantifiers. More specifically, Szymanik and Zajenkowski [23] argued that the computational identification of generalized quantifiers using semantic automata is relevant to cognitive verification processes of natural language quantifiers. In their experiment, participants were asked to judge whether quantified sentences were true of pictures containing 15 objects with different properties. Response times for the verification tasks were significantly longer in proportional quantified sentences (*more than half* and *less than half*) than in standard quantified sentences (*all* and *some*). This indicates that interpretation of proportional quantifiers requires more cognitive effort than interpretation of standard quantifiers, which is in accord with the computational model of generalized quantifiers.

In sum, proportional and non-proportional quantifiers show a different logical and cognitive behavior. The main question we are concerned with in this paper is whether such a difference between two kinds of quantification also appears in diagrammatic reasoning. To approach to this question, we will start, in the next section, with discussing how diagrams can represent sentences containing proportional quantifiers.

3 Diagrams of Proportional Quantifiers

3.1 A Problem in Diagrammatic Representation of *Most*

Euler diagrams represent sets of objects in terms of circles or closed curves and represent the inclusion and exclusion relations between sets by combining circles. Note here that circle sizes are irrelevant to the understanding of inclusion and

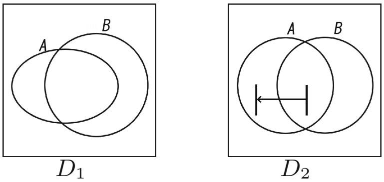


Fig. 1. A diagram for *Three-fourths of A are B* (D_1) and Rescher-Gallagher's Venn diagram with an arrow convention for *Most A are B* (D_2).

exclusion relations between sets i.e., the relations at the level (i) in our classification of set-theoretic relations discussed in Sect. 1. By contrast, the size differences play an important role in expressing the proportional relations between sets, i.e., the relations at the level (ii) in our classification.

As an example of a diagram in which the proportionality plays a role, consider the proportional quantifier *three-fourths* and its diagrammatic representation D_1 in Fig. 1. Here the area of the $A\overline{B}$ region is 1/4 of the total area of the A region and the area of the AB region is 3/4. Thus, this diagram corresponds to the sentence *Three-fourths of A are B*. We can find such a use of diagrams in the seminal work on conditionals by Adams [1], where the probabilities of conditionals are described by proportions of subregions in Euler-style diagrams. Also, in more recent developments of diagrammatic logic, the notion of "area-proportionality" is formalized using the weight values assigned to regions of Venn and Euler diagrams ([6]; see also [22]).

However, this line of extension of Euler diagrams is inadequate as a diagrammatic representation of the proportional quantifier *most*. In diagrammatic representations of *most*, one has to visualize the fact that the area of one region is greater than that of another region, without specifying the particular values of the areas of the relevant regions. This is an instance of the *over-specifity* property of diagrams discussed in details in [20,21]. What is necessary to capture the intended meaning of *most* is a natural device that indicates the proportionality relation between two regions and at the same time leaves underspecified the relation between the regions in question and the other regions in the diagram.

Rescher and Gallagher [15] overcame this problem by introducing to Venn diagrams the conventional device of the arrow to indicate that the extension of one region is less than that of another region. An example is shown as D_2 in Fig. 1. In this diagram, the AB region means the regions that can be extended, and the $A\overline{B}$ means the region that can be reduced. Thus even if we do not know the exact ratios of the regions' areas, we can extract the information that $\mathbf{A} \cap \mathbf{B}$ is greater than $\mathbf{A} - \mathbf{B}$.

By using the framework of Venn diagrams and introducing the conventional device of arrows, Rescher and Gallagher's diagrams succeed in avoiding the over-specificity problems inherent to proportional diagrams. However, giving up the idea of using the proportion of regions to indicate the proportionality of sets makes their diagrams less intuitive and hence more difficult to understand in

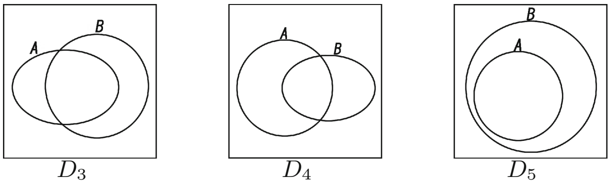


Fig. 2. Proportional diagrams for *Most A are B* (D_3), *Most A are not B* (D_4), and *All A are B* (D_5)

actual use. We can say that Venn diagrams with the arrow conventions are hybrid in that they combine a concrete circle-based form and an abstract convention in terms of arrows to represent the relational meaning of *most*.

For the diagrammatic representation of proportional quantifiers, we prefer to preserve the idea that the proportion of regions indicates the proportional relation between sets. In our view, diagrams that can be actually available to users are written on a paper or displayed on a PC monitor, with concrete forms. For this reason, in our experimental study, we do not adopt Venn diagrams with arrow conventions but instead use Euler diagrams whose regions have different areas, with the understanding that the sizes of the relevant regions whose proportionality is in focus are fixed and the other parts can be freely extended or reduced. We call the diagrams used in our experiments *proportional diagrams*.

In the actual scenes of the experiment, proportional diagrams were provided not as mere pictorial images, but rather as instances of general diagrams for *most* by instructing their syntax and semantics (see Appendix 2).[1] For example, *Most A are B* is represented by D_3 of Fig. 2, where (i) the proportion of AB region by $A\overline{B}$ region is specified as a *greater than* (>) relation with the tentative ratio of 2 : 1, and (ii) the proportion of BA region by $B\overline{A}$ region is unknown; the tentative ratio is set to 1 : 1, in order to restrain the invalid inference from *Most A are B* to *Most B are A*. D_4 of Fig. 2 corresponds to *More A are not B*, where the AB region and $A\overline{B}$ region are set to a 1 : 2 proportion and the BA region and $B\overline{A}$ region are set to a 1 : 1 proportion. D_5 of Fig. 2 corresponds to *All A are B*, where the area of the AB region is equal to that of the $\overline{A}B$ region and the invalid inference from *All A are B* to *Most B are A* is restrained.

Before moving on to discussing how proportional diagrams can be used in reasoning, two remarks are in order here. First, the system presented here works for pair-wise syllogistic inferences, where each sentence (premise and conclusion) has only two terms; however, without introducing a further convention, the system does not generalize to more complex cases. We made this simplifying assumption because we focused on the effectiveness of proportional diagrams in actual syllogistic inferences. To generalize the system of proportional diagrams, one would need a syntactic device to distinguish partially overlapping circles

[1] In addition, we can adopt a method using size-scalable diagrams in which object sizes can be changed from a default. Sato et al. [19] reported that the use of size-scalable diagrams in logical reasoning reduced the interfering effect of diagram layout.

expressing indeterminacy, as is common in Euler diagrams and Venn diagrams, from partially overlapping circles indicating the special proportionality relation expressing the meaning of *Most A are B*. One may use the Rescher-Gallagher's convention of using arrows or, for that matter, any syntactic convention to indicate the proportionality relation between two circles. Such a generalization of proportional diagrams and its empirical evaluation are left for future research.

Secondly, it is worth mentioning that Euler diagrams for proportional quantification used here can be naturally formalized within the framework of the relation-based approach to formalization of Euler diagrams [12,14]. According to this approach, a relation between sets such as inclusion and exclusion relations is taken as a primitive to define a diagram at an abstract level. Then a semantics and diagrammatic proof system can be provided in a similar way to the system of natural logic [13]. It is natural to add to this relation-based framework a relation corresponding to *Most A are B* (see [8] for a related study). We leave the detailed formal treatment of proportional diagrams for another occasion.

3.2 Reasoning with Proportional Diagrams

Reasoning with the proportional quantifier *most* has been studied in some depth by logicians having interest in natural language inferences (for early works, see [2]). Within syllogistic fragments, which are sometimes called as "plurative syllogisms", some researchers have proposed decision procedures with diagrams: Venn diagrams with arrow in [15], as stated above, and Lewis Carroll diagrams in [9]. More recently, the "natural logic" approach to natural language inference has combined natural language semantics based on the generalized quantifier theory, as seen in Sect. 2, and proof theory, to provide modern reconstructions of syllogistic reasoning. Endrullis and Moss [8] developed a proof system of syllogistic reasoning with *all*, *some*, and *most*.

Consider the following four arguments with *most*:

- *All BA, Most CB* (AU1); therefore, *Most CA* (U)
- *All AB, Most C not B* (AW2); therefore, *Most C not A* (W)
- *No BA, Most CB* (EU1); therefore, *Most C not A* (W)
- *No AB, Most CB* (EU2); therefore, *Most C not A* (W)

Here the label A stands for a sentence with *All*, E for a sentence with *No*, U for a sentence with *Most*, and W for a sentence with *Most-not*. The conclusion or hypothesis (beginning with *therefore*) is *entailed* by the premises in each argument. In other words, if all the premises are true, then the conclusion also is necessarily true. Therefore, each argument is a valid inference. Indeed, except for trivial cases (e.g., conclusion sentences converted from *All CA* to *Most CA*, and from *No CA* to *Most C not A*), valid syllogisms involving *most* and *most-not* comprise only the above four arguments (see [25]; for the sake of simplicity, *some* is not included here).

For invalid arguments, furthermore, non-entailment relations between premises and hypothesis can be generated in two ways (cf. [11], Chap. 5). First is a *contradiction*. Taking the AU1 syllogism above as an example, the hypothesis *Most C not A* contradicts the premises; if the premises are true, then the hypothesis cannot be true. Second is a *consistency* (*compatibility*). In the AU1 syllogism, the hypothesis *All CA* or *No CA* is consistent with the premises, in that if the premises are true, then the hypothesis may or may not be true. (See Appendix 1 for more details; note that not all valid/invalid syllogisms were used here.)

The extent to which ordinary people correctly make inferences with *most* in a typical sentential format is not well understood. If the interpretations of *most* are computationally and cognitively more complex than the interpretations of *all* (see Sect. 2), it is natural to speculate that inferences with *most* are also more complex than inferences with *all*. However, there is no empirical support for this speculation. It may be more accurate to say that there are no experimental studies which cover all of our tasks using *most*, *most-not*, *all*, and *no*. As a notable exception, Chater and Oaksford [5] employed the AU1U syllogism in their experiment, with a resulting accuracy rate of 85 %, and Geurts and van Der Slik [10] included inferences with *most* in more extended syllogisms with multiple quantifiers, for example, inferences from *Most A played against more than two B* and *All B were C* to *Most A played against more than two C*. Changes of *most* to *every* made no difference between them in participants' performances.

The solving process for reasoning tasks with proportional diagrams is essentially the same as that for standard syllogistic tasks with Euler diagrams [17]. Tasks for logical reasoning with diagrams typically consist of sentences (premises and a conclusion), and diagrams corresponding to the premise sentences, as shown in Figs. 3 and 4. The processes of unifying diagrams and extracting the information from sentences are illustrated by arrows.

Figure 3 shows the cases of an extended syllogism having the premises *All BA* and *Most CB*. In (1), (2) and (3), the first premise *All BA* is represented by D_1, and the second premise *Most CB* by D_2. There are two possible configurations of circles C and A in unifying the premise diagrams D_1 and D_2. In the first unified diagram D_3, the CA region is larger than the $C\overline{A}$ region. From this diagram we can extract the information $|\mathbf{C} \cap \mathbf{A}| > |\mathbf{C} - \mathbf{A}|$ (i.e., *Most CA*). In the second unified diagram D_4, circle C is totally included in circle A, thus we can extract the information $\mathbf{C} \subseteq \mathbf{A}$ (i.e., *All CA*).

Let us see how to solve the inferences in (1), (2) and (3) in turn. For (1), which is an instance of AU1U syllogism, one can start with extracting the information from the hypothesis *Most CA*. Then one can test whether this bottom-up information matches the top-down information extracted from the unified diagrams. Given that *All CA* implies *Most CA*, we can match the information from the hypothesis to both the information from D_3 and that from D_4. Thus, we can correctly judge that the hypothesis *Most CA* is entailed by the premises.

Regarding (2) (AU1W syllogism), the hypothesis has the form *Most C not A*. The information that can be read off from this hypothesis does not match the information from D_3 nor the information from D_4. Thus, we can judge that the hypothesis *Most C not A* contradicts the premises.

Regarding (3) (AU1A syllogism), the hypothesis has the form *All CA*. The information from the hypothesis matches information from D_4, but not from D_3 (*Most CA* of D_3 does not imply *All CA*)[2]. We can thus conclude that the given hypothesis *All CA* may be true or may be false. Thus, we can judge that the hypothesis *All CA* is consistent with the premises.

The above strategies that test whether the top-down information extracted from the unified diagrams matches the bottom-up information of the hypothesis sentence are common to standard syllogistic tasks using Euler diagrams. Figure 4 shows some examples. The syllogisms in (1) and (2) have the premises *All AB* and *No CB*. In this case, the unification of premise diagrams D_1 and D_2 produces the unique configuration D_3, corresponding to *No CA*. In the case of AE2E syllogism in (1), the hypothesis *No CA* matches the information *No CA* from the unified diagram D_3, leading to the correct answer (entailment). In the case of AE2A syllogism in (2), the hypothesis *All CA* does not match the information *No CA* from D_3. Accordingly, we can judge that the premises contradict the hypothesis. The syllogism in (3) (EA3E syllogism) has the premises *No BA* and *All BC* and the consistent conclusion *No CA*. There are four possibilities for the relationships between circles C and A: D_6, D_7, D_8, and D_9. The hypothesis matches D_6 but not D_7, D_8 or D_9. This kind of syllogism with no valid conclusion can actually be solved by enumerating multiple possibilities to unify the premise diagrams (see [18] for more details).

3.3 Predictions

Based on the analyses so far, we make two predictions.

(1) Reasoning with the proportional quantifier *most* (*most-not*) is more difficult and effortful than reasoning with standard quantifier *all* (*no*). This is true for participants who use only linguistic material as well as those who are also given proportional diagrams.

(2) The proportional diagrams improve the accuracy and speed of performances not only in reasoning with standard quantifiers but also in reasoning with proportional quantifiers, including all modes of entailment, contradiction, and consistency.

In this study, diagrams are added to sentential tasks of reasoning. This situation requires participants to handle both sentences and diagrams; i.e., they must do two jobs (cf. [16]). Nevertheless, if performance is faster with diagrams than without diagrams, this is evidence for the effectiveness of diagrams in reasoning. By contrast, even if performance is more accurate with than without diagrams, this does not necessarily count as evidence for the effectiveness of diagrams, because only the tasks with diagrams contain additional information on diagrams. Therefore, response-time data as well as accuracy data will be used to evaluate the effect of diagrams.

[2] Note here that the existence of $C\overline{A}$ region is a counter-example to the argument of the AU1A syllogism. See Takemura [24] for a formal specification of counter-example construction with Euler diagrams.

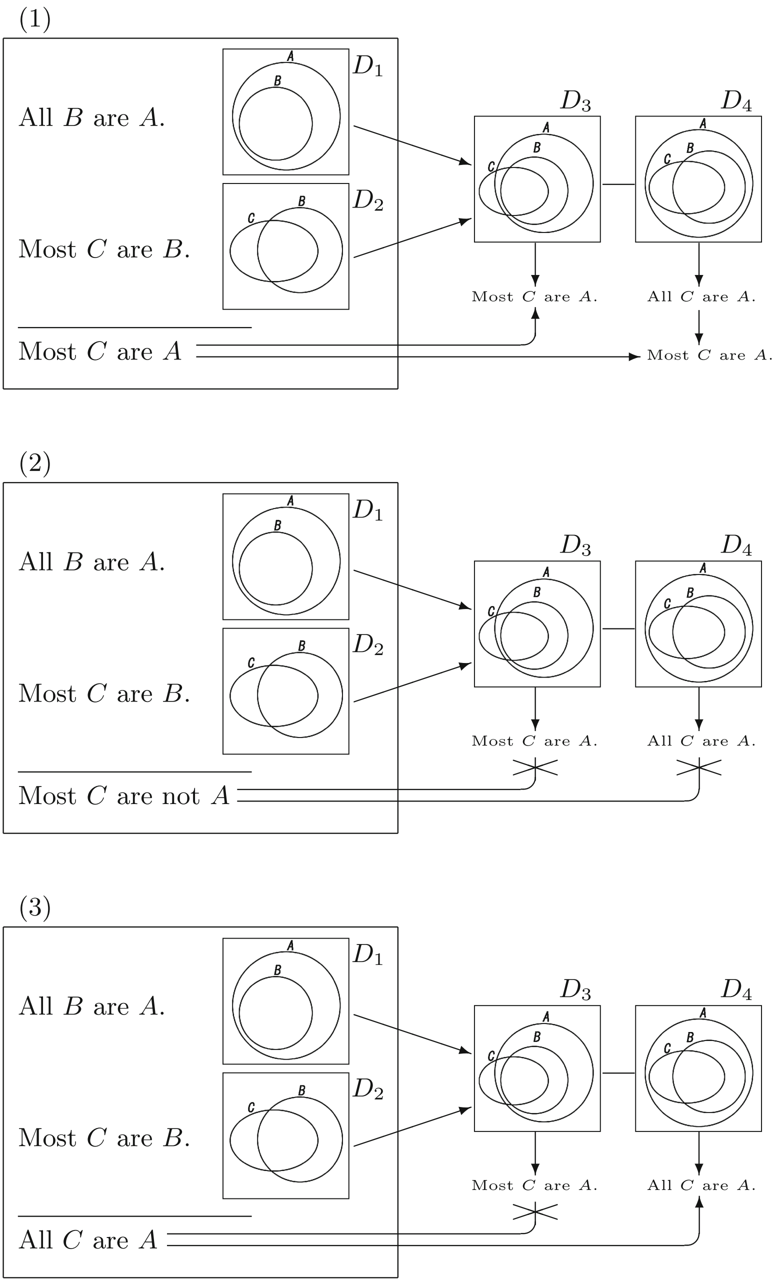


Fig. 3. Solving processes for reasoning tasks using proportional diagrams: (1) AU1U syllogism: entailment, (2) AU1W syllogism: contradiction, (3) AU1A syllogism: consistency

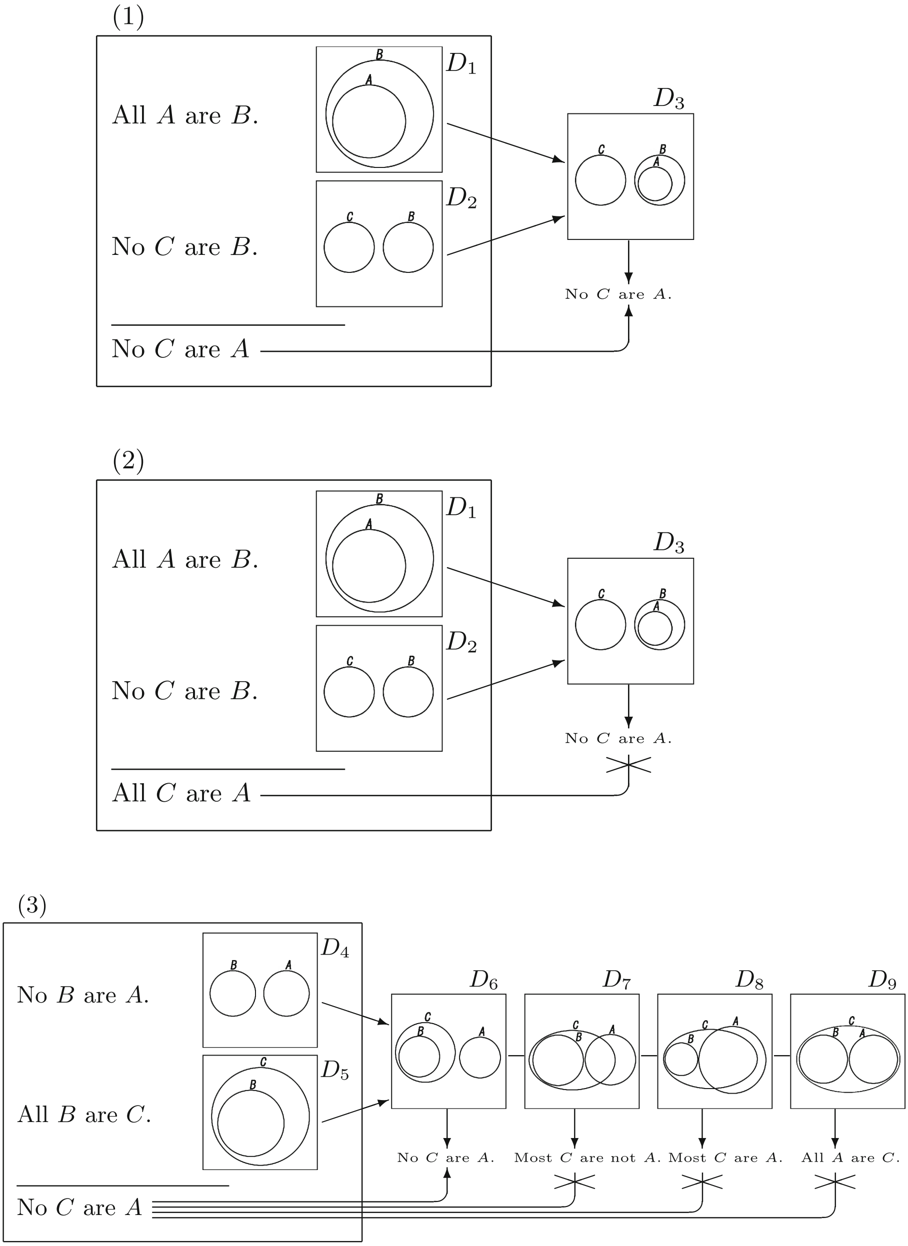


Fig. 4. Solving processes for reasoning tasks using Euler diagrams: (1) AE2E syllogism: entailment, (2) AE2A syllogism: contradiction, (3) EA3E syllogism: consistency

4 Experiment

4.1 Method

Participants. Forty-two undergraduate and graduate students from the University of Tokyo were recruited by means of an advertisement posted on campus. The mean age was 20.33 ($SD = 2.09$) with a range of 18–28 years. All participants gave informed consent and were paid for their participation. The Ethics Review Committee of the Graduate School of Arts and Sciences at the University of Tokyo approved all procedures in this experiment. The participants were Japanese-speaking students, and the sentences and instructions were provided in Japanese. None had any prior training in syllogistic logic. Participants were divided into two groups: a Diagrammatic group ($N = 21$), in which diagrams were used, and a Linguistic group ($N = 21$), in which diagrams were not used.

Materials. We presented 39 items: 17 standard syllogisms and 22 non-standard syllogisms (see Appendix 1 for the list of syllogisms). The sentences of standard syllogisms were universally quantified sentences either of the form *All A are B* or *No A are B*. The sentences of non-standard syllogisms were proportional quantified sentences of the forms *Most A are B* and *Most A are not B*. As shown in Fig. 5, the participants were presented with two premises and a hypothesis (conclusion) on a PC monitor and were asked to answer the question of *If the following two premises are true, is the hypothesis also true?*, by selecting a response from a list of three options: 1. *Hypothesis is true* (i.e., entailment). 2. *Hypothesis is false* (contradiction). 3. *Neither 1 nor 2: Hypothesis may or may not be true* (consistency). The premises in 9 syllogisms (4 non-standard) entail the hypotheses and the premises in 13 syllogisms (7 non-standard) contradict the hypotheses, and the premises in 17 syllogisms (11 non-standard) are consistent with the hypotheses. The quantified sentences included three properties, color (red or blue) for A terms, shape (square or round) for B terms, and striped pattern (horizontal or vertical) for C terms.

Procedures. The experiment was conducted individually. First, the Diagrammatic group only was given two pages of instructions on the meaning of Euler diagrams, but they did not receive any instructions about how to manipulate diagrams when solving syllogisms (for details, see Appendix 2). Second, both groups were given two pages of instructions regarding on three types of entailment relationships between premises and conclusion: "entailment", "contradiction" and "consistency" each with an example (for details, see Appendix 2). The participants were asked to press, as quickly and accurately as possible, a button with the number representing their answer. The 39 reasoning tasks were presented in random order. There was no time limit.

4.2 Results and Discussion

Accuracy rates (numbers of correct answers) for the non-standard syllogisms were significantly lower than those for the standard syllogisms. This tendency

If the following two premises are true,
is the hypothesis also true?

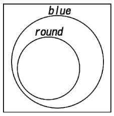
Premise 1 All round objects are blue.

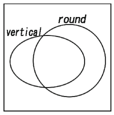
Premise 2 Most vertical objects are round.

Hypothesis Most vertical objects are blue.

1. Hypothesis is true.
2. Hypothesis is false.
3. Neither 1 nor 2: Hypothesis may or may not be true.

If the following two premises are true,
is the hypothesis also true?

Premise 1 All round objects are blue. 

Premise 2 Most vertical objects are round. 

Hypothesis Most vertical objects are blue.

1. Hypothesis is true.
2. Hypothesis is false.
3. Neither 1 nor 2: Hypothesis may or may not be true.

Fig. 5. Examples of reasoning tasks (AU1U syllogism) for the Linguistic group (left) and the Diagrammatic group (right)

was common to both the Linguistic group (73.4 % vs. 84.9 %, $z = 2.78$, $P = 0.0055$, Wilcoxon test) and the Diagrammatic group (81.6 % vs. 92.4 %, $z = 2.80$, $P = 0.0051$). Response times (for correctly answered items) for the non-standard syllogisms were also significantly longer than those for the standard syllogisms. This tendency was also common to both the Linguistic group (26.98 s vs. 21.05 s, $t(20) = 5.074$, $p < 0.01$) and the Diagrammatic group (20.13 s vs. 14.19 s, $t(20) = 6.084$, $p < 0.01$). Thus, syllogistic reasoning with proportional quantifiers *most* is relatively difficult and effortful, supporting our first prediction.

Table 1 shows accuracy rates and response times (for correct answers only) in the Linguistic and Diagrammatic groups. The results for each syllogistic type are shown in Appendix 1. For the non-standard syllogisms, accuracy rates (numbers of correct answers) were substantially higher in the Diagrammatic group than in the Linguistic group (81.6 % vs. 73.4 %), but there was no significant difference between them ($U = 146.5$, $P = 0.0610$, Mann-Whitney Test). Accuracy rates for the standard syllogisms were significantly higher in the Diagrammatic group than in the Linguistic group (92.4 % vs. 84.7 %, $U = 146.5$, $P = 0.0422$). More detailed analyses of entailment, contradiction, and consistency were conducted. For the non-standard syllogisms, accuracy rates were significantly higher in the Diagrammatic group than in the Linguistic group for entailment (90.5 % vs. 80.9 %, $U = 142.5$, $P = 0.0233$) and contradiction (88.4 % vs. 80.3 %, $U = 137$, $P = 0.0275$). There was no significant difference for consistency (74.0 % vs. 66.2 %, $U = 177$, $P = 0.2669$). For the standard syllogisms, there was no significant difference for each condition of entailment ($P = 0.2681$), contradiction ($P = 0.3425$) and consistency ($P = 0.0586$).

Response times of correctly answered items were logarithmically transformed and subjected to t-tests. Response times were significantly shorter in the Diagrammatic group than in the Linguistic group for standard syllogisms (14.19 s vs. 21.05 s, $t(40) = 3.506$, $p < 0.01$) and for non-standard syllogisms (20.13 s vs.

Table 1. Accuracy rates and response times (correct answer only) for standard syllogisms and non-standard syllogisms in the Linguistic group and Diagrammatic group

	Entailment	Contradiction	Consistency	Total
Standard syllogisms				
Linguistic group	92.4 % 20.38 s	92.1 % 18.37 s	71.4 % 24.79 s	84.9 % 21.05 s
Diagrammatic group	96.2 % 12.74 s	95.2 % 13.77 s	86.5 % 15.55 s	92.4 % 14.19 s
Non-standard syllogisms				
Linguistic group	80.9 % 22.16 s	80.3 % 23.48 s	66.2 % 31.90 s	73.4 % 26.98 s
Diagrammatic group	90.5 % 16.68 s	88.4 % 17.96 s	74.0 % 23.89 s	81.6 % 20.13 s

26.98 s, $t(40) = 2.526$, $p < 0.05$). In the following analyses, we excluded the participants' data where there was no correct answer in each condition of entailment, contradiction, and consistency. In the standard syllogisms, response times were significantly shorter in the Diagrammatic group than in the Linguistic group for entailment (12.74 s vs. 20.38 s, $t(39) = 5.566$, $p < 0.01$), contradiction (13.77 s vs. 18.37 s, $t(39) = 2.205$, $p < 0.05$), and consistency (15.55 s vs. 24.79 s, $t(39) = 3.887$, $p < 0.01$). In the non-standard syllogisms, response times were significantly shorter in the Diagrammatic group than in the Linguistic group for entailment (16.68 s vs. 22.16 s, $t(39) = 2.095$, $p < 0.05$), contradiction (17.96 s vs. 23.48 s, $t(39) = 2.393$, $p < 0.05$), and consistency (23.89 s vs. 31.90 s, $t(39) = 2.127$, $p < 0.05$). The results shown in Table 1 suggest that proportional diagrams tend to be effective in syllogistic reasoning with the proportional quantifier *most*, consistent with our second prediction.

5 Conclusion and Future Work

As seen in Sect. 2, previous work has revealed that *most* sentences are computationally and cognitively complex in interpretation or verification. Little attention, however, has been paid to *most* sentence inferences; further, not much is yet known about the kind of diagrams that work well in such inferences. In this study, we showed that syllogistic reasoning with proportional quantifiers *most* is cognitively complex yet can effectively be supported by Euler-style diagrams that represent the proportionality relationships between sets in terms of area-proportionality. In particular, our result indicates that difference between *all* and *most* in the complexity of comprehension also reflects the complexity of reasoning tasks in both linguistic and diagrammatic formats.

Future research should explore the interaction of proportional diagrams and diagrams for existential quantifiers. For instance, it is well-discussed that *Most A are B* and *Most A are C* entail *Some B are C* [9]. It is clear that diagrams expressing the proportionality relationship in an intuitive way can help deriving such an inference from *most* to *some*. However, combining proportional diagrams with diagrams asserting the non-emptiness of a set is not a trivial task; in addition to the generalization of proportional diagrams mentioned in Sect. 3.1, we leave for future research how to set up a more expressive representation system for proportional diagrams.

In the study of visualization and graphics, it has been discussed that perceptual judgements of the relative sizes of areas is relatively difficult and effortful [7]. A detailed investigation on perception of proportional diagrams is also left for future work.

As we saw in Sect. 2, by merging the theoretical and empirical findings on various kinds of quantifiers, we can explore the relationship between logical, computational and cognitive aspects of human reasoning. Imposing a constraint on the possible ways in which we can reason by means of diagrams can contribute to this direction of research and serve as a fruitful way to capture the complexity of reasoning tasks.

Appendix 1: The Results of Each Task

Table 2. Accuracy rates and response times (correct answer only) for 39 syllogisms in the Linguistic group (left) and Diagrammatic group (right)

Premises	Entailment	Contradiction	Consistency
All BA, All CB (AA1)	All CA (A) 100 % 100 % 15.59 s 14.68 s	No CA (E) 100 % 100 % 14.75s 13.18 s	-
All AB, All CB (AA2)	-	-	No CA (E) 52.4 % 90.5 % 25.93 s 17.87 s
All BA, All BC (AA3)	-	No CA (E) 85.7 % 95.2 % 17.34 s 17.95 s	All CA (A) 80.9 % 90.5 % 16.88 s 30.10s
All BA, No CB (AE1)	-	-	No CA (E) 71.4 % 90.5 % 25.23 s 32.55 s
All AB, No CB (AE2)	No CA (E) 85.7 % 95.2 % 21.93 s 22.55 s	All CA (A) 85.7 % 95.2 % 23.59 s 21.73 s	-
All BA, No BC (AE3)	-	-	no CA (E) 95.2 % 85.7 % 23.95 s 12.28 s
All AB, No BC (AE4)	No CA (E) 81.0 % 90.5 % 29.46 s 27.06 s	All CA (A) 90.5 % 90.5 % 25.11 s 15.94 s	-
No BA, All CB (EA1)	No CA (E) 100 % 100 % 16.48 s 13.88 s	All CA (A) 100 % 100 % 15.53 s 10.72 s	-
No AB, All CB (EA2)	No CA (E) 95.2 % 95.2 % 21.29 s 9.77 s	All CA (A) 90.5 % 90.5 % 19.50 s 15.92 s	-
No BA, All BC (EA3)	-	-	No CA (E) 66.7 % 80.9 % 27.83 s 25.22 s
No AB, All BC (EA4)	-	-	No CA (E) 61.9 % 80.9 % 30.28 s 19.67 s
All BA, Most CB (AU1)	Most CA (U) 90.5 % 95.2 % 17.85 s 22.79 s	Most C not A (W) 90.5 % 95.2 % 16.48 s 15.89 s	All CA (A) 57.1 % 76.2 % 22.23 s 19.52s
All AB, Most CB (AU2)	-	All CA (A) 61.9 % 66.7 % 24.78 s 31.06 s	Most CA (U) 85.7 % 90.5 % 26.56 s 24.49 s
All BA, Most BC (AU3)	-	-	Most C not A (W) 61.9 % 66.7 % 23.16 s 35.39 s
All BA, Most C not B (AW1)	-	No CA (E) 61.9 % 90.5 % 26.29 s 38.56 s	Most C not A (W) 71.4 % 76.2 % 24.87 s 22.92 s
All AB, Most C not B (AW2)	Most C not A (W) 61.9 % 80.9 % 24.11 s 19.12 s	Most CA (U) 90.5 % 85.7 % 25.70 s 12.87 s	No CA (E) 52.4 % 80.9 % 27.26 s 33.68 s
All BA, Most B not C (AW3)	-	-	Most CA (U) 76.2 % 76.2 % 36.51 s 20.29 s
All AB, Most B not C (AW4)	-	-	Most C not A (W) 71.4 % 71.4 % 50.49 s 23.45 s
No BA, Most CB (EU1)	Most C not A (W) 100 % 90.5 % 25.72 s 17.67 s	Most CA (U) 95.2 % 90.5 % 19.69 s 20.26 s	No CA (E) 52.4 % 71.4 % 19.42 s 13.37 s
No AB, Most CB (EU2)	Most C not A (W) 71.4 % 95.2 % 21.46 s 39.25 s	Most CA (U) 90.5 % 95.2 % 20.89 s 16.95 s	No CA (E) 57.1 % 71.4 % 35.34 s 30.15 s
No BA, Most BC (EU3)	-	-	Most C not A (W) 81.0 % 61.9 % 27.60 s 15.60 s
No AB, Most BC (EU4)	-	All CA (A) 71.4 % 95.2 % 36.81 s 18.59 s	Most CA (U) 61.9 % 71.4 % 31.32 s 35.92 s

Appendix 2: Instructions Used in Experiment

See: http://abelard.flet.keio.ac.jp/person/sato/index/appendix_d16.pdf

References

1. Adams, E.W.: The Logic of Conditionals: An Application of Probability to Deductive Logic. Springer, Dordrecht (1975)
2. Altham, J.E.J.: The Logic of Plurality. Methuen, London (1971)
3. Barwise, J., Cooper, R.: Generalized quantifiers and natural language. Linguist. Philos. **4**, 159–219 (1981)
4. van Benthem, J.: Essays in Logical Semantics. Reidel, Dordrecht (1986)
5. Chater, N., Oaksford, M.: The probability heuristics model of syllogistic reasoning. Cogn. Psychol. **38**, 191–258 (1999)
6. Chow, S., Ruskey, F.: Drawing area-proportional Venn and Euler diagrams. In: Liotta, G. (ed.) GD 2003. LNCS, vol. 2912, pp. 466–477. Springer, Heidelberg (2004)
7. Cleveland, W.S., McGill, R.: Graphical perception: theory, experimentation, and application to the development of graphical methods. J. Am. Stat. Assoc. **79**, 531–554 (1984)
8. Endrullis, J., Moss, L.S.: Syllogistic logic with "most". In: de Paiva, V., de Queiroz, R., Moss, L.S., Leivant, D., de Oliveira, A. (eds.) WoLLIC 2015. LNCS, vol. 9160, pp. 124–139. Springer, Heidelberg (2015)
9. Geach, P.T.: Reason and Argument. University of California Press, Berkeley (1976)
10. Geurts, B., van Der Slik, F.: Monotonicity and processing load. J. Seman. **22**, 97–117 (2005)
11. MacCartney, B.: Natural Language Inference. Ph.D. thesis, Stanford University (2009)
12. Mineshima, K., Okada, M., Takemura, R.: A diagrammatic reasoning system with Euler circles. J. Logic Lang. Inf. **21**, 365–391 (2012)
13. Mineshima, K., Okada, M., Takemura, R.: A generalized syllogistic inference system based on inclusion and exclusion relations. Stud. Logica **100**, 753–785 (2012)
14. Mineshima, K., Sato, Y., Takemura, R., Okada, M.: Towards explaining the cognitive efficacy of Euler diagrams in syllogistic reasoning: a relational perspective. J. Vis. Lang. Comput. **25**, 156–169 (2014)
15. Rescher, N., Gallagher, N.A.: Venn diagrams for plurative syllogisms. Philos. Stud. **16**, 49–55 (1965)
16. Sato, Y., Masuda, S., Someya, Y., Tsujii, T., Watanabe, S.: An fMRI analysis of the efficacy of Euler diagrams in logical reasoning. In: VL/HCC 2015, pp. 143–151. IEEE Press (2015)
17. Sato, Y., Mineshima, K.: How diagrams can support syllogistic reasoning: an experimental study. J. Logic Lang. Inf. **24**, 409–455 (2015)
18. Sato, Y., Wajima, Y., Ueda, K.: An empirical study of diagrammatic inference process by recording the moving operation of diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 190–197. Springer, Heidelberg (2014)
19. Sato, Y., Wajima, Y., Ueda, K.: Visual bias of diagram in logical reasoning. In: CogSci 2014, pp. 2342–2347. Cognitive Science Society, Austin (2014b)

20. Shimojima, A.: On the Efficacy of Representation. Ph.D. thesis, Indiana University (1996)
21. Shimojima, A.: Semantic Properties of Diagrams and Their Cognitive Potentials. CSLI Publications, Stanford (2015)
22. Stapleton, G., Rodgers, P., Howse, J.: A general method for drawing area-proportional Euler diagrams. J. Vis. Lang. Comput. **22**, 426–442 (2011)
23. Szymanik, J., Zajenkowski, M.: Comprehension of simple quantifiers: empirical evaluation of a computational model. Cogn. Sci. **34**, 521–532 (2010)
24. Takemura, R.: Counter-example construction with Euler diagrams. Stud. Logica **103**, 669–696 (2015)
25. Thompson, B.: Syllogisms using "few", "many", and "most". Notre Dame J. Form. Logic **23**, 75–84 (1982)

Extensions of Euler Diagrams in Peirce's Four Manuscripts on Logical Graphs

Ahti-Veikko Pietarinen(✉)

Tallinn University of Technology, Tallinn, Estonia
ahti-veikko.pietarinen@ttu.ee

Abstract. Charles Peirce's important manuscript on Euler diagrams (Ms 479, 1903) was partially printed in the *Collected Papers* in 1933 (CP 4.347-371). That transcription omitted many paragraphs, figures and important variants of the main text, and diagrams were reproduced misleadingly or imprecisely. Another important and wholly unpublished paper of his (Ms 481, 1896-7) presents a novel extension of Euler's diagrams for negative terms. Third, among the discarded pages of a published article (Ms 1147, 1901) we find a variant on logical graphs suggesting similar extensions. Ms 855 (1911) is yet another unpublished note in which Peirce deals with existentials and shading. The present paper restores Peirce's original drawings from these four manuscripts and explains their main innovations. As Euler diagrams were not designed to reason about relative terms, Peirce's interest was not in their mathematical application or problem solving but in showing what the basic elements of syllogistic reasoning are.

Keywords: Euler diagrams · Venn diagrams · History of diagrams · Peirce · Logical graphs · Syllogism

1 Ms 479 ("On Logical Graphs", 1903)

Peirce's 1903 manuscript on logical graphs is known for its proposal of extending the Euler-Venn diagrams to the representation of existential statements [1, 2]. It is also known for its harsh criticism of the fundamentals of the Euler-Venn construction, especially its lack of relative terms and the diminished expressive power. His paper was meant as the second part of a massive book-length treatise, entitled *Logical Tracts No. 2* (Ms 492). The *Tracts*, of which over 400 manuscript pages has been preserved at the Houghton Library [6], was Peirce's companion when preparing his famous 1903 *Lowell Lectures* on Existential Graphs (EG).

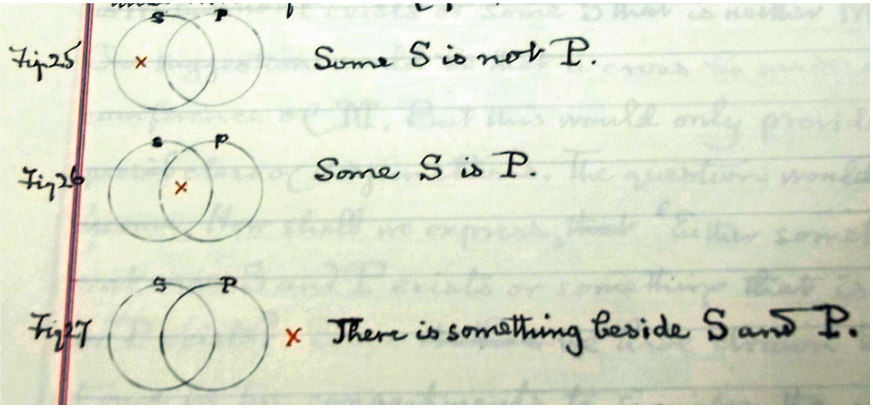
Ms 479 proposes a host of improvements over the traditional Euler-Venn diagrams. Among them is Peirce's suggestion how to represent the negation of the copula of inclusion, where the copula of inclusion is a spatial relation of "enclosing only what is enclosed by" ("All Ss are Ms"), by introducing a *dot* • (or alternatively a *cross* ×) to "represent some existing individual" (CP 4.349). Further, if the dot or a cross rests on a closed curve that isolates a boundary of the zone (compartment) of an Euler diagram, Peirce took it to mean that "it is doubtful on which side it belongs" (ibid.). A mark on the boundary thus represents logical disjunction. If two or more marks lie on the

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 139–154, 2016.
DOI: 10.1007/978-3-319-42333-3_11

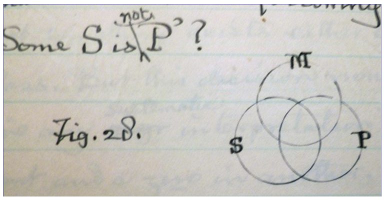
boundary of the same compartment there is, Peirce notes, “nothing that prevents their being identical” (ibid.). The same holds for marks on the same zone. Thus more than one mark on the same zone does not mean that more than one individual exists.

Peirce lists five shortcomings of the basic system of Euler diagrams.[1] The second is that the system is limited to express only either that something does not exist or else leaves one entirely ignorant of whether something exists or not. In Peirce’s terms, the system “*cannot affirm the existence* of any description of an object” (CP 4.356). Peirce’s remedy, similar to what Venn incidentally had proposed in his 1883 review of *Studies in Logic* (SiL; [15]), is to draw ×, in red, in compartments to signify that “something of the corresponding description occurs in the universe” (CP 4.359). (Ms 855, see Sect. 4, suggests that Peirce never saw this review.)

The relevant examples are the following three forms of propositions (Figs. 25–27):[2]



The denial of such existence is expressed by replacing × by circles O.



Further conventions are needed to interpret cases with three or more compartments, as well as when × lies in one compartment and O in another. For the first case (e.g. as in Fig. 28 on the right) Peirce notes that there may be an ambiguity between stating the existence of

[1] These five shortcomings of the theory of Euler diagrams are (i) its inadequacy of dealing with every syllogistic form; (ii) that it “cannot affirm the existence of any description of an object”; (iii) that it fails to represent disjunctions in the general case; (iv) that it “affords no means of expressing any other than dichotomous” information; and (v) that these diagrams fail to exhibit “relational or abstractional” kind of reasoning, meaning that the system “has no vital power of growth beyond the point to which it has been carried” (Ms 491; Ms 479).

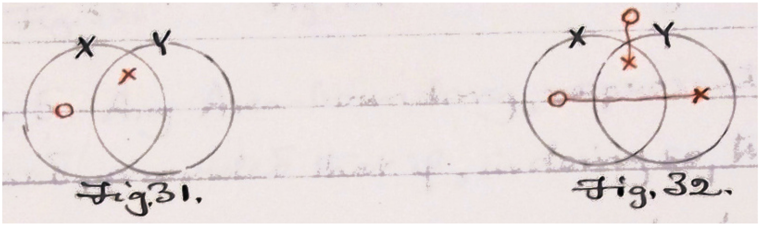
[2] For reasons of space, we omit reproducing Figs. 1–24, which mostly are examples of ordinary Euler diagrams and Venn’s modifications (see CP 4.350-8). Figures 25–36, 54–66 in Sect. 1 and in Variants 1 and 2 are photographic images from Peirce’s original manuscripts deposited at the Harvard Houghton Library, slightly enhanced digitally for readability.

some S that is M but not P, and the existence of some S that is neither M nor P. He suggests that disconnected signs in different compartments are to be read conjunctively, and that connected signs in different compartments are to be read disjunctively. The denial of existence is thus not only a substitution of O for ×: one also needs to correctly reverse the connections and disconnections between the crosses.

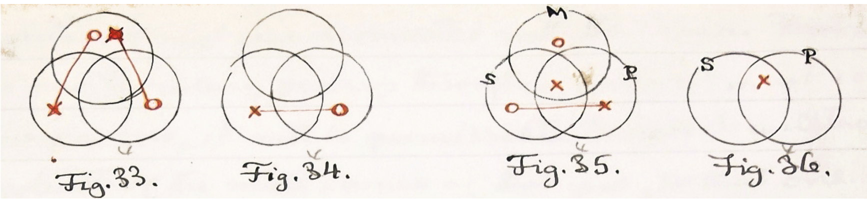
Two opposite signs that rest on the same compartment mean that the proposition is absurd (Fig. 29) – unless they coincide, in which case they annul each other (Fig. 30).

Peirce also takes crosses that lie on the boundary of compartments to be equivalent to cases in which they occur on both sides of that circumference. Thus *logical disjunctions* can be denoted. A similar method, using connected lines between the dots, was subsequently and independently reinvented in the form of *spider diagrams* [7].

Peirce then presents the rule ("Rule 2. Any sign of assertion can receive any accretion", CP 4.362) by which the assertion in Fig. 31 ("All X are Y and some X are Y") can be transformed into the assertion in Fig. 32 ("All X are Y or some non-X are Y and some X are Y or all non-X are Y"):

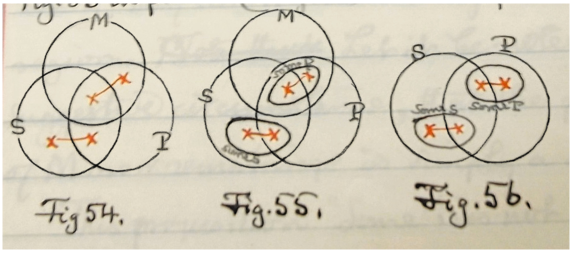


Moreover, his Rule 4, which concerns the attachment and detachment of crosses and zeros, entitles one to infer Fig. 34 from Fig. 33, and Fig. 36 from Fig. 35:

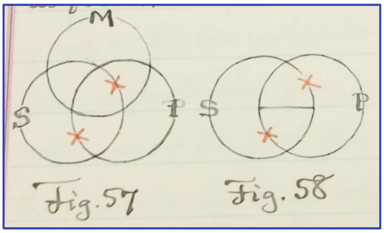


Six in total, these rules to transform Euler diagrams grow in complexity. Peirce proposes simplifications to complex rules and presents examples of syllogisms that work according to these simplified rules (Figs. 39–56 in Ms 479, here omitted, see [2]).

The typography of these diagrams must be exactly right. Figure 56, for instance, was sloppily reproduced in CP 4.363 (there Fig. 54) and may have left the false impression that the two tokens of × in the intersection are connected, which they are not:[3]



Next, Peirce discusses how one is to express *spurious propositions*, such as that in Fig. 56, which states that "Some S is not some P". Instead of these connected crosses, he suggests placing × on the boundaries. Thus the graphs in Figs. 57 and 58 express the same proposition. Here an important discovery emerges. In Fig. 58 the horizontal line in the intersection tells us that the two × must not become connected, that is, they must not appear in the same region. That line segment, which is what remains from the lower boundary of the circle M in Fig. 57, serves the role analogous to the *sign of negation*: it means that the two tokens of × separated by it are not identical. The proposition in Fig. 58 thus expresses that there are at least two non-identical individuals, that is, at least two individuals exist.

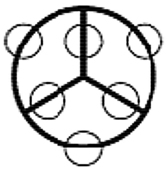


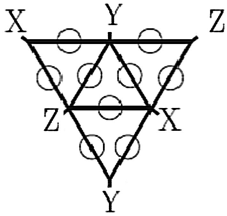
This way of denoting negation by a linear separation device is notably similar to what became the sign of negation in his 1896 EGs [6, 8, 14]: a linear separation depicted as one-dimensional encircling of the graph severed from the sheet. To assert that at least two individuals exist, EGs use these encirclements to separate two ends (corresponding to two assertions of existence) of a line of identity, thus making that line discontinuous on the two-dimensional sheet.

To express that there are at least n individuals existential graphs we proceed as follows. The EG below left is Peirce's graph that there are at least four individuals in

[3] Other unfortunate typographical inaccuracies in the CP were using the letter *x* for crosses and the number 0 for circles, as well as using capital letters as labels placed outside the circles. Peirce wrote the letters *on* the circles. He might have adopted this convention from Johann Christian Lange's 1712 *Nucleus Logicae Weisianae*, the work which Peirce, just as Venn, thought "anybody familiar with such literature the title proclaims it to be a work by [Christian] Weise probably with a running commentary or copious notes by Lange" (Ms 479: 16). Weise's *Nucleus Logicae* was indeed originally published, Peirce remarks, as a small booklet in 1691 and edited and expanded into an over 800-page volume published 1712 by Weise's student Lange.

the universe; the EG on the right that there are at least six (the variables X, Y, Z denote, according to Peirce, "(presumably) unknown individuals", Ms 484, 1898):[4]

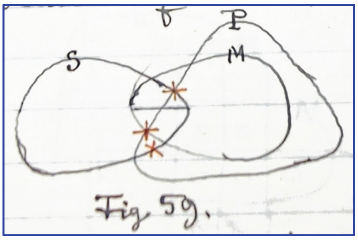


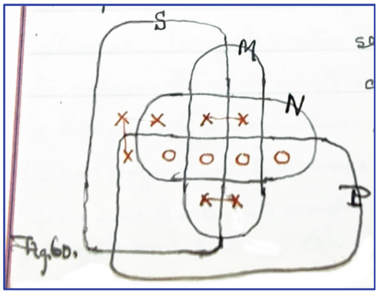


Immediately preceding his invention of the method of EGs, the idea of the line denoting *both* the existence and the denial was in operation in Peirce's 1896 "positive system of graphs" (MS 482, alt. variant). Here are two examples, in his hand:

:"All wise men are virtuous"
"Everything is either wise or virtuous":

Peirce is thus using the linear separator from his 1896 EG to extend the theory of Euler-Venn diagrams in 1903, not vice versa.





To return to the Euler-Venn diagrams, in order to express particular propositions of third degree, Peirce states that the graph in Fig. 59 will not do, since those two lower crosses × adjacent to the same compartment may refer to the same individual. Unlike for the two topmost crosses, no line separates the two × that border that same lower compartment.[5] Peirce's solution is to draw propositions asserting the existence of at least three individuals in the manner depicted in Fig. 60, with no cross remaining unconnected.[6]

The question naturally arises whether a systematic method can be found that generalizes Euler diagrams to n-degree propositions. In countable universes, numerical statements can be represented by increasing the number of disjuncts. Peirce addresses

4 Generically, count the number of connected systems of lines separated by such circles.

5 This figure, numbered Fig. 57 in CP 4.364 was published there with a different but equivalent orientation.

6 A reviewer suggests a general result may be that representing n individuals requires $n + 1$ contours. Peirce's simplification that follows here suggests that n contours suffice.

this question in connection to what he takes to be the third imperfection of the system of Euler circles: its inadequacy to deal with disjunctions in the general case. For example, when it comes to the assertions of disjunctions of conjunctions (disjunctive normal forms), diagrams soon become cluttered. An example is the graph in Fig. 61, which expresses that "Either some A is B and everything is either A or B, or else all A is B and some B is not A".

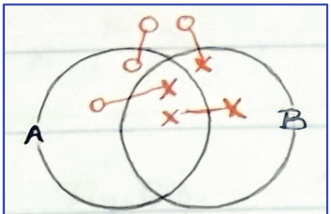


Fig. 61.

Peirce's suggested simplification is to encircle every diagram within yet another circle, and then, for perspicuity, to compartmentalize those 'mother circles' into rectangles (Fig. 62). These mother circles represent the *universes of discourse* (in his terms "Universe of Hypothesis"). One now dispenses with drawing complex connecting lines as in Fig. 61: now every compartment in Fig. 62 represents *a possible case of the diagram being true*.

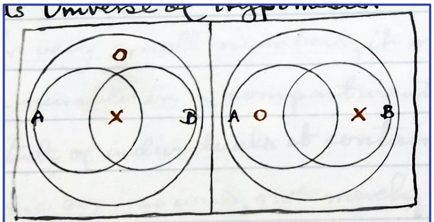


Fig. 62.

In this manner, an Euler diagram expressing a complex proposition becomes a *sequence* of Euler diagrams, each rectangle or a *frame* depicting one of the disjuncts of that proposition. For instance, the frame on the left of Fig. 62 expresses that "Some A is B and that there is nothing else in the universe of discourse in that occasion that is either A or B", while what is included within the frame on the right expresses that "Every A is B and that there are Bs that are not As", just as in Fig. 61.

The box notation, though emerging along different historical paths and without the idea of further framing the mother circles to form sequences of diagrams, became the standard usage by which the universe of discourse is denoted in theories of Euler and Venn diagrams. In speaking of different universes Peirce adds a distinctive modal flavour – not unprecedented though as it is in these very same months of 1903 that he was developing his systems of modal logic in the gamma part of EGs.

At any event, the consequence of this remedy to simplify logical disjunctions is that, as in Fig. 63, diagrams encircled by a boundary that denotes the universe of discourse (here with a label A) can now be used to express propositions of the fourth degree ("There are at least four individuals that are As"). Thus its dual, the denial of the proposition that there are at least four individuals that are As, is expressed in Fig. 64 ("There are not as many as four As").

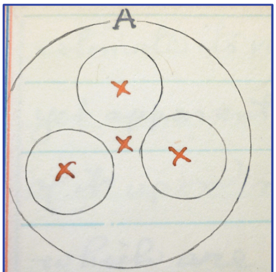


Fig. 63.

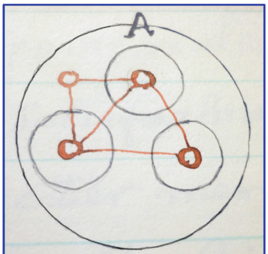
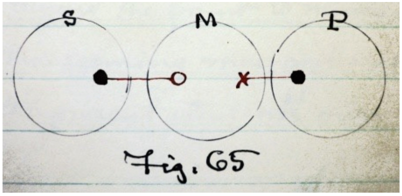
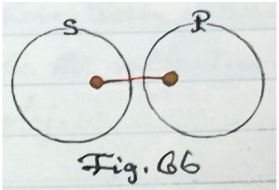


Fig. 64.

It is this method of encircling entire Euler diagrams that Peirce observes to give rise to a generalization that allows one to express *arbitrary numerical propositions*. He did not develop these possibilities any further, however. Nor did he proceed to study the sequences of frames of Euler diagrams. We may surmise that this idea is connected to, or is an anticipation of, his "moving pictures of thought" [14, 15] that characterise EGs, especially the application of rules of transformation as a continuous animation of deductive inferences.

Finally, Peirce introduces a *blot* serving as the 'wild card': it denotes either of the two marks so that any token means "either a cross or a circle". Thus the inference from Figs. 65 to 66 is not an inference rule but a *generalized inference schema*:





An instance of that schema is for instance the valid syllogism of Barbara.

Peirce ends the main sequence of Ms 479 with a note that there is the fifth, and a "fatal defect" of the system of Euler diagram, namely that "it has no vital power of growth beyond the point to which it has been here carried" (CP 4.370). The ceiling is set by the plane geometry limiting its expressive power to first-order logic with monadic predicates. To extend the system into an analysis of relative propositions, multitudes or abstractions, that is reasoning along the lines of general algebra of logic or EGs seemed to him beyond possibility. Constraint diagrams are an example of a modest increase in expressivity up to dyadic relations [9].

Variant 1. Euler Diagrams with Six Conventions. In an alternative, unpublished sequence of Ms 479 Peirce describes another extension of Euler diagrams with six conventions to set up that system. It begins with the convention that describes the *universe of discourse*: the subject mutually well-understood between the "drawer" and the "interpreter" of the diagrams. The sheet on which these diagrams are drawn consists of "different possible points" of a "certain individual subject".

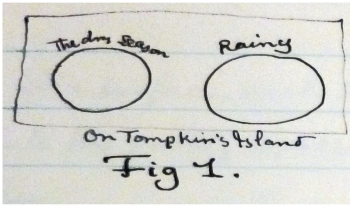
Second, ovals drawn on the sheet are connected with an assertion about that universe of discourse, and are denoted by a capital letter. Any point outside of the oval represents a *possible state of things* in which the assertion would be true of that universe. Any point inside the oval represents a possible state of things in which the corresponding assertion would be false. This suggests an interesting early semantic approach to the modal features exhibited in Euler diagrams.

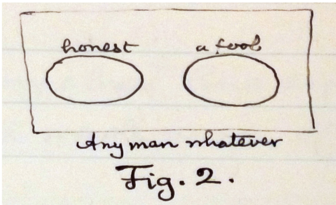
Third, these compartments may be *shaded*, and every point on a shaded area means *non-existent individual*.[7] Fourth, any point coloured in *red ink* means that the

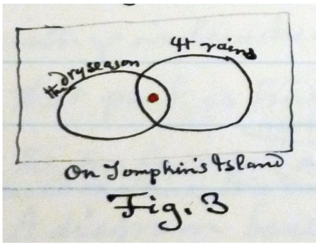
[7] This is a common convention that was reintroduced, in a different sense, in [10], in order to strike out all constituents of a certain region and not to demolish a particular individual.

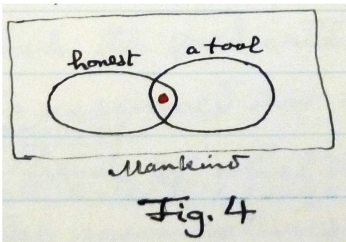
corresponding *state of things is* "*realized* at some time or in some reference". Fifth, a *point coloured red and lying on the boundary* of an oval means that "it is not determinate" whether the realized state of things is represented by a point inside of the boundary or outside of that boundary. That is, we can represent *disjunctions* similarly as he did with his other extensions of Euler diagrams. The twist is this model-theoretic gloss on the realized and unrealized states of things. Sixth, a *box*, which Peirce instructs is to be drawn in green ink, marks the *limits of the sheet*.

Peirce's only examples in these few draft pages are the following. In Fig. 1 the universe of discourse of the diagram is the state of weather in Tompkins Island. The proposition on the left is "It is the dry season" and the proposition on the right is "It is raining". The diagram thus asserts that "It is always either the dry season or is raining on Tompkins Island (or both)". Figure 2 below means that every man is either honest or fool. Figure 3 means that on Tompkins Island it sometimes does not rain though it is not in the dry season. Figure 4 means that some man is neither a fool nor honest.



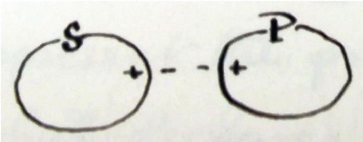


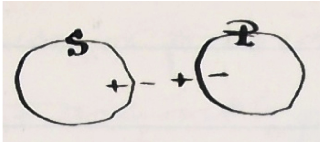




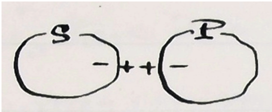
Variant 2. Euler Diagrams on a Sphere. In another and likewise previously unpublished variant of Ms 479 Peirce presents another, surprising generalization: Euler diagrams drawn on a *sphere*. He tells that "there is no particular appropriateness in drawing the diagrams on a plane surface rather than on a sphere", since "the collection of all beings is as a definite a whole as the collection of all mortal beings; and thus the sphere is rather the more appropriate". On the universe of a sphere, the difference between a dot inside or outside of the oval vanishes. What is the inside (the positive area) and the outside (the negative area) of the ovals has therefore to be fixed at the outset. Instead of two possible positions for the ovals (one inside of the other or else disjoint) there are *three* possible relative positions for two non-intersecting ovals: one in which the uncovered surface of the sphere lies in the middle of the two ovals, one in which the surface lies outside of an outer oval, and one in which two parts of the surface lie outside of the two ovals each.

As an example of these cases, Peirce drew the following three diagrams, in which he denotes by ± the inside/outside of a circle on a sphere:





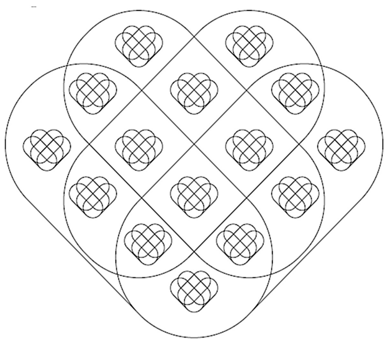
"There are no saints that are perfect" "There are no saints that are not perfect"



"There is nothing that is neither saint nor perfect", that is, "Everything is either saint or perfect (or both)"

Peirce prefigures, albeit in a somewhat different sense, the later works e.g. by [11] (see p. 15 on H. C. Smith) that extended Venn-like diagrams on spheres. Another recent extension studies Venn–Euler diagrams as three-dimensional closed surfaces or solids [12]. The two are not equivalent approaches, since $n - 1$-spheres and n-balls are very different manifolds with different topological properties.

In this second variant, too, Peirce mentions the second defect of Euler diagrams, namely that the ordinary system only represents propositions that express the non-existence of individuals. Euler diagrams can express universal propositions only via expressing non-existence of exceptions. For example "All saints are perfect" means that "There are no saints that are not perfect". But Euler circles do not capture propositions expressing "There is a saint that is not perfect". Peirce's alternative sequence does not proceed to explore remedies to this second defect any further, however.



This variant of Ms 479 ends with Peirce recapitulating the history of the invention of Euler diagrams, including Lambert's linear diagrams [18]. He notes that "no real improvement upon the system was made until Mr. Venn, in 1880, removed the first of the above mentioned defects by simply shading those compartments of the figure that correspond to combinations to be represented as nonexistent". Venn also suggested, Peirce continues to note, to draw ellipses instead of circles and to draw them in fours, in threes, and in pairs. For the number of terms exceeding four, Peirce proposes the figure above in which the diagrams for fours are iterated within one big diagram of fours, amounting to a Venn diagram for eight terms drawn on a two-dimensional surface, thus dispensing with introducing additional dimensions or other gimmicks (Ms 479: 13, alt. variant).[8]

2 Ms 481 ("On Logical Graphs", 1896–7)

One more hitherto unacknowledged gem from the Peirce archives is a 10-page paper he wrote in late 1896 or early 1897. In this paper Peirce studies 22 examples of Euler diagrams that have a quite non-standard planar form. The rest of Ms 481, which we do not explore here, contains examples of his other logical graphs, especially the system of existential graphs that co-evolved in this very same paper.

The improvement on Euler diagrams that Peirce suggests is to draw them as zones the boundaries of which may have *convex-concave shapes*. The meaning of that shape is that the terms that face the concave (non-convex) side of such zones are *negative terms*. For instance, while in Fig. 1 the meaning of the diagram is that "Any man there may be is mortal", in Fig. 2 the asserted proposition is "No man is a quadruped man". The meaning of the diagram in Fig. 3 is "Everything is either just or wicked":[9]

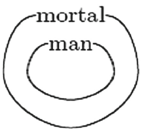


Fig. 1

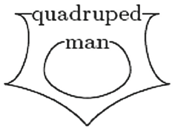


Fig. 2

Fig. 3

There is no assertion of the existence of things in these diagrams; only the assertion of the non-existence of exceptions to the relevant classes. That is, in Fig. 1 there are no non-mortal men, in Fig. 2 there are no quadruped men, and in Fig. 3 no non-wicked injustice.

In order to assert existence Peirce adds dots to the diagrams. Figure 4 asserts that "there is a soul", and Fig. 5 that "there is something besides money":

Fig. 4

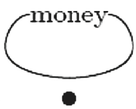


Fig. 5

He then represents 12 syllogisms (without captions) by means of this method: four universal ones employing combinations of nested regular (convex) and non-convex

[8] I thank a reviewer for a highly relevant reference [17] on projections of Euler diagrams, similar to this anticipation of Peirce's. The iterated diagram was omitted from the publication of Ms 479 in CP, as indeed was another large diagram that applies iteration to a complex example taken from Ladd-Franklin (SiL: 58–61). Peirce explains: "In order to illustrate the method I will apply it (without any preconsideration at all) to the following problem by Mrs. Franklin […]" (Ms 479: 59). A three-page explanation of this example then follows. Appendix reproduces the image of this large diagram.

[9] Another reviewer assumes these concave curves lack the kind of iconicity or well-matchedness that Peirce might have desired them to have. One way of viewing them could be denoting 'cuts' in a hyperbolic space, which might then restore their presumed iconic character.

ovals (Figs. 6–9), and eight particular ones employing combinations of dots and regular and non-convex ovals that also may overlap (Figs. 10–17):

Fig. 6 *Fig. 7* *Fig. 8* *Fig. 9*

Fig. 10 *Fig. 11* *Fig. 12* *Fig. 13*

Fig. 14 *Fig. 15* *Fig. 16* *Fig. 17*

Peirce also suggests a treatment of spurious syllogisms (Figs. 18–21): just as he later does in Ms 479, the dots can be placed on the boundaries of regular surfaces. This 'fence-sitting' represents "uncertainty whether the existing individual spoken of lies within or without the class that oval represents" (Ms 481). That is, at least one of the zones adjacent to the dot must be non-empty. Peirce's notation can thus express disjunctions that dispense with lines familiar from spider-diagrams that connect the dots between the interiors and exteriors of different zones:[10]

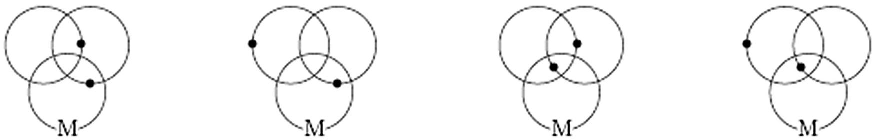


Fig. 18 *Fig. 19* *Fig. 20* *Fig. 21*

These figures are, Peirce adds, "useful in teaching". They show when one extreme is outside and the other inside the middle term, which is where "the force of the [syllogistic] reasoning" lies (Ms 481; cf. Ms 479, alt. pages).

3 Ms 1147 ("Logical Graphs", 1901)

Among the unpublished pages drafted for Baldwin's *Dictionary of Philosophy and Psychology* entry "Logical Graphs" (Ms 1147, 1901), we find Peirce having another take on non-convex Euler diagrams he had introduced in 1896–7. Now he explains the notation and how to denote existence in them as follows:

[10] We surmise that all disjunctions can be expressed by these boundary dots, if the shading of zones to deny existence is also used (see Sect. 4).

In order to represent the negation of the copula of inclusion, a dot may be drawn to represent some existing individual enclosed by one oval but not by the other. Thus, Fig. 1 does not show whether any existing thing is enclosed by A, or not. For it may be invisible. It only shows that if anything is enclosed by A, it must be enclosed by B. But Fig. 2 shows something enclosed by A and not by B. On this system negation is, of course, represented by the reverse side of the enclosure. If by 'enclosed' we mean being on the concave side, we may draw Fig. 1 as in Fig. 3; and Fig. 4 will signify that everything is either A or B.

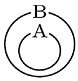
Fig. 1


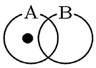
Fig. 2


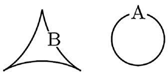
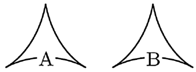
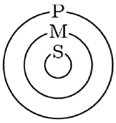
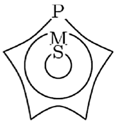
Fig. 3


Fig. 4




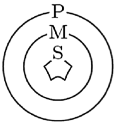
Any M is P,
Any S is M;
∴ Any S is P.

Fig. 4



No M is P,
Any S is M;
∴ No S is P.

Fig. 5

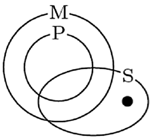


Any M is P,
Everything is either S or M,
∴ Everything is either S or P.

Fig. 6

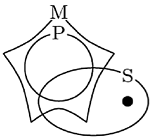
Everything is either M or P,
No S is M;
∴ Any S is P.

Fig. 7



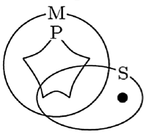
Any P is M,
Some S is not M;
∴ Some S is not P.

Fig. 8



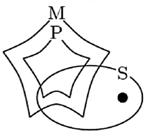
No M is P,
Some S is M;
∴ Some S is not P.

Fig. 9



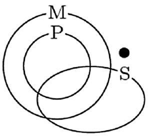
Everything is either M or P,
Some S is not M;
∴ Some S is P.

Fig. 10



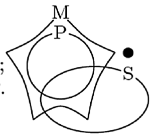
Any M is P,
Some S is M;
∴ Some S is P.

Fig. 11



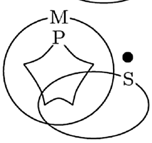
Any P is M,
Something is neither S nor M;
∴ Something is neither S nor P.

Fig. 12



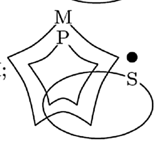
No M is P,
Some M is not S;
∴ Something is neither S nor P.

Fig. 13



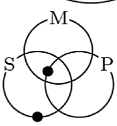
Everything is either M or P,
Something is neither S nor M;
∴ Some P is not S.

Fig. 14



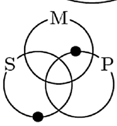
Any M is P,
Some M is not S;
∴ Some P is not S.

Fig. 15



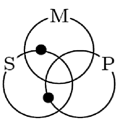
Something is neither M nor P,
Some S is M;
∴ Some S is not some not-P.

Fig. 16



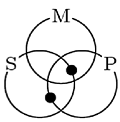
Something is neither M nor P,
Some M is not S;
∴ Some not-S is not some not-P.

Fig. 17



Some M is not P,
Some S is not M;
∴ Some S is not some not-P.

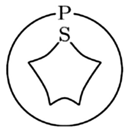
Fig. 18



Some M is P,
Some S is not M;
∴ Some S is not some P.

Fig. 19

He then notes that it adds "to the power of this method to substitute for 'enclosed by', *being on the concave side of"*. An example is the graph on the right, meaning "Everything is either S or P".



Now Peirce is able to represent all 16 syllogisms in these non-convex Euler diagrams, thus repairing the very first defect of the system (Figs. 4–19).[11]

4 Ms 855 ("A Logical Criticism of the Articles of Religious Faith", 1911)

We also find Euler-Venn diagrams in the third draft of the second section of this unpublished set of manuscripts, whose overall purpose was to show the nature of the justification of different kinds of reasoning. (Peirce struggles to find a compelling argument that there is no fourth kind of reasoning besides retroduction, deduction and induction). He is led to remark on the representation of existentials in logical graphs.

Peirce begins with the two examples of Alpha EGs:

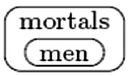


Fig. 1

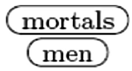


Fig. 2

There is no affirmation of existence of any *thing* in these graphs: Fig. 1 denies that any non-mortal man exists; Fig. 2 denies that any mortal man exists. Peirce attributes to Venn the improvement of Euler's diagrams in which one could understand Fig. 1 to *affirm* "the existence of mortals not men, of mortals that are men, and of beings not mortal" and that "three analogous affirmations might mistakenly be read into Fig. 2." What Venn did was that he *shaded* the compartments "representing the possible class that he meant to deny", so that Fig. 3 "can only mean 'No man is mortal'; Fig. 4, only 'No man is immortal'; Fig. 5, only 'None but men are mortals'; and Fig. 6, only 'Immortals, if there be any, are all men'":

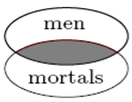


Fig. 3

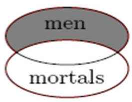


Fig. 4

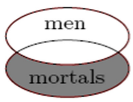


Fig. 5

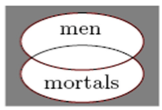


Fig. 6

Peirce writes that also a second type of generalization was in the offing to Venn but that "it did not occur to Mr. Venn that he ought to use a second kind of shading or tinting to mark any kind of thing whose existence (or possibility) he meant to affirm. Thus, he could but did not have Fig. 7 assert that 'Some man is mortal'; Fig. 8, that 'Some man is immortal'; Fig. 9, that 'Something besides men is mortal'; and Fig. 10, that 'Something is neither a man nor mortal'":

[11] There are variations concerning the 12 schemas of syllogisms in Ms 481 and the 16 schemas in Ms 1147: Figs. 10, 11, 14 and 15 from Ms 481 are not represented in this latter table.

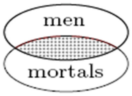


Fig. 7

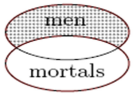


Fig. 8

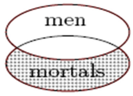


Fig. 9

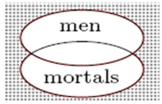


Fig. 10

His comment suggests that he either never learned about or else had forgotten Venn's 1883 review of SiL [16] in which Venn proposes this second type of shading.[12]

From these considerations Peirce moves on to note what to him is the most serious defect of the system of representation by Euler or Venn diagrams: that there is no way of representing *dependent quantification* in those systems. His example is the following pair of sentences:

> There is a man who loves whatever woman there may be.
> There is no woman who has not some man who loves her.

In other words, Euler diagrams cannot assert universal-existential statements or existential-universal statements with two-place relations, which is the crucial element of the true meaning of quantifiers: to express their relative scopes. It is the relative scope phenomenon that lends first-order logic its expressive power.

5 Conclusions

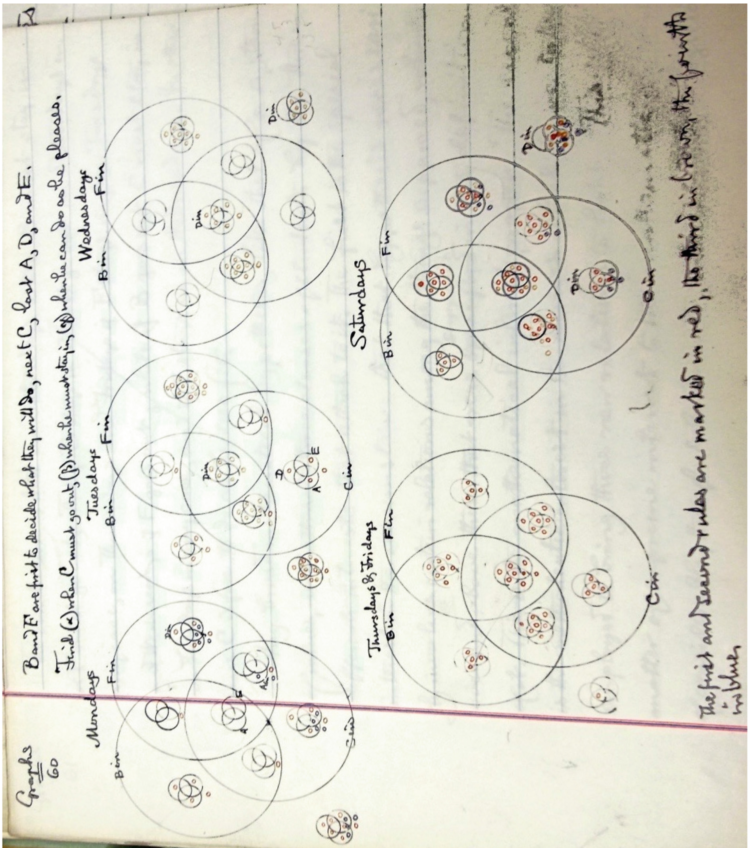
In these four and mostly unpublished manuscripts Peirce manages to present a number of novel and anticipatory ideas how to overcome the five defects he took to haunt ordinary Euler diagrams. Whether his ideas are all sound, or whether they can be combined to yield new consistent systems, perhaps with increased expressive power equal to or greater than monadic first-order logic, remains to be seen. Peirce himself concluded that "the chief interest of non-relative deductive logic is not of the mathematical kind. That is to say, it does not lie in deducing necessary conclusions from assuming hypotheses nor in discovering methods for making such deductions" (Ms 479: 58, alt. variant). He might have not thought very highly of the prospects of Euler diagrams, which might have given him some additional reason to pursue methods of logical graphs that express relative assertions and quantification free from the prevailing restrictions of Euler circles [19, 20]. The relationship between the two was to him as wide as that of between the "dichotomic mathematics of non-relative logic" ("one dull chapter" and "the very most rudimentary that mathematics can be" – a Boolean algebra) and the "mathematics of plane geometry" ("an inexhaustible Proteus"). His quaint analogue was that of "the works of Mother Goose" in comparison to "Voltaire" (Ms 479: 55, alt. variant).

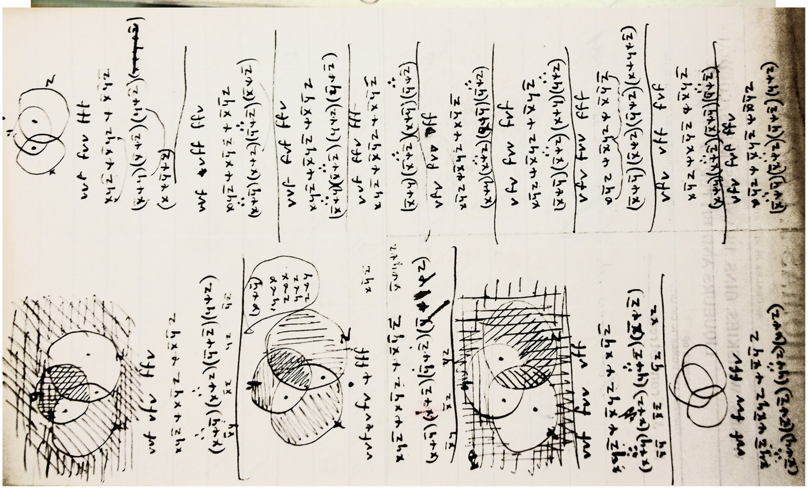
Acknowledgments. Supported by projects PUT267 (Estonian Research Council) and 127335 (Academy of Finland): *Diagrammatic Mind: Logical and Cognitive Aspects of Iconicity*; 2014–15 Foreign Experts Program of State Administration of Foreign Experts Affairs, China; and the 2012 Harvard Houghton Library Fellowship (Principal Investigator A.-V. Pietarinen).

[12] See [1]. However, a few draft pages exist in which Peirce uses up to five different types of shadings (e.g. Ms L 76, written on the back of a letter, not microfilmed). Their semantics remains veiled. See Appendix, second figure.

A Appendix

The images below are from Ms 479 (top/right) and Ms L 76 (bottom/left).





References

1. Moktefi, A., Pietarinen, A.-V.: On the diagrammatic representation of existential statements with Venn and Euler diagrams. J. Logic Lang. Inf. **24**, 361–374 (2015)
2. Shin, S.-J.: The Logical Status of Diagrams. Cambridge University Press, Cambridge (1994)
3. Peirce, C.S.: Manuscripts. Charles Sanders Peirce Papers (MS Am 1632). Houghton Library, Harvard University. Catalogued in Robin, R.S. (1967). Annotated Catalogue of the Papers of Charles S. Peirce. Amherst: University of Massachusetts Press [Ms] (1839–1914)
4. Peirce, C.S. (ed.): Studies in Logic. By Members of the Johns Hopkins University. Little, Brown, and Company, Boston (1883). [SiL]
5. Peirce, C.S.: Collected Papers of Charles Sanders Peirce, 8 vols., Hartshorne, C., Weiss, P. (eds.), vols. 1–6, Burks, A.W. (ed.), vols. 7–8. Harvard University Press, Cambridge (1931–1958). [CP]
6. Peirce, C.S.: Logic of the Future: Peirce's Writings on Existential Graphs. Pietarinen, A.-V. (ed.) (2016, to appear)
7. Howse, J., Stapleton, G., Taylor, H.: Spider diagrams. J. Comput. Math. **8**, 145–194 (2005)
8. Roberts, D.D.: The Existential Graphs of Charles S. Peirce. The Hague, Mouton (1973)
9. Stapleton, G., Howse, J., Taylor, J.: A decidable constraint diagram reasoning system. J. Logic Comput. **15**, 975–1008 (2005)
10. Couturat, L.: The Algebra of Logic. Open Court, Chicago (1914)
11. Edwards, A.W.F.: Cogwheels of the Mind: The Story of Venn Diagrams. JHU Press, Baltimore (2004)
12. Rodgers, P.J., Flower, J., Stapleton, G.: Introducing 3D Venn and Euler diagrams. In: Proceedings of the 3rd International Workshop on Euler Diagrams (2012)
13. Moktefi, A., Bellucci, F., Pietarinen, A.-V.: Continuity, connectivity and regularity in spatial diagrams for N terms. In: Burton, J. Choudhury, L. (eds.), Diagrams, Logic and Cognition, CEUR Workshop Proceedings, vol 1132, pp. 23–30 (2013)
14. Pietarinen, A.-V.: Existential graphs: what a diagrammatic logic of cognition might look like. Hist. Philos. Logic **32**(3), 265–281 (2011)
15. Pietarinen, A.-V.: Moving pictures of thought II: graphs, games, and pragmaticism's proof. Semiotica **186**, 315–331 (2011)
16. Venn, J.: Review of studies in logic. Mind **8**, 594–603 (1883)
17. Gil, J., Howse, J., Kent, S., Taylor, J.: Projections in Venn-Euler diagrams. In: IEEE Symposium on Visual Languages, pp. 119–126 (2000)
18. Bellucci, F., Moktefi, A., Pietarinen, A.-V.: Diagrammatic autarchy: linear diagrams in the 17th and 18th century. In: Burton, J., Choudhury, L. (eds.) Diagrams, Logic and Cognition, CEUR Workshop Proceedings, vol. 1132, pp. 31–35 (2013)
19. Pietarinen, A.-V.: Peirce and the logic of image. Semiotica **192**, 251–261 (2012)
20. Bellucci, F. Pietarinen, A.-V.: From Mitchell to Carus: 14 years of logical diagrams in the making. Trans. Charles S. Peirce Soc. (2016, in press)

Diagrams and Education

Hint, Instruction, and Practice: The Necessary Components for Promoting Spontaneous Diagram Use in Students' Written Work?

Emmanuel Manalo[1(✉)] and Yuri Uesaka[2]

[1] Graduate School of Education, Kyoto University, Kyoto, Japan
emmanuel.manalo@gmail.com
[2] Graduate School of Education, The University of Tokyo, Tokyo, Japan
y_uesaka@p.u-tokyo.ac.jp

Abstract. This study investigated the efficacy of providing a hint, instruction, and practice in promoting spontaneous diagram use in the written work of 21 students undertaking an undergraduate course in education. The course required the students to regularly produce for homework a one-page explanation of what they had learned. In the first few weeks of the course, they rarely included diagrams in their explanations. Following a hint to use diagrams (provided as comment/feedback on their homework), diagram use significantly increased. When instruction in effective use of diagrams was provided, the level of diagram use maintained but did not increase. However, when practice in using diagrams was additionally provided, further significant increases in diagram use followed, which maintained over the ensuing weeks of the course. These findings suggest that to spontaneously use diagrams in their written work, students need to be provided a combination of advice, instruction, and practice in such use.

Keywords: Spontaneous diagram production · Written communication · Strategy use advice and encouragement · Diagram use instruction · Skills practice

1 Introduction

The research literature concerning the use of diagrams in communicative situations indicates that such use is efficacious [1–4]. When both verbal representations (such as text on a printed page or words spoken by a teacher) and visual representations (such as illustrations or other forms of diagrams shown on a page, board, or screen) are appropriately used in conveying a message, both the verbal and visual channels of the message recipient's working memory are engaged, making it more likely that the intended message would be understood. In simple terms, the message recipient not only reads or hears the content of the message, but also sees what it might 'look like.' When integrated, the meaning of what has been read/heard *and seen* could make understanding of the intended message easier. For example, it would likely be easier to grasp the structure of a topic if it is not only written or spoken about but also shown in terms of a schematic diagram.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 157–171, 2016.
DOI: 10.1007/978-3-319-42333-3_12

Despite the apparent usefulness of including diagrams in communicating information to others, there is one serious problem: students generally lack spontaneity in using diagrams in such communication [5–7]. There is not a great deal of research that has been conducted regarding this problem, but what research has revealed about the factors that influence student diagram use in communication – particularly written communication – is outlined in the following subsection.

1.1 Factors that Influence Diagram Use in Communicative Situations

One important finding is that the intended audience of the communication makes a difference as to whether diagrams would likely be used [5, 7]. More specifically, students are more likely to include diagrams when writing notes for their own selves, and less likely to include diagrams when writing explanations for other people. Manalo and Uesaka [7] suggested that a possible reason for this is that diagrams may be perceived as serving more useful functions in writing notes for oneself (e.g., summarization of main points, connection of key ideas). In contrast, diagram use may be viewed as more risky when producing explanations for others: such use could lead to misunderstanding as diagrams tend to leave out non-essential details, and they demand a greater degree of interpretation on the part of the audience. It is also possible that students view diagrams as less 'formal' than words when explaining what they know in academic contexts. Such a view could arise because important means for conveying knowledge – such as essays, reports, and test answers – explicitly require writing in words. However, diagrams are at most optional for such products, and may be considered as belonging more to the planning stage rather than the final product.

Another important finding is that some individual- and task-related factors influence the likelihood of diagram inclusion in written communication [6, 7]. The reason is that these factors affect the cognitive processing cost associated with diagram production. One example of an individual factor is language proficiency: when students have to use a foreign language to explain information they have learned, they are less likely to employ diagrams – especially if their proficiency in that language is low. This may seem counter-intuitive in that one would imagine that students would more likely resort to the use of diagrammatic representations if they have to use a language they are not so proficient in (i.e., to compensate for what they might find difficult to explain in that language). However, there is limited processing capacity in working memory [8, 9] and when students have to use a language they lack proficiency in, production of text in that language depletes the cognitive processing resources in working memory to the extent that insufficient resources remain for the production of any diagrams. Manalo and Uesaka [6] reported evidence for this: Japanese university students' proficiency in English was found to significantly correlate with their use of diagrams when explaining what they had learned in English, but not in Japanese.

Where task-related factors are concerned, an example is the imageability of the information that needs to be explained (i.e., how easy or difficult it is to imagine). Manalo and Uesaka [6, 7] reported findings that when students have to explain information of low imageability, they are less likely to use diagrams. The reason is essentially the same as for the previously mentioned language proficiency finding:

constructing diagrams to represent information that is hard to imagine demands more cognitive processing resources in working memory, and is therefore less likely to be undertaken because there may be inadequate resources for it.

In the area of math word problem solving, an instructional intervention that has been found to improve students' spontaneous diagram use is the provision of teacher verbal encouragement to use diagrams and practice in drawing diagrams [10]. Uesaka, Manalo, and Ichikawa reported that students who had been provided both encouragement and practice in drawing (in addition to regular instruction in problem solving) subsequently showed the highest improvement in spontaneous diagram use [10]. They explained this finding in terms of verbal encouragement helping students to appreciate the value of diagram use in problem solving, and practice in drawing developing students' procedural knowledge in constructing appropriate diagrams. This explanation is congruent with previous arguments that student learning strategy use depends on their knowing that those strategies would be useful, as well as their knowing how and when to use those strategies [11, 12]. However, previous research had not examined whether encouragement to use diagrams and practice in using diagrams would similarly be effective in increasing spontaneous diagram use in communicative situations.

One intervention that has been shown in previous research to be effective in promoting students' spontaneous diagram use in communicative situations is peer interaction. Uesaka and Manalo reported that when students were required to verbally explain information they had learned to peers in interactive learning situations, they spontaneously drew more diagrams in the process of explaining (more so than students in a control condition where they had to similarly explain, but in a non-interactive manner) [13]. Uesaka and Manalo explained that interaction facilitates awareness of the usefulness of diagrams in such communicative situations: through feedback and questions that the explainer's interlocutor provides during the interaction process, the explainer comes to realize the limitations of using words alone, and the need to use other representations – particularly diagrams – to successfully convey the content of the explanation. The finding of this study confirms the importance of perceiving the value of diagram use if students are to spontaneously use diagrams in their communicative efforts.

However, even though peer interaction has been found to be effective in promoting spontaneous diagram use in communication *while students were in the process of interacting,* no evidence has been found that such diagram use transfers to other subsequent communication tasks. In fact, Manalo et al. [14] reported that despite a spontaneous increase in student diagram use during an interactive peer explanation phase in their study, diagram use reverted to previously low levels in a subsequent (non-interactive) explanation writing task. The reason for this transfer failure is important to understand as self-regulation in learning requires that students are able to apply their knowledge at crucial times during learning performances [15].

There are two possible reasons for the lack of spontaneous diagram use in the subsequent explanation writing task. One is that, from the peer interactive explanation session, the students could have acquired a more task-specific knowledge that "diagrams are useful *when verbally explaining in an interactive manner to others,*" rather than the more abstract, general, and transferrable knowledge that "diagrams are useful *when explaining information to others.*" The other possible reason is that, even if they had acquired the more abstract knowledge about the usefulness of diagrams in explaining, many of the students might have lacked the necessary skills in constructing the appropriate diagrams for the explanations they were writing.

1.2 Problem Statement and Overview of the Present Study

The main challenge addressed in the present study was how to promote students' spontaneous use of diagrams in written communication – particularly when explaining information to others. A secondary challenge was to design an intervention that would have ecological validity – in other words, an intervention that would work not only in an experimental situation, but also in real educational contexts.

The interventions used in the present study aimed at directly addressing issues that have been identified in previous research as likely impediments to spontaneous diagram use. Thus, to address the possibility that students might not realize the value of incorporating diagrams in their written work, a hint about the usefulness of diagrams was provided by the instructor in the form of individual written feedback on explanations that students produced. To address the possibility that students might be deficient in knowledge about diagram use for enhancing the communicative effectiveness of written work, instruction on such use was provided. And to address the possibility that students might lack skills in constructing the appropriate diagrams to use when explaining various kinds of information, practice was provided in such construction.

The second challenge concerning ecological validity was addressed by conducting the study described here within a real undergraduate course in education studies in a national university in Japan. The course is taught entirely in English, and the majority of students who take the course are Japanese, for whom English is a foreign language. During the semester when this study was conducted, some international students were also enrolled in the course, but all students had English as a second or foreign language. Apart from covering various theories, concepts, and research in education, the goals of the course include the development of students' communicative competence. Thus, course conduct incorporates activities requiring oral and written output from students (e.g., discussions, written exercises) to facilitate the development of such competence. One such activity is for students to complete a one-page written explanation homework task each week, in which they are asked to explain what they have learned in the course during that particular week to an imaginary student who does not know anything about the contents of the course. The interventions in this study focused on students' spontaneous use of diagrams in that homework assignment.

The main hypothesis tested in this study was that the provision of a hint to enhance perception about the usefulness of diagrams, instruction to improve knowledge about effective use of diagrams, and practice to develop skills in constructing diagrams would result in significant increases in students' spontaneous use of diagrams in explanations they write. A related second hypothesis was that, while enhancing students' perception about the usefulness of diagrams and improving their knowledge about effective use of diagrams would result in some students using diagrams more spontaneously, it would not be until the students receive practice in the construction of appropriate diagrams that the majority would evidence the desired spontaneity in diagram use. This hypothesis was based on previous findings suggesting that perception of usefulness (indicated by increased diagram use during interactive explanations with peers [13, 14]), and knowledge about effective use of diagrams (indicated by diagram use in notes that students had taken for their own selves [14]) may not be enough to promote spontaneity in diagram use *when constructing written explanations for others.* Practice may additionally be necessary as students may lack skills in constructing diagrams that they could be sufficiently confident about in terms of enhancing the effectiveness of explanations they write for other people.

A third hypothesis tested in this study was that, from beginning to end of semester, students would evidence improvements in their spontaneous diagram use in both note taking and explanation writing as measured by their performance in tasks (pre- and post-intervention tests) that are different from the one they receive the intervention in (i.e., their weekly explanation writing task). A related fourth hypothesis was that, while in the pre-intervention test students might evidence higher diagram use in note taking compared to explanation writing, such a difference would no longer be present in their post-intervention test (i.e., diagram use in explanation writing would increase to the extent that it would no longer be lower than in note taking). Previous research has shown that students tend to use more diagrams when taking notes for their own selves compared to when writing explanations for others [5, 7], so it would be interesting to examine whether the interventions used in this study might be sufficient to reduce or eliminate that difference in use.

2 Method

2.1 Participants

The participants were 21 undergraduate students taking an introductory course in education studies (aged approximately 19–20 years; females = 11; Japanese = 13, other nationalities = 8). Faculty ethics committee approval was obtained for the conduct of this study. The students were provided written and verbal explanations at the beginning of the course that some of the work they produce would be analyzed for research and course development purposes. They were given an option of having their work excluded from such analyses, but all students provided written consent for use of their work.

2.2 Materials and Procedure

Pre- and Post-Intervention Tests. The course that the students were taking comprised a total of 14 weekly 90-min class sessions. At the end of the first and the thirteenth class sessions, the students were given a reading/note taking and explanation-writing task as ‘independent’ pre- and post-intervention tests (i.e., ‘independent’ in the sense that these had nothing to do with the regular content of the course). These were administered to obtain measurements of the students’ use of diagrams in note taking and explanation writing, with the use of materials that could be experimentally controlled (in contrast to the regular materials used in the course, over which experimental control was deemed inappropriate). Two short English passages, both just under 600 words in length, were used as reading materials: one about the jigsaw classroom, and the other about theory of mind. These topics were selected because they were similar to the kinds of topics dealt with in the course, but were not included as part of the course. Care was taken in preparing these reading materials to make them as equivalent as possible. Approximately half of the students were randomly given one passage, while the other half received the other passage, in the pre-intervention test. The students then received the other passage they had not read in the post-intervention test.

The procedure used in administering the pre- and post-intervention tests were the same. The students were given 10 min to read and take notes from the passage they were assigned. They received an A4-size sheet of paper on which to take notes, and they were informed that they could use their notes in an explanation task that would follow, but that they would not be able to refer back to the passage they were reading. The students were then given 10 min to produce an explanation of the passage they had read, imagining that their audience was another student who knew nothing about that topic. After this, they were given five questions to answer. The first two required responses on 5-point Likert-type scales, and asked about prior knowledge concerning, and ease/difficulty in understanding, the passage they had read. The other three questions were to assess their comprehension of the passage and required short, written answers.

Weekly Explanation Homework and Interventions Used. As mentioned in the introduction section, the course required students to complete and submit an explanation homework task each week (except in weeks 1, 6, and 14). This homework required students to explain the most important points they had learned from the class session that week. They were asked to imagine that their reader was another student who knew nothing about the topics covered in the course. They were also informed that the explanation should be sufficient on its own (i.e., the reader should understand it without having to be provided additional verbal explanation). The students received an A4-size sheet of paper to write their explanation on. The homework was collected the following week for instructor feedback, and returned the week after that.

It should be noted that no marks or grades were given for each homework task sheet that the students completed, only written comments about the quality and adequacy of the explanation they produced. However, the students were required to include those sheets in their portfolio (for submission at semester end), which was allocated 40 % of the total course grade. In the grading rubrics for that portfolio, marks were allocated for satisfactory completion and quality features of the homework tasks. However, no mention was made in those rubrics of diagrams, or of expectations for students to include diagrams. Thus, diagram use in the explanation homework tasks was neither an explicit requirement, nor a feature directly linked to marks or grades.

The interventions used in the present study were (1) a hint about the usefulness of diagrams in writing explanations, (2) instruction in the effective use of diagrams for such explanations, and (3) practice in the construction of diagrams to use in explaining various kinds of information. These interventions were provided at key stages during the weeks of the semester to find out their effect on student diagram use.

The hint was provided as a comment that "including diagrams could make your explanations easier to understand" (the same wording was used for all students). This was written, together with any other comments, on the bottom of students' homework task sheets. (The sheets with feedback were returned to students individually during class, and students were encouraged and given a brief amount of time during class to read over the feedback they had received.) Provision of the hint was staggered so that some of the students (randomly selected) received it earlier than others. For those who received the hint earlier, it was provided on their week 3 homework, which was submitted in week 4 and returned in week 5; thus, any effects that the hint could have had would have been evident from their week 5 homework. For the students who received the hint later, the corresponding weeks were: hint given on week 5 homework, submitted in week 7 (as the students worked on a project in week 6), and returned in week 8; thus, any effects would have been evident from their week 8 homework.

All students received the instruction on effective use of diagrams in week 10; thus any effect of that instruction should have been evident from their week 10 homework. Approximately 20 min instruction was provided toward the end of the class session, covering reasons for using diagrams (i.e., to help clarify own understanding of the information to be explained, and because research has shown that people learn better from words and pictures than from words alone [e.g., 3]), and ways to use diagrams in explanations (i.e., to illustrate, provide an overview or structure, show process or cause-and-effect relationships, and compare or contrast). Each of these reasons and ways was explained and examples of the kinds of diagrams referred to were shown. However, the students were not given an opportunity during the week 10 class session to practice constructing diagrams.

Practice in constructing diagrams was provided during the week 11 class session. Approximately 30 min toward the end of the class was allocated to this. First, the instructor quickly reviewed the key points from the instruction about diagram use provided in week 10. Then students were given a photocopy of the week 3 explanation homework they had earlier submitted. This particular homework was selected because not a single student included a diagram for it. The students were provided a new sheet with instructions to consider and draw diagrams they might be able to include to make their explanation easier to understand. The topic covered in the week 3 class session was early childhood education, and a few examples of diagrams students produced during the practice session in week 11 are shown in Fig. 1. During the session, the instructor was available to provide comment and/or feedback, and students could briefly discuss their newly constructed diagrams with other students.

In addition to their usual explanation homework, the students were also assigned an additional homework task in week 11, which was to construct one diagram for each of the ways diagrams could be used in explanations using any of the topics/materials that had been covered in the course up to that time. This homework was assigned to give the students additional practice in constructing diagrams, and would have likely required at least 30 min of their time to complete. Examples of the diagrams that one student produced for this homework are shown in Fig. 2.

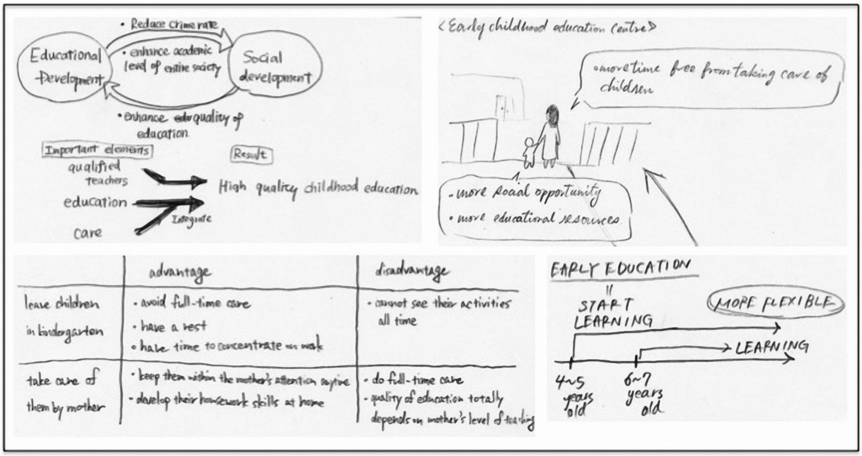


Fig. 1. Examples of diagrams that students produced during the week 11 class practice

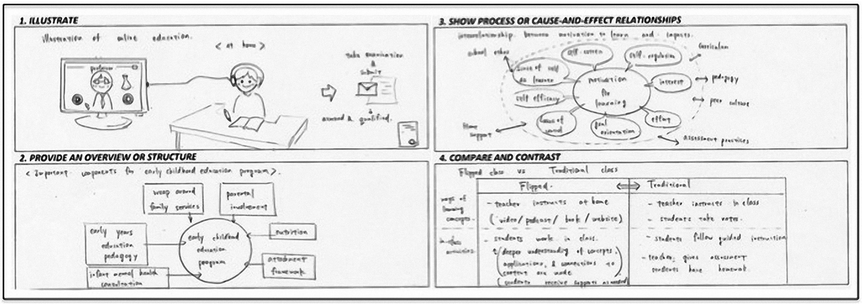


Fig. 2. Examples of diagrams that one student produced for week 11 practice homework

No intervention was provided in weeks 12 and 13, so the explanation homework that students submitted during those weeks' classes were examined for maintenance of any spontaneous diagram use they might have acquired as a consequence of the interventions provided in the preceding weeks.

For the sake of clarity, Table 1 shows the intervention phases and the homework tasks that were categorized under those phases.

Table 1. Weekly homework task numbers belonging to the different phases of the study, according to whether the hint was provided early or later

Hint provision	Baseline	After hint	After instruction	After instruction + practice	Maintenance
Early	2, 3, 4	5, 7, 8, 9	10	11	12, 13
Later	2, 3, 4, 5, 7	8, 9	10	11	12, 13

2.3 Analysis

The students' homework sheets were examined to determine whether the students used a diagram in their explanations. Use of at least one diagram was scored as 1, and no diagram as 0 (the number of diagrams used was not taken into consideration in scoring). For the purposes of this study, a diagram was defined as any representations produced by the students, other than representations in the form of words, sentences, or numbers on their own. For example, drawings and charts counted as diagrams, as did arrows and similar symbols when these were used to link three or more concepts. Analysis focused on whether the interventions made a difference to the proportions of students using diagrams in their homework over the course of the semester.

For the notes and explanations that the students produced in the pre- and post-intervention tests, similar scoring (i.e., to determine whether or not a diagram was used) was applied. The proportions of students using diagrams in their notes and explanations at pre-intervention and at post-intervention were then compared.

The first author and a research assistant with no vested interest in the outcomes of this study independently carried out data scoring. The kappa coefficient values for inter-rater agreement were .92 for the homework data and .85 for the pre- and post-intervention tests data, both of which represent almost perfect agreement [16].

3 Results

3.1 Did the Interventions Have an Effect on Students' Diagram Use in Their Homework?

An analysis of variance (ANOVA) was carried out on the students' diagram use data in their homework tasks, with timing of hint provided (early, later) and intervention phase (baseline, after hint, after instruction, after instruction + practice, maintenance) as independent variables (between-participant and within-participant, respectively). The authors had earlier agreed on a criterion of "no more than three missing assignments" for any student's data to be included in the analysis and, based on this decision, one student's data was excluded from this analysis.

The results revealed a significant effect due to phase, $F(4, 72) = 12.07$, $p < .001$. Figure 3 shows the mean proportions of student diagram use in each of those phases. The effects due to the timing of hint and the interaction were both not significant.

Simple main effects analysis using Ryan's method (with the significance level set at .05) revealed significant differences in pairwise comparisons between all the phases, except (i) between "after hint" and "after instruction", and (ii) between "after instruction + practice" and "maintenance". These results indicate that provision of the hint significantly increased students' diagram use in their homework. However, the provision of instruction did not add any further significant increases to the level of diagram use already achieved following the hint provision. It was not until practice was additionally provided that further significant increases in diagram use ensued. This level of diagram use was maintained over the remaining two weeks of the semester. These findings lend support to the first two hypotheses posed in this study.

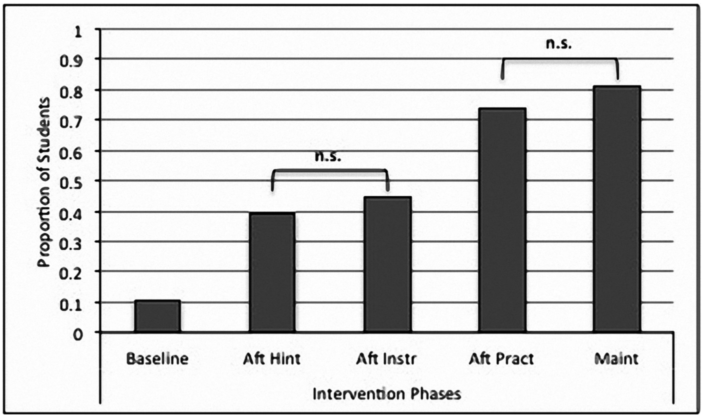


Fig. 3. Mean proportions of student diagram use during the intervention phases

3.2 Did the Students' Diagram Use in Their Written Work Increase from Beginning to End of the Semester?

An ANOVA was also carried out on the students' diagram use data in the pre- and post-intervention tests. Passage order (jigsaw classroom or theory of mind passage given at pre-intervention), time (pre-intervention, post-intervention), and tasks (note taking, explanation writing) were the independent variables, with passage order being a between-participant variable, and the other two being within-participant variables. Two students' data were excluded from this analysis as they were absent for the post-intervention test.

The results revealed a significant time effect ($F(1, 17) = 4.18$, $p < .001$), and a marginally significant interaction effect between time and task ($F(1, 17) = 3.99$, $p = .062$). The effect due to passage order was not significant.

The significant effect due to time indicates that the students used more diagrams at post-intervention compared to pre-intervention. Simple main effects analysis of the interaction between time and task revealed that, at pre-intervention, the students' diagram use in note taking was significantly higher than in the explanations they produced, $F(1, 34) = 5.14$, $p = .030$. However, the difference between note taking and explanation writing was no longer significant at post-intervention, $F(1, 34) = .233$, $p = .633$. These differences can clearly been seen in Fig. 4. The simple main effects analysis also revealed that diagram use in note taking significantly increased from pre- to post-intervention ($F(1, 34) = 7.533$, $p = .010$), as did diagram use in explanation writing ($F(1, 34) = 30.671$, $p < .001$). These findings lend support to the third and fourth hypotheses posed in this study.

The students' responses to the questions asked in the pre- and post-intervention tests indicated that the students had limited prior knowledge about the topics of the passages, but they understood most of their content. Overall performance in the comprehension questions was high (range of means for the questions = 70–100 % correct) confirming that the students mostly understood the content of those passages.

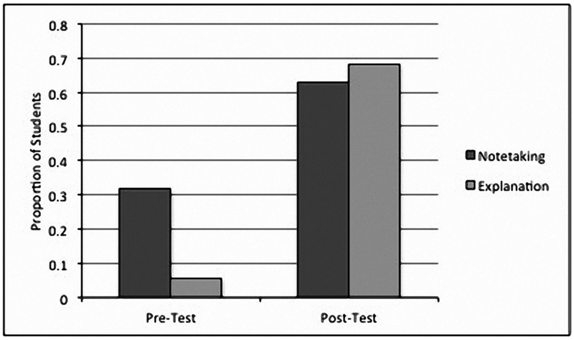


Fig. 4. Mean proportions of student diagram use in note taking and explanation writing in the pre- and post-intervention tests

4 Discussion

The hypotheses tested in this study were supported by the results. The interventions were effective in increasing diagram use in the students' explanation writing, as evidenced by the significant effect of intervention phase, which confirmed the first hypothesis. As Fig. 3 shows, the majority of students did not include diagrams in the explanations they produced until after both instruction *and practice* had been provided, confirming the second hypothesis about the importance of practice in promoting spontaneity in diagram use. The third hypothesis was also confirmed: significant improvements in diagram use were observed not just in the students' homework but also in their post-intervention test performance. Finally, in the students' post-intervention test performance, diagram use in note taking and in explanation writing was found to be equivalent (which was not the case in their pre-intervention test performance), suggesting a change in students' perceptions about the relative value of including diagrams in notes and in explanations – and confirming the fourth hypothesis.

4.1 Why the Interventions Worked

As noted in the introduction section, the interventions used in this research aimed at directly addressing issues that had previously been identified as likely impediments to diagram use. Those issues were failure to realize the value of incorporating diagrams in written work, deficiency in knowledge about the use of diagrams for enhancing the communicative effectiveness of written work, and inadequacy of skills for constructing diagrams that may be deemed useful [7, 13, 14]. Thus, the success of the interventions used in this study can be explained in terms of reducing or eliminating barriers that students may encounter in diagram use.

The hint provision might in effect have provided students with two of the three sorts of knowledge about strategies that Paris et al. [11] considered as necessary for invoking learning strategies: knowing *that,* and knowing *when.* The instructor-provided hint could have made students realize *that* diagrams could be useful *when* attempting to write explanations for other people. Although such knowledge may sound obvious to

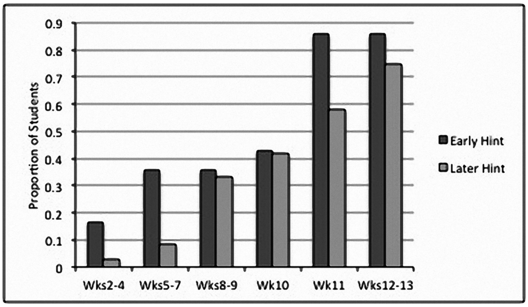


Fig. 5. Mean proportions of diagram use by the "early hint" and "later hint" students over the weeks of the semester

diagrams researchers, it may not be as obvious to the majority of students as academic socialization mainly emphasizes the use of verbal/textual representations in conveying to others knowledge that has been acquired – such as in tests, and in reports and other forms of assignment [5].

The effect of timing of the hint provided was not found to be significant in the statistical analysis undertaken. The analysis compared the overall diagram use of the students who received the hint early and those who received it later. Thus, it would make sense that their overall diagram use would be equivalent, otherwise the groups could be considered as dissimilar or even non-comparable. However, as shown in Fig. 5, the hint provision produced the predicted increases in diagram use among the early-hint and later-hint groups in weeks 5–7 and weeks 8–9, respectively.

The provision of instruction as part of the interventions provided was intended to address one of the key reasons for failure to use learning strategies that Garner [12] identified: knowledge deficiencies. Students may be aware of certain learning strategies or even that those strategies are supposed to be effective, but if they are deficient in their knowledge about how those strategies can be used, those students are unlikely to use the strategies. Hence, students may know about diagrams and their usefulness in learning situations but, for those students to actually use diagrams, they first need to know how they can use diagrams in target learning situations.

An interesting and somewhat unexpected finding in the present study was that the provision of instruction did not result in any further increase in student diagram use beyond what had already been attained following the hint provision (see Fig. 3). The most likely explanation for this is that instruction may have provided students with useful semantic knowledge about diagram use in explanations, but not the procedural knowledge necessary for them to confidently apply that semantic knowledge to their own work. This explanation is supported by the finding that, when practice in constructing diagrams was later provided, a significant increase in the proportion of students who used diagrams followed. However, to avoid any possible misunderstanding, it should be stressed here that *practice on its own* – without instruction – would also likely be insufficient. Without the corresponding instruction, students would lack essential semantic knowledge, and any practice they undertake would lack focus on the variety of ways for effectively using diagrams.

The third component of the intervention – practice – was crucial in that it provided students with opportunities to apply instruction they had received and/or knowledge they already possessed about useful ways to incorporate diagrams in explanations. The third sort of knowledge that Paris et al. [11] considered necessary for students to use learning strategies was knowing *how*. The findings of this study suggest that knowing how has two vital components: knowledge about how the strategy can be used, and skills about how that knowledge can be utilized in target situations. Without the latter, students may not spontaneously use a strategy: they may hesitate or desist in using the strategy as it could be too troublesome to use [cf. 17], and they could end up making mistakes in using it. Practice, however, promotes the development of procedural knowledge (i.e., knowing what to do). Thus, acquiring the necessary procedural knowledge would likely result in making the prospect of using the strategy appear less troublesome and less fraught with potential pitfalls.

Prior research has revealed that cognitive processing cost could also influence students' spontaneity in using diagrams in communicative situations [6, 7]. The instruction and practice components of the interventions used in the present study probably contributed to reducing the processing cost involved in diagram production. Semantic knowledge about how to use diagrams acquired through the instruction component, and procedural knowledge about how to construct diagrams developed through the practice component, likely made it less cognitively costly to think about and construct the diagrams that could assist in clarifying the explanations the students were writing.

4.2 Transfer to Other Tasks

A very important finding in the present study was that increases in students' spontaneous diagram use were observed, not just in their weekly homework tasks (where the interventions were implemented), but also in the post-intervention test administered toward the end of the semester. This suggests transfer of spontaneity in diagram use from the homework situation to the test situation. Although the tasks involved in these were similar (e.g., explanation was required in both homework and the explanation writing component of the test), there were also important dissimilarities: the post-intervention test was conducted under time constraint, and it also included note taking (in which increased diagram use was also observed).

The significant increase in spontaneous diagram use in note taking and explanation writing in the post-intervention test is also important because those tests were independent and experimentally controlled. With the real materials that the students were learning in class, it was difficult and ethically problematic to impose such control. It was therefore possible that factors like imageability (which can affect diagram use [6, 7]) varied between the materials covered in the class sessions each week. Thus, to be able to verify the increase in students' spontaneous diagram use, using that independent and experimentally controlled measure was crucial from a research perspective.

Furthermore, although this has not been reported in the results section of this paper as it was found through a subsequent post hoc analysis, transfer was also observed in the students' final test writing: 13 out of the 21 students (62 %) still used diagrams in some of the test question answers they produced. That final test differed in format from the students' weekly homework, and it was held several weeks after the last intervention had been provided. This finding is therefore indicative not only of transfer to a somewhat different written explanation task, but also of maintenance of the intervention effect over a longer period of time.

A crucial question to address in future research is whether the diagrams that students spontaneously produce truly enhance the communicative quality of their written work. This question was deemed outside the scope of the current paper because of time and publication-length constraints. However, one indication that diagram use did enhance the quality of students' written work is that, in the final test mentioned above, the students who used diagrams scored significantly higher (mean = 18.69, SD = 1.25) than the students who did not (mean = 15.88, SD = 2.75), $t(19) = 3.23$, $p = .004$. The test questions were scored solely on correctness and quality of the answers produced, and no points were allotted to inclusion of diagrams. The higher score of those who used diagrams therefore suggests better quality answers, possibly as a consequence – at least in part – of incorporating the diagrams. Again, this was found through a later post hoc analysis and therefore not included in the main results.

4.3 Conclusion

The most important finding in the present study is that it is possible to promote spontaneity in students' diagram use in written communication – not in an experimental setting, but in a real educational context. The components of the interventions used (hint, instruction, and practice) appeared to have brought about the desired, significant change in the students: from almost none of them using diagrams in their written work at the beginning of the semester, to almost all of them employing diagrams in their production of such work at the end of the semester.

Diagram use is efficacious in many educational and daily life contexts and, as such, it is generally considered important for students to acquire the skills necessary for such use. Few research studies however have addressed the question of how to promote spontaneity in students' diagram construction and use – even though in reality there would be few daily life situations where people would find diagrams supplied to them for their use. The present study developed and tested one viable method for promoting student spontaneity in diagram use in the area of communication. The authors hope that the successful outcomes reported here would stimulate further research into this important but largely neglected aspect of diagram use.

Acknowledgments. This research was supported by a grant-in-aid (15H01976) received from the Japan Society for the Promotion of Science. The authors would like to thank Tatsushi Fukaya for his helpful suggestion concerning the research design, and Christine Li for her assistance in data processing.

References

1. Ainsworth, S.: DeFT: a conceptual framework for considering learning with multiple representations. Learn. Instr. **16**, 183–198 (2006)
2. Ainsworth, S., Prain, V., Tytler, R.: Drawing to learn in science. Science **333**, 1096–1097 (2011)
3. Mayer, R.E.: Multimedia Learning, 2nd edn. Cambridge University Press, New York (2009)
4. Mayer, R.E.: The promise of multimedia learning: using the same instructional design methods across different media. Learn. Instr. **13**, 125–139 (2003)
5. Manalo, E., Ueaska, Y., Kriz, S., Kato, M., Fukaya, T.: Science and engineering students' use of diagrams during note taking versus explanation. Educ. Stud. **39**, 118–123 (2013)
6. Manalo, E., Uesaka, Y.: Elucidating the mechanism of spontaneous diagram use in explanations: how cognitive processing of text and diagrammatic representations are influenced by individual and task-related factors. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 35–50. Springer, Heidelberg (2012)
7. Manalo, E., Uesaka, Y.: Students' spontaneous use of diagrams in written communication: understanding variations according to purpose and cognitive cost entailed. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 78–92. Springer, Heidelberg (2014)
8. Baddeley, A.D.: Working Memory. Oxford University Press, Oxford (1986)
9. Baddeley, A.D.: Human Memory. Allyn & Bacon, Boston (1998)
10. Uesaka, Y., Manalo, E., Ichikawa, S.: The effects of perception of efficacy and diagram construction skills on students' spontaneous use of diagrams when solving math word problems. In: Goel, A.K., Jamnik, M., Narayanan, N.H. (eds.) Diagrams 2010. LNCS, vol. 6170, pp. 197–211. Springer, Heidelberg (2010)
11. Paris, S.G., Lipson, M.Y., Wixson, K.K.: Becoming a strategic reader. Contemp. Educ. Psychol. **8**, 293–316 (1983)
12. Garner, R.: When children and adults do not use learning strategies: toward a theory of settings. Rev. Educ. Res. **60**, 517–529 (1990)
13. Uesaka, Y., Manalo, E.: How communicative learning situations influence students' use of diagrams: focusing on spontaneous diagram construction and protocols during explanation. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 93–107. Springer, Heidelberg (2014)
14. Manalo, E., Uesaka, Y., Sheppard, C.: Can interactive communication promote students' use of diagrams in explaining what they have learned? In: Papers Presented at the EARLI (European Association for Researchers on Learning and Instruction) Conference, 25–29 August. Cyprus University of Technology, Limassol, Cyprus (2015)
15. Zimmerman, B.J.: Becoming a self-regulated learner: an overview. Theory Pract. **41**, 64–70 (2002)
16. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. Biometrics **33**, 159–174 (1977)
17. Uesaka, Y., Manalo, E., Ichikawa, S.: What kinds of perceptions and daily learning behaviors promote students' use of diagrams in mathematics problem solving? Learn. Instr. **17**, 322–335 (2007)

Promoting Multi-perspective Integration as a 21st Century Skill: The Effects of Instructional Methods Encouraging Students' Spontaneous Use of Tables for Organizing Information

Yuri Uesaka(✉), Mika Igarashi, and Rei Suetsugu

Graduate School of Education, The University of Tokyo, Tokyo, Japan
{y_uesaka,mika81,suetsugu}@p.u-tokyo.ac.jp

Abstract. Integrating multiple perspectives when constructing an argumentation about a topic which has both arguments and counter-arguments is a very important 21st Century skill. In this study, we examined whether using tables for organizing information would be supportive of this argumentation process. Participants were 56 8th-grade students participating in a 5-day summer course. Pre- and post-assessments were administered at the beginning and end, and participants made oral presentations, which were video recorded for analysis. Participants were assigned to one of three conditions. In two of the conditions, participants were encouraged to use tables for organizing information and, in one of those conditions, participants were additionally provided exercises that required identification of problems in videos showing bad examples of argumentation. Results showed that participants encouraged to use tables spontaneously used more diagrams and constructed more argumentation in which multiple perspectives were integrated than others without such encouragement. Participants provided the exercise in problem identification also showed higher recognition of the value of using diagrams.

Keywords: Multi-perspective integration · Table · Diagram · Instructional methods

1 Introduction

1.1 The Necessity of Integrating Multiple Perspectives for Surviving in the 21st Century

In discussions about skills necessary for surviving in the 21st Century, argumentation competence is considered to be one of the most important skills for students to develop [e.g., 1]. Although, in the traditional curriculum, the idea has been that students should learn knowledge and skills that are parts of the academic world, the more recent curriculum innovation movement has changed this idea. New curricula started to be designed based on the competencies that students need to acquire in a future society [e.g., 2]. Key competences proposed by the OECD (Organisation for Economic

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 172–186, 2016.
DOI: 10.1007/978-3-319-42333-3_13

Co-operation and Development) are also congruent with this idea, and they include the competence of being able to use tools interactively (e.g., language, diagrams, and technology) [3]. However, in reality, concrete instructional methods have not been sufficiently proposed for how those competencies could be developed.

Among the many aspects of argumentation skills, the skill to integrate different perspectives about a topic which has both arguments for and against it, is the most crucial. The 21st Century is predicted to be a VUCA (volatile, uncertain, complex, and ambiguous) society, which means that people will meet many problems for which the solutions are not clear [4], and people should find out how to deal with those problems through discussion and collaboration. To survive in such a society characterized by a need for people to collaboratively solve problems, students should acquire a new skill to take other people's perspectives – even criticisms about their personally held views – into their own perspectives to enhance the quality of their arguments.

Although it is important for students to have the skill to integrate multiple perspectives, some evidence coming from the research area of argumentations suggests that students do not sufficiently possess such skills. For example, many studies that have focused on students' argument skills have revealed that students tend to show a "my-side bias" [5], which means that students tend to collect evidence supporting their own ideas, and they also tend to neglect other people's perspectives and criticisms [e.g., 6, 7]. This suggests that students do not sufficiently acquire competence in understanding different perspectives, integrating such perspectives, and proposing a new idea that overcomes the limitations of their previous ideas.

When considering instructional methods for improving the skill of integrating multiple perspectives, the use of diagrams has been demonstrated to be an effective strategy [e.g., 8, 9]. It can be expected that students should use such a strategy by themselves even without any instruction from teachers or experimenters. If students do not use diagrams by themselves, then diagrams will never contribute to solving problems in real life settings, which is strongly expected in the context of 21st century skills development. However, as described later in more detail, previous proposed instructions for enhancing students' quality of argumentation including multi-perspective integration do not sufficiently contribute to the promotion of the spontaneous use of diagrams by students themselves. Thus this study developed an instructional method for promoting students' spontaneous use of diagrams, and for improving multi-perspective integration.

1.2 How to Promote Multi-perspective Integration: The Effect of Diagrams

This section expands on the idea of multi-perspective integration in this paper. When deliberating on a topic that has both arguments and counter-arguments, understanding several perspectives and then creating a new idea integrating those perspectives is a very important skill, and this is what multi-perspective integration means in the current paper. For example, when people discuss to decide whether infant push chairs should be folded up when they ride on trains, which is a real topic discussed in Japan, several ideas exist. One idea is that infant push chairs should be folded up as these are so space consuming and can be troublesome especially during rush hours. Another idea however is that infant push chairs can ride on trains without being folded up in order to meet the needs of mothers who usually bring very heavy things relating to their babies (which their infant push chairs can help them carry). Although traditional argumentation skills often refer to the importance of supporting ideas to make arguments convincing to people, a more important skill is understanding different ideas and integrating those to make an argument that could satisfy people from both sides of the argument. Examples of new ideas in this case would be that infant push chairs should be folded only during rush hours, and that train companies should prepare two types of train cars: one in which people do not have to fold up infant push chairs all the time, and one in which people have to fold up infant push chairs – thus, people can select which type of train car to use according to their preferences.

Although conducting an argumentation integrating multiple perspectives is an important skill for living successfully in a society in which collaboration is necessary, this type of argumentation has not been dealt with well in the context of argumentation research. This is because most of the argumentation researchers have focused on the argumentation which is based on Toulmins' model [10]. Toulmin's model is a kind of schema, in which six elements are specified that are necessary to make appropriate claims: claims, data, warrant, backing, qualifier, and rebuttal. This schema is effective when people construct an argument to support their own claim. However, the question of how to integrate other perspectives and criticisms from other people is not included. School instructional practices have also neglected focusing on multi-perspective integration as a skill. For example, debate is the most common way used in schools to enhance skills of argumentation. This technique is also one in which people keep one position and consider how to validate their claim. Therefore, it does not include the perspective that students should integrate arguments for and against ideas and should propose a new idea in which the limitations of both sides have been resolved.

Although a very limited number of studies has targeted integrating multiple perspectives, exceptional studies have been conducted by Nussbaum and his research group [e.g., 8, 9, 11]. Nussbaum [11] emphasized the importance of argument-counterargument integration and classified three types of strategies used for integration: a weighing strategy, a synthesis strategy, and a refutation strategy. The second strategy, creative synthesis, is congruent with the concept of multi-perspective integration proposed in this paper. Nussbaum [8] noted that, "The use of creative synthesis as an argumentation move is neglected in most current argumentation models" (p. 551). Although Nussbaum and his research group [8, 9] focused on all three strategies, this paper focuses only on synthesis strategy (here referred to as multi-perspective integration).

Another important suggestion from the Nussbaum series of studies is the efficacy of diagrams for multi-perspective integration. Nussbaum [8] proposed the Argumentation Vee Diagrams (AVDs) which are shaped like the letter V, and showed that they were effective for integrating perspectives in written argumentation. Another study, conducted by Nussbaum and Edwards [9], also showed that the effects of this type of diagram were enhanced when the critical questions which were induced by Walton's [12] dialogue theory were included in the instruction. Those series of studies from Nussbaum's group provide strong evidence demonstrating the efficacy of diagram use in integrating multiple perspectives.

1.3 Tables as Thinking Tools for Multi-Perspective Integration

Although this paper does not intend at all to deny the effects of the studies conducted by Nussbaum' research group, several new perspectives may be able to add to the findings of those studies. The first of such perspectives is that students should develop their skills of using diagrams as thinking tools instead of viewing diagrams as instructional tools of teachers. In more concrete terms, in both studies conducted by Nussbaum [8] and Nussbaum and Edwards [9], students used those diagrams in the classroom after finishing the discussion with peers, and their performance in argumentation was not sufficiently high when neither the AVDs nor classroom discussion were provided by a teacher. This means that participants in their study did not sufficiently acquire the skills of using diagrams as their thinking tools and used diagrams more as teacher's instructional tools. Thus a teacher should develop the attitude and knowledge of students for spontaneously using diagrams as their own thinking tools even when a teacher is not present to encourage their use.

The second perspective is the effectiveness of "ubiquitous diagrams". Using diagrams that many people are familiar with because they are used in many situations in the culture are particularly valuable because people find them easy to understand and use as thinking tools. We call this kind of diagrams "ubiquitous diagrams", and Cartesian graphs, flow charts, and tables are included in this category. When viewed from this perspective, although AVDs are good for considering argument-counterargument integration, an AVD is a diagram which has a very specific function and not all people possess the background knowledge to be able to proficiently use them.

One proposal to take into account this view is to use tables instead of AVDs for promoting the construction of an argument integrating multiple perspectives. Tables are typical of "ubiquitous diagrams" as these are used in many situations and many people share the background knowledge for understanding and constructing them. For example, textbooks in schools and newspapers includes many tables. Tables would also be effective when people are deciding their own position about a topic that has both arguments and counter-arguments. The mechanism why tables promote the integration of multi-perspectives would be as follows: Tables can summarize the advantages and disadvantages of both sides, and at a glance can provide an understanding of the range of ideas and can therefore contribute to developing the user's own ideas. For example, when considering whether infant push chairs should be folded up in trains or not, a student can construct a table, which shows what kinds of good points and criticisms exist on each side. Students would be able to develop more easily a new idea, and they would be able to resolve criticisms by using tables and understanding the whole range of ideas about the topic. The shape of tables is a bit different from AVDs but the information included in the diagrams are not so different, and tables are more familiar than AVDs for many students, so tables may also be effective in promoting multi-perspective integration. And tables are more effective when encouraging students to use diagrams spontaneously.

This idea that using tables in instruction to promote multi-perspective integration is new from the view point of diagrams research. Although the efficacy of tables has been shown in many studies, tables have not been applied to develop multi-perspective integration. In addition, no examination had been undertaken of whether instruction for school students to promote the spontaneous use of tables improve multi-perspective integration. Thus, in this study, a new instructional method was developed to promote multi-perspective integration among students by enhancing their use of tables for organizing information.

Another perspective that the current study can contribute to this research area is that the students themselves should develop awareness about the efficacy of tables as a thinking tool for constructing multi-perspective integration. Uesaka et al. [14] demonstrated with an experiment that if students did not have explicit knowledge about the efficacy of diagram use, they will not use diagrams spontaneously in novel situations. This finding suggests that, in order to promote the spontaneous use of diagrams, not only using effective diagrams but also providing instruction or tasks to make students understand the value of diagrams is important.

To promote awareness about the efficacy of tables as a thinking tool for constructing an argument among students, the "thinking-after-instruction" approach in class [15] would be effective. This approach consists of four phases: teacher's instruction, checking comprehension, deepening understanding, and self-evaluation. This design principle has been proposed based on cognitive psychology, and it emphasizes students' conscious awareness about the meaning and points of what they learn. The approach is more actively used in learning with conceptual understanding, but there are some examples in school practices of its use in the area of skills learning. In the case of skills learning, the four phases above are conducted as follows. Firstly, during the teacher's instruction phase, a teacher instructs the class about the points that students should pay attention to, using good examples. Secondly, during the checking comprehension phase, students are provided some tasks to check whether they sufficiently comprehend the teacher's instruction. For example, some badly performed examples are provided and students are asked to identify with their peers what points are not appropriate. Thirdly, during the deepening understanding phase, students are provided other higher level tasks to deepen their understanding such as conducting a kind of argumentation with a topic. Finally, during the final self-evaluation phase, students are asked to verbally articulate and write down what they have understood in this class and what they have not yet understood. In skills learning, even if students cannot perform a skill well in class, recognizing which points they should pay attention to in the next learning opportunity would enhance their performance. The current study used this design when developing a new instructional method to improve students' multi-perspective integration.

1.4 Overview and Hypothesis of the Current Studies

This paper focused on developing a new instructional method to improve students' multi-perspective integration skill by enhancing the use of tables for organizing information. To examine the effects of the proposed instructional method, a 5-day special summer course for 8th-grade students was conducted in a university. Pre- and post-instructional assessments were administered on the first and the last days. In the assessment sessions, participants were provided a topic and asked to present their ideas, which were video recorded for analysis. In addition, other tasks were administrated to evaluate how much students were aware about the points for conducting argumentations, in which awareness about the efficacy of diagram use was also measured. Between assessment sessions, instructional sessions were provided and participants were assigned to one of three conditions. The first group of participants were provided two additional treatments, in which they were encouraged to use tables for organizing information and also provided exercises that required identification of problems in videos showing bad examples of argumentation. The second group of participants received one additional treatment of encouragement in using tables for organizing information. The third group of participants did not receive any additional treatment. The exercise to identify problems in a bad example of argumentation, which was only included in the first group, was conducted as the "checking comprehension" phase and the other groups did not have this phase included. It means that only the first group received the full set of "thinking-after-instruction" approach in class.

In this study, it was hypothesized that if students received an instruction including encouragement to use tables with the explanation of the efficacy of using tables, they would conduct multi-perspective integration better than students who did not receive such instruction. In addition, using the full set of "thinking-after-instruction" approach in class was hypothesized to contribute toward students' better awareness of the efficacy of diagram use, compared to students who did not receive the full set of "thinking-after-instruction" approach in class. Based on these hypotheses, the following two predictions were made. Firstly, in the two conditions where participants received the additional treatment of being encouraged to use tables, participants would use tables more in assessments and, as a result of it, they would perform better in conducting multi-perspective integration. Secondly, in the one condition where participants also received an exercise to identify problems in a bad argumentation example, participants would evidence better recognition of the efficacy of diagram use compared to participants in the other conditions. To examine those hypotheses, contrast analysis was used instead of general multiple comparison ANOVA, as the clear hypotheses that one group would be better than the other groups existed. If these predictions were confirmed with the data gathered, these hypotheses would be considered validated.

2 Method

2.1 Participants and Experimental Design

Fifty-six 8th-grade students in junior high schools from two wards in Japan participated voluntarily in this study. Participants were randomly assigned to one of three conditions: a condition with additional instruction about the use of tables and identifying problems exercise (we will call this group the Table Instruction (TI) + Identification Exercise (IE) condition), a condition with additional instruction about the use of tables (we call this group the Table Instruction (TI) condition), and a condition without any additional treatment (we call this this group the Control condition).

Students who did not participate in the whole course were excluded from the analyses. As a result, the total number of the participants included in the analyses was 40 (the number of the participants in the TI+IE condition was 13, the number of the participants in the TI condition was 13, and in the Control condition it was 14).

2.2 Materials

Materials Used in Pre- and Post-instructional Sessions. In this study, several tasks were administered in two assessment sessions (pre-instructional assessment session and post-instructional assessment session). Among several tasks administered during these session, this paper focuses on two tasks for the purposes of the study it describes. The tasks reported in this paper are a performing argumentation task and a general beliefs task. The performing argumentation task was conducted to assess students' actual skills in argumentation. Three controversial topics that were all familiar to students were prepared (e.g., whether schools should permit students to bring their mobile phone to school or not). Each participant was provided one of three topics and expected to present his/her own idea after 10 minutes of preparation. Their performances were video recorded for analysis. A supplemental document was prepared for each topic that included typical examples of ideas on both sides of the argument and several diagrams relating to the topic. When conducting this task, sheets of paper with a board that can be used for note-taking or for showing materials to an audience were also provided to each participant. In addition, the general beliefs task was administered to evaluate participants' ideas about important points when conducting argumentation. A4 sheets of paper were provided for participants to write down what they considered important with the use of bullet points.

Materials Used in Classroom Instructional Sessions. In the classroom instructional sessions, four controversial topics were prepared (e.g., whether high schools should permit students to hold part time jobs or not). For each topic, a supplementary document was also provided. Sheets of A4-size paper and a board were also provided to participants and they could use those as they wanted.

2.3 Procedure

Questionnaires Conducted by Mail and Assignment of Participants to Conditions. The experimental course was conducted as a special summer program in the University of Tokyo. About one month before starting the course, participants and their parents were sent mail-based questionnaires that included items asking about their academic achievement in five main subjects according to five achievement levels (1 was the lowest and 5 was the highest). Sending back the mail-based questionnaires was an obligation attached to participating in this course. Based on the information received, students were randomly assigned to one of three conditions after balancing their academic achievements. In more concrete terms, participants were ranked by using their total score (calculated by adding up their achievement levels in the five main subjects), and each 3 participants from the top were randomly assigned to one of the three conditions. When analyzing data statistically, data derived from this kind of assignment should be analyzed as paired data. However, some students failed to attend all 5 days of the course and were deleted from the subsequent analyses. If the other data corresponding to missing data were deleted, the number of data that could be analyzed would have been reduced to a non-viable level for statistical analysis. Thus, all participants who attended all days of the course were included, and data were treated as non-paired data in the analyses that were undertaken.

The First Day of the Experimental Class (Pre-instructional Assessment Session). Each condition's participants joined a class together. Each day's class lasted 50 min. On the first day, the main purpose of the day was conducting pre-instructional assessment. After an instructor briefly explained the purpose of this special course and conducted some activities to develop rapport between students, several tasks were conducted as pre-instructional assessment that took about 30 minutes. In administering the argumentation task, participants were divided into small groups that comprised three students in each group, and each participant was provided one of three topics and expected to present their own idea in front of the other two after 10 minutes of preparation. It was announced that performance would be video recorded. When preparing their presentation, a supplemental document was provided to participants. If participants perceived it as necessary, they could use the supplementary document as they wished during both preparation and presentation. They were also provided with A4 sheets of paper and an A4-size board with a clip that can be used as a flip board during a presentation. They were permitted to use the board and paper when preparing and presenting their ideas. The general beliefs task was administered after performing the argumentation task to evaluate participants' ideas about important points when conducting argumentation. Participants were asked to write down points about what they considered important when presenting their ideas on a topic that had both arguments for and against it. They were encouraged to write down as much as possible using bullet points. Three minutes were provided to answer this question.

The Second Day (Instructional Assessment Session). From the second day to the fourth day, the instructional sessions were provided. The second and third days mainly targeted argumentation skills, and the fourth day targeted presentation skills. Differences in manipulations existed mainly on the second and third days.

The second day's topic was thinking skills for argumentation. Participants in all conditions were instructed the following two main points with activities: Firstly, they were encouraged to summarize the advantages and disadvantages of different perspectives before deciding their attitude about a topic, and secondly they are encouraged to integrate perspectives so as to overcome limitations on both sides of the argument. When teaching the main message of the class, a teacher demonstrated how to use those points concretely with an example, which was about whether infant push chairs should be folded or not on trains. After the teacher provided instruction about those points, a new topic about whether high schools should permit students to hold part-time jobs or not was provided and all participants were asked to construct an argument by using the points taught in collaboration with their peers. At the end of the class, participants were encouraged to self-evaluate about what they had understood and what they had failed to understand.

The differences in manipulations between the conditions on the second day were as follows. In the Control group, as a way of summarizing the advantages and disadvantages of different perspectives, writing down as much as possible using bullet points was taught. In contrast, in the TI+IE and the TI conditions, participants were instructed to use a table for summarizing the advantages and disadvantages of different perspectives. In addition, when undertaking the exercise on the topic of high school permission for part time jobs, the TI+IE condition and the TI condition participants were provided the frame of a table to fill in the information in their work sheet. In the Control group participants were not provided such a frame, and the same size but blank sheet of paper that participants could use freely was instead given.

Participants in the TI+IE condition were additionally provided an identifying problems exercise. For this exercise, a teacher made video clips from about 30 seconds to 1 minutes, in which a teacher conducted argumentation badly without paying attention to points the students had received instruction about in the class.

The Third Day (Instructional Assessment Session). On the third day, the topic was the structure of argumentation to convince audiences. On this day, it was emphasized to all participants in all conditions that people should provide the reason why their idea could be considered as being better than other ideas. This is because students tend to just summarize the advantages and disadvantages of both sides and conclude without explaining why they reached a particular conclusion. How participants could add the reason to their arguments was taught with a concrete example. In addition, participants were taught about the importance of showing the information source if they referred to other sources like the supplementary document. After those points had been covered, the students engaged in activities in which they prepared argumentation for a topic that they would present on the fourth day. A rubric sheet for presentation evaluation which would be used on the fourth day was provided. At the end of the class, an opportunity for self-evaluation was also provided.

The differences of the manipulations on that day among the conditions were as follows. Firstly, in the TI+IE condition and the TI condition, participants were instructed that they should utilize a table when explaining the reason for the goodness and the background of their idea. In addition, when preparing their argumentation, participants were verbally encouraged by the teacher to use a table. The TI+IE condition participants were additionally provided the identifying problems exercise. Video clips that a teacher performed were provided and participants were expected to point out problems. There were argumentations in the videos in which a teacher conducted an argumentation without explaining the reasons of a proposal or the pertinent background, and in which a teacher conducted an argumentation without referring to the sources that were used.

The Fourth Day (Instructional Assessment Session). The fourth day's topic was that of presentation. Participants in all conditions were instructed the following four important points: using diagrams, speaking with a loud voice, maintaining eye contact, and speaking without reading a draft. After the instruction, participants practiced to present with peers in small groups, during which participants presented for about 30 seconds in front of an iPad (for recording) and checked immediately with peers whether the instructed points were realized in their presentation. Participants, afterwards, presented all their ideas in front of other students. During the presentation, a peer evaluated by using a sheet of paper containing rubrics and another peer recorded their performance with an iPad. The difference in the manipulation was limited, but in the TI+IE condition and in IT condition, participants were instructed that a table for organizing information to decide their views was also usable when presenting their ideas to others. Other manipulations were all the same among the three conditions on that day.

The Fifth Day of the Experimental Class (Post-instructional Assessment Session). On the fifth day, post-instructional assessment was conducted. Several tasks including performing the argumentation task and the general beliefs task were administered. The procedures used were the same as during the pre-instructional assessment.

3 Results

3.1 Analysis of Performing Argumentation Task

Analysis of Participants' Argumentation. Before starting the analysis, participants' responses in the performing argumentation task were coded. Students' arguments were transcribed, and coded according to whether they conducted a multi-perspective integration argument or not. Examples of participants' transcripts coded as "multi-perspective integration" and "not multi-perspective integration" are shown in Table 1, which contains translations from Japanese to English. The average inter-rater agreement was found to be 88.8%, which the present authors deemed as satisfactory.

The analysis was conducted with the following two steps as the present study hypothesized that a table instruction would enhance participants' multi-perspective

Table 1. Examples of participants' argumentation

Example of argumentation coded as "multi-perspective integration"	Well, I thought about whether schools should permit students to bring their mobile or not. My opinion, my opinion is that schools should permit students to bring their mobile to the school. An advantage of giving permission is that students can make contact with their parents in case of an emergency, which is A's idea[a]. On the other hand, a disadvantage of giving permission is that students may use their mobile during class. An advantage of not giving permission is that students don't use their mobile during class, which is related to E's idea[a]. A disadvantage of not giving permission is that students cannot make contact with their parents in case of an emergency and they may end up at a loss. Please look at this table. In a survey, 70 % of the participants aged in their 20 s to 40 s agreed with the idea of giving permission. I agree with the idea of conditional permission. The condition is…during class…the condition I would like to propose is that a school prohibits students from using their mobile except for during recess. Also, they can bring their mobile to the school only if their parents permit them to do so. If a student's mobile is stolen, his/her parents will be responsible for that so trouble between children will not be caused, I think. That's why I referred to parents' permission. I agree with the idea of giving permission. That's all.
Example of argumentation coded as "not multi-perspective integration"	I'd like to talk about whether schools should permit students to bring their mobile or not. It is the best advantage of giving permission for students to use their mobile as a supplementary study tool such as a dictionary and a calculator. A disadvantage is…that students may break rules when they use their mobile during class. Well, according to a survey, using their mobile during class, they text messages or do other things. An advantage of prohibiting students from bringing their mobile to the school is that their mobile will not be lost or stolen in the school. A disadvantage of prohibition is…that students cannot confirm the safety of themselves in case of an emergency. Therefore, I think it is OK for students to bring their mobile to the school if they use it as a dictionary or such a useful tool. However, basically, I'm against the idea of giving permission, because probably…students can solve problems such as mathematical problems without solving them on their own by using their mobile as a calculator.

[a] "A's idea" and "E's idea" are included in a provided supplemental document.

Table 2. Ratios of participants conducting multi-perspective integration (SDs)

Condition	Number of participants	Pre-instructional assessment (SDs)	Post-instructional assessment (SDs)
The TI+IE condition	13	0.31 (0.48)	0.62 (0.51)
The TI condition	13	0.39 (0.51)	0.77 (0.44)
The control condition	14	0.07 (0.27)	0.36 (0.50)

integration. Firstly, the percentages of conducting multi-perspective integration in each condition were compared by contrasting two groups. The first group consisted of participants from the two conditions that received instruction about the use of tables (the TI +IE and the TI conditions), and the second group consisted of participants in the one condition that did not receive such instruction (the Control condition). Secondly, the difference between participants in the two groups that received instruction about the use of tables (the TI+IE and the TI groups) was examined.

Based on the idea described above, the percentages of conducting multi-perspective integration in the post-instructional analysis were compared with a contrast analysis between the groups that received table instruction (TI+IE and TI) and the group that did not (Control). The results revealed that the effect of the instruction to encourage using a table was statistically significant ($t_{(38)} = 2.09$, $p = .04$), and participants in the first group evidenced a higher percentage of conducting multi-perspective integration than those in the second group (see Table 2). On the other hand, the difference of percentages of conducting multi-perspective integration between the TI+IE condition and the TI condition was not statistically different ($t_{(38)} = 0.81$, $p = .42$). To confirm the equivalence among conditions at the start of the study, the same analysis was conducted on the pre-instructional assessment scores. The results revealed no significant effects.

Analysis of Participants' Table Use. Participants' responses on a board which was given to them during their performance of the argumentation task were coded in terms of whether a table was used or not. In other words, their responses were coded as argued "with a table" or "without a table". An example is shown in Fig. 1 (left side) that was coded as arguing "with a table". Another example is also shown in Fig. 1 (right side) that was coded as arguing "without a table". The average inter-rater agreement was found to be 95%, which the present authors deemed as satisfactory.

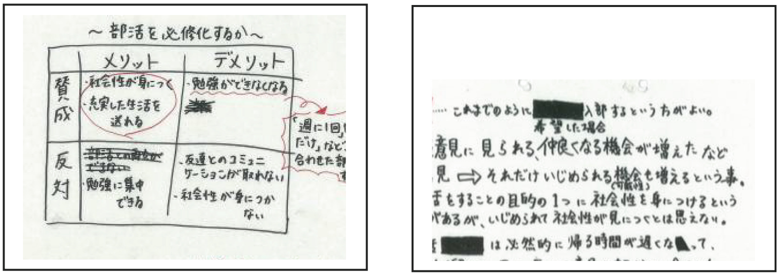


Fig. 1. Examples of participants' responses on a board (On the left side is an example of "with a table", and on the right side is of "without a table")

Table 3. Ratios of participants using a table in assessment

Condition	Number of participants	Pre-instructional assessment(SDs)	Post-instructional assessment(SDs)
The TI+IE condition	13	0.00 (0)	0.85 (0.38)
The TI condition	13	0.08 (0.28)	0.69 (0.48)
The control condition	14	0.00 (0)	0.00 (0)

The percentages of using a table for argumentation in the post-instructional analysis were analyzed using contrast analysis. Conditions with instruction encouraging table use (the TI+IE and the TI conditions) and the condition not provided such instruction (the Control condition) were compared. The result revealed a statistically significant difference ($t_{(38)} = 6.68$, $p < .01$). Participants in the two conditions with table encouragement had higher percentages of table use than participants in the condition without such instruction (see Table 3). In contrast, the table use percentages of participants in the two conditions provided table use encouragement instruction were not statistically different ($t_{(38)} = 1.13$, $p = .27$). To confirm the equivalence among conditions at the start of the study, the same analysis was conducted on the data gathered during pre-instructional assessment. There was no statistically significant difference found between the conditions.

3.2 Analysis of General Beliefs Task

Participants' responses to this task were also coded. The average inter-rater agreement was found to be 97.5%, which the present authors deemed as satisfactory. The percentage of participants who referred to the efficacy of diagrams were analyzed using contrast analysis. The analysis was a bit different from the analysis of participants' argumentation and a table use, as the hypothesis was that the participants who belonged to the condition with table use encouragement instruction and identification problems exercise would articulate more their beliefs about such efficacy than participants in the other two conditions. Thus firstly, the TI+IE condition was contrasted with the other two groups, and the difference between the other two groups was examined later. The percentages are summarized in Table 4.

Based on the idea described above, the percentages of participants writing down the efficacy of diagram use in the general beliefs task conducted in the post assessment was counted and compared between the TI+IE condition and the other two conditions (the TI condition and the Control condition). The result showed that the difference was statistically significant ($t_{(38)} = 2.35$, $p = .02$). In contrast, the difference between the TI condition and the Control condition was not statistically significant ($t_{(38)} = 0.52$, $p = .61$).

Table 4. Ratios of participants declaring the efficacy of diagrams in assessment

Condition	Number of participants	Pre-instructional assessment(SDs)	Post-instructional assessment(SDs)
The TI+IE condition	13	0.00 (0)	0.54 (0.52)
The TI condition	13	0.00 (0)	0.23 (0.44)
The control condition	14	0.07 (0.27)	0.14 (0.36)

To confirm the equivalence among the conditions, the same analyses were applied to the data of the general beliefs task conducted during the pre-instructional assessment. No significant differences were found among the three conditions.

4 Discussion

4.1 The Effects of Manipulations on Students' Performance and Beliefs

The key findings in the present study were the following two, which both supported the predictions induced from the hypotheses. The first key finding was that students in the two conditions with table encouragement instruction conducted multi-perspective integration more than students in the Control condition. More frequent use of tables in the two conditions with table instruction, compared to the Control condition, was also confirmed. It suggests that participants who received the instruction encouraging table use utilized more tables spontaneously in the assessment session even without receiving instruction to utilize tables during that assessment session. Those participants also conducted multi-perspective integration more than participants in the Control group who did not received the instruction encouraging them to use tables.

The second key finding was that students in a condition provided the exercise in identifying problems recognized more the efficacy of diagram use compared to the other two conditions without this exercise. It means that the performance of conducting multi-perspective integration by the "table instruction plus identifying problem exercise" condition and the "only table instruction" condition was at the same level, but the beliefs of participants in those conditions were different. The participants in the condition with table instruction *and* identifying problem exercise used diagrams with clearer awareness about the effects of diagrams. Although this study did not examine it sufficiently, awareness about the effect of diagrams has a possibility of contributing to further utilization of diagrams in other situations not covered in the class. To clarify the effect of the participants' awareness about the efficacy of diagrams, it is necessary to conduct follow-up studies, which should include further transfer tasks in addition to those that were used in this study.

4.2 Implications for Diagrams Research Areas

The paper includes the following two new perspectives that had not been examined sufficiently in previous research. One is that of cultivating students' attitudes for using diagrams as their thinking tools. Previous research such as studies conducted by Nussbaum and Edwards [9] used diagrams as instructional tools of the teacher and did not sufficiently explore students' attitude about the use of diagrams in situations where teachers do not encourage their use. Not only the research groups of Nussbaum, but many other studies conducted in the diagram research area have examined the effect of diagrams by instructing participants to use diagrams. As a result, those studies have not examined sufficiently students' competencies in using diagrams as thinking tools without instruction to use diagrams from an instructor or a teacher. However, when considering the competencies necessary to survive in the 21st Century, people's attitude to use diagrams by themselves spontaneously even without instruction to use them would be desirable. Those aspects had been examined in previous research concerning diagram use [e.g., 13, 16], but were examined with materials about math word problem solving. For example, Uesaka et al. [13] examined factors necessary to promote spontaneous use of diagrams in mathematical problem solving with experimental classes, and revealed that enhancing perceptions about the efficacy of diagram use and improving skills in drawing diagrams are important in promoting spontaneity in use. Those factors were also embedded in this study. This paper explored these comparatively new perspectives in diagram research focusing on argumentation.

The second perspective is “ubiquitous diagrams” and transferability of diagram use. In the area of diagrams research, many specific diagrams have been proposed and those are very powerful in specific situations. The AVD diagrams proposed by Nussbaum were also this type of diagrams. Another dimension that should be included as views in diagram research might be utilizing “ubiquitous diagrams”. Usually a certain level of cognitive cost is necessary to master a type of diagram up to the level that it would promote problem solving. So specific diagrams entail costs at the starting period of using those diagrams. In contrast, some types of diagrams which are often used in many situations (“ubiquitous diagrams”) do not entail such acquisition costs at the starting point. Also, these types of diagrams are effective thinking and communication tools as background knowledge to read/comprehend and construct such diagrams is already shared among many people. The table is one type of such diagrams. Ubiquitous diagrams including other types of diagrams such as flow charts and Cartesian graphs have high transferability, which means they are easier to use in other situations. Although the value of specific diagrams cannot be denied, generality and transferability of diagrams is a perspective that had not been emphasized adequately in previous diagrams research.

4.3 Implications for Education and Directions for Future Research

The competence to conduct multi-perspective integration can become a new goal in argumentation education. As described in the introduction, the most popular strategy to enhance argumentation skills is debating. However, students cannot acquire sufficiently the attitude and the skills of integrating perspectives that contain both arguments and counter-arguments even after they have experienced debating. For developing the skills useful for effectively solving problems in the 21st Century, new types of education for the development of students’ competence in conducting multi-perspective integration should be included. This study also demonstrated the efficacy of the “thinking-after-instruction approach” in class to achieve this goal. Provision of education to promote development of the necessary attitudes and skills for multi-perspective integration would be useful in almost all educational contexts.

One of the necessary perspectives that can be integrated in future research is to promote critical thinking skills among students so that they can effectively evaluate the validity of data and integrate those into their argumentations. In this study, a supplemental document was provided and participants could use it as they wanted. Data was included in supplementary files and some of the data were not robust. For example, data that came from a limited number of participants or from a specific group. However, none of the participant referred to limitations regarding those points and this fact suggests a weakness in the participants' critical thinking. Although the course described in this study did not cover the competence of critical thinking skills, one of the future directions is covering such aspects of competence.

Acknowledgments. This work was supported by JSPS KAKENHI Grant Numbers 15H01976 and 15K13126. The authors would like to thank Emmanuel Manalo of Kyoto University for supporting and advancing this research and paper.

References

1. Hilton, M.: Exploring the Intersection of Science Education and 21st Century Skills: A Workshop Summary. The National Academies Press, Washington, D.C. (2010)
2. An Action Canada Task Force Report: Future Tense: Adapting Canadian Education Systems for the 21st Century (2013)
3. Rychen, D.S., Salganik, L.H.: Key Competencies for a Successful Life and a Well-Functioning Society. Hogrefe & Huber, Göttingen (2003)
4. Taguma, M.: Redefining key competencies for the 21st century: new perspective for 2030 in OECD. Paper Presented in a Symposium “Constructing Educational Models for the 21st Century” at the University of Toyo, 5 December 2015. (in Japanese)
5. Perkins, D.N.: Postprimary education has little impact on informal reasoning. J. Educ. Psychol. **77**, 562–571 (1985)
6. Wolfe, C.R.: Individual difference in the “myside bias” in reasoning and written argumentation. Written Commun. **29**, 477–501 (2012)
7. Baron, J.: Myside bias in thinking about abortion. Think. Reason. **1**, 221–235 (1995)
8. Nussbaum, E.M.: Using argumentation Vee Diagrams (AVDs) for promoting argument-counterargument integration in reflective writing. J. Educ. Psychol. **100**, 549–565 (2008)
9. Nussbaum, E.M., Edwards, O.V.: Critical questions and argumentation stratagems: a framework for enhancing and analyzing students' reasoning practices. J. Learn. Sci. **20**, 443–488 (2011)
10. Toulmin, S.: The Uses of Argument. Cambridge University Press, New York (1958)
11. Nussbaum, E.M., Schraw, G.: Promoting argument-counterargument integration in students' writing. J. Exp. Psychol. **76**, 59–92 (2007)
12. Walton, D., Reed, C., Macagno, F.: Argumentation Schemes. Cambridge, New York (2008)
13. Uesaka, Y., Manalo, E.: How communicative learning situations influence students' use of diagrams: focusing on spontaneous diagram construction and protocols during explanation. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 93–107. Springer, Heidelberg (2014)
14. Uesaka, Y., Manalo, E., Ichikawa, S.: The effects of perception of efficacy and diagram construction skills on students' spontaneous use of diagrams when solving math word problems. In: Goel, A.K., Jamnik, M., Narayanan, N. (eds.) Diagrams 2010. LNCS, vol. 6170, pp. 197–211. Springer, Heidelberg (2010)
15. Ichikawa, S.: Manabu Iyoku to Sukiru wo Sodateru [Promoting Students' Motivation and Learning Skills] Shougakukan, Tokyo (2004)
16. Uesaka, Y., Manalo, E., Ichikawa, S.: What kinds of perceptions and daily learning behaviors promote students' use of diagrams in mathematics problem solving. Learn. Instr. **17**, 322–335 (2007)

Using Diagrammatic Drawings to Understand Fictional Spaces: Exploring the Buendía House in Gabriel García Márquez's *One Hundred Years of Solitude*

Sarah Pérez-Kriz and Ricardo F. Vivancos-Pérez(✉)

George Mason University, Fairfax, VA, USA
{skriz, rvivanco}@gmu.edu

Abstract. Fictional spaces described in literary texts are unique in that their spatial information is always underspecified, leaving readers with the task of filling in the unspecified details. While ways in which readers fill in unspecified elements have been proposed, very little empirical work has been done to examine the process. This paper presents an empirical study of how readers fill in unspecified details of fictional spaces described in literary texts. We asked readers to draw diagrams of the Buendía family home in the novel *One Hundred Years of Solitude* by Gabriel García Márquez. Diagrams were analyzed to inventory the spatial elements that were depicted across several readers. The results indicate that readers fill in the spatial details of fictional spaces using their own culturally-specific understanding of similar real world spaces, and that narrative events also assign prominence and detail to certain areas within a fictional space. Using diagrams to understand fictional spaces is discussed from a pedagogical perspective.

Keywords: Fictional spaces · Latin American literature · Cultural schemas · Spatial memory

1 Introduction

Fictional spaces depicted in literature are, by default, experienced differently than real world spaces. Fictional spaces are quite unique; a reader has access to nothing more than linguistic (written) information about the space, whereas real world spaces are commonly experienced phenomenologically, offering visuospatial information and proprioceptive feedback. Perhaps more importantly, fictional spaces are always underspecified [1, 2]. This means that only parts of a fictional space are described in a literary text, offering the reader a small number of specific spatial details of the space. Readers are left to fill in the other spatial details of a fictional space, as necessary.

How do readers fill in spatial details about fictional spaces when they are not provided in the literary text? Busselle and Bilandzic [1] propose that readers rely on schemas from previous experience to create a model of what they call the "story world." More specifically, McVee et al. [3] argue that readers apply schemas that are embodied, transactional, and culturally informed to make interpretations about what

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 187–193, 2016.
DOI: 10.1007/978-3-319-42333-3_14

they read. In the case that readers do not have a culturally appropriate schema that matches the sociocultural context of the literary text, they may apply the schemas they do have available, which are likely based on their previous real world experiences. This means that there may be quite a mismatch between the world depicted in a literary text and the knowledge structures that readers apply when they read and reason about the fictional world.

Furthermore, research on narrative comprehension suggests that readers mentally index narrative events based on several dimensions, including the time and space in which they occur [4]. This means that in addition to applying cultural schemas, readers may utilize narrative events to help them create details of the fictional space.

Methodologically, it is somewhat difficult to empirically study fictional spaces because one must identify a way to measure how and when readers fill in spatial details of the fictional space. Complicating matters further is the fact that everyday readers do not have a reason to construct a complete understanding of a fictional space, and so filling in spatial details that were not specified in the text is not a mandatory activity for most readers. Even students, who are required to do a much more in-depth reading than average "pleasure" readers, are rarely asked to think about the unspecified elements of a text's fictional world.

We propose a methodology for empirically studying fictional spaces—asking readers to draw diagrams of the spatial layout. Drawing a diagram that depicts the layout of a fictional space requires a reader to commit to filling in the unspecified spatial details of the space. We can then evaluate the diagrammatic drawings to learn more about the details that readers choose to use. By identifying trends across many readers' diagrams, we can infer that common details are due to readers' similar cultural schemas and shared sociocultural experiences.

Below we present a study that asked students to diagram a fictional house depicted in a popular Latin American novel. We chose Gabriel García Márquez's acclaimed novel *One Hundred Years of Solitude* [5] for two reasons. First, because the novel is set in a time and place that North American readers have not experienced personally, it should be quite evident if readers rely on their modern-day experiences living in North America to fill in the details of the fictional space. (Although narrative time and setting are not exactly specified, it is clear the novel takes place in Latin America during the mid 19th and early 20th centuries in the fictional village of Macondo.) Second, García Márquez never allowed the novel to be made into a film. Therefore, there are no widespread visual interpretations of the key fictional spaces in the text.

2 Method

2.1 Participants

American students in an undergraduate general education literature class on global magical realism at a public U.S. university were invited to participate in the study; participation was voluntary and was not a requirement of the course. Thirteen students agreed to participate in the study. The students ranged in class status from freshmen to seniors. All reported reading at least three-fourths of the novel. None of the participants had lived in Latin America. The class had read and discussed the English translation of *One Hundred Years of Solitude* [5] for three weeks prior to the study.

2.2 Procedure

The students who agreed to participate in the study were given a packet of materials. On the top of the first page, the students were instructed: “Please draw the layout of the Buendía family house when the banana company arrives in Macondo (Chap. 12). Please label your drawing.” On the top of the second page, the prompt read: “Please draw the layout of Macondo after the banana plantation is established. (Chap. 12). Please label your drawing.” The third page was a questionnaire that asked whether they finished the book and how much they had read, whether they had lived in Latin America, and ratings of their skills as a reader. Participants were given 15 min to complete the study. They did not have access to the novel during the study.

2.3 Analyses

An initial analysis of the drawings indicated that most participants misunderstood the prompt that asked them to draw the town of Macondo. Out of the 13 participants, three left this page blank and four drew a layout of the house rather than the town. Because the prompt was confusing to many participants, we excluded this question from our analysis and focused only on readers’ diagrams of the house.

Two analyses of the Buendía house layout drawings were conducted. First, in order to examine the spatial accuracy of students’ memory-based drawings, we determined whether readers correctly represented the spatial relations that *were* described in the novel by comparing the diagrams to three excerpts from the text that explicitly stated the relative locations of certain rooms in the house (all italics ours):

(a) **The relative location of José Arcadio Buendía’s (and later Aureliano’s) laboratory/workshop.** “José Arcadio Buendía spent the long months of the rainy season shut up in *a small room that he had built in the rear of the house* so that no one would disturb his experiments” [5:4].

(b) **The relative location of Melquíades’ room.** “When Úrsula undertook the enlargement of the house, she had them build him *a special room next to Aureliano’s workshop, far from the noise and bustle of the house…*” [5:71].

(c) **The layout of the entrance of the house, described when José Arcadio’s trail of blood runs through the house.** “[the blood] made a right angle at the Buendía house, *went in under the closed door, crossed through the parlor,* hugging the walls so as not to stain the rugs, *went on to the other living room, made a wide curve to avoid the dining-room table, went along the porch with the begonias… and went through the pantry and came out in the kitchen…*” [5:132].

Second, to evaluate whether readers use schemas to fill in spatial details that were not specified in the narrative, we looked for common architectural features that were depicted across many of the participants’ drawings. This included features such as whether the house had one or two stories and the placement of the yard and outdoor areas in relation to the house structure.

3 Results and Discussion

A total of nine house diagrams were analyzed. Although 13 readers participated in the study, one person left the page blank and three participants drew artistic renderings of the house that contained no layout information.

3.1 Accuracy of Spatial Details Described in the Text

Six of the nine drawings specifically indicated the placement of the laboratory/workshop, and all six were in line with how the room was described in the novel, placing it in a remote location of the house, such as in a corner or as a separate room jutting off of a main wall. As is shown in Fig. 1A, one participant drew the laboratory/workshop on the second level, which is a way of indicating the remoteness of the room. (Note that in Fig. 1A the room is labeled as a study.)

Of the five participants who included Melquíades' room in their drawings, four of them placed the room in a remote location. Meanwhile, only one participant (see Fig. 1B) placed the room quite centrally, which does not match how the room was described in the novel. The text also specified that Melquíades' room was built adjacent to the laboratory/workshop. Although only three participants drew both rooms, all three placed the rooms adjacent to one another. (See Figs. 1A and C). Taken together, these results show that when participants included the laboratory/workshop or Melquíades' room in their drawings, they placed these rooms in locations that matched the spatial configuration described in the novel.

The diagrams' configuration of the front of the house, on the other hand, did not match the flow of the rooms described in the novel (i.e., parlor, living room, then dining room). Six of the nine drawings indicated a front door or porch, and were therefore included in the analysis. None of the participants' drawings included a parlor. One drawing (see Fig. 1B) shows an unlabeled room at the entrance of the house. Three of the six diagrams placed the dining room at the entrance of the house and two indicated that the living room would be the first room entered.

That several diagrams show the dining room at the entrance of the house suggests that memory of spatial details of a fictional space can be influenced by narrative events. The dining room was enlarged in Chap. 12, and throughout the novel there are numerous descriptions of guests coming to the house to dine. The dining room was a space in which many narrative events took place, as opposed to the parlor, which was rarely used by the family and was not included in any of the students' diagrams. This lends support for the claim that narrative events can shape readers' representations of the fictional space, even when concrete spatial detail is provided in the text.

Additional evidence that narrative events shape the spatial details of readers' drawings can be found in the inclusion of the chestnut tree in the students' diagrams (See Fig. 1A–C). The chestnut tree plays an important role in several events in the

A

B

C

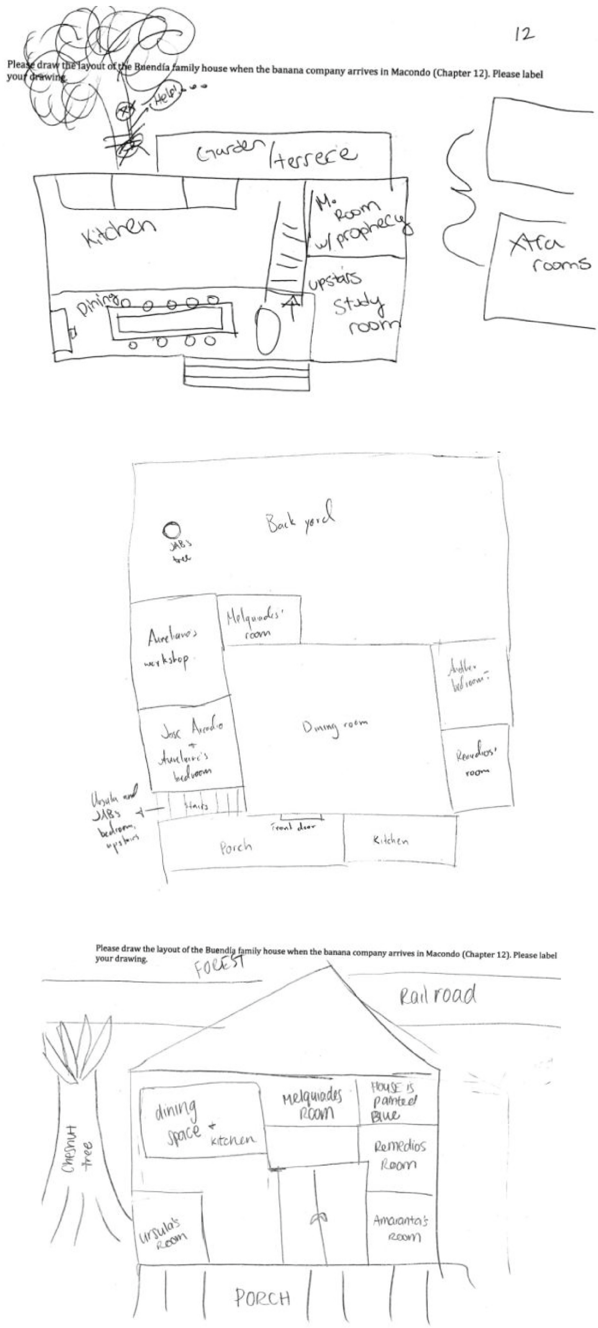


Fig. 1. A–C Three students' diagrams of the Buendía house

narrative. When the patriarch of the family, José Arcadio Buendía, loses his mind, he is tied to the chestnut tree and lives the rest of his days in the courtyard. Various family members interact with him while he is tied to the tree, and the tree serves as his designated space within the Buendía home. Thus, the inclusion and placement of the dining room and chestnut tree in the diagrams suggest that places that figure prominently into narrative events may be more privileged in memory than places that are less relevant to the narrative storyline.

3.2 Cultural Interpretations of the Buendía House

Readers often included architectural features of modern single-family North American homes in their drawings, suggesting that they applied cultural schemas and expectations about familiar domestic dwellings when interpreting the spatial layout of the fictional Latin American courtyard house. For instance, two participants drew the Buendía house with two levels, even though the novel never described a second story. Furthermore, one student's drawing, shown in Fig. 1C, places the matriarch and patriarch's bedroom on the second story, where the master bedroom is traditionally located in North American houses. This contradicts the novel, as the matriarch was described as often hiding valuables by burying them in the earth underneath her bed.

Another indication that readers interpreted the fictional house within their own cultural expectations was the single-family home "framework" that two of the participants included in their diagrams. As is shown in Fig. 1B, these participants framed their layout drawing in a prototypical square-triangle/house-roof structure. Moreover, one of the excluded drawings that depicted only the outside of the house also clearly showed a prototypical American single family home, complete with a chimney and a trail of smoke.

Finally, a third indication of the use of cultural schemas to interpret the layout of the house was the finding that none of the participants represented the layout of the house as a prototypical courtyard house, despite the courtyard being routinely referenced throughout the novel. Although the exact spatial relationship between the courtyard and the house was not specified, and the style of courtyard houses can differ widely, there are certain identifying architectural features of a courtyard house, namely that the house has an outdoor area that is somewhat contained and private as opposed to being outward-facing or street-facing [6]. However, none of the diagrams indicate a relationship between the house and the outdoor areas that remotely resembles a courtyard house, even in its most liberal sense.

The findings from the analysis of common architectural features of the drawings are in line with previous proposals that readers use cultural schemas to fill in some of the basic details of the narrative world [1, 3]. As we saw with the second-story placement of the matriarch's bedroom, the application of cultural schemas sometimes led to the inclusion of details that were contrary to what was stated in the novel. Because houses are part of a larger societal and cultural structure [7], it is not surprising that readers' diagrams of a fictional house resemble the types of real world houses that they encounter in their daily lives.

4 Conclusion

The results of the study indicate that readers fill in the spatial details of fictional spaces using their own culturally-specific understanding of similar real-world spaces, and that narrative events can assign prominence and detail to certain areas within a fictional space. The results also suggest that readers retain some memory of the spatial details specified in the text, but perhaps these details must be augmented with additional focus elsewhere in the narrative in order for readers to depict those details correctly. In the case of the entrance to the house, even though the spatial details were specified in the text, the parlor did not play into many narrative events, and readers chose to place a room that was more prominent at the entrance of the house.

As a proof-of-concept study, the results we obtained suggest that asking readers to draw diagrams of fictional spaces is a viable methodology for empirically investigating how readers fill in unspecified details. While this has implications for research in psychology and literary studies, we should not overlook the value of this methodology from a pedagogical perspective. Diagrammatic drawing exercises in the literature classroom can help an instructor understand students' cultural expectations and biases. The concreteness of the drawings can be an excellent starting point for discussions that can further evaluate students' interpretations of a literary text. If students do not have the adequate sociocultural knowledge for the literary text at hand, reviewing their diagrammatic drawings of the text's fictional spaces can be a first step to helping them build the knowledge structures that they lack. Although at first glance it may seem that diagrams have limited utility in a literature course, this study has shown how diagrammatic drawings illuminate readers' interpretations of a literary text.

References

1. Busselle, R., Bilandzic, H.: Fictionality and perceived realism in experiencing stories: a model of narrative comprehension and engagement. Commun. Theory **18**, 255–280 (2008)
2. Zoran, G.: Towards a theory of space in narrative. Poetics Today **5**(2), 309–335 (1984)
3. McVee, M.B., Dunsmore, K., Gavelek, J.R.: Schema theory revisited. Rev. Educ. Res. **75**(4), 531–566 (2005)
4. Zwaan, R., Langston, M., Graesser, A.: The construction of situation models in narrative comprehension: an event-indexing model. Psychol. Sci. **6**(5), 292–297 (1995)
5. García Márquez, G.: One Hundred Years of Solitude. Harper Perennial Modern Classics, New York (1970). (Translated from the Spanish by G. Rabassa)
6. Rapoport, A.: The nature of the courtyard house: a conceptual analysis. TDSR **18**(2), 57–72 (2007)
7. Wilk, R.: The built environment and consumer decisions. In: Kent, S. (ed.) Domestic Architecture and the Use of Space, pp. 34–42. Cambridge University Press, Cambridge (1990)

Design Principles for Diagrams

An Investigation into OWL for Concrete Syntax Specification Using UML Notations

Anitta Thomas[1(✉)], Aurona J. Gerber[2,3], and Alta van der Merwe[2]

[1] School of Computing, University of South Africa, The Science Campus, Florida Park, Johannesburg, South Africa
thomaa@unisa.ac.za

[2] Department of Informatics, University of Pretoria, Pretoria, South Africa
{aurona.gerber,alta.vdm}@up.ac.za

[3] Center for Artificial Intelligence Research (CAIR), CSIR Meraka, Pretoria, South Africa

Abstract. The Web Ontology Language OWL is a prominent ontology language for specifying ontologies. Although OWL ontologies are well-used for representing and reasoning about knowledge in various domains, they are sparsely studied for visual language specification. The work in this paper, therefore, explores OWL for visual language specification by specifying the concrete syntax of selected UML class diagram notations in an ontology. The selected diagram notations are specified as spatial configurations of primitive elements and qualitative base spatial relationships of Region Connection Calculus-8 (RCC-8). Furthermore, the automated reasoning features of ontology reasoners are investigated to verify the completeness and the correctness of the specification. The verification results indicate that the given specification needs to be revised to support applications to draw the selected notations. The value of such a specification in supporting a semantic diagram interpretation application is demonstrated using the automated instance classification feature of ontology reasoners.

Keywords: UML notations · Concrete syntax specification · RCC-8 · OWL · Ontology · Ontology reasoner

1 Introduction

In Computing, diagrams are studied as elements of visual languages and an aspect studied about visual languages is its specification. In general, a visual language specification aims to capture the syntax and semantics of the relevant language [22]. Moreover a syntax specification of a visual language can focus on two different aspects; the abstract syntax and the concrete syntax, where the concrete syntax deals with the visual representation of the concepts in the abstract syntax [1,23]. The term concrete syntax is used throughout this paper even though such syntax can also be referred to as layout syntax [9], visual syntax [31] or token syntax [16]. Although a clear distinction between the abstract and

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 197–211, 2016.
DOI: 10.1007/978-3-319-42333-3_15

the concrete syntax or the syntax and the semantics is not applicable to all visual languages [22], in general, a visual language specification deals with one or more of these three aspects; abstract syntax, concrete syntax and semantics.

To specify a visual language symbolically, a specification technique is used. There are numerous specification techniques used for visual language specifications. Two such specification techniques are graph grammars [9,19] and Description Logics (DLs) [10,11]. There are many factors that influence the choice of a specification technique. Adequate expressiveness [21], computational efficiency [21], existing knowledge and expertise [22] and availability of supporting tools are a few such factors.

OWL is the World Wide Web Consortium (W3C) semantic web language for authoring ontologies [28]. OWL ontologies represent explicit domain knowledge to semantically enrich the Web. Such ontologies are widely used for knowledge representation in numerous disciplines such as biology, medicine, geography, astronomy and agriculture [24]. OWL is a machine readable form of an expressive DL and thus OWL ontologies make use of ontology reasoners [24]. An area where OWL ontologies is under-explored is the field of visual language specification. It is worthwhile to explore OWL for visual language specifications because such specifications can be utilized within the context of semantic web applications. Moreover, there exists numerous tools to support OWL ontology development, which can be leveraged for developing visual language specifications.

There are many aspects to explore about OWL as a specification technique. This paper explores two such aspects: how can the concrete syntax of diagram notations of a visual language be specified in an OWL ontology? What are the inherent features of OWL and OWL ontology reasoners that can be exploited to support visual language specifications and applications? To answer these questions, this work uses selected notations of a set of UML class diagram constructs for concrete syntax specification, verification and application. Although UML notations are only used as sample notations in this paper, they were chosen because UML is widely-known [18] and UML, in general, does not have a formal concrete syntax specification [26].

In this work, the concrete syntax specification defined in an ontology (hereafter referred to as a concrete syntax ontology) is envisaged to support technical applications that can reason about the visual structure of diagrams. An application to assist visually impaired users in deciphering a raster or a vector image of an UML diagram is an example of such an application. It should be noted that the proposed concrete syntax ontology is not for reasoning about the models represented in UML diagrams (as in [3,18]) and thus, this work is not concerned about model translations from UML to OWL (as in [7,33]). This study only uses UML notations to explore OWL as a specification technique and thus no comparison between OWL and other existing specification techniques is made in this work.

This work contributes to the current knowledge of visual languages in three different ways: it explores OWL as a concrete syntax specification technique, it explores OWL reasoner features to verify concrete syntax specification and to support diagram interpretation applications and it provides a formal encoding of selected UML class diagram notations. The former two aspects make generic contribution to the area of visual language specification.

This paper is structured into eight sections. Section 2 is a background section that includes brief background information on OWL, selected UML class diagram notations, and RCC-8. Section 3 presents a brief discussion of related work. In Sect. 4, the selected UML class diagram notations are modeled as spatial configurations of primitive elements and spatial relationships of RCC-8 along with their respective encoding in the OWL ontology. Criteria for completeness and correctness to verify the concrete syntax specification are discussed in Sect. 5. In Sect. 6 the concrete syntax given in Sect. 4 is evaluated for completeness and correctness using the automated reasoning features of ontology reasoners, where applicable. In Sect. 7 a conceptual application of the concrete syntax ontology within a diagram interpretation context is presented. Section 8 concludes with a summary and the intended value of the contributions.

2 Background

2.1 OWL

An OWL ontology models a domain using classes, properties, instances and data values [15]. A class represents a set of objects, a property describes a possible relationship between objects, instances describe the objects itself and a data value links an instance to a specific data type [14]. OWL provides a rich set of constructs such as union, intersection and negation to describe classes and characteristics such as transitivity, symmetry and reflexivity to describe properties. Due to the compositional nature of OWL, complex classes can be described using other classes in the ontology [14].

An ontology reasoner is a key aspect in working with ontologies. An ontology reasoner is used to check the correctness as well as to infer new knowledge based on what is described in the ontology. In other words, it helps in detecting inconsistencies as well as in maintaining the class hierarchies by inference based on the explicitly stated information in the ontology. The automated reasoning capabilities of an ontology reasoner are important in maintaining correct ontologies [14].

2.2 Selected UML Class Diagram Notations

UML is managed by the standards consortium Object Management Group and it has a specification, the current version being UML 2.4.1 [26,27]. UML can be used to model both structural and behavioral features of an application [27] and it can be used in the design, analysis, implementation and documentation phases of software applications [26,27]. Among the numerous diagrams of UML to support structural and behavioral modeling, a class diagram is used to represent a structural model of an application which typically consists of classes and the relationships among classes, using both graphical and textual notations [27].

UML class diagram constructs are specified in the Classes package of the UML 2.4.1 specification [27]. This package includes fifty six constructs that can be used to represent object oriented models using class, package and object diagrams. The set of UML constructs that are considered in this work are *Class*, *Interface*, *Association*, *Aggregation*, *Composition*, *Dependency*, *Generalization*, *Realization* and *InterfaceRealization*. Figure 1 lists these nine UML constructs and their ten notations used in this paper. Based on these selected notations all constructs except two (*Association* and *Interface*) have one notation each. Further details of these constructs including their semantics can be found in [27].

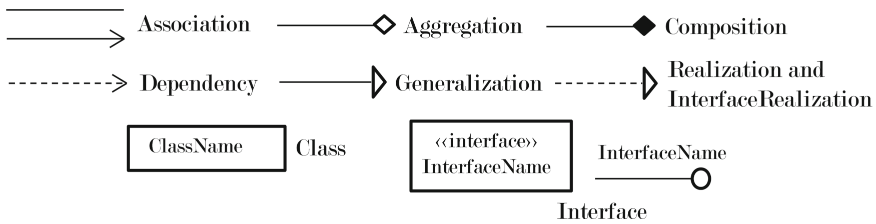


Fig. 1. Selected UML class diagram notations [27]

The constructs and notations in Fig. 1 were selected because they are identified as typical constructs and notations in UML class diagrams (pages 147 to 150 in [27]). The notations depicted in Fig. 1 are not the only notations for the selected nine constructs and similarly, these nine constructs are not the only constructs with notations that can be used in a class diagram. However to limit the scope of this paper a set of UML notations is selected as in Fig. 1.

2.3 RCC-8

RCC-8 represents space qualitatively and it excludes numerical representation and computation of space. As one of the most commonly used calculus for qualitative spatial representation and reasoning, RCC-8 is used in areas such as Geographic Information System, engineering design, robotic navigation, biology, qualitative document structure recognition and visual language specification [4].

RCC-8 contains a set of eight base spatial relationships. These spatial relationships are disconnected (DC), externally connected (EC), partially overlapping (PO), equals to (EQ), tangential proper part of (TPP), non-tangential proper part of ($NTPP$), inverse tangential proper part of (TPP^{-1}) and inverse non-tangential proper part of ($NTPP^{-1}$) [29]. These eight relationships are mutually exclusive and exhaustive meaning that two regions in a given space satisfy exactly one of these spatial relationships [5]. Depictions of these eight spatial relationships are given in Fig. 2.

OWL has been extensively studied for both spatial representation and reasoning using RCC-8 [2,12,20,30]. Realizing complete RCC-8 spatial reasoning in

OWL is not straightforward due to some inherent limitations of OWL [12,13,20]. An example of such a limitation is the lack of support for complex role inclusion axioms required to encode the entries of the composition table of RCC-8 [12]. On the other hand, there have been successful encoding of RCC-8 reasoning in OWL using Semantic Web Rule Language [2,20].

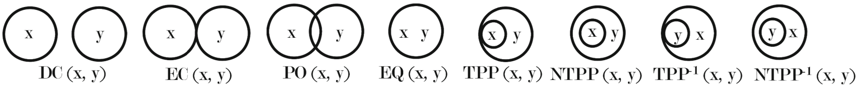


Fig. 2. Depictions of the eight spatial relations in RCC-8 [29]

3 Related Work

Visual language specification techniques can be broadly classified into grammar-based, logic-based and algebraic-based formalisms [22]. Since OWL is a logic-based ontology language, it can be categorized as a specification technique in the logical formalism. Within the logical formalism, DLs were explored for visual syntax specification [22]. For example, a general DL formalism has been previously used to specify the concrete syntax of entity-relationship (ER) diagram constructs and syntactical constructs of Pictorial Janus [22]. The concrete syntax specification of ER diagrams was then used in DL systems CLASSIC and LOOM to automate diagram reasoning to realize a syntax-directed diagram editor that can validate diagrams [11]. The concrete syntax of Pictorial Janus was used to formalize its semantics, which was also used to realize a diagram editor that verifies the semantics of diagrams of Pictorial Janus [10]. Although DL provide the logical foundations for OWL [24], the use of OWL ontologies as visual language specifications is still under-explored. As a W3C standard, OWL is the knowledge representation language to realize semantic web [28] and due to the standardization, it has extensive tool support. Thus authoring a visual language specification in OWL can make use of the existing OWL tools and such a specification is desired in potential semantic web applications.

OWL ontologies are used to represent knowledge about visual concepts such as shapes and graphical concepts. Two examples of such generic shapes ontologies are discussed in [25,32]. Such generic ontologies do not in general encode concrete syntax of visual languages.

RCC-8 has been previously used for visual language specification [4]. RCC-8 was used to model the syntax of the visual programming language Pictorial Janus. The syntax of Pictorial Janus is presented pictorially and thus RCC-8 was used to model the spatial configurations of various syntactical constructs of this programming language. The specification technique used for specifying the visual syntax of Pictorial Janus is many-sorted logic [8]. The study in [8] indicates that RCC-8 can indeed be used for modeling the concrete syntax of diagrammatic notations.

4 Concrete Syntax Specification of UML Notations

In this section the use of OWL to model the concrete syntax of UML class diagram notations is explored. The selected UML notations are modeled using primitive elements and spatial relationships as in [8,10,11,31]. The details of the primitive elements and the spatial relationships for the selected UML notations and how they are specified in the OWL ontology are discussed in the next Subsects. 4.1 and 4.2. The specification of the UML notations is given in Sect. 4.3.

4.1 Modeling and Encoding Primitive Elements

The selected UML notations are composed of nine primitive elements namely arrow, circle, filled diamond, unfilled diamond, line, dotted line, rectangle, string and triangle. These nine primitive elements are depicted in Fig. 3.

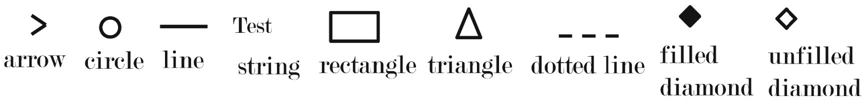


Fig. 3. Primitive elements of the selected UML notations

These nine primitive elements are specified as nine OWL primitive classes in the ontology. The class names representing these primitive elements are *Arrow*, *Circle*, *DiamondFilled*, *DiamondUnfilled*, *Line*, *DottedLine*, *Rectangle*, *String* and *Triangle*. These nine classes are added as subclasses of an OWL class *Primitives*. The modeling approach used to model the primitive elements can be classified as the non-attributed representation of graphical objects as stated in the visual language literature [6,21].

4.2 Modeling and Encoding Spatial Relationships

Eight spatial relationships in RCC-8 are added as object properties in the OWL ontology. The spatial relationships DC, EC, PO, EQ, TPP, $NTPP$, TPP^{-1}, $NTPP^{-1}$ are encoded in the concrete syntax ontology as object properties *isDisconnectedTo*, *isExternallyConnectedTo*, *isPartiallyOverlappingTo*, *isEqualTo*, *isTangentialProperPartOf*, *isNonTangentialProperPartOf*, *hasTangentialProperPartOf* and *hasNonTangentialProperPartOf* respectively. The modeling approach used to describe the spatial relationships can be seen as the connection-based representation [6].

In this work RCC8 is primarily used for spatial representation. Even though realizing spatial reasoning in RCC-8 is not the primary focus of this work, some semantics of RCC-8 spatial relationships is captured using OWL object property characteristics mentioned in Sect. 2.1. Specifically this work uses the symmetric object property characteristic to encode the fact that *isDisconnectedTo*, *isExternallyConnectedTo*, *isEqualTo* and *isPartiallyOverlappingTo* are symmetrical, and *hasTangentialProperPartOf* and *hasNonTangentialProperPartOf* are inverse of *isTangentialProperPartOf* and *isNonTangentialProperPartOf* respectively [2].

4.3 Specification of UML Class Diagram Notations

The nine UML constructs are modeled as nine OWL classes namely *UMLClass*, *Interface*, *Association*, *Aggregation*, *Composition*, *Dependency*, *Generalization*, *Realization* and *InterfaceRealization* in the ontology. The concrete syntax of the selected notations is then specified in the respective OWL class definitions. These nine OWL classes are included as subclasses of *UMLConstructs*, a sibling class of *Primitives* (mentioned in Sect. 4.1).

Class. A *Class* is represented using a rectangle and a string inside the rectangle where the string indicates the class name. The spatial configuration of the selected notation of *Class* can then be modeled using rectangle, string and RCC-8. This spatial configuration is specified in the ontology as:

```
Class: UMLClass
EquivalentTo: Rectangle and
(hasNonTangentialProperPartOf exactly 1 String)
SubClassOf: UMLConstructs
```

Interface. One notation of *Interface* uses a rectangle and two strings inside the rectangle and the second one uses a line, a circle and a string. Thus the spatial configurations of *Interface* can be modeled using rectangle, line, circle, string and RCC-8, which is specified in the ontology as:

```
Class: Interface
EquivalentTo: Rectangle and
(hasNonTangentialProperPartOf exactly 2 String),
Line and (isDisconnectedTo some String)
and (isExternallyConnectedTo some Circle)
SubClassOf: UMLConstructs
```

Association. An *Association* is a link between two *Class*es, which is represented using a line and an arrow, or a line. The spatial configurations of *Association* can thus be modeled using a line, an arrow and two classes as:

```
Class: Association
EquivalentTo: Line and (isExternallyConnectedTo some UMLClass)
and (isExternallyConnectedTo some (Arrow
and (isExternallyConnectedTo some UMLClass))),
Line and (isExternallyConnectedTo min 2 UMLClass)
SubClassOf: UMLConstructs
```

Aggregation. An *Aggregation* connects two *Class*es using a line and an unfilled diamond. The spatial configuration of *Association* is thus specified as:

```
Class: Aggregation
EquivalentTo: Line and (isExternallyConnectedTo some UMLClass)
and (isExternallyConnectedTo some (DiamondUnfilled
and (isExternallyConnectedTo some UMLClass)))
SubClassOf: UMLConstructs
```

Composition. A *Composition* is a relationship between two *Class*es represented using a line and a filled diamond. The spatial configuration of *Composition* is thus specified as:

```
Class: Composition
EquivalentTo: Line and (isExternallyConnectedTo some UMLClass)
and (isExternallyConnectedTo some (DiamondFilled
and (isExternallyConnectedTo some UMLClass)))
SubClassOf: UMLConstructs
```

Dependency. A *Dependency* is indicated using a dotted line and an arrow between two *Class*es. Therefore the spatial configuration of *Dependency* is specified as:

```
Class: Dependency
EquivalentTo: LineDotted and (isExternallyConnectedTo
some UMLClass) and (isExternallyConnectedTo
some (Arrow and (isExternallyConnectedTo some UMLClass)))
SubClassOf: UMLConstructs
```

Generalization. A *Generalization* is indicated using a line and a triangle between two *Class*es. Therefore the spatial configuration of *Generalization* is specified as:

```
Class: Generalization
EquivalentTo: Line and (isExternallyConnectedTo some UMLClass)
and (isExternallyConnectedTo some (Triangle
and (isExternallyConnectedTo some UMLClass)))
SubClassOf: UMLConstructs
```

Realization. A *Realization* is indicated between two *Class*es using a dotted line and a triangle. Therefore the spatial configuration of *Realization* is specified as:

```
Class: Realization
EquivalentTo: LineDotted and (isExternallyConnectedTo
some UMLClass) and (isExternallyConnectedTo
some (Triangle and (isExternallyConnectedTo some UMLClass)))
SubClassOf: UMLConstructs
```

InterfaceRealization. An *InterfaceRealization* is indicated between a *Class* and an *Interface* using a dotted line and a triangle. Therefore the spatial configuration of *InterfaceRealization* is specified as:

```
Class: InterfaceRealization
EquivalentTo: LineDotted and (isExternallyConnectedTo
some UMLClass) and (isExternallyConnectedTo
some (Triangle and (isExternallyConnectedTo some Interface)))
SubClassOf: UMLConstructs
```

5 Verification Criteria for Concrete Syntax Specification

In this section criteria to verify the concrete syntax specification given in Sect. 4.3 are discussed. The notations depicted in Fig. 1 are distinct, which means that a notation should only map to one UML construct. However the mapping of a UML construct to a unique notation is only valid for seven of these constructs as *Interface* and *Association* have two notations each. Even though the notations for *InterfaceRealization* and *Realization* are the same, the former UML construct is a link between an *Interface* and a *Class* but the latter connects two *Class*es, thus resulting in two distinct spatial configurations. The following two general criteria are identified to verify the given concrete syntax specification in Sect. 4.3:

Completeness: A specification is complete if each UML construct has an encoded concrete syntax in the concrete syntax ontology. This definition of completeness is adapted from [1].

Correctness: A specification is correct if an encoded spatial configuration maps exactly to one UML construct represented in the concrete syntax ontology. This definition of correctness is also adapted from [1]. Another correctness criterion is to check whether the encoded spatial configurations in the ontology always lead to acceptable notations. This criterion evaluates whether the spatial configurations of the notations have been modeled correctly.

6 Evaluation of the Concrete Syntax Specification

In this section the use of the automated reasoning features of ontology reasoners is explored to verify the concrete syntax specification given in Sect. 4.3 according to the criteria given in Sect. 5. The concrete syntax specification given in Sect. 4.3 is complete based on the fact that the OWL classes representing UML constructs are defined classes. i.e. these nine OWL classes have definitions that encode the concrete syntax of the relevant notations.

To ensure that the concrete syntax specification is correct based on the first criterion for correctness, the automated reasoning feature of the ontology reasoner can be utilized. Specifically the OWL class feature *disjointness* can be used to ensure that each notation is encoded distinctly, which ensures that a notation maps uniquely to a UML construct. When two OWL classes are disjoint the sets of objects represented by these classes are disjoint. Class disjointness is not a reasoner service but an OWL class descriptor that needs to be explicitly stated for the classes in an ontology [15]. If the reasoner, however, infers that two classes cannot be disjoint based on the explicit and inferred knowledge, then it will be highlighted as a contradiction in the ontology.

To demonstrate that class disjointness ensure unique spatial configuration for each of the OWL classes presented in Sect. 4.3, two separate cases are considered. The first case is for the OWL classes (examples: *UMLClass* and *Interface*) that are defined as specializations of the same OWL class (example: *Rectangle*) representing a primitive element. The second case is for OWL classes that are defined as specializations of different OWL classes (examples: *UMLClass* and *Association*) representing different primitive elements (examples: *Rectangle* and *Line*).

Case 1: Consider three OWL classes $C1$, $C2$ and P in the ontology where $P \sqsubseteq Primitives$, $C1 \sqsubseteq UMLConstructs$, $C2 \sqsubseteq UMLConstructs$. $C1$ and $C2$ are two different classes representing two separate UML constructs in the ontology in Sect. 4.3. If *C1* and *C2* are defined as specializations of P, then $C1 \sqsubseteq P$ and $C2 \sqsubseteq P$. Since $C1$ and $C2$ are defined using P, they are expressed as $C1 \equiv P \sqcap (exp1)$ and $C2 \equiv P \sqcap (exp2)$. For $C1$ and $C2$ to be disjoint, the sets of objects represented by $P \sqcap (exp1)$ and $P \sqcap (exp2)$ do not have objects in common. Since both $C1$ and $C2$ are defined as specializations of P, the expressions $exp1$ and $exp2$ must represent distinct spatial configurations.

Case 2: Consider four OWL classes $C1$, $C2$, $P1$ and $P2$ in the ontology where $P1 \sqsubseteq Primitives$, $P2 \sqsubseteq Primitives$, $C1 \sqsubseteq UMLConstructs$, $C2 \sqsubseteq UMLConstructs$. $C1$ and $C2$ are two different classes representing two separate UML constructs in the ontology in Sect. 4.3. If *C1* and *C2* are defined as specializations of $P1$ and $P2$ respectively, then $C1 \sqsubseteq P1$ and $C2 \sqsubseteq P2$. Since $C1$ and $C2$ are specializations of $P1$ and $P2$ respectively, they are expressed as $C1 \equiv P1 \sqcap (exp1)$ and $C2 \equiv P2 \sqcap (exp2)$. For $C1$ and $C2$ to be disjoint, the sets of objects represented by $P1 \sqcap (exp1)$ and $P2 \sqcap (exp2)$ do not have objects in common. If $P1$ and $P2$ are disjoint, then $C1$ and $C2$ represent distinct spatial configurations provided that $C1$ and $C2$ cannot be defined as specializations of $P2$ and $P1$ respectively.

To check the correctness of the concrete syntax specification, the eight spatial relationships were defined distinct from one another, eight of the subclasses of the OWL class *Primitives* were marked *disjoint* from one another and the nine OWL subclasses of *UMLConstructs* representing the nine UML constructs were also marked *disjoint* from one another. Moreover, selected OWL classes defined under *UMLConstructs* and *Primitives* were also made disjoint from one another. For instance, the OWL class *UMLClass* is made disjoint from every subclass of *Primitives* except the OWL class *Rectangle*, as *UMLClass* is defined as a *Rectangle*, which means that an instance of *Rectangle* can be an instance of *UMLClass* as well. Invoking the reasoner after including the disjoint class descriptor for the OWL classes as stated above does not indicate any contradiction in the concrete syntax ontology in Sect. 4.3. In the absence of a contradiction the encoded spatial configurations in the concrete syntax ontology given in Sect. 4.3 must be distinct.

The second criterion of correctness can be checked by counter-examples of notations that satisfy the spatial configuration specified in the ontology but differs from the intended notation. For example, Fig. 4 demonstrates an incorrect notation of *Interface*, which satisfy the spatial configuration specified for this construct in the concrete syntax ontology. However the notation in Fig. 4 cannot be considered as an interface as the string representing the interface name is not placed 'near' to the line and circle.

Fig. 4. Incorrect notation for *Interface*

The fact that the visual syntax specification in Sect. 4.3 does not satisfy one of the criteria discussed in Sect. 5 indicates that the specification is not suitable to be used in technical applications designed to validate and draw UML notations. This unsuitability is because of the fact that the given specification is not precise enough. To be useful in such applications, the current specification of UML notations needs to be revised. For example, to specify that the string that represents the interface name should be placed 'near' the circle and line requires a spatial relationship that cannot be expressed using RCC-8 [22]. However, as demonstrated in the next section, the given concrete syntax ontology can be used in an application that interprets a *valid* UML class diagram by mapping lower level graphical data to higher-level semantic UML concepts. A *valid* class diagram in this context refers to a diagram that uses only the notations depicted in Fig. 1.

7 Concrete Syntax Ontology to Support Diagram Interpretation

Automated diagram interpretation can be seen as a form of image interpretation [11]. The process of extracting higher-level semantic concepts from low-level graphical data in image interpretation [17] is also applicable in an automated diagram interpretation application. The concrete syntax ontology presented in this paper can be used for inferring UML constructs using the automated reasoning feature, *instance checking*, of the ontology reasoners provided that a class diagram is described using the primitive elements and the spatial relationships [22] given in Sects. 4.1 and 4.2.

Instance checking is an automated reasoner service [15], which makes use of the explicitly stated and inferred information about the classes and instances in an ontology to determine whether an instance belongs to a class. a is inferred to be an instance of class A if a satisfies the class descriptions of A [15]. Instance checking can be utilized in a diagram interpretation application where the UML constructs need to be interpreted from a valid class diagram.

Consider the UML class diagram in Fig. 5(a), which depicts four *Class*es (A, B, C, D), one *Generalization* (between A and B), one *Aggregation* (between B and C) and one *Association* (between C and D). To utilize instance checking to interpret the class diagram in Fig. 5(a), fourteen instances, which represent fourteen primitive elements, along with their OWL object properties, representing the spatial relationships were encoded in the ontology. Given below is a sample entry in the ontology for one of these fourteen instances, the string A in Fig. 5(a), which is encoded as an instance named a of OWL class *String*. For ease of reference Fig. 5(b) annotates the UML class diagram in Fig. 5(a) with unique identifiers for the primitive elements used in the ontology.

```
Individual: a
Types: String
Facts: isDisconnectedTo  lineAB, isDisconnectedTo  lineDC,
isDisconnectedTo  rectC, isDisconnectedTo  d,
isDisconnectedTo  diamond, isDisconnectedTo  lineBC,
isDisconnectedTo  c, isDisconnectedTo  rectB,
isDisconnectedTo  triangle, isEqualTo  a, isDisconnectedTo  arrow,
isDisconnectedTo  b, isDisconnectedTo  rectD
isNonTangentialProperPartOf  rectA
```

Not all primitive elements have to be encoded in such detail (as listed above for *String* instance a) as the reasoner can infer many spatial relationships between instances. For example, based on the description of instance a of the OWL primitive class *String*, the reasoner will infer that *lineAB*, *lineDC*, *rectC*, *d*, *diamond*, *lineBC*, *c*, *rectB*, *triangle*, *arrow*, *b* and *rectD* have *isDisconnectedTo* object relationship with a because *isDisconnectedTo* is described to be

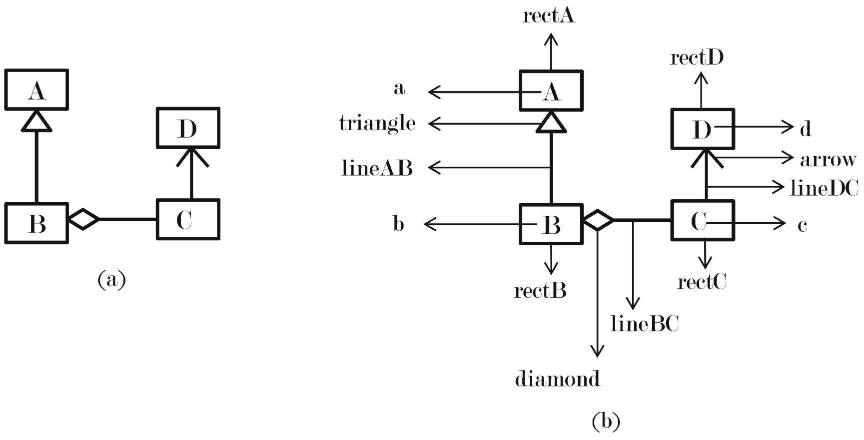


Fig. 5. (a) A sample UML class diagram (b) UML class diagram in (a) annotated with identifiers used in the concrete syntax ontology

symmetric in the concrete syntax ontology (Sect. 4.2). Currently the encoding of instances (representing primitives) is done manually, however, ideally it should be done using a software tool.

When the reasoner is invoked with these fourteen instances in the concrete syntax ontology, it correctly identifies four instances of *UMLClass*, one instance of *Association*, one instance of *Aggregation* and one instance of *Generalization*. The correct interpretation of the UML constructs for the class diagram in Fig. 5(a) indicates that the concrete syntax ontology is a feasible technique to support technical applications to realize diagram interpretation.

Admittedly, the class diagram given in Fig. 5(a) is rather simple using only a subset of notations described in Sect. 2.2. To deal with complex class diagrams, more UML constructs and notations have to be incorporated in the OWL ontology.

8 Conclusion

In this work, the concrete syntax of selected UML notations were modeled using the base spatial relationships of RCC-8. The spatial configurations of the UML notations were then specified in an ontology, where an OWL class is defined for each UML construct. Furthermore criteria for verifying the completeness and correctness of the visual syntax specification were presented. This work also demonstrates how the automated reasoning features of the ontology reasoners can be leveraged to verify a concrete syntax specification. The verification results indicate that the given concrete syntax specification is not precise enough to be used in all possible diagram processing applications. Nevertheless, how a concrete syntax ontology can be used within a diagram interpretation application is discussed.

Given our results, it can be concluded that an OWL ontology is a feasible technique for concrete syntax specification provided that the spatial configurations are modeled using classes, properties, objects and data types as required by OWL. An advantage of using an OWL ontology is the use of ontology reasoners to verify the syntax specification and to support technical applications for diagram interpretation. Using OWL is of value because it allows the reuse of mature software artifacts including OWL ontology editors and reasoners for concrete syntax specification. Using OWL as a specification technique may also lead to the use of visual language specifications in semantic web applications.

Although this work focused on a limited number of notations, it is also of value to the general field of visual language specification. The process of modeling diagram notations, specifying the concrete syntax in an OWL ontology and utilizing the ontology reasoners as demonstrated in this work can be applied to another visual language as well.

References

1. Baar, T.: Correctly defined concrete syntax for visual modeling languages. In: Wang, J., Whittle, J., Harel, D., Reggio, G. (eds.) MoDELS 2006. LNCS, vol. 4199, pp. 111–125. Springer, Heidelberg (2006)

2. Batsakis, S., Petrakis, E.G.M.: SOWL: a framework for handling spatio-temporal information in OWL 2.0. In: Bassiliades, N., Governatori, G., Paschke, A. (eds.) RuleML 2011 - Europe. LNCS, vol. 6826, pp. 242–249. Springer, Heidelberg (2011)
3. Berardi, D., Cali, A., Calvanese, D., Giacomo, G.D.: Reasoning on UML class diagrams. Artif. Intell. **168**, 70–118 (2005)
4. Cohn, A.G., Renz, J.: Qualitative spatial representation and reasoning. In: Frank van Harmelen, V.L., Porter, B. (eds.) Handbook of Knowledge Representation, pp. 551–596. Elsevier (2007)
5. Cohn, A., Bennett, B., Gooday, J., Gotts, N.: Qualitative spatial representation and reasoning with the region connection calculus. GeoInformatica **1**(3), 275–316 (1997)
6. Costagliola, G., Delucia, A., Orefice, S., Tortora, G.: A framework of syntactic models for the implementation of visual languages. In: Proceedings of IEEE Symposium on Visual Languages, pp. 58–65. IEEE (1997)
7. Gasevic, D., Djuric, D., Devedzic, V., Damjanovi, V.: Converting UML to OWL ontologies. In: Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters, WWW Alt. 2004, pp. 488–489. ACM, New York (2004)
8. Gooday, J., Cohn, A.: Visual language syntax and semantics: a spatial logic approach. In: Proceedings of AVI 1996 Workshop on Theory of Visual Languages (1996)
9. Grunske, L., Winter, K., Yatapanage, N.: Defining the abstract syntax of visual languages with advanced graph grammars: a case study based on behavior trees. J. Vis. Lang. Comput. **19**(3), 343–379 (2008)
10. Haarslev, V.: Formal semantics of visual languages using spatial reasoning. In: Proceedings of the 11th IEEE International Symposium on Visual Languages, pp. 156–163, September 1995
11. Haarslev, V.: Using description logic for reasoning about diagrammatical notations. In: Padgham, L. (ed.) Proceedings of the International Workshop on Description Logics, pp. 124–128 (1996)
12. Hogenboom, F., Borgman, B., Frasincar, F., Kaymak, U.: Spatial knowledge representation on the semantic web. In: IEEE Fourth International Conference on Semantic Computing (ICSC), pp. 252–259. IEEE (2010)
13. Hogenboom, F., Frasincar, F., Kaymak, U.: A review of approaches for representing RCC8 in OWL. In: Proceedings of the 25th ACM Symposium on Applied Computing (SAC 2010), pp. 1444–1445. ACM (2010)
14. Horridge, M., Drummond, N., Jupp, S., Moulton, G., Stevens, R.: A practical guide to building OWL ontologies using protégé 4 and CO-ODE tools. University of Manchester (2009)
15. Horrocks, I., Parsia, B., Sattler, U.: OWL 2 Web Ontology Language: Direct Semantics, 2nd edn. World Wide Web Consortium, Cambridge (2012)
16. Howse, J., Molina, F., Shin, S.-J., Taylor, J.G.: On diagram tokens and types. In: Hegarty, M., Meyer, B., Narayanan, N.H. (eds.) Diagrams 2002. LNCS (LNAI), vol. 2317, pp. 146–160. Springer, Heidelberg (2002)
17. Hudelot, C.: Towards a cognitive platform for semantic image interpretation; application to the recognition of biological organisms. Ph.D. thesis, University of Nice - Sophia Antipolis (2005)
18. Khan, A.H., Porres, I.: Consistency of UML class, object and statechart diagrams using ontology reasoners. J. Vis. Lang. Comput. **26**, 42–65 (2015)

19. de Lara, J., Vangheluwe, H.: Defining visual notations and their manipulation through meta-modelling and graph transformation. J. Vis. Lang. Comput. **15**(34), 309–330 (2004)
20. Marc-Zwecker, S., de Bertrand de Beuvron, F., Zanni-Merk, C., Ber, F.L.: Qualitative spatial reasoning in RCC8 with OWL and SWRL. In: Proceedings of the Knowledge Engineering and Ontology Development Conference (KEOD 2013) (2013)
21. Marriott, K., Meyer, B.: On the classification of visual languages by grammar hierarchies. J. Vis. Lang. Comput. **8**(4), 375–402 (1997)
22. Marriott, K., Meyer, B., Wittenburg, K.B.: A survey of visual language specification and recognition. In: Marriott, K., Meyer, B. (eds.) Visual Language Theory, pp. 5–85. Springer-Verlag New York Inc., New York (1998)
23. Minas, M.: Syntax definition with graphs. Electron. Notes Theor. Comput. Sci. **148**(1), 19–40 (2006)
24. Motik, B., Cuenca Grau, B., Sattler, U.: Structured objects in OWL: representation and reasoning. In: Proceedings of the 17th International Conference on World Wide Web, WWW 2008, pp. 555–564. ACM, New York (2008)
25. Niknam, M., Kemke, C.: Modeling shapes and graphics concepts in an ontology. In: Hastings, J., Oliver Kutz, M.B., Borgo, S. (eds.) Proceedings of the First Interdisciplinary Workshop on SHAPES (2011)
26. Object Management Group: Information technology - Object Management Group Unified Modeling Language (OMG UML), Infrastructure, April 2012
27. Object Management Group: Information technology - Object Management Group Unified Modeling Language (OMG UML), Superstructure, April 2012
28. OWL 2 Web Ontology Language Document Overview, 2nd edn. http://www.w3.org/TR/owl2-overview/. Accessed 2 Dec 2015
29. Renz, J. (ed.): Qualitative Spatial Reasoning with Topological Information. LNCS (LNAI), vol. 2293. Springer, Heidelberg (2002)
30. Stocker, M., Sirin, E.: PelletSpatial: a hybrid RCC-8 and RDF/OWL reasoning and query engine. In: Proceedings of OWL: Experiences and Directions 2009 (OWLED 2009) (2009)
31. Thomas, A., Gerber, A.J., van der Merwe, A.: Visual syntax of UML class and package diagram constructs as an ontology. In: Fred, A., Dietz, J., Aveiro, D., Liu, K., Filipe, J. (eds.) Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015), vol. 2, pp. 17–28. SCITEPRESS (2015)
32. Vasilakis, G., Garcia-Rojas, A., Papaleo, L., Catalano, C.E., Robbiano, F., Spagnuolo, M., Vavalis, M., Pitikakis, M.: A common ontology for multi-dimensional shapes. In: Proceedings of the International Workshop on Multimedia Annotation and Retrieval Enabled by Shared Ontologies, pp. 31–43 (2007)
33. Zedlitz, J., Jörke, J., Luttenberger, N.: From UML to OWL 2. In: Lukose, D., Ahmad, A.R., Suliman, A. (eds.) KTW 2011. CCIS, vol. 295, pp. 154–163. Springer, Heidelberg (2012)

Using One-Dimensional Compaction for Smaller Graph Drawings

Ulf Rüegg(✉), Christoph Daniel Schulze,
Daniel Grevismühl, and Reinhard von Hanxleden

Deptartment of Computer Science, Kiel University, Kiel, Germany
{uru,cds,dag,rvh}@informatik.uni-kiel.de

Abstract. We use the technique of one-dimensional compaction as part of two new methods tackling problems in the context of automatic diagram layout: First, a post-processing of the layer-based layout algorithm, also known as Sugiyama layout, and second a placement algorithm for connected components with external extensions. We apply our methods to dataflow diagrams from practical applications and find that the first method significantly reduces the width of left-to-right drawn diagrams. The second method allows to properly arrange disconnected graphs that have hierarchy-crossing edges.

1 Introduction

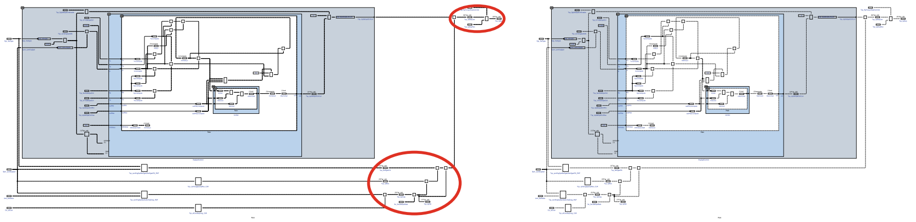
Automatically drawing graph-based visual models has gained more and more acceptance over the past years, with industrial tools starting to incorporate automatic layout facilities, be it semi-automatic or fully-automatic, to support model-driven engineering or interactive browsing of models [2]. Example tools are *LabVIEW* (National Instruments), *EHANDBOOK* (ETAS), *Simulink* (The MathWorks, Inc.), and *Ptolemy* (UC Berkeley).

For applications where *hierarchical dataflow diagrams* are used, the layout techniques have continuously been improved to handle most of the peculiarities of this type of diagram [13]. Still, further improvements are required regarding the compactness of the resulting drawings [6]. In this paper we show how the simple technique of *one-dimensional compaction* can be used to significantly improve the compactness of dataflow diagrams drawn with state-of-the-art methods (see Fig. 1 for a result). While we motivate our contributions from the perspective of dataflow diagrams, they are not restricted to this type of diagram. The presented methods are implemented as part of the open-source Eclipse Layout Kernel (ELK)[1].

One-dimensional compaction is a well-known technique to minimize the area occupied by a set of objects in the plane. As opposed to the NP-hard two-dimensional compaction problem, it can be solved efficiently in time $O(n \log n)$, n being the number of objects [8]. Given a set of rectangles $\mathcal{R}$, where every rectangle is of the form $r = (r_x, r_y, r_w, r_h)$, one seeks for a set of rectangles

[1] http://www.eclipse.org/elk.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 212–218, 2016.
DOI: 10.1007/978-3-319-42333-3_16



(a) **(b)**

Fig. 1. Illustration of our first contribution. (a) An automatically drawn dataflow diagram with the layer-based layout methods by Schulze et al. [13]. Circled nodes are pushed to the right due to the method's nature. (b) The same diagram after our post-processing. The diagram's width is reduced by about 16 % and the average edge length is reduced by over 50 %.

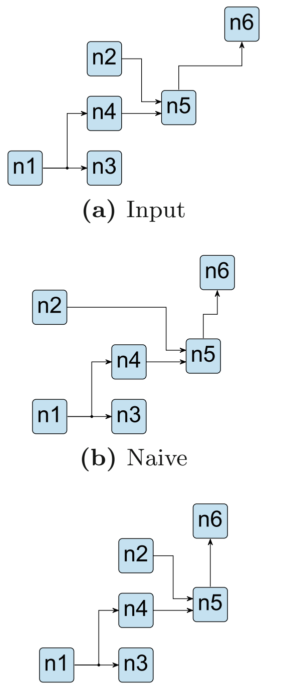
$\mathcal{R}'$ by changing x-coordinates only such that no pair of rectangles overlaps, the relative positioning is preserved, and the overall width w is minimized, with $w = |\ \max_{r' \in \mathcal{R}'}\ (r'_x + r'_w)\ -\ \min_{r' \in \mathcal{R}'}\ r'_x\ |$.

In the remainder of this paper we outline our contributions. Further implementation details, how to address the peculiarities of dataflow diagrams, and possible strategies for improvement, can be found in an accompanying technical report [12].

2 Layer-Based Drawings

In 1981, Sugiyama et al. described the structure of a successful methodology to draw directed graphs in the plane [14]. It is known under various names such as Sugiyama-style layout, hierarchical layout, and layer-based layout [7]. Essentially, it consists of five consecutive phases: (1) *cycle breaking* makes cyclic graphs acyclic by reversing edges, (2) *layering* assigns nodes to indexed layers such that edges always connect layers of lower index to higher index, (3) *crossing minimization* aims at reducing the number of edge crossings, (4) *node coordinate assignment* determines explicit y-coordinates for nodes, and (5) *edge routing* determines paths for edges and assigns x-coordinates to nodes. Most literature in this context assumes that nodes are of the same size. However, this is not the case with most practical applications, and it has been observed that the compactness of diagrams suffers in the presence of significant size differences [3,6]. Existing methods to tackle this issue either result in unpleasant drawings or increase the complexity of subsequent steps of the approach [3,9,10]. A common idea is to assign large nodes to multiple layers, for instance, by splitting them into multiple small chunks. The crossing minimization phase then has to keep edges from crossing nodes, and the node coordinate assignment has to assert that all chunks receive the same y-coordinate. Nonetheless, the problem becomes more and more imminent with diagram exploring approaches where nodes sizes may differ by factors of 10 or even 100 [2,6].

Recently, Schulze et al. presented several extensions to the layer-based approach to handle dataflow diagrams with *ports* (explicit attachment point of edges on a node's perimeter) and *orthogonally* routed edges [13], see Fig. 1. Working with these kinds of diagrams, we observed that scenarios where wide nodes (shaded background) prevent more compact placements are quite common. The problem illustrated in Fig. 1a is that the layer-based approach assigns nodes rigidly to layers and no pair of connected nodes may be placed in the same layer, thus pushing the set of small nodes in the lower right to the right. Here, one-dimensional compaction allows to reduce the diagram's overall width by breaking the rigid layering and pushing everything as far as possible to the left. During the process, vertical segments of orthogonally routed edges may be regarded as rectangles with either zero or very small width. Since the compaction procedure can be applied to the final drawing, after the traditional layer-based approach has finished completely, no additional complexity is added to any of the layer-based phases.



(a) Input

(b) Naive

(c) Desired

Fig. 2. Example of applying one-dimensional compaction to an input graph (a) with different objectives of minimal width and minimal edge length in (b) and (c).

Diagrams as the one seen in Fig. 1 can be formalized as *directed hypergraphs*, which are pairs $HG = (V, H)$. V is a set of nodes and $H \subseteq (P(V) \times P(V))$ a set of hyperedges that are connected to nodes via one of the nodes' *ports*. Schulze et al. represent each hyperedge $h = (S, T) \in H$ by a set of edges, i. e. for every pair $s \in S$ and $t \in T$ a directed edge $e = (s, t)$ is introduced. This allows to use known layer-based methods without the requirement to specifically address hyperedges. Both nodes and edges can carry labels that contribute to their bounding boxes. Additionally, a drawing must adhere to certain spacings between nodes and edges. After applying standard layer-based techniques, one-dimensional compaction can be applied to HG by transforming it into a set of rectangles $\mathcal{R}$: (1) the bounding box of every node $v \in V$ is added as a rectangle to $\mathcal{R}$. (2) For every vertical segment of an edge $e \in H$, we add a rectangle with corresponding height and unit width to $\mathcal{R}$. To guarantee enough room for edge labels they can be added to the set of rectangles as well. Still, to get satisfying results in practice, several subtleties have to be addressed. First, prescribed spacing values between diagram elements have to be maintained. This can be done by either enlarging the rectangles in $\mathcal{R}$ or by adding minimum separation constraints to the edges of the constraint graph. Second, with the extensions by Schulze et al. [13], edges are allowed to connect to the northern and southern border of a node. Consider the edge $e = (\mathsf{n5}, \mathsf{n6})$ in Fig. 2. Such vertical segments are *grouped* with the corresponding node during

Table 1. Results of applying one-dimensional compaction to layer-based drawings of dataflow diagrams. LR stands for left compaction followed by right compaction, and EL stands for compaction aiming for short edges. $\bar{n}$ and $\bar{e}$ denote the average number of nodes and edges. $\bar{w}$ denotes the average width after compaction in percent of the original width, $\bar{el}$ the average edge length. Standard deviations are given in brackets.

	$\bar{n}$	$\bar{e}$		$\bar{w}(\%)$	$\bar{el}(\%)$
EHANDBOOK	25.4 [15.0]	30.8 [18.3]	LR	83.7 [11.9]	78.2 [15.4]
			EL	85.3 [11.2]	76.6 [16.2]
Ptolemy	15.7 [7.4]	19.6 [11.5]	LR	93.6 [8.2]	88.3 [13.8]
			EL	94.3 [7.4]	87.1 [13.3]

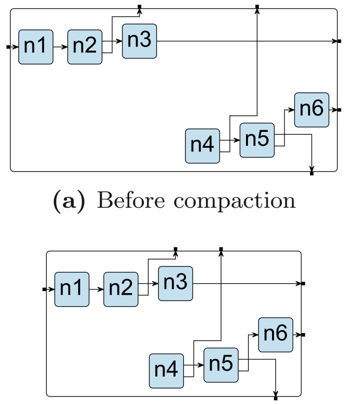
compaction, which prevents them from detaching. Third, edge lengths are not considered: compare the position of n2 in Fig. 2b and c. While this problem has been discussed in the literature [8], we suggest two simple solutions specifically tailored for graph layout: (a) Having compacted to the left, fix the positions of nodes that have no outgoing edge in the original graph, and execute another compaction pass to the right. This preserves minimum width with some edge length reduction. (b) Add the edges of HG to the constraint graph and use the adapted network simplex algorithm presented by Gansner et al. [4] to find a placement with minimum edge length. More details on all three points can be found in the accompanying technical report [12].

Our main goal is to improve on diagrams that occur in practice. Our evaluation set consists of 69 diagrams from the commercial interactive model browsing solution EHANDBOOK[2] and a subset of 529 diagrams shipping with the academic Ptolemy project [11]. Both diagram types are hierarchical: nodes can contain further nodes, i. e. sub-diagrams. We extracted such sub-diagrams and evaluate them separately, which is feasible since the layout algorithm considers every sub-diagram separately anyway. The results of applying our method can be seen in Table 1. We measured values for both compaction strategies previously mentioned: subsequent left-right compaction with node locking (LR) and minimizing edge length (EL). The average width of the EHANDBOOK and Ptolemy drawings decreased by about 16 % and 6 %, the edge lengths decreased by 22 % and 12 %. No significant difference can be observed between the two compaction strategies. Still, since edges that can obviously be shortened are immediately noticed by users (cf. Fig. 2), we suggest to use compaction with edge length minimization. Executed on an Intel i7 2 GHz CPU and 8 GB memory laptop, both methods finish in well under 10 ms for up to 100 nodes. For up to 1000 nodes EL's execution time increases significantly, which is expected since the network simplex algorithm is used. Still, it finishes in under 0.6 s. Therefore all setups are fast enough for applications that involve user interaction.

[2] http://www.etas.com/de/products/ehandbook.php.

3 Connected Components with External Extensions

When a diagram consists of multiple sub-graphs that are not connected among each other, see Fig. 3 for a simple example, the problem arises to place the sub-graphs in the plane such that little space is used. Each sub-graph can be approximated by its bounding box and the problem can be formulated as a rectangle packing problem. However, such problems are often NP-complete [8] and rectangles may be poor approximations. Freivalds et al. and Goehlsdorf et al. discuss relevant related work and present heuristics for the problem based on a *polyomino* representation, which approximates every sub-graph using squares on a grid [1,5]. The approaches work well for flat diagrams. With dataflow diagrams, a node can contain a sub-graph and nodes of the sub-graph can be connected to nodes on other hierarchy levels via so-called *external ports* on the hierarchical node's perimeter. When placing the sub-graphs in the plane these edges have to be considered. They are not allowed to cross other sub-graphs. This cannot be prevented using the previously mentioned methods. Furthermore, sub-graphs should be placed such that the overall length of external edges is as small as possible.



(a) Before compaction

(b) After compaction

Fig. 3. Placing a diagram's sub-graphs must assert that no external edge crosses a subgraph.

To better approximate a sub-graph, we construct its *rectilinear convex hull* and split it into a set of rectangles. Both can be done in $O(n \log n)$ time using a scanline method, where n is the number of points used to represent the area covered by a sub-graph in the first case, and the number of corners of the rectilinear convex hull in the second case.

Let $\mathcal{C}$ be a set of *components*, see also Fig. 4. Each component $c_i \in \mathcal{C}$ is a tuple $c_i = (\mathcal{R}_i, \mathcal{E}_i)$, where $\mathcal{R}_i$ is a non-empty set of *rectangles* and $\mathcal{E}_i$ is a (possibly empty) set of *external extensions*. Rectangles are 4-tuples (see Sect. 1). The k-th rectangle of c_i is r_i^k. We assume that all rectangles of the same component somewhere touch alongside their border. An external extension $e_i^l = (d_i^l, \delta_i^l, \epsilon_i^l)$ of a component c_i is a triple of a direction $d_i^l \in \{n, e, s, w\}$, an offset δ_i^l relative to r_i^0, and a width ϵ_i^l. The offset and the width describe an extension clockwise, i. e. for a south extension, the offset is its right-most point and the width points to the left. Intuitively it represents a line or a strip attached to the border of a rectangle which extends infinitely into the specified direction. See Fig. 4 for an illustration. We say an extension (d, δ, ϵ) is *horizontal* if $d \in \{w, e\}$ and *vertical* if $d \in \{n, s\}$. A set of components $\mathcal{C}$ is considered *proper* if no pair of components overlaps and no external extension overlaps a component.

Using compaction to minimize the area of a set of components requires a proper set of components to start with. We use the *cell packing* algorithm [12] that turns a (possibly improper) set of components into a proper one by calculating sensible x and y-coordinates for all rectangles.

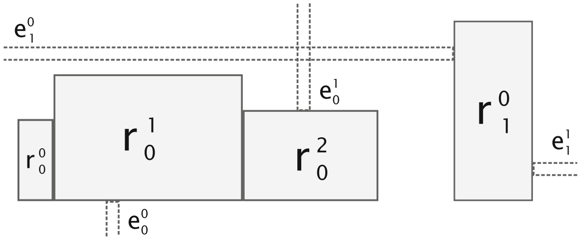


Fig. 4. The diagram shows two components c_0 and c_1. c_0 consists of three rectangles and two external extensions and c_1 consists of a single rectangle and two extensions. The external extensions e_1^0 and e_0^1 are allowed to overlap since one is vertical and the other one is horizontal. They are not, however, allowed to overlap with any of the rectangles.

For compacting layer-based drawings as described in Sect. 2, it is sufficient to compact along the x-dimension only. This time, however, it is necessary to compact in both dimensions, which is possible by continuously applying one-dimensional compaction in alternating dimensions and directions until no further, or little, progress is made. For a given set of components $\mathcal{C}$ we construct a grouped constraint graph (cf. [12]). Each component is represented by a group and the component's rectangles are added to it. The external extensions are converted into finite rectangles: each external extension is cut at the point where it intersects with the bounding box surrounding all components of $\mathcal{C}$. After each compaction pass these lengths have to be adjusted to prevent components from permuting.

Obviously, horizontal and vertical extensions that represent rectangles cannot be present at the same time during one-dimensional compaction since the set of rectangles may not be valid. Remember that external extensions are allowed to overlap with each other but the representing rectangles are not allowed to overlap. Still, it is important that the horizontal extensions are considered during vertical compaction, to prevent nodes from overlapping with external extensions; the same is true for vertical extensions during horizontal compaction. The independent application of horizontal and vertical compaction allows to use two different sets of rectangles depending on the compaction direction: $\mathcal{H} = \mathcal{R} \cup \{(d, \delta, \epsilon) \in \mathcal{E} : d \in \{n, s\}\}$ for horizontal compaction and $\mathcal{V} = \mathcal{R} \cup \{(d, \delta, \epsilon) \in \mathcal{E} : d \in \{e, w\}\}$ for vertical compaction.

4 Final Remarks

In this paper we show how one-dimensional compaction can be applied to two problems from the field of automatic diagram layout, more specifically, layer-based drawings and placement of disconnected graphs. We tested our methods with dataflow diagrams from practice and found that the width of layer-based drawings can significantly be reduced and that they allow disconnected graphs with hierarchy-crossing edges that are part of hierarchical graphs to be placed.

Acknowledgements. This work was supported by the German Research Foundation under the project *Compact Graph Drawing with Port Constraints* (ComDraPor, DFG HA 4407/8-1).

References

1. Freivalds, K., Dogrusoz, U., Kikusts, P.: Disconnected graph layout and the polyomino packing approach. In: Mutzel, P., Jünger, M., Leipert, S. (eds.) GD 2001. LNCS, vol. 2265, pp. 378–391. Springer, Heidelberg (2002)
2. Frey, P., von Hanxleden, R., Krüger, C., Rüegg, U., Schneider, C., Spönemann, M.: Efficient exploration of complex data flow models. In: Proceedings of Modellierung 2014, Vienna, Austria, March 2014
3. Friedrich, C., Schreiber, F.: Flexible layering in hierarchical drawings with nodes of arbitrary size. In: Proceedings of the 27th Australasian Conference on Computer Science (ACSC 2004), pp. 369–376. Australian Computer Society Inc. (2004)
4. Gansner, E.R., Koutsofios, E., North, S.C., Vo, K.-P.: A technique for drawing directed graphs. Softw. Eng. **19**(3), 214–230 (1993)
5. Goehlsdorf, D., Kaufmann, M., Siebenhaller, M.: Placing connected components of disconnected graphs. In: 6th International Asia-Pacific Symposium on Visualization, pp. 101–108, February 2007
6. Gutwenger, C., von Hanxleden, R., Mutzel, P., Rüegg, U., Spönemann, M.: Examining the compactness of automatic layout algorithms for practical diagrams. In: Proceedings of the Workshop on Graph Visualization in Practice (GraphViP 2014), Melbourne, Australia, July 2014
7. Healy, P., Nikolov, N.S.: Hierarchical drawing algorithms. In: Tamassia, R. (ed.) Handbook of Graph Drawing and Visualization, pp. 409–453. CRC Press (2013)
8. Lengauer, T.: Combinatorial Algorithms for Integrated Circuit Layout. Wiley, New York (1990)
9. Misue, K., Eades, P., Lai, W., Sugiyama, K.: Layout adjustment and the mental map. J. Vis. Lang. Comput. **6**(2), 183–210 (1995)
10. North, S.C., Woodhull, G.: Online hierarchical graph drawing. In: Mutzel, P., Jünger, M., Leipert, S. (eds.) GD 2001. LNCS, vol. 2265, pp. 232–246. Springer, Heidelberg (2002)
11. Ptolemaeus, C. (ed.): System Design, Modeling, and Simulation Using Ptolemy II. Ptolemy.org (2014). http://ptolemy.eecs.berkeley.edu/books/Systems/
12. Rüegg, U., Schulze, C.D., Grevismühl, D., von Hanxleden, R.: Using one-dimensional compaction for smaller graph drawings. Technical report 1601. Kiel University, Department of Computer Science, April 2016. ISSN: 2192–6247
13. Schulze, C.D., Spönemann, M., von Hanxleden, R.: Drawing layered graphs with port constraints. J. Vis. Lang. Comput. Spec. Issue Diagram Aesthetics Layout **25**(2), 89–106 (2014)
14. Sugiyama, K., Tagawa, S., Toda, M.: Methods for visual understanding of hierarchical system structures. IEEE Trans. Syst. Man Cybern. **11**(2), 109–125 (1981)

On Comments in Visual Languages

Christoph Daniel Schulze(✉), Christina Plöger, and Reinhard von Hanxleden

Deptartment of Computer Science, Kiel University, Kiel, Germany
{cds,cpl,rvh}@informatik.uni-kiel.de

Abstract. Visual languages based on node-link diagrams can be used to develop software and usually offer the possibility to write explanatory comments. Which node a comment refers to is usually not made explicit, but is implicitly clear to readers through placement and content. While automatic layout algorithms can make working with diagrams more productive, they tend to destroy such implicit clues because they are not aware of them and thus do not preserve the spatial relationship between diagram elements. Implicit clues thus need to be inferred and made explicit to be taken into account by layout algorithms. In this paper, we improve upon a previous paper on the subject [9], introducing further heuristics that aim to describe relations between comments and nodes. These heuristics mainly help to reduce the number of attachments of comments that should not be attached to anything. We also derive propositions on how developers of visual languages should integrate comments.

1 Introduction

Visual languages are in widespread use for developing software, either in addition to or at the expense of more traditional text-based languages. Languages such as ASCET (ETAS Group) or LabVIEW (National Instruments) allow developers to define software systems using *node-link diagrams* such as the one in Fig. 1: *nodes* (or *actors*) are entities that consume and produce data, which are transmitted between nodes through the *links* connecting them. Most visual languages support comments, usually in the form of special kinds of nodes that display text. However, finding out which node a comment refers to can be a challenge because of the two-dimensional nature of positioning them. Some languages solve this problem by allowing developers to explicitly attach comments to diagram elements and visualizing the attachment with a line. But not every language supports this feature, and not every developer uses it if it does. Developers often seem to rely on other, more implicit clues instead, such as the distance between comments and nodes. This falls into the category of secondary notation [5].

For diagrams to be understandable in the first place, their elements have to be carefully placed on the drawing area. Layout algorithms can reduce this effort by positioning nodes and routing edges automatically. However, unless a comment is explicitly attached to the node it refers to, the layout algorithm has no knowledge of their relation and they may end up in vastly different places

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 219–225, 2016.
DOI: 10.1007/978-3-319-42333-3_17

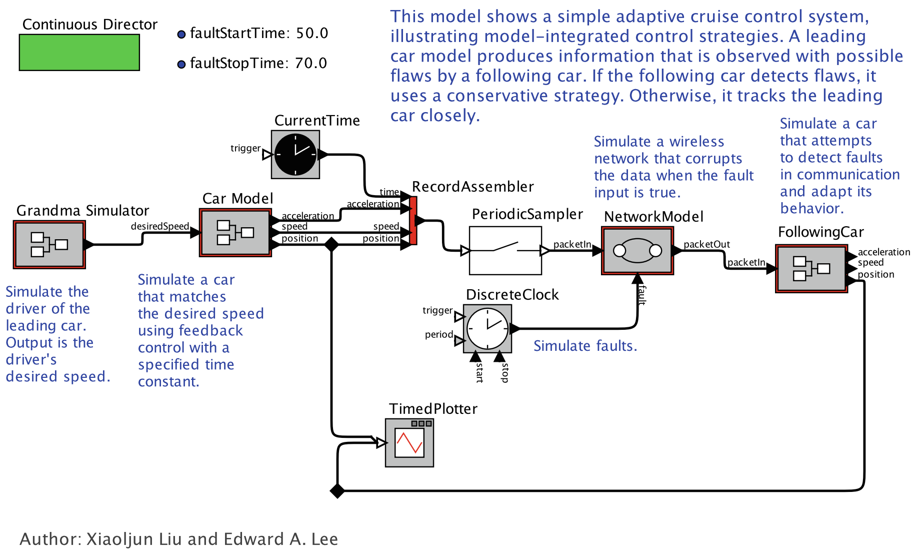


Fig. 1. A small node-link diagram as laid out manually using the Ptolemy language.

in the final layout. This wreaks havoc with the implicit clues that would have allowed a viewer to understand which node a comment refers to. This problem mainly applies to *layout creation algorithms*, which calculate new layout from scratch, as opposed to *layout adjustment algorithms*, which try to clean up an existing layout while preserving spatial relationships [4]. However, many layout algorithms used in practice fall into the former category. To use them in spite of their problems with comments, it is necessary to infer attachments between comments and nodes to make them explicit for the layout algorithm. This is what we introduced as the *comment attachment problem* in previous work [9].

Contributions. In previous work, we evaluated comment attachment based on the distance between comments and nodes [9]. This resulted in a lot of comments that should have been left unattached, but were attached to nodes by the presented algorithm, which is what we call *spurious attachments*. Here, we introduce a number of additional heuristics that ultimately serve to reduce the number of spurious attachments. We evaluate them on a set of Ptolemy diagrams and draw conclusions on how to properly integrate comments into visual languages. Note that for this paper, we limit outselves to attachments between comments and nodes and leave attachments between comments and other elements, such as edges and ports, for future work. More details are available in a technical report [7].

Use Case. Ptolemy is a visual language based on node-link diagrams developed at UC Berkeley [6] that supports comments in the form of nodes that contain text. The KIELER Ptolemy Browser[1] allows users to browse through a Ptolemy

[1] http://rtsys.informatik.uni-kiel.de/kieler.

model along with its submodels in a single window by dynamically expanding or collapsing nodes that contain further models, which requires automatic layout algorithms. The layout algorithm we use [8] is a layout creation algorithm based on the hierarchical layout method first introduced by Sugiyama et al. [10]; we therefore need comment attachment to keep comments close to the nodes they implicitly refer to.

Ptolemy ships with a set of demo models intended to showcase different models of computation, actors, and development techniques. 348 of them, created by 40 different developers, will serve as our main data set throughout this paper. Overall, the models averaged 21.4 nodes as well as 3.1 comments per model, of which 182 (about 17 %) refer to a specific node.

Related Work. We are not aware of any studies on how developers use comments in visual languages. However, the usage of documentation systems such as Javadoc have been studied, for example by Kramer [3], but the results do not seem to be applicable to our domain: Javadoc has clear rules on what comments refer to, which visual languages usually lack.

To the best of our knowledge, our previous paper is still the only one on the subject of inferring comment attachments in visual languages [9]. Eichelberger recognizes that comments can relate to different elements (or none at all) in UML class diagrams [2]. Other work based on textual languages, for example by Buse and Weimer on automatically augmenting Javadoc comments [1], also requires knowledge about relations between comments and code. However, the attachment rules for documentation in textual languages are usually clearly defined, not as ambiguous as in visual languages. With comment attachment, applications such as automatic handling of documentation may become viable for visual languages as well.

Outline. In the next section, we will introduce and discuss our comment attachment heuristics. We will then evaluate them and discuss the results in Sect. 3, trying to derive suggestions for developers of visual languages. We conclude the paper in Sect. 4 with open topics for future research.

2 Heuristics

We can distinguish two categories of comments: *node comments* refer to a specific node while *non-node comments* do not. Non-node comments can be further divided: *title comments* contain the title of a diagram, *author comments* contain the names of a diagram's authors, and *general comments* contain general information about a diagram not specific to any one node. The goal of any *automatic attachment algorithm* is to attach every node comment to the node it refers to while leaving non-node comments unattached.

In the following, we will introduce the basic idea of each heuristic. There are two kinds of heuristics: *filters* aim to recognize different types of non-node comments to prevent them from being attached to anything, and *regular heuristics* try to recognize node comments and attach them to the node they refer to.

Detailed evaluations of how well each heuristic describes the usage of comments in out main data set can be found in the accompanying technical report [7].

Font Size Filter. Text documents usually start with a title set in a larger font size than the rest of the text. One may well hypothesize title comments in diagrams to be set in a larger font size as well. The filter thus finds the set of comments with the largest font size. If the set only contains a single comment, it selects that as the title comment and thus filters it out, provided that it is not the only comment on the uppermost hierarchy level and that its font size exceeds the default font size.

Text Prefix Filter. Visual languages usually do not support special comment types for non-node comments, but it seems reasonable to hypothesize that they will often start with similar phrases. In our data set, these are phrases such as "This model" (general comments) and "Author:" (author comments). The text prefix filter thus filters out comments that start with such prefixes.

Area Filter. It seems reasonable to assume that general descriptions of what a program does will often be longer— and therefore larger— than more specific comments. The area filter thus filters out a comment if its area exceeds a certain threshold. We were, however, unable to find a good threshold value above which all comments can be considered non-node comments.

Node References Heuristic. If the name of a node appears in a comment, we consider this to be a *node reference*. If a comment contains such a node reference, it seems sensible to assume that it should be attached to that node. This ceases to be true once further references occur in the comment: since our use case only allows comments to be attached to a single node, we consider such comments to be general comments. If a node name appears exactly as is in the text of a comment, the heuristic attaches the two unless the comment contains the names of other nodes as well.

Distance Heuristic. The distance between a comment and a node may be the most obvious heuristic and was already examined in our first paper on the subject [9]. The hypothesis here is that the node a comment refers to is the one closest to the comment. The heuristic thus finds the node closest to a given comment and attaches the two unless their distance exceeds a predefined threshold.

Alignment Heuristic. In graphic design, alignment between elements is used as a means to establish a relationship between them. It seems reasonable to assume that comments are aligned to the node they should be attached to. For a given comment, the heuristic thus finds the node best aligned to it, possibly restricted to nodes within a certain maximum distance, and attaches the two unless the alignment exceeds a predefined threshold. As it turns out, though, alignment is actually not a very good predictor for attachments.

2.1 Discussion

Based on analyses we performed, it seems that established conventions such as a big font size or how the list of authors is to be included in a diagram work best for comment attachment. Other heuristics work to an extent (node references, distance) or have little predictive value (area, alignment). A considerable share of the information that helps link comments to nodes still seems to be in a comment's text, which is a lot harder to analyze.

There are two limiting factors to this analysis. First, the data set is smaller than we would like it to be. The number of diagrams is comparatively low (348), as is the number of authors that produced the diagrams (40). Also, the number of comments actually attached in our manual attachment (182 out of 1078, 17 %) is not that high. The second and more severe problem is that all diagrams were created as demonstration models for the Ptolemy tool to help explain how certain actors or models of computations are used and how to develop using Ptolemy; the heuristics that work well for this particular set of diagrams are not guaranteed to work well for another set. In fact, from looking at diagrams produced with other languages and by other developers it seems that we may not find a universally applicable set of rules for comment attachment. We feel confident, however, that our heuristics are a good starting point to analyse the usage of comments in visual languages.

3 Evaluation

To evaluate our heuristics, we compared the automatic attachments computed for our data set against a manual attachment. That attachment was produced by inspecting each comment and determining which node, if any, it refers to. If that was not clear, the model was removed from our data set.

During the evaluation, we look at which node each comment is attached to in both attachments. There are four cases:

Correct. A comment has the same attachment in both attachments.
Changed. A comment is attached to different nodes.
Lost. A comment is attached to a node in the manual attachment, but is not attached to anything in the automatic attachment.
Spurious. A comment is not attached to any node in the manual attachment, but is attached to a node in the automatic attachment.

Attaching comments based only on the distance heuristic (see Fig. 2a) yields results similar to previous results [9]: as the distance threshold is increased, the overall error rate increases mainly because the number of spurious attachments increases. The number of lost comments decreases as more comments are attached to nodes.

To reduce the number of spurious attachments, Fig. 2b shows the results of applying the two most unproblematic filters (the font size and text prefix filters) as well as the node reference and distance heuristics: if the node reference heuristic finds an attachment, that attachment is applied; otherwise, the distance heuristic is invoked. This significantly reduces the number of spurious attachments as the distance threshold is increased. More importantly, however, the node reference heuristic causes fewer lost attachments, at the expense of

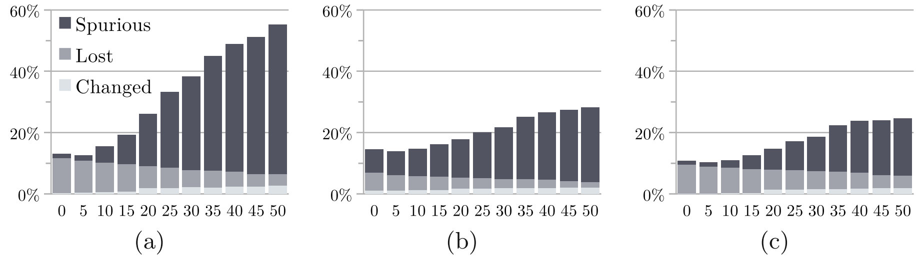


Fig. 2. Results of comparing automatic attachments to the manual attachment, all involving the distance heuristic subject to different threshold values. (a) Distance heuristic only [9]. (b) Font size filter, text prefix filter, node reference heuristic, and distance heuristic. (c) As (b), but with a maximum distance threshold of 30 imposed on the node reference heuristic and with the area filter with a conservative setting.

more spurious attachments in the lower threshold areas. We think that this is a worthwhile tradeoff, since a node attached to a comment by the node reference heuristic is at least mentioned in the comment.

In an attempt to further reduce the number of spurious attachments caused by the node reference heuristic, Fig. 2c shows the results of applying a distance threshold of 30 to the node reference heuristic, and of engaging the area filter with a very conservative setting. This decreases the amount of spurious attachments and the error rate overall to about 10 % at best, at the expense of fewer found correct attachments as the number of lost attachments increases.

3.1 Discussion

The effectiveness of comment attachment largely depends on a good configuration of the heuristics. These results suggest that comment attachment should be replaced by proper support for explicit attachments in visual languages. However, comment attachment stays relevant for browsing scenarios similar to our use case, for languages that do not provide explicit attachments, or when users do not make use of them.

The latter problem seems most relevant to integrating comments into visual languages. A lack of proper attachments can prevent tool developers from making more advanced features available, such as good automatic layout, semantic reasoning, or even generating documentation. The best solution may be twofold. First, provide different kinds of comments, such as general comments, author comments, and node comments. Dragging a node comment onto the drawing area could then include displaying "attachment lines" that indicate which node the tool will interpret the comment to refer to, thus forcing explicit attachments.

The addition of such features to existing tools offers another area of application for comment attachment. Opening diagrams that do not make use of explicit attachments would trigger comment attachment and present the user with an automatically inferred attachment that they can then modify.

4 Conclusion

Building on our previous work on the subject, we introduced more heuristics to describe comment usage that varied in how well they perform. We used those heuristics to improve upon previous attachment results, mainly by keeping more non-node comments from being attached. Use cases include automatic layout as well as semantic reasoning about programs written in visual languages, as has already been explored in the context of textual languages.

Regarding future work, it first seems necessary to analyse the usage of comments in more visual languages and to compare the results. It also seems worthwhile to survey users of visual languages as to whether they use any deliberate conventions when writing and placing comments. Second, the attachment framework as well as the heuristics will have to be extended to support comments attached to diagram elements other than nodes. And third, comments sometimes describe a whole group of nodes. It seems extremely hard to infer such group attachments, but this intuition needs confirmation.

References

1. Buse, R.P., Weimer, W.R.: Automatic documentation inference for exceptions. In: Proceedings of the 2008 International Symposium on Software Testing and Analysis, ISSTA 2008, pp. 273–282. ACM (2008)
2. Eichelberger, H.: Aesthetics and automatic layout of UML class diagrams. Ph.D. thesis, Bayerische Julius-Maximilians-Universität Würzburg (2005)
3. Kramer, D.: API documentation from source code comments: a case study of Javadoc. In: Proceedings of the 17th Annual International Conference on Computer Documentation, SIGDOC 1999, pp. 147–153. ACM (1999)
4. Misue, K., Eades, P., Lai, W., Sugiyama, K.: Layout adjustment and the mental map. J. Vis. Lang. Comput. **6**(2), 183–210 (1995)
5. Petre, M.: Cognitive dimensions beyond the notation. J. Vis. Lang. Comput. **17**(4), 292–301 (2006)
6. Ptolemaeus, C. (ed.): System Design, Modeling, and Simulation Using Ptolemy II. Ptolemy.org (2014). http://ptolemy.eecs.berkeley.edu/books/Systems/
7. Schulze, C.D., Plöger, C., von Hanxleden, R.: On comments in visual languages. Technical report 1602, Christian-Albrechts-Universität zu Kiel, Department of Computer Science, April 2016. ISSN: 2192–6247
8. Schulze, C.D., Spönemann, M., von Hanxleden, R.: Drawing layered graphs with port constraints. J. Vis. Lang. Comput. Spec. Issue Diagram Aesthetics Layout **25**(2), 89–106 (2014)
9. Schulze, C.D., von Hanxleden, R.: Automatic layout in the face of unattached comments. In: Proceedings of the IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC 2014), Melbourne, Australia, July 2014
10. Sugiyama, K., Tagawa, S., Toda, M.: Methods for visual understanding of hierarchical system structures. IEEE Trans. Syst. Man Cybern. **11**(2), 109–125 (1981)

Bistable Perception and Fractal Reasoning

Keith McGreggor(✉) and Ashok Goel

College of Computing, Georgia Institute of Technology, Atlanta, GA, USA
keith.mcgreggor@gatech.edu, ashok.goel@cc.gatech.edu

Abstract. A visual percept is deemed bistable if there are two potential yet mutually exclusive interpretations of the percept between which the human visual system cannot unambiguously choose. Perhaps the most famous example of such a bistable visual percept is the Necker Cube. In this paper, we present a novel computational model of bistable perception based on visual analogy using fractal representations.

1 Bistable Perception

A visual percept is deemed bistable if there are two potential yet mutually exclusive interpretations of the percept between which the human visual system cannot unambiguously choose. Alternation between the available interpretations appears to happen in a spontaneous and stochastic manner [3, 6]. Perhaps the most famous example of a bistable visual percept is the Necker Cube, as originally illustrated by Necker, reproduced in Fig. 1 [7]. As Necker notes, although the figure is drawn to indicate that the solid angle labeled A should be seen as closest and the solid angle X should be seen as furthest, one's perception of the figure will shift involuntarily to cause the opposite interpretation.

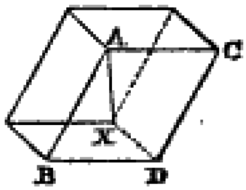


Fig. 1. The Necker cube [7].

Psychologists, cognitive scientists and neuroscientists have been fascinated by the phenomenon and potential causes of bistable perception. Computational models of the Necker Cube problem also have been developed including localist connectionist models [8] and autoassociative models [2].

Recently, we have explored problems of visual analogy via reasoning condoned by a fractal representation of the visual information [4, 5]. Here, we present how an algorithm based upon the fractal representation would perform when considering the Necker Cube problem.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 226–231, 2016.
DOI: 10.1007/978-3-319-42333-3_18

2 The Necker Cube as a Fractal Visual Analogy Problem

Fractal representations are dependent upon a particular partitioning scheme [5]. In this experiment, we wish to vary the nature of the partitioning scheme, examining the relationships at a variety of levels of abstraction. Our motivation is to examine the certainty (or rather, uncertainty) in conclusively determining an interpretation of the Necker Cube at those various levels of abstraction, for a bistable perception would be one which would remain uncertain regardless of the level of abstraction with which the problem is regarded.

We set up our experiment in the following manner. First, we created a very exact rendition of the Necker Cube. This target cube image, which we label as NC, is shown in Fig. 2. We then created, from that original drawing, three sets of alternative visual interpretations of the cube, each set containing two interpretation choices: C_1, an image with the forward face lowermost, and C_2, an image with the forward face uppermost.

Each set maintained the same isometric projection as the target cube, but in each set, a visual cue was embedded to suggest which face was forward. In Set 1, a technique known as "haloed lines" was used [1]. In Set 2, the edges that are to be interpreted as "behind" are rendered in a slightly different color. In Set 3, the occluded edges are removed entirely, leaving an impression of a solid cube. Figure 4 illustrates the alternative pairs we created.

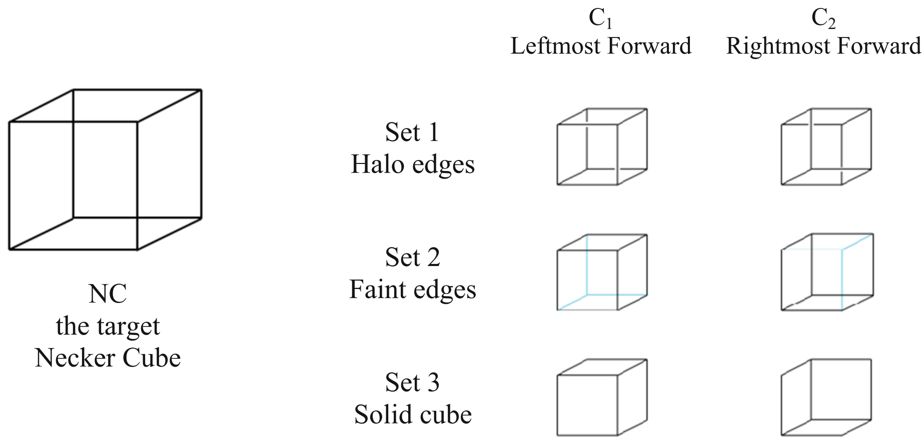


Fig. 2. The Necker cube and sets of alternative interpretations.

Since our algorithm computes the similarity between visual analogies and uses fractal representations, we created for each set with three fractal relationships. The first relationship (R) was between the target Necker Cube (NC) and itself (to establish a self-referential identity). The second relationship (R_1) was between NC and the test set's C_1 image, and the third (R_2) was between NC and the test set's C_2 image. To which of the two relationships, R_1 or R_2, is the R relationship most similar?

PROBLEM SEGMENTATION

By examination, the set of interpretations are individual images.

Let NC be the target Necker cube image.
Let C := { C_1, C_2, ... } be the set of individual interpretations.

RELATIONSHIP DESIGNATIONS

Let R be a relationship, determined by a fractal representation, as follows:
R ← MutualFractal(NC, NC)

ABSTRACTION LEVEL PREPARATION

Let d be the maximum pixel dimension of any image in the set { NC } ∪ C.

Let δ be the abstraction decrement value where $1 \leq \delta \leq d$.

Let A := { a_1, a_2, ... } represent an ordered range of abstraction values where
$a_1 \leftarrow d$, and $a_i \leftarrow (a_{i-1} - \delta)$ $\forall$ i, $2 \leq i$ and $a_i > 2$

The values within A constitute the grid values to be used when partitioning the problem's images.

Algorithm 1. The Fractal Necker Algorithm, preparatory stage.

The algorithm for calculating Necker analogies for each set, the Fractal Necker algorithm, is given below. In the preparatory phase, a Necker Cube problem is first segmented into its component images (the target image NC, and the collection of interpretation images). Next, the algorithm determines the relationship between NC and itself, expressed as a mutual fractal representation. Then, a range of abstraction levels is determined. The abstraction levels are determined to be a partitioning of the given images into gridded sections at a prescribed size and regularity. In this experiment we wished to note the circumstances under which the algorithm would prefer one or the other alternative interpretation of the target Necker Cube.

In the exploratory and re-representation phase, the algorithm concludes by determining the confidence in the answers at each level of abstraction [4]. Thus for each level of abstraction, the relationship R is re-represented into that partitioning. Then, for each of the candidate images, a potentially analogous relationship is determined and a similarity value is calculated. The balance of the fractal algorithm, using the deviation from the mean of these similarities, continues through a variety of levels of abstraction, looking for an unambiguous answer that meets a specified confidence value. However, for our experiment, we wanted to examine the confidence values present at all levels of abstraction: we did not wish for the algorithm to halt if one of the interpretations exceeded the confidence threshold. Thus, we set the confidence level artificially high (100 %). This caused the algorithm to proceed to calculate similarity values at all levels of abstraction.

Given NC, C, R, and A as determined in the preparatory stage,
determine which $c \in C$ best interprets NC.

DEFINITIONS

Let E be a real number which represents the number of standard deviations beyond which a value's answer may be judged as "confident"

Let Similarity(X,Y) be the Tversky similarity metric for sets X and Y, where $0.0 \leq \text{Similarity}(X,Y) \leq 1.0$ and $\text{Similarity}(X,X) = 1.0$.

EXECUTION

For each abstraction $a \in A$:

- Re-represent each fractal representation $r \in R$ according to abstraction a
- Let $S \leftarrow \varnothing$
- For each answer image $C_i \in C$:
  $R_i \leftarrow$ MutualFractal(NC, C_i) according to abstraction a
  $S \leftarrow S \cup \{ \text{Similarity}(R, R_i) \}$
- Set $n \leftarrow |S|$
- Set $\mu \leftarrow$ mean (S)
- Set $\sigma_\mu \leftarrow$ stdev (S)/$\sqrt{n}$
- Set $D \leftarrow \{ D_1, D_2, \ldots D_n \}$ where $D_i = (S_i-\mu)/\sigma_\mu$
- Generate the set $Z := \{ Z_i \ldots \}$ such that $Z_i \in D$ and $Z_i > E$
- If $|Z| = 1$, stop and return the answer image $C_i \in C$ which corresponds to Z_i

Algorithm 2. The Fractal Necker Algorithm, execution stage.

3 Results

We ran the algorithm on each of the three sets given above. As indicated above, the algorithm calculated similarity values for all of the available levels of abstraction, beginning with the coarsest image resolution (200 × 200) and proceeding in a regular fashion down to the very finest (5 × 5). At each level of abstraction, the similarity value for each of the possible interpretations is calculated, using the Tversky formula [9], and set alpha to 1.0 and beta equal to 0.0. From those values, the algorithm calculated the mean and standard deviation, and then calculated the deviation and confidence for each answer.

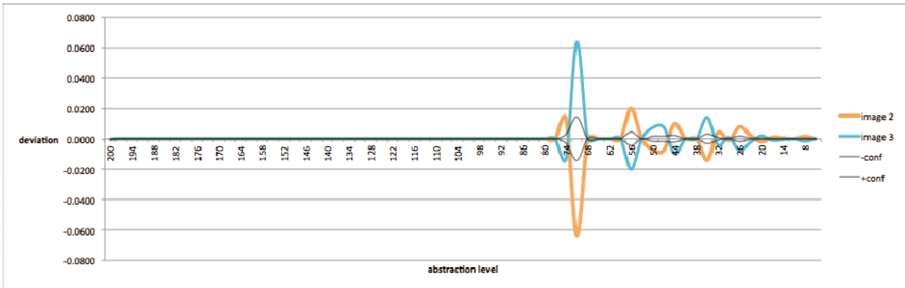


Fig. 3. Deviation oscillations for Set 1.

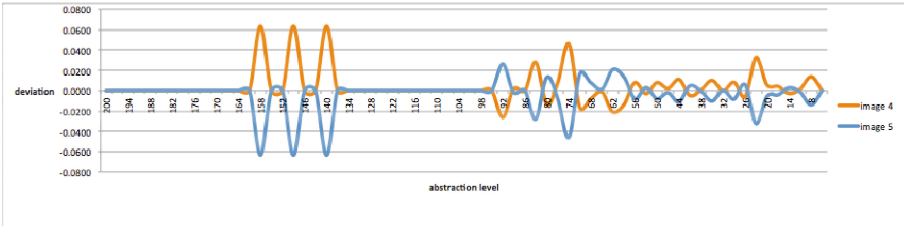


Fig. 4. Deviation oscillations for Set 2.

Intriguingly, the algorithm showed a clear instability in its ability to choose between either of the alternative interpretations for each set of the Necker problems tested. In fact, in no case was there any preference for either interpretation, which was determined unambiguously, even though the confidence values for the interpretation exceeded that corresponding to a confidence of 95 % for a sample set of two. Thus, we claim that in this experiment our computational model of the Necker Cube problem, based upon images represented fractally, exhibits bistable perception.

The charts in Figs. 3, 4 and 5 plot the deviation of the interpretation similarity values against the level of abstraction, from coarsest to finest, for each of the sets, plainly showing the oscillation between interpretations. We note that there are occasional oscillations in the deviation at some coarse levels of abstraction, particularly apparent in Sets 2 and 3. Then, a regular pattern of oscillation appears to occur in each

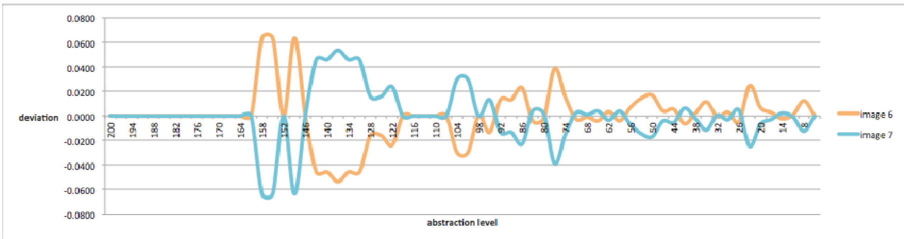


Fig. 5. Deviation oscillations for Set 3.

set after the abstraction level dips below 100 × 100. We attribute some of this to the manner in which the fractal representation is calculated: at a given partitioning level, an empty, temporary image buffer is calculated which is an even multiple of the partitioning in both directions, and then the image is composited into the center of that temporary buffer prior to the calculation of the fractal representation.

The oscillations present in Sets 2 and 3 suggest this, but Set 1's chart does not. Our interpretation is that the haloed line effect used in Set 1 is not a remarkable feature within the image until the partitioning reaches a lower limit; thus, in Set 1 the deviations remain almost perfectly flat for much of the coarse abstractions.

4 Implications

We have described that fractal encoding can generate the features for calculating relational similarity at multiple levels of resolution. We have demonstrated that these properties of fractal representations enable a parsimonious explanation of bistable perception in the Necker Cube problem. We showed that even when presented with sets of potential interpretations with varying visual cues, the fractal model of bistable perception exhibits an inability to determine an unambiguous and significant interpretation of the Necker cube's orientation. To us, this suggests that the analogical reasoning afforded by the fractal representation and illustrated via the Fractal Necker algorithm may offer insights into the gestalt perceptual capabilities of humans.

References

1. Appel, A., Rohlf, J., Stein, A.: The haloed line effect for hidden line elimination, vol. 13. ACM (1979)
2. Kawamoto, A., Anderson, J.: A neural network model of multistable perception. Acta Psychol. **59**, 35–65 (1985)
3. Kogo, N., Galli, A., Wagemans, J.: Switching dynamics of border ownership: a stochastic model for bi-stable perception. Vis. Res. **51**(18), 2085–2098 (2011)
4. McGreggor, K., Goel, A.: Confident reasoning on the ravens progressive matrices test. In: Proceedings of the 25th AAAI Conference on Artificial Intelligence, Quebec (2014)
5. McGreggor, K., Kunda, M., Goel, A.: Fractals and ravens. Artif. Intell. **215**, 1–23 (2014)
6. Nagao, N., Nishimura, H., Matsui, N.: A neural chaos model of multistable perception. Neural Process. Lett. **12**(3), 267–276 (2000)
7. Necker, L.: On an apparent change of position in a drawing or engraved figure of a crystal. Lond. Edinb. Philos. Mag. J. Sci. **1**, 329–337 (1832). (Third Series)
8. Rumelhart, D., Smolensky, P., McClelland, J., Hinton, G.: Schema and sequential thought processes in PDP models. In: Rumelhart, D., McCleland, J. (eds.) Parallel Distributed Processing, pp. 7–57 (1986)
9. Tversky, A.: Features of similarity. Psychol. Rev. **84**(4), 327–352 (1977)

Diagrams Layout

Who, Where, When and with Whom? Evaluation of Group Meeting Visualizations

Simone Kriglstein[1(✉)], Johanna Haider[1], Günter Wallner[2], and Margit Pohl[1]

[1] Vienna University of Technology, Vienna, Austria
simone.kriglstein@tuwien.ac.at, {johanna.haider,margit}@igw.tuwien.ac.at
[2] University of Applied Arts Vienna, Vienna, Austria
guenter.wallner@uni-ak.ac.at

Abstract. Visualizing time-dependent and location-based data is a challenging problem but highly relevant for areas like intelligence analysis, traffic control, or social network analysis. In this context, we address the problem of visualizing meetings between persons, groups of persons, vehicles, or other entities. However, the temporal dimension inherent in such data makes traditional map representations less well suited for this kind of problem as they easily become cluttered. To overcome this issue we developed a modified map representation and three alternative representations (two matrix-based visualizations and one based on Gantt charts). An empirical evaluation comparing these four visualizations and assessing correctness, recognition rates of groups, and subjective preference indicates that the alternative visualizations perform significantly better than the map-based representation when meetings need to be identified. In addition, we identify specific strengths and weaknesses of the investigated visualizations and propose design considerations.

Keywords: Information visualization · Map · Matrix · Gantt charts

1 Introduction

The analysis of spatio-temporal data to identify patterns of movements (e.g., movements of persons or groups of persons, vehicles, etc.) is of great interest in several domains such as location-based social networks analysis (e.g., [5,15]), crime analysis (e.g., [4]), or movement pattern analysis (e.g., [3]). However, analysts and decision-makers are often confronted with an enormous amount of multidimensional information that can be extracted from spatio-temporal data. In order to make these large amounts of data manageable there is a need to develop effective visualization methods. Especially in case of the analysis of group meetings not only questions relating to *when* and *where* are of interest but also more complex questions such as *who* meets with *whom*. For representing spatial data, map representations are very popular to show different types of geographic information in an intuitive way. Such maps can support analysts to deduce associations and connections from the spatio-temporal data (e.g., population density versus recreation areas) [17]. However, the temporal dimension is problematic

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 235–249, 2016.
DOI: 10.1007/978-3-319-42333-3_19

with map representations as a single map usually only represents a single slice in time. Yet, this temporal aspect is vital for understanding meeting patterns of various people.

Addressing this challenge, we designed alternative visualizations in the course of two projects to facilitate the analysis of location data of a group of individuals. It was important for us to find a possibility to answer *where*, *when*, *who* and with *whom* questions with a map representation but we also investigated alternative representations: two matrix-based visualizations and one adapted from Gantt charts. Matrix visualizations simultaneously show the relationships between multiple variables while at the same time being free from occlusions [6] whereas Gantt charts are usually well suited to show activities displayed against time. Therefore, we believe that these representations could be suitable alternatives to the traditional map representation. To assess and compare these representations – henceforth referred to as *Map*, *Gantt*, *Matrix*, and *Augmented Matrix* – with respect to correctness, recognition rates of groups, individual preference, and their utility for identifying and interpreting meetings of persons we conducted an empirical evaluation with 24 subjects. Our results show that the three alternative visualizations (*Augmented Matrix*, *Matrix*, *Gantt*) perform better than the *Map* representation and have also been preferred by the participants for identifying meetings. However, although *Augmented Matrix*, *Matrix*, and *Gantt* have their strengths, the results also indicate certain weaknesses. Based on the results, we derive design considerations for designing time-dependent location-based visualizations. The goal of this study is not to evaluate a specific system, but to provide general information on what kind of visualization is appropriate for identifying meetings between entities.

2 Related Work

The map is the predominant visualization for spatial data, as the timeline is for temporal data. Nevertheless, the visualization of spatio-temporal data is a challenging problem [8] since a map easily gets cluttered when several points in time are represented on a single map. To compare multiple time slices several maps can be displayed as small multiples, time can be represented by adding a third dimension, or animations can be used to visualize changes between the different slices in time. However, these solutions can be suboptimal. For example, while small multiples offer the analyst multiple time slices at once the number of maps which can be displayed simultaneously is limited by the available screen space. Moreover, small multiplies can make it difficult to understand how a map evolves over time [7]. Although these issues can be addressed by employing animation, evaluation studies have shown that animation can lead to confusion when too many data points move simultaneously which, in turn, can cause analysts to miss relevant information (cf. [13]). Many geo-visualization tools (e.g., [12]) use the concept of *Space Time Cubes* [2] to visualize time-dependent movement data inside a cube and where the height axis is used to represent time. However, in the visualization research community there is some discussion about the possible implications of 3D representations (see, e.g., [1]). Especially if spatio-temporal data needs to be analyzed with maps it seems that 2D visualizations are less error-prone for simple tasks (cf. [11, 14]) compared to 3D representations.

There exist only a few approaches for the representation of movement data where the map is not the main visualization. Two such approaches are described below. The first, concerned with visualizing movement data of players of mobile outdoor games and proposed by Orellana et al. [18], shares some similarities with the Gantt charts we used in our study. They especially emphasize the interaction between different players such as meetings of two or more players. The second, proposed by Shen and Ma [20], augments adjacency matrices with a path visualization to facilitate path finding. In our study, we use Gantt charts as well as matrices as possibilities to represent meetings. Gantt charts are commonly used to support planning activities [1] and are especially suited for the representation of the interactions between different processes. This is a problem which can be compared to the meeting of several people at the same point in time. In general, Gantt charts are appropriate and usable for the representation of events in time [16]. Matrix-based representations are also a possibility to show meetings of several persons. Ghoniem et al. [6] compared matrix-based visualizations to a node-link visualization and found out that matrices are preferable when graphs become bigger than 20 nodes. On the other hand, matrices make it more difficult to detect paths, an issue addressed by Shen and Ma [20] (see above) whose approach influenced our *Augmented Matrix* visualization. Kessell and Tversky [10] conducted a study which has some similarities to our approach. They also address the issue of the combination of space, time and persons/objects. They argue that matrices and line visualizations are highly appropriate for this kind of tasks. Their results are that in general matrices are preferable to line visualizations. However, there are some significant differences between their study and our approach in that our tasks are much more exploratory in nature and to support that we used more complex data sets. Nevertheless, there are some similarities between the results of the two studies and the visualizations used.

To sum up, existing literature on visualizing time-oriented data indicates that 3D, animation, and small multiples can be error-prone and difficult to analyze. Possible alternatives are 2D visualizations based on a timeline metaphor as described by Kessell and Tversky [10].

3 Time-Dependent Location-Based Visualization

For our study we have chosen four visualizations – *Map*, *Gantt*, *Matrix*, and *Augmented Matrix* (see Fig. 1) – that use different characteristics to encode two major attributes of time-dependent location data: a) the position of a person and b) the point in time associated with this position. The four visualizations are briefly described below.

Map. The presence of people at a certain location is marked by a circle located at the $x-$ and $y-$coordinates associated with the location on a two-dimensional map. Each person is represented with a unique color. If a location is visited multiple times or by different persons then the circle is split into evenly sized sectors, with each sector representing one visit of a person and being colored based on the color associated with that person. In other words, if a person visits the same location multiple times then that person also occupies as many sectors. The number of visits to a location is also reflected by the size of the circle. Next to each sector the time period of the stay is depicted. In addition, the name of the location is displayed next to each circle. For example, in the map shown in Fig. 1(a) Person 1 was at location *B* from 6:00 to 8:00 and at location *A* from 5:00 to 6:00, at the same time frame when Person 2 was there.

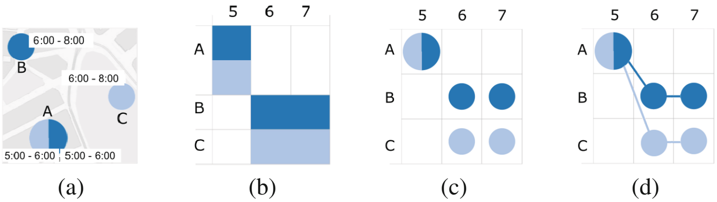


Fig. 1. Comparison between the tested visualizations *Map* (a), *Gantt* (b), *Matrix* (c), and *Augmented Matrix* (d). *A*, *B*, and *C* denote locations and 5, 6, and 7 denote hourly time intervals. Different persons are represented by different colors: ■ Person 1 ■ Person 2

Gantt. Gantt charts [22] help to convey schedules by illustrating the start and finish dates of tasks of a project. We adapted this method for the representation of time-dependent location data. The x-axis represents one-hour intervals and locations are listed alphabetically along the y-axis. Each person is represented by a colored bar and the position and length of the bar reflects the time of arrival and departure, i.e., the duration of stay of a person at a location. For example, in Fig. 1(b) Person 1 is at location *B* between 6:00 and 8:00 and Person 2 is at location *C* between 6:00 and 8:00. Persons in the same location are drawn beneath each other (cf. location *A* from 5:00 to 6:00 in Fig. 1(b)).

Matrix. Matrix visualizations [23] are a further promising visualization to depict the presence and strengths of relationships between different variables in a compact way. In our case each row corresponds to a single location and each column represents a one-hour interval. The presence of a person at a specific location i during a specific hour j is marked by placing a circle at cell (i, j). If multiple persons are present at the same location during the same interval then the circle is divided into equally sized parts. For example, in Fig. 1(c) Person 2 is at location *C* from 6:00 until 8:00 whereas both, Person 1 and Person 2, have been at location *A* from 5:00 to 6:00.

Augmented Matrix. Matrix visualizations have the advantage to be free of visual clutter but are not well suited for path finding (cf. [6]). To overcome this limitation Shen and Ma [20] augmented adjacency matrices with a path visualization to facilitate the easy tracing of paths. Influenced by their approach we extended the above described matrix representation with lines to explicitly show the movements of persons between different locations (see Fig. 1(d)). These lines are color-coded to show which person has changed the location.

4 Empirical Evaluation

The goal of our study was (a) to clarify the advantages and disadvantages of *Map*, *Gantt*, *Matrix*, and *Augmented Matrix* for the purpose of identifying meetings of persons and (b) to assess correctness and subjective preference. For this purpose, the study aimed to address the following research questions:

RQ1 - Correctness: Which visualization is interpreted more correctly by the participants?

RQ2 - Recognition Rate: Have the properties of the groups (specifically, the number of people in a group, number of meetings, and the total amount of time a group spent together) an influence on the recognition rate?

RQ3 - Preference: Do participants prefer the *Map*, *Gantt*, *Matrix*, or *Augmented Matrix*?

RQ4 - Saliency: Which groups are perceived as salient by the participants? Are there differences between *Map*, *Gantt*, *Matrix*, and *Augmented Matrix*?

At this point we should also note that the emphasis of this study is to test possible visualizations independently of interactions. Testing visualizations and interactions in conjunction would confuse the relationship between independent and dependent variables. It would not be possible to clearly assess whether it is the visualization or the interaction which produces positive or negative effects. We intend to study interactions influencing the perception and interpretations of meetings in the future. Another possibility to improve such visualizations would be to compute meetings algorithmically and provide a list of these meetings to the users. A problem in this context is that if there are many meetings the list will be quite confusing. Nevertheless, it could be a valuable addition to such visualizations to provide such a list. Again, we did not add this feature because we did not want to confound the results.

Tasks. For our study, different types of tasks were designed to identify strengths and weaknesses of the different visualizations. These tasks should reflect common questions that can occur during the analysis of groups and their movement patterns (e.g., how long does a group stay at the same location; do groups repeatedly change locations to meet). We generalized these tasks in a way that they are also applicable in and relevant to other domains. In summary, the following three tasks were developed to investigate our research questions:

Task 1 - Duration: Participants were asked to identify groups that met in the same constellation[1] for more than one hour at one location.

Task 2 - Location Changes: Participants had to list groups that met in the same constellation (see footnote 1) at different locations for at least one hour each.

Task 3 - Salient Groups: Participants were instructed to list up to three groups which they perceived as striking in some sort of way and to provide an explanation why they thought these groups were special.

Test Cases. The dataset we used consisted of four days and contained visits to 26 different locations (labeled *A-Z*) from 12 people (color-coded). In addition, we ensured that each person is at least once at the same time at the same location as another person. The colors for encoding the different persons were selected based on a color-scheme for qualitative data from ColorBrewer [9]. For each day a *Map*, a *Gantt*, a *Matrix*, and

[1] That is, groups that consist of exactly the same persons. If a person joins or leaves a group then it is considered a different group.

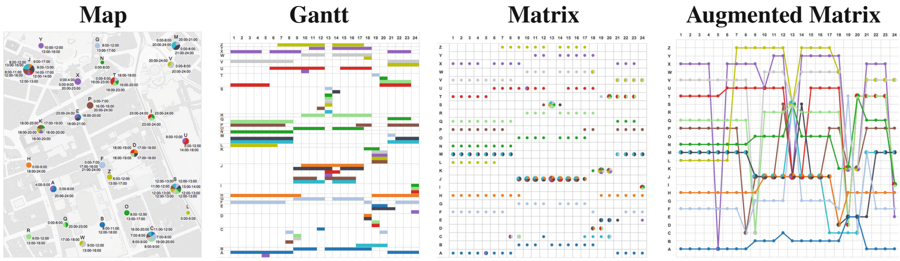


Fig. 2. One of the four test cases which were used in the study (Test Case 1).

an *Augmented Matrix* visualization was created by using Tableau[2] (see Fig. 2). Based on these four days we built four different test cases: Test Case 1 and Test Case 2 consisted of the visualizations of a single day, whereas Test Case 3 and Test Case 4 were composed of the visualizations of two consecutive days displayed next to each other. The number of correct groups to identify varied for the different test cases and between Task 1 and Task 2, resulting in a total of 60 groups. In addition, the number of people (from 2 to 5 people, $M = 2.5$, $SD = 0.74$), the number of meetings (from 1 to 7, $M = 1.83$, $SD = 0.94$), and the duration of the meetings (ranging from 1 to 8 hours, $M = 2.2$, $SD = 1.55$) varied from group to group.

Procedure. To address our research questions, we decided to use an online questionnaire[3] which was created using LimeSurvey[4]. We decided on a survey in order to reach a broader audience since this way the participants could partake in the study from their home anytime without time pressure. The survey included (a) closed questions using, for example, checkboxes, (b) open-ended questions to offer the participants the possibility to explain their decisions, and (c) rank-ordered questions to rank their preferences. It started with general questions (including age, gender, and familiarity with visualizations on a five-point scale) followed by questions concerning the presented visualizations. These visualization questions included the three tasks which each participant had to solve with each of the four visualizations for each of the four test cases (yielding 48 questions in total, within-subject design). Each visualization and test case was presented on a single page which contained all three tasks. If the test case consisted of two days then the visualizations were juxtaposed horizontally. Test cases and visualizations were presented in an arbitrary sequence. Finally, participants were asked to rank the visualizations by preference. We also conducted a pre-test with five participants to ensure that questions were understandable and tasks could be completed within a reasonable time. As participants reported loss of concentration and focus during the pre-test we split the survey into two parts of one hour each (consisting of Test Case 1 and 2 and Test Case 3 and 4, respectively), with the second part being conducted one week after the first part.

[2] http://www.tableau.com/ (Accessed: January, 2016).
[3] Complete list of questions and test cases: http://igw.tuwien.ac.at/groupviz/.
[4] http://www.limesurvey.org (Accessed: January, 2016).

Sample. The invitation to participate in the online survey was sent via email to Computer Science students who were considered to have at least basic knowledge with visualizations. In total, 24 participants between 23 and 40 years ($M = 27.6, SD = 3.6$) responded to the survey. One third of the participants was female and 16 participants were male. More than half of the respondents considered themselves to be highly ($n = 3$) or very familiar ($n = 11$) with visualizations, seven claimed to be moderately familiar, and only three said to be slightly or not at all familiar. None of the participants reported color blindness. All 24 participants completed the first part of the survey and 18 of them also completed the second part.

5 Results

The following results are based on quantitative analysis and a qualitative content analysis [19] which was applied to the responses to the open-ended questions.

5.1 RQ1 - Correctness

Correctness was measured by the number of correctly identified group meetings. To analyze the effect of the factors VISUALIZATION (V), TEST CASE (TC), and TASK (T) on the number of correctly identified groups in Task 1 and Task 2, a $4 \times 4 \times 2$ ($V \times TC \times T$) repeated measures ANOVA was conducted. Participants who did not fill out the second part of the survey were excluded from this analysis, yielding a sample size of 18 participants. Furthermore, as the number of correct groups differed from test case to test case and from task to task the number of correctly identified groups was normalized to the range $[0..1]$ by dividing the individual values by the total number of groups for the respective test case and task, yielding the fraction of correctly identified groups. Please note, that the values reported in the following are thus fractions and not absolute values.

The results show a significant main effect of VISUALIZATION ($F(3,51) = 44.64, p < .001$, $\eta_p^2 = .724$). Post hoc tests using Bonferroni correction revealed that the participants performed significantly poorer with the *Map* ($M = 0.303, SE = 0.038$) representation than with the *Gantt* ($M = 0.540, SE = 0.026$), *Matrix* ($M = 0.585, SE = 0.034$), and *Augmented Matrix* ($M = 0.605, SE = 0.031$) representations ($p < .001$ in each case). The main effect of TEST CASE was also significant ($F(1.99, 33.79) = 25.74, p < .001$, $\eta_p^2 = .602$, Greenhouse-Geisser adjusted). Participants performed best on *TC 1* ($M = 0.589, SE = 0.026$), followed by *TC 2* ($M = 0.54, SE = 0.027$) and *TC 3* ($M = 0.491, SE = 0.032$) and lastly *TC 4* ($M = 0.411, SE = 0.035$), that is participants generally performed better on the smaller datasets than on the larger datasets. The analysis also revealed a significant main effect of TASK ($F(1,17) = 116.54, p < .001$, $\eta_p^2 = .873$), indicating that participants performed considerably better on Task 1 ($M = 0.691, SE = 0.027$) than on Task 2 ($M = 0.325, SE = 0.036$).

Among the interactions effects only the interaction between VISUALIZATION and TASK was significant ($F(3,51) = 6.114, p = .001$, $\eta_p^2 = .265$). In case of Task 1 participants performed the best with the *Augmented Matrix* ($M = 0.788, SE = 0.032$) and the *Matrix* ($M = 0.783, SE = 0.035$) representation, followed by the *Gantt*

Source	SS	df	MS	F	η_p^2	p
V	8.396	3	2.799	44.640	.724	< **.001**
TC†	2.480	1.988	1.248	25.739	.602	< **.001**
T	19.322	1	19.322	116.536	.873	< **.001**
TC × T	0.116	3	0.039	1.527	.082	.219
TC × V	0.229	9	0.025	1.071	.059	.388
T × V	0.691	3	0.230	6.114	.265	**.001**
TC × T × V	0.154	9	0.017	0.894	.050	.533

significant p-values are highlighted in bold, †Greenhouse-Geisser adjusted
SS (Sum of Squares), MS (Mean Square), η_p^2 (Partial Eta squared)
V (Visualization), TC (Test Case), T (Task)

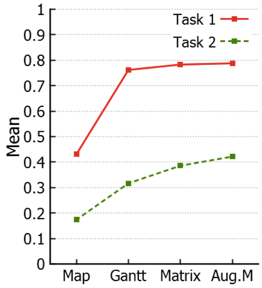


Fig. 3. Left: Summary table of the repeated measure ANOVA results (correctness). Right: Interaction graph for TASK × VISUALIZATION.

($M = 0.762, SE = 0.029$) chart and lastly the *Map* ($M = 0.431, SE = 0.049$) visualization. Examining the fraction of correctly identified groups by visualization for Task 2 shows an identical order with *Augmented Matrix* ($M = 0.422, SE = 0.044$) being best, followed by *Matrix* ($M = 0.386, SE = 0.044$), *Gantt* ($M = 0.317, SE = 0.042$), and *Map* ($M = 0.174, SE = 0.037$). However, the differences between the representations are slightly more pronounced and the average number of correctly identified groups is generally lower for each visualization than in Task 1.

Decomposing the interaction effect by VISUALIZATION with a simple effects test confirmed that the differences between Task 1 and Task 2 are statistically significant for each visualization ($p < .001$ in each case). Decomposing the interaction effect by TASK showed that participants performed significantly worse with the *Map* representation than with the other three visualization for both tasks ($p < .001$ for each comparison). However, in case of Task 2 the difference between the *Augmented Matrix* and *Gantt* visualization was also significant ($p = .02$), indicating an advantage of the *Augmented Matrix* for identifying repeated meetings of the same group compared to the *Gantt* chart. All other two-way interactions and the three-way interaction were not significant. Figure 3 summarizes the results of the repeated measures ANOVA.

5.2 RQ2 - Recognition Rate

Next, we assessed if the properties of the groups, that is, number of people in a group, number of meetings per group, and whether the total amount of time a group spent together had an influence on the recognition rate – the proportion of people who correctly reported the group – of the group. To evaluate the relationships between the above mentioned properties and the recognition rate – and due to non-normality of the data – a Spearman rank correlation has been conducted separately for each visualization and task (see Table 1).

In case of Task 1 the results revealed moderate to strong, statistically significant, correlations between the total amount of time a group spent together and the recognition rate for each of the four visualizations. In case of the *Matrix*, *Augmented Matrix*, and *Gantt* visualizations this can be explained by the fact that groups that meet longer are visually more evident (e.g., pie charts appear consecutively, bars are longer) and thus

Table 1. Spearman rank correlations between recognition rate and total amount of time a group spent together (Δt), number of people in a group (N_p), and number of meetings (N_m) for Task 1 and Task 2.

	Task 1			Task 2		
Recognition rate	Δt	N_p	N_m	Δt	N_p	N_m
Matrix	$.663^{**}$	$-.123$	$.657^{**}$	$.447^{*}$	$-.276$	$-.076$
Augmented matrix	$.792^{**}$	$-.022$	$.733^{**}$	$.663^{**}$	$-.350$	$.082$
Gantt	$.497^{**}$	$-.137$	$.292$	$.329$	$-.262$	$-.047$
Map	$.631^{**}$	$-.185$	$.503^{**}$	$.402^{*}$	$-.054$	$-.012$

*significant at $p < .05$, **significant at $p < .001$

easier to spot. In addition, the number of meetings correlated positively with groups being recognized for all but the *Gantt* representation. Given that Task 1 asked for identifying groups who meet at least for two hours at any location, multiple meetings increased the chance for recognizing a group. Task 2 showed similar positive correlations among duration and recognition rate for all representations, except for the *Gantt* chart. This was, in some way, unexpected as longer bars should also be beneficial for detecting groups who meet at different locations. In addition, the number of meetings did not positively influence the recognition of groups anymore. This, however, had to be expected since almost all of the groups contained in the datasets for Task 2 met at maximum two times. The number of people a group consists of, on the other hand, had no influence in neither of the visualizations and tasks, most likely because the number of persons was quite limited in most of the cases (in 53 out of 60 cases, the groups consisted of a maximum of three people).

To further investigate the correlation among duration and recognition rate which was not significant in the case of the *Gantt* chart we examined the recognition of individual groups more closely and could observe that groups with a white space between people of the group were seemingly more difficult to recognize than groups without white space. This observation was confirmed by a paired sampled t-test ($N = 18$) comparing the average recognition rate of groups with and without white space, $t(17) = -8.792, p < .001$, $d = 2.07$. Groups without white space were on average easier to identify ($M = 0.61, SD = 0.11$) than those containing white space in at least a single meeting ($M = 0.35, SD = 0.16$). This is in line with the Gestalt law of proximity that *things that are close together are perceptually grouped together* [21, p. 189]. In case of the *Gantt* chart the white space sometimes suggested that people belong to different groups although actually belonging to the same (see Fig. 4, left).

Comparing the recognition rate of individual groups between the *Matrix* and *Augmented Matrix* visualization, the auxiliary lines of the *Augmented Matrix* improved the recognition of 50.0 % of the groups but also impeded the identification in 38.3 % of the cases (no change in the remaining 11.7 %). Although these differences were not statistically significant at the .05 level except in a single case (assessed via a McNemar test separately for each group) there is a tendency that meetings are missed if there are too many lines in its proximity, or to put it differently, if the lines increase the visual clutter in the area surrounding the meeting. However, lines also improved the discovery of groups, usually in cases where the lines were more distinguishable or where

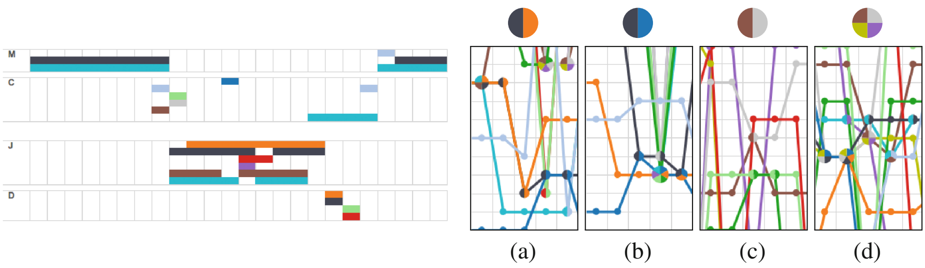


Fig. 4. Left: White space between people of the same group hindered group identification in the *Gantt* representation. For example, group meeting at *M* and *C* (top) was only recognized by 18 % of the participants in Task 2 as opposed to group meeting at *J* and *D* (bottom) which was identified by 47 %. Both groups met twice for one hour each at different locations. Right: Examples of cases where auxiliary lines improved (a, b) or impaired recognition of groups (c, d). The group under consideration is depicted above each example.

the convergence of lines was more readily apparent. By way of example, Fig. 4 (right) shows some groups where the auxiliary lines improved or reduced the recognition rate by at least 0.2.

5.3 RQ3 - Preference

Preference refers to the subjective preference ratings of the participants. Analysis of the preference ranking (1 = best, 4 = worst) of the visualizations after the small datasets (that is, Test Case 1 and 2) using a Friedman test ($N = 24$) showed a significant overall difference among the rankings ($\chi^2(3) = 41.15, p < .001, W = .572$). The *Matrix* visualization ranked best with a mean rank of 1.5, followed by the *Augmented Matrix* (2.08), *Gantt* (2.63), and lastly the *Map* (3.79) representation. Post hoc analysis with Wilcoxon signed-rank tests with a Bonferroni correction applied showed significant differences between each pair of visualizations except between the *Augmented Matrix* and *Gantt* representation ($Z = -1.626, p = .104$). Analysis of the ranking after the large datasets (Test Case 3 and 4) gives almost identical results. Again, the Friedman test ($N = 18$) was significant ($\chi^2(3) = 43.13, p < .001, W = .799$) with mean ranks of 1.22 (*Matrix*), 2.06 (*Augmented Matrix*), 2.78 (*Gantt*), and 3.94 (*Map*). However, in contrast to above, the difference between *Augmented Matrix* and *Gantt* representation was also significant ($Z = -2.422, p = .015$).

From the explanations provided by the participants it was apparent that the *Matrix* visualization was mostly appreciated for providing a structured way to solve Task 1 and Task 2, with participants describing the visualization as clear (6/6)[5], well readable (4/3), and mentioning that it offers a good overview (4/5). However, four participants noted that location changes are hard to see (4/0). We got contradictory responses to the *Augmented Matrix* representation: some participants noted that the lines are distracting (2/2) and clutter the visualization (7/7) while other participants found the lines useful to see relations and track changes (9/5). This ambivalence is also reflected in the impact of the auxiliary lines on the recognition rate of groups (see RQ2 above). With respect

[5] number of statements after the small/large test cases.

Table 2. Categories developed from the qualitative content analysis along with the number of statements falling within each category.

Visualization	Duration	Group Size	Repetition	Color	Position	Proximity	Σ
Augm. Matrix	65	42	41	35	24	5	212
Matrix	49	28	27	33	19	5	161
Gantt	40	28	20	19	18	2	127
Map	11	39	17	16	11	3	97
Σ	165	137	105	103	72	15	597

1 ▭ 65 Statements

to the *Gantt* representation most participants indicated that the locations were hard to discern due to the thin separating lines between locations (4/8) and that the irregular heights of the rows were irritating (6/3). On the other hand, some participants liked the *Gantt* chart because of its readability (5/0), although this quality was not mentioned explicitly after the large test cases anymore. A minority of the participants also ranked the *Gantt* higher because they did not like the pie charts in the other visualizations (2/1).

The poor result of the *Map* can be mostly attributed to the fact that participants found it difficult to infer actual meetings of groups as the meeting times were not readily apparent but rather had to be calculated from the labels (13/10) which, subsequently, also complicated the identification of recurring group meetings. For example, one participant criticized that *"you have to read out the times, as opposed to the other types where you can derive the times directly from the axis"* while another found that *"the need to compare all the labels makes it hard to figure out if people really meet or if they are just at the same location at different times"*. However, a few participants, though admitting that the huge amount of labels made it difficult to solve Task 1 and Task 2, liked the geospatial representation (1/3).

5.4 RQ4 - Saliency

In Task 3 participants were asked to identify any conspicuous groups (saliency). A qualitative content analysis [19] of the participants' responses to Task 3 was conducted by three researchers in an iterative process. In a first round keywords were extracted from the comments which were then analyzed to form categories. Based on the keywords the statements were then assigned to the corresponding categories. Please note, that this means that a comment could be assigned to more than one category. Finally, the comments within each category were counted. In the end, the comments provided by the participants were categorized into six categories. Table 2 lists the number of statements – summarized across all four test cases – falling within each category grouped by visualization. In the following we will discuss each category in more detail.

Duration. Most descriptions of salient groups with the *Augmented Matrix*, *Matrix*, and *Gantt* representation contained statements about the time a group spent together (approx. 1/3 of all statements made with each visualization). Longer meetings were more often highlighted with these three visualizations in comparison to short meetings, more specifically (from in total 165 statements) participants made 135 statements highlighting the long duration of a meeting as opposed to just 23 statements being concerned with short meetings. This, however, may well be because longer meetings were easier to recognize than shorter ones (see RQ2). Worth noting, but not surprising given that participants found it very troublesome to calculate the time-spans of meetings from the labels (cf. RQ3), is the comparable low number of statements about the duration of group meetings with the *Map* representation.

Group Size. In total terms, group size was the second most mentioned feature for considering a group as salient. Most noticeable is the large fraction of statements (approx. 38.6 %) concerning group size with the *Map* representation: as the exact durations of meetings were hard to infer participants were seemingly more focused on this aspect. Summarized across all visualizations, larger groups were reported much more often (119 statements) than smaller groups (6 statements) whereas participants usually referred to groups with more than three people as large and groups consisting of two people as small.

Repetition. Most statements (approx. 39 %) referring to repeated meetings of a group were made with the *Augmented Matrix* visualization, most likely because the auxiliary lines made it easier to trace the route of the group's members – an impression which is also reflected in some of the comments made by the participants who found the lines helpful to track changes (cf. RQ2). In 32 cases participants stressed that the repetitive meetings of a group are interesting because they met at different locations, while groups meeting at the same place have only been emphasized in 15 statements. In the other cases, participants did not specify why they considered the repetition as salient. Again, most statements highlighting different locations (15 out of 32) have been made with the *Augmented Matrix* visualization.

Color. Around 14 % to 20 % of the statements per visualization contained some sort of reference to the colors used to depict the individual persons. Participants, for example, emphasized groups because of the color combination in general (28 statements) or because of the contrast between the colors (29 statements).

Position. The position of the groups in the visualization has also been a contributing factor if groups were perceived as salient or not, with around 10.9 % to 13.6 % of statements per visualization explicitly referring to the position as reason for being salient. This was particularly an issue for the *Augmented Matrix*, *Matrix*, and *Gantt* representation since especially groups which were located in the center of the visualization (29 statements) or on the left side (11 statements) were considered to stand out.

Proximity. Proximity to other groups played a minor role for participants to consider a group as salient. In case of the *Augmented Matrix*, *Matrix*, and *Gantt* representation this is in line with our expectations as the order of the locations in these visualizations do not reflect the geographical distance. However, in case of the *Map* visualization this low number is a bit surprising but is most likely attributable to the nearly equidistant spacing between the locations.

6 Discussion

Visualizations. The results show that, in general, the three alternative visualizations (*Augmented Matrix*, *Matrix*, *Gantt*) are better suited for identifying meetings than the *Map* representation (Task 1 and Task 2). Furthermore, the *Map* was clearly the least preferred visualization. Based on the literature review we expected this result, since maps seem not to be an appropriate visualization to present several points in time. More detailed analysis indicates that there are some differences between the *Augmented Matrix*, *Matrix*, and *Gantt* representation. In case of Task 2 (identifying groups meeting at different locations) there is a significant difference between the *Augmented Matrix* and the *Gantt* chart. This corresponds to the preference rankings where the *Gantt* representation was rated third. One of the reasons that the *Gantt* chart did not perform as well as the other two alternative visualizations is probably that the white space between the bars of the *Gantt* chart were sometimes misleading. Participants preferred the *Matrix* visualization, followed by the *Augmented Matrix*. Reasons given for the high ranking of the *Matrix* were that it enabled the participants to solve the tasks in a structured way. The *Augmented Matrix* was criticized because of the use of lines which sometimes led to clutter. However, some participants appreciated the auxiliary lines as they made it easier to see relations and changes of locations. This is probably related to the result that the lines sometimes made the identification of meetings easier and sometimes more difficult. Although the impact of these lines on overall correctness was not significant as shown by the ANOVA analysis it would make sense to investigate this issue in more detail based on the comments of the participants and on the recognition rate of groups.

Tasks. The results show that the visualizations are better suited for the duration detection task (Task 1) than for detecting location changes of groups (Task 2). A reason for that can be that Task 1 which asks for the duration of meetings is simpler than the task of detecting two or more meetings of the same group of persons (Task 2). In both tasks, groups with longer meetings were easier to identify than shorter ones most likely because they were visually more evident.

Test Cases. The study also shows that the participants performed better on the smaller datasets (Test Case 1 and 2) than on the larger datasets (Test Case 3 and 4). Although we expected this result, it was interesting to see that there was no significant interaction between the test cases and tasks nor between test cases and visualizations. Furthermore, the test cases did not influence the participants' subjective preference.

One issue to keep in mind is that we conducted the study with students and not with experts. As mentioned above the investigation of effectiveness and utility for group detection was the primary goal of this study. This depends more on the mechanisms of human cognition and less on domain-specific knowledge. We thus decided to choose students as sample since it is difficult to reach a large number of experts which also have the time to participate in a study that takes around two hours. Finally, when interpreting the results one should consider that the colors may have influenced the recognition rates of groups as the results of the qualitative content analysis suggest a certain effect of the colors on which groups were perceived as salient. However, by using an established color scheme and using different color combinations for different groups we tried to mitigate such effects.

7 Design Considerations

Based on the above discussed results we propose the following design considerations:

Use map in combination with other visualizations: It may be useful to combine a map with one of the other visualizations, for example, in a coordinated multiple view setting.

Use auxiliary lines carefully: We suggest to use auxiliary lines with caution, for example, only on a certain subset of entities.

Avoid white space between entities of the same group: It may be beneficial to provide possibilities to rearrange the order of entities in the visualization to ensure that entities belonging to the same group are displayed next to each other.

Show duration of meetings explicitly: Showing the time explicitly improved the identification of (recurring) group meetings.

Consider the influence of visual properties: The qualitative content analysis revealed an influence of certain properties of the visualization on which groups were perceived as salient or not.

8 Conclusion

We conducted an empirical evaluation to identify appropriate visualizations to support users in the activity of identifying meetings of persons, vehicles, or other entities. As expected, map visualizations did not perform very well. The other visualizations which we used performed significantly better. Based on the research results we also developed recommendations for the design of these visualizations. There are several issues which were not addressed in this study. Among them is the question which kinds of interactivity should be adopted to support the users. It is, for example, likely that the possibility to select interesting cases could assist users in solving their tasks. In addition, studies with domain experts are necessary to clarify in which areas the visualization we suggested are most beneficial.

Acknowledgments. The research leading to these results in the VALCRI project has received funding from the European Union 7th Framework Programme (FP7/2007-2013) under grant agreement no FP7-IP- 608142, to Middlesex University and Partners. Simone Kriglstein was supported by CVAST (funded by the Austrian Federal Ministry of Science, Research, and Economy in the exceptional Laura Bassi Centres of Excellence initiative, project nr: 822746).

References

1. Aigner, W., Miksch, S., Schumann, H., Tominski, C.: Visualization of Time-Oriented Data. Springer, London (2011)
2. Andrienko, G., Andrienko, N., Hurter, C., Rinzivillo, S., Wrobel, S.: Scalable analysis of movement data for extracting and exploring significant places. IEEE Trans. Vis. Comput. Graph. **19**(7), 1078–1094 (2013)

3. Andrienko, G., Andrienko, N., Mladenov, M., Mock, M., Pölitz, C.: Discovering bits of place histories from people's activity traces. In: Proceedings of the VAST 2010, pp. 59–66. IEEE (2010)
4. Boba Santos, R.: Crime Analysis With Crime Mapping, 3rd edn. SAGE Publications, Thousand Oaks (2013)
5. Chorley, M.J., Whitaker, R.M., Allen, S.M.: Personality and location-based social networks. Comput. Hum. Behav. **46**, 45–56 (2015)
6. Ghoniem, M., Fekete, J.D., Castagliola, P.: A comparison of the readability of graphs using node-link and matrix-based representations. In: Proceedings of the INFOVIS 2004, pp. 17–24. IEEE (2004)
7. Gleicher, M., Albers, D., Walker, R., Jusufi, I., Hansen, C.D., Roberts, J.C.: Visual comparison for information visualization. Inf. Vis. **10**(4), 289–309 (2011)
8. Guo, D., Chen, J., MacEachren, A.M., Liao, K.: A visualization system for space-time and multivariate patterns (vis-stamp). IEEE Trans. Vis. Comput. Graph. **12**(6), 1461–1474 (2006)
9. Harrower, M., Brewer, C.A.: Colorbrewer.org: an online tool for selecting colour schemes for maps. In: Dodge, M., Kitchin, R., Perkins, C. (eds.) The Map Reader, pp. 261–268. Wiley, Hoboken (2011)
10. Kessell, A.M., Tversky, B.: Visualizing space, time, and agents: production, performance, and preference. Cogn. Process. **12**(1), 43–52 (2011)
11. Kjellin, A., Pettersson, L.W., Seipel, S., Lind, M.: Evaluating 2D and 3D visualizations of spatiotemporal information. ACM Trans. Appl. Percept. **7**(3), 19:1–19:23 (2008)
12. Kraak, M.J.: The space-time cube revisited from a geovisualization perspective. In: Proceedings of the ICC 2003. The International Cartographic (2003)
13. Kriglstein, S., Pohl, M., Smuc, M.: Pep up your time machine: recommendations for the design of information visualizations of time-dependent data. In: Huang, W. (ed.) Handbook of Human Centric Visualization, pp. 203–225. Springer, New York (2014)
14. Kristensson, P.O., Dahlback, N., Anundi, D., Bjornstad, M., Gillberg, H., Haraldsson, J., Martensson, I., Nordvall, M., Stahl, J.: An evaluation of space time cube representation of spatiotemporal patterns. IEEE Trans. Vis. Comput. Graph. **15**(4), 696–702 (2009)
15. Li, N., Chen, G.: Analysis of a location-based social network. In: Proceedings of the CSE 2004, pp. 263–270. IEEE (2009)
16. Luz, S., Masoodian, M.: Comparing static gantt and mosaic charts for visualization of task schedules. In: Proceedings of the IV 2011, pp. 182–187. IEEE (2011)
17. MacEachren, A.M.: How Maps Work - Representation, Visualization, and Design. Guilford Press, New York (2004)
18. Orellana, D., Wachowicz, M., Andrienko, N., Andrienko, G.: Uncovering interaction patterns in mobile outdoor gaming. In: Proceedings of the GEOWS 2009, pp. 177–182. IEEE (2009)
19. Schreier, M.: Qualitative Content Analysis in Practice. SAGE Publications, Thousand Oaks (2012)
20. Shen, Z., Ma, K.L.: Path visualization for adjacency matrices. In: Proceedings of the EUROVIS 2007, pp. 83–90. Eurographics Association (2007)
21. Ware, C.: Information Visualization: Perception for Design. Morgan Kaufmann Publishers Inc., Burlington (2004)
22. Wilson, J.M.: Gantt charts: a centenary appreciation. Eur. J. Oper. Res. **149**(2), 430–437 (2003)
23. Wu, H.M., Tzeng, S., Chen, C.H.: Matrix visualization. In: Wu, H.M., Tzeng, S., Chen, C.H. (eds.) Handbook of Data Visualization, pp. 681–708. Springer, Heidelberg (2008)

The Perception of Clutter in Linear Diagrams

Mohanad Alqadah[1(✉)], Gem Stapleton[1], John Howse[1], and Peter Chapman[2]

[1] University of Brighton, Brighton, UK
{m.alqadah1,g.e.stapleton,john.howse}@brighton.ac.uk
[2] Edinburgh Napier University, Edinburgh, UK
p.chapman@napier.ac.uk

Abstract. Linear diagrams are an effective way of representing sets and their relationships. The topological and graphical properties of linear diagrams can affect perceived relative levels of clutter. This paper defines four different measures of clutter for linear diagrams. Participants in an empirical study were asked to rank linear diagrams according to their perception of clutter. We analyzed the correlation between how the clutter measures ranked linear diagrams compared to the overall ranking derived from the participants' perceptions. We concluded that the clutter measure which counts the number of line segments best matches participants' perception.

Keywords: Linear diagrams · Clutter · Diagram comprehension

1 Introduction

Representing information using diagrams can have huge benefits, but only if the diagrams themselves are effective. One aspect of the effectiveness of diagrammatic communication is related to clutter. If diagrams appear cluttered then their visual appeal and ability to support end-users with understanding the represented information can be reduced. Hence, there is clearly a need to theoretically understand what it means for diagrams to be cluttered and the impact of clutter on task performance. This paper is concerned with clutter in linear diagrams, recently shown to be superior to Euler and Venn diagrams when users perform set-theoretic tasks [1] and to linguistic representations of syllogisms [6].

A linear diagram consists of horizontal line segments drawn parallel to the x-axis. Each set is represented by the line segments that share their y-coordinate. For example, Fig. 2 represents five sets using five line segments. Line segments for different sets can occupy the same vertical space, known as an *overlap*. Where an overlap contains line segments for sets $A_1, \ldots, A_n$ and does not contain line segments for $B_1, \ldots, B_m$, the overlap represents the information that $A_1 \cap \ldots \cap A_n \cap \overline{B_1} \cap \ldots \cap \overline{B_m}$ is non-empty. Moreover, if a set intersection is not represented by an overlap then that set intersection is empty. For example, in Fig. 1 the left-most overlap contains line segments for the sets Animals and Zebras but not Cats, Lions or Tigers. Thus, $\text{Animals} \cap \text{Zebras} \cap \overline{\text{Cats}} \cap \overline{\text{Lions}} \cap \overline{\text{Tigers}} \neq \emptyset$.

There is currently no understanding of what constitutes a cluttered linear diagram. This paper sets out to address the question of how to measure perceived

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 250–257, 2016.
DOI: 10.1007/978-3-319-42333-3_20

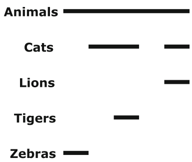


Fig. 1. A linear diagram.

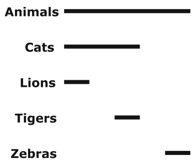


Fig. 2. A linear diagram.

clutter in linear diagrams. We introduce four measures of clutter in Sect. 2. The design of our experiment to determine whether any of these four measures correlate with users' perception of clutter is described in Sect. 3. We present the analysis and results in Sect. 4 and conclude in Sect. 5. The diagrams used in the study, along with the raw data collected, can be found at https://sites.google.com/site/msapro/phdstudythree.

2 Four Measures of Clutter for Linear Diagrams

Our first measure, called the *contour score* (CS), is based on a measure of clutter established when counting zones in Euler diagrams [2]. We can easily adapt this measure to linear diagrams as overlaps in linear diagrams directly correspond to zones in Euler diagrams. The CS for linear diagrams is computed as follows: each overlap contributes n to the contour score, where n is the number of lines in the overlap. For example, each diagram in Fig. 3 has six overlaps, identified by the use of grid lines, and have a CS of 11. In these four diagrams, the overlaps are annotated with their contribution to the contour score under the diagram, indicated by the CS label. These diagrams represent *the same information* as each other and have *the same CS*. However, they are syntactically different and, so, there may be differences in how people perceive their relative levels of clutter.

Our first new measure of clutter specifically designed for linear diagrams is the *line score* (LS): each set contributes n to the line score, where n is the number of line segments that are used to represent that set. For example, in Fig. 3, the diagrams (v1) and (v2) both have a LS of 8; the sets are annotated with their contribution to the LS under the column labelled LS. However, the

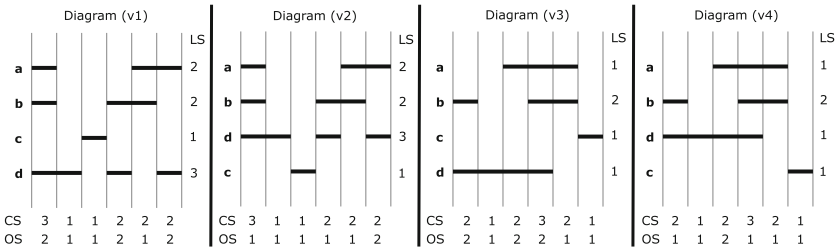


Fig. 3. Different measures of clutter in linear diagrams.

other two diagrams both have a LS of 5; the LS is lower because the overlaps are in a different order, leading to fewer line breaks.

While LS can be altered by changing the order of the overlaps, we also have the choice of reordering the horizontal lines in the diagram. Such reordering changes the visual appearance of the diagram and could also be linked to perceived clutter. Therefore, our third new measure of clutter, called the *overlap score* (OS), captures the ‘vertical clutter’ in linear diagrams; LS can be thought of as capturing the ‘horizontal clutter’. The OS of linear diagrams can be computed from the blocks of lines in the overlaps: each overlap contributes n to the overlap score, where n is the number of blocks of lines in the overlap. For example, in Fig. 3, the diagrams (v1) and (v3) both have an OS of 9. The first overlap in (v1) has two blocks of lines: the two lines for sets ‘a’ and ‘b’ are a block of lines, then there is no line for the set ‘c’, and finally a further block, comprising a single line, for the set ‘d’. Each overlap is annotated with its contribution to the overlap score. The diagrams (v2) and (v4) both have an OS of 7. The overlap score is lower than (v1) and (v3) because of the different orders of the represented sets.

Our last clutter measure is the combined score of the LS and the OS of linear diagrams, which we call the *line-and-overlap score* (LOS). This clutter measure is designed to capture both vertical clutter and horizontal clutter in linear diagrams, should this be perceived. Formally, the LOS for linear diagrams is the sum of the LS and OS scores for the linear diagram. For example, in Fig. 3, the LOSs are as follows: (v1) 17; (v2) 15; (v3) 14; and (v4) 12.

3 Experiment Design

The study consisted of four tasks, each of which required participants to rank 12 linear diagrams, with a ranking of 1 being least cluttered and 12 being most cluttered. Joint rankings were permitted. The first three tasks fixed the number of sets being represented to 4, 6, and 8 respectively. This allowed us to establish perceived relative clutter when the number of sets did not change. The final task included linear diagrams with 5, 6 and 7 sets (four linear diagrams for each number of sets). This allowed us to establish whether any of the clutter measures were effective at differentiating diagrams with perceived differences in clutter as the number of sets increased. For each task, a primary design feature of the set of 12 diagrams we used is that the different clutter measures give rise to different rankings of the diagrams. This was important for us to gain insight into the relative effectiveness of the clutter measures.

The 12 linear diagrams generated for task 1 were divided into three sets of four diagrams. The four diagrams in each set represented the same information as each other (so they had the same overlaps) but with different layouts. These layouts varied the (horizontal) order of the overlaps and the (vertical) order of the sets. This meant that the line scores and overlap scores varied, whereas the contour score was necessarily fixed. The contour scores did vary between the three sets of diagrams, however. The 12 diagrams were designed to have

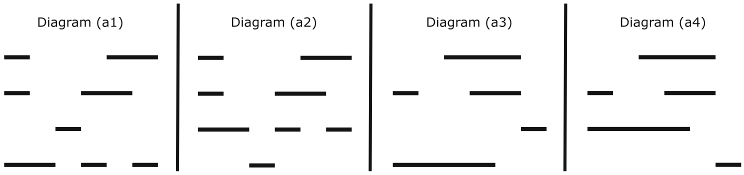


Fig. 4. Linear diagrams with four sets that were used in the study.

three CS clutter levels and six levels of clutter for both of LS and OS (and, consequently, six levels for LOS). Therefore, each clutter measure in the task ranked the 12 diagrams in a different order, allowing us to compare the derived diagram rankings with an overall ranking derived from the participants' rankings. For example, diagrams d1.1 to d1.4 form the first set of diagrams and can be seen in Fig. 3 (note that they were presented to the participants on separate sheets of paper and, as in Fig. 4, without the diagram names being shown). Tasks 2 and 3 have the same design as task one but with 6- and 8-set diagrams.

Table 1. The characteristics of the diagrams.

Diagram number (Task 1)	**d1.1**	**d1.2**	**d1.3**	**d1.4**	**d2.1**	**d2.2**	**d2.3**	**d2.4**	**d3.1**	**d3.2**	**d3.3**	**d3.4**
Number of overlaps	6	6	6	6	8	8	8	8	10	10	10	10
CS score	11	11	11	11	18	18	18	18	24	24	24	24
LS score	8	8	5	5	11	11	7	7	11	11	6	6
OS score	9	7	9	7	12	11	12	11	14	12	14	12
LOS score	17	15	14	12	23	22	19	18	25	23	20	18
Diagram number (Task 2)	**d1.1**	**d1.2**	**d1.3**	**d1.4**	**d2.1**	**d2.2**	**d2.3**	**d2.4**	**d3.1**	**d3.2**	**d3.3**	**d3.4**
Number of overlaps	18	18	18	18	22	22	22	22	26	26	26	26
CS score	60	60	60	60	67	67	67	67	85	85	85	85
LS score	26	26	14	14	35	35	16	16	42	42	19	19
OS score	32	25	32	25	41	37	41	37	46	41	46	41
LOS score	58	51	46	39	76	72	57	53	88	83	65	60
Diagram number (Task 3)	**d1.1**	**d1.2**	**d1.3**	**d1.4**	**d2.1**	**d2.2**	**d2.3**	**d2.4**	**d3.1**	**d3.2**	**d3.3**	**d3.4**
Number of overlaps	24	24	24	24	30	30	30	30	36	36	36	36
CS score	103	103	103	103	127	127	127	127	134	134	134	134
LS score	48	48	24	24	69	69	26	26	68	68	32	32
OS score	50	46	50	46	70	53	70	53	79	66	79	66
LOS score	98	94	74	70	139	122	96	79	147	134	111	98
Diagram number (Task 4)	**d1**	**d2**	**d3**	**d4**	**d5**	**d6**	**d7**	**d8**	**d9**	**d10**	**d11**	**d12**
Number of sets	5	5	5	5	6	6	6	6	7	7	7	7
Number of overlaps	11	12	13	14	15	16	17	18	19	20	21	22
CS score	20	28	36	44	40	48	56	64	60	68	76	84
LS score	15	8	18	10	24	11	29	16	33	19	41	21
OS score	17	14	23	20	30	26	35	32	44	38	45	48
LOS score	32	22	41	30	54	37	64	48	77	57	86	69

The diagrams for task 4 consisted of 5-, 6-, and 7-set linear diagrams. Each clutter measure ranked the 12 diagrams in a different order. This was deemed important because the four different rankings allowed us to find out which clutter measure correlates most strongly with the overall participants' ranking. Moreover, for each measure, no pair of diagrams had the same clutter score. This meant that each measure totally ordered the diagrams. For each of the four tasks, Table 1 provides details on the diagrams in terms of number of overlaps present and the clutter scores arising from each of the four measures.

4 Experiment Execution, Analysis and Results

Initially eight participants (6 M, 2 F, ages 18–55) took part in a pilot study. The pilot study was successful and the participants finished the four tasks in less than one hour. As no changes were deemed necessary, the pilot data was carried forward for analysis with the data collected in the main study phase. A further 52 participants were recruited, giving a total of 60 participants (46 M, 14 F, ages 18–55). All the participants were staff or students from the University of Brighton; none of them were members of the authors' research group.

To test the effectiveness of the four clutter measures, for each task we first derived an overall ranking of the 12 diagrams from the participants' rankings. Consistent with other researchers who studied diagram complexity, for instance [2,3,5], the best clutter measure was identified by the Pearson correlation test on the participants' preference data. We performed a correlation analysis between the overall participants' ranking and that derived from each clutter measure. We viewed a measure of clutter as accurate if there was a significant correlation (at 5 %) between the clutter measure ranking and the overall participants' ranking.

For task 1 after collecting all the participants' orderings of the 12 diagrams, we calculated an overall participants' ranking using a Friedman test to estimate the median ranking for each diagram. This was then converted the estimates into an overall participants' ranking. Table 2 shows the overall participants' ranking for task 1 in the appropriate row. The results for task 1 are given in the first row of Table 3 which shows the correlation coefficients and, in brackets, the *p*-values; bold typeface indicates significance. We can see, therefore, that the strongest significant correlation is with the line score. In addition, the line-and-overlap score is significantly correlated whereas the contour score and the overlap score are not.

Table 2 shows the rankings of the 12 diagrams for tasks 2 and 3 alongside the overall participants' ranking. Table 3 shows the correlation coefficients and the *p*-values. For both tasks, the strongest significant correlation is between the overall participants' ranking and the LS measure, with LOS also being significant.

For task 4, Table 2 shows the rankings of the 12 diagrams for task 4 alongside the overall participants' ranking. Table 3 shows the correlation coefficients and the *p*-values. As with the other three tasks, the strongest significant correlation is between the overall participants' ranking and the LS measure. However, for this task all clutter measures significantly correlate with the overall participants' ranking.

Table 2. Clutter rankings and participants' ranking for task 1–4.

Task 1	**d1.1**	**d1.2**	**d1.3**	**d1.4**	**d2.1**	**d2.2**	**d2.3**	**d2.4**	**d3.1**	**d3.2**	**d3.3**	**d3.4**
CS ranking	2.5	2.5	2.5	2.5	6.5	6.5	6.5	6.5	10.5	10.5	10.5	10.5
LS ranking	7.5	7.5	1.5	1.5	10.5	10.5	5.5	5.5	10.5	10.5	3.5	3.5
OS ranking	3.5	1.5	3.5	1.5	8.5	5.5	8.5	5.5	11.5	8.5	11.5	8.5
LOS ranking	4	3	2	1	10.5	9	7	5.5	12	10.5	8	5.5
Participants' ranking	5	8	2	1	12	10.5	7	6	10.5	9	4	3
Task 2	**d1.1**	**d1.2**	**d1.3**	**d1.4**	**d2.1**	**d2.2**	**d2.3**	**d2.4**	**d3.1**	**d3.2**	**d3.3**	**d3.4**
CS ranking	2.5	2.5	2.5	2.5	6.5	6.5	6.5	6.5	10.5	10.5	10.5	10.5
LS ranking	7.5	7.5	1.5	1.5	9.5	9.5	3.5	3.5	11.5	11.5	5.5	5.5
OS ranking	3.5	1.5	3.5	1.5	7.5	5.5	7.5	5.5	11.5	7.5	11.5	7.5
LOS ranking	6	3	2	1	10	9	5	4	12	11	8	7
Participants' ranking	7.5	7.5	1	2	9	10	3	4	11	12	5	6
Task 3	**d1.1**	**d1.2**	**d1.3**	**d1.4**	**d2.1**	**d2.2**	**d2.3**	**d2.4**	**d3.1**	**d3.2**	**d3.3**	**d3.4**
CS ranking	2.5	2.5	2.5	2.5	6.5	6.5	6.5	6.5	10.5	10.5	10.5	10.5
LS ranking	7.5	7.5	1.5	1.5	11.5	11.5	3.5	3.5	9.5	9.5	5.5	5.5
OS ranking	3.5	1.5	3.5	1.5	9.5	5.5	9.5	5.5	11.5	7.5	11.5	7.5
LOS ranking	6.5	4	2	1	11	9	5	3	12	10	8	6.5
Participants' ranking	7	8	1	2	9	10.5	4	3	10.5	12	6	5
Task 4	**d1**	**d2**	**d3**	**d4**	**d5**	**d6**	**d7**	**d8**	**d9**	**d10**	**d11**	**d12**
CS ranking	1	2	3	5	4	6	7	9	8	10	11	12
LS ranking	4	1	6	2	9	3	10	5	11	7	12	8
OS ranking	2	1	4	3	6	5	8	7	10	9	11	12
LOS ranking	3	1	5	2	7	4	9	6	11	8	12	10
Participants' ranking	4	1	5	2	9	3	10	6	11	7	12	8

Table 3. Correlations between clutter measures and perception, by task.

	CS	LS	OS	LOS
Task 1	0.311 (0.325)	**0.948 (0.000)**	0.374 (0.232)	**0.798 (0.002)**
Task 2	0.474 (0.120)	**0.991 (0.000)**	0.381 (0.222)	**0.851 (0.000)**
Task 3	0.459 (0.133)	**0.942 (0.000)**	0.361 (0.249)	**0.867 (0.000)**
Task 4	**0.615 (0.033)**	**0.993 (0.000)**	**0.839 (0.001)**	**0.958 (0.000)**

Table 3 shows that both of LS and LOS measures were significantly correlated to the overall participants' ranking in all four tasks. To establish whether LS is significantly more correlated than LOS we used the Fisher r-to-z transformation which converts correlations into a normally distributed measure. Then we used a Z-test to see whether LS is significantly more correlated than LOS. The calculated values of z for the four tasks were as follows: 3.84, 7.69, 2.32, and 4.83 respectively. A one-tailed test yields p-values of 0.0001, 0.0000, 0.0102, and 0.0000 respectively which are all less than 0.05. Therefore the LS measure is significantly more correlated with the overall participants' ranking than the LOS measure.

In summary for each of the four tasks, both the line score and the line-and-overlap score were significantly correlated with participants' perception of clutter in linear diagrams. In each case, however, there was a significantly stronger correlation with the line score. This is unsurprising as, at least for tasks 1 to 3, the overlap score did not yield a diagram ranking that was significantly correlated with the overall participants' ranking. In particular, the overlap score correlation coefficient for these three tasks was quite low, demonstrating that there was little relationship at all. This indicates why adding the overlap score to the line score, yielding the line-and-overlap score, resulted in a weaker correlation.

Recall that task 4 was the only task to include a variety of numbers of sets. This design feature allowed us to gain insight into whether the clutter measures were able to distinguish differences in perceived clutter as the number of sets varied. Interestingly, task 4 (and only task 4) yielded data where all four measures were significantly correlated with perceived clutter. However, the strongest correlation was still with the line score. This indicates that simply comparing the number of line segments present in linear diagrams effectively reflects perceived levels of clutter regardless of the number of sets being visualized: linear diagrams with fewer line segments are perceived to be less cluttered.

5 Conclusion

This paper has provided an understanding of how people perceive clutter in linear diagrams. By considering the syntax of linear diagrams, and how it can be altered through reordering overlaps and sets, we identified four potential measures of clutter, namely: the contour score (generalized from similar research on Euler diagrams), the line score, the overlap score and the line-and-overlap score. Through empirical research, we established that the line score significantly correlates with perceived clutter, regardless of the number of sets present in linear diagrams. In summary, the relative number of line segments present in linear diagrams accurately predicts perceived relative levels of clutter.

The results of our research tell us that reducing the number of line segments in linear diagrams reduces perceived clutter. Techniques already exist for reducing the number of line segments, as implemented in the linear diagram generator used to create the diagrams in our study [4]. A key future research goal is to establish the impact of clutter in linear diagrams on user task performance.

References

1. Chapman, P., Stapleton, G., Rodgers, P., Micallef, L., Blake, A.: Visualizing sets: an empirical comparison of diagram types. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 146–160. Springer, Heidelberg (2014)
2. John, C., Fish, A., Howse, J., Taylor, J.: Exploring the notion of 'clutter' in Euler diagrams. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) Diagrams 2006. LNCS (LNAI), vol. 4045, pp. 267–282. Springer, Heidelberg (2006)

3. Purchase, H.C., Freeman, E., Hamer, J.: An exploration of visual complexity. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 200–213. Springer, Heidelberg (2012)
4. Rodgers, P., Stapleton, G., Chapman, P.: Visualizing sets with linear diagrams. ACM Trans. Comput.-Hum. Interact. (TOCHI) **22**(6), 27 (2015)
5. Rosenholtz, R., Li, Y., Nakano, L.: Measuring visual clutter. J. Vis. **7**(2), 17 (2007)
6. Sato, Y., Mineshima, K.: How diagrams can support syllogistic reasoning: an experimental study. JoLLI **24**(4), 409–455 (2015)

People, Place, and Time: Inferences from Diagrams

Barbara Tversky[1(✉)], Jie Gao[1], James E. Corter[1], Yuko Tanaka[2], and Jeffrey V. Nickerson[3]

[1] Teachers College, Columbia University, New York, NY 10027, USA
btversky@stanford.edu,
jgao@ets.org, jec34@columbia.edu
[2] Nagoya Institute of Technology, Nagoya, Aichi 4668555, Japan
tanaka.yuko@nitech.ac.jp
[3] Stevens Institute of Technology, Hoboken, NJ 07030, USA
jnickerson@stevens.edu

Abstract. Keeping track of things as they move in space and time is a task common to scientists, marketers, spies, coaches, and more. Visualizations of complex information aid drawing inferences and conclusions but there are many ways to represent data. Here we show that the kinds of inferences people draw depend on the kind of visualization, boxes in tables or lines in graphs. Lines link and boxes contain; they both direct attention and create meaning.

Keywords: Diagrams · Information visualization · Inference · Data displays

1 Introduction

People are always on the move. So are other living things, and even inanimate things, not just tangibles like packages, airplanes, and lava but also slang, fashion, music, rumors. Tracking and understanding movements of things in space and time is a task shared by scientists, historians, football coaches, paparazzi, marketers, physicians, spies, Facebook, event planners, police, culture mavens, advertisers, gossip columnists, friends, and more. The movements of beings and things are valuable data to be explained by theories. Why do people or things cluster in one place or avoid another? Why did person X see Y and then Z? Why did they meet there? Why is this place popular at one time and not at another? Speculating about the movements of people or things over time is endlessly fascinating, and the number of queries, hypotheses, and explanations that can be generated enormous.

Making sense of complex data is made easier by organizing it spatially into diagrams. Diagrams are composed of simple geometric forms, dots, lines, boxes, and more that both carry meaning and direct attention (e.g., [6, 7]). Lines direct attention by drawing the eye from place to place. Lines create meaning by conveying relationships, connections from one place to another, as in route maps or networks or line graphs. Boxes also direct attention, by bringing the eye to the contents of the boxes. Boxes are containers, they enclose one set of elements and separate them from elements in other boxes. Boxes create meaning by creating categories. They indicate that everything

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 258–264, 2016.
DOI: 10.1007/978-3-319-42333-3_21

within the box is similar, sharing features, and different from everything outside the box. Lines and boxes, like other simple geometric marks, carry meaning. They alter conclusions, inferences, and interpretations.

Lines and boxes should also bias data exploration and inferences from displays of people, place, and time. Previous research evaluated production, preference, and performance of displays of people, place, and time [3]. When asked to create ways to keep track of movements of people, most participants created matrices or tables; a minority connected people over time with lines. Preference by other participants followed the same pattern. Overall, matrices with people as cell entries and time and place in rows and columns respectively were most commonly produced and preferred. This format has good foundations. Place and time are fixed, immutable, but people can move from cell to cell. Performance was assessed by the time to verify many kinds of inferences from the data. Lines facilitated inferences about time, but all other kinds of inferences were faster from tables.

Displays of data are frequently used for data exploration, to search for patterns and generate hypotheses and inferences about the underlying processes. The form and structure of the display can influence exploration, hypotheses, inferences and conclusions (e.g., [2, 9]); Here, we investigate the roles of lines and boxes in the spontaneous generation of inferences. Because lines connect people over time, lines should bias conclusions and inferences about people, and secondarily about time. Boxes emphasize their contents, the confluence of people, place, and time, and should support a greater variety of conclusions and inferences. We deliberately used an open-ended task to capture the full range of inferences that people might generate from these types of displays.

2 Method

2.1 Participants

Eighty-one people, 39 men, participated through Amazon's Mechanical Turk website aged from 18 to 75 (mean = 30.9). Most had some college (85.4 %) and were native English speakers (93.4 %).

The stimuli (Figs. 1 and 2) were taken from [3]. The line diagram was drawn to be comparable to the box diagram. Both displayed the locations of four students at four times of day with time horizontal, place vertical, and people as cell entries. For the boxes (Fig. 1) people were represented as color-coded dots inside the boxes. For the lines (Fig. 2), people were represented as colored lines going from cell to cell. The lines intersect in the middle of the box when the person was in that location at that time; otherwise they cross the boxes at the edges. There were very few errors, so participants seemed to understand the displays.

2.2 Procedure

Participants were randomly assigned to box or line graphs. For both, the first screen, (Fig. 3) showed a different example, a bar graph, with several possible inferences:

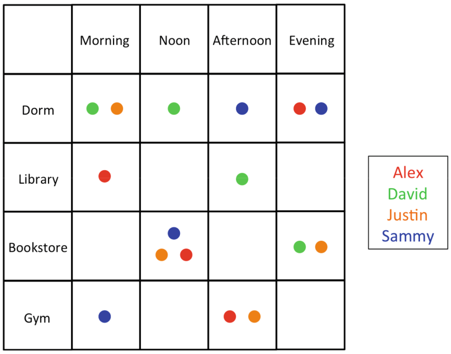


Fig. 1. The *box* stimulus display

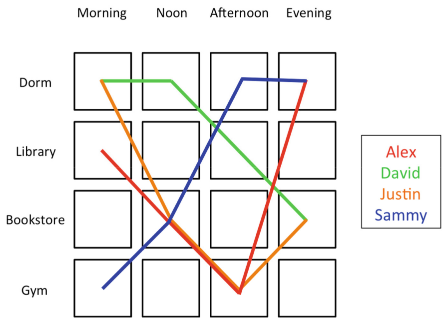


Fig. 2. The *line* stimulus display

"The population in both California and New York grew from 1900 to 2000. In 1900, the population was greater in New York than in California. In 2000, the population was greater in California than in New York." Then either the box or line diagram was presented and participants were directed: "Please study the following graph and use the space below to draw as many inferences as possible." After this task, participants were asked for demographic information.

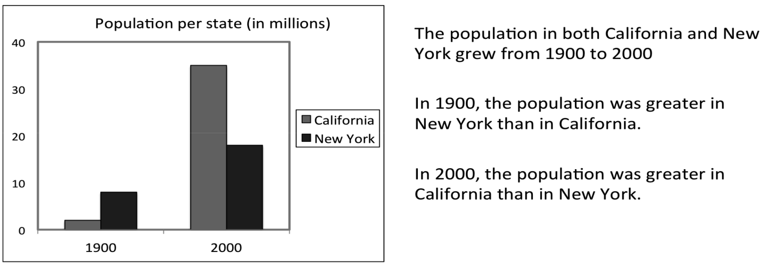


Fig. 3. Example used in the instructions.

3 Results

Because the data were presented graphically, all statements are inferences. They were coded and analyzed in two ways: the primary and secondary organizer used; and the number of different types of statements/inferences produced. The first and second authors coded. There were very few disagreements, and those cases were readily settled through discussion.

3.1 Primary and Secondary Organizers

The organizers were *Time People*, or *Location*. Here is an example with *People* as primary organizer and *Time* as secondary organizer:

> *"**David** went to the dorm in the morning, stayed at the dorm until noon, went to the library at the afternoon, and ended up at the bookstore at evening. **Justin** went to the dorm in the morning, the bookstore at noon, the gym in the afternoon and back to the bookstore at evening. **Alex** went to the library in the morning, the bookstore at noon, the gym at the afternoon and to the dorm at evening. **Sammy** went to the gym in the morning, to the bookstore and noon, to the dorm at the afternoon, and stayed at the dorm until the evening."*

The results are shown in Figs. 4 and 5. The distribution of primary organizers differs between the two conditions, $\chi 2$ (2; n = 81) = 5.815, p = .043; the secondary organizers did not: $\chi 2$ (2; n = 81) = 2.489, p = .288. For both, the primary organizer was *People*, then *Location*, then *Time*. *People* dominated far more for lines than for boxes. The dominant secondary organizer was *Time* for lines and *Place* for boxes.

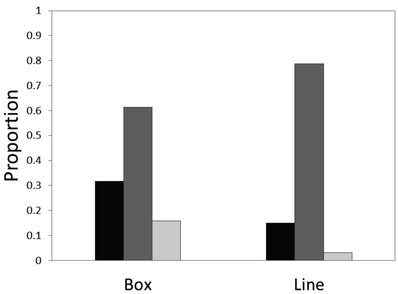


Fig. 4. Distribution of primary organizer by condition.

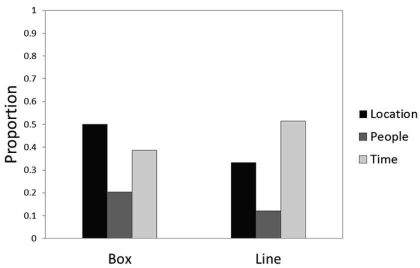


Fig. 5. Distribution of secondary organizer by condition

3.2 Number of Statements/Inferences

Careful examination of the protocols revealed the following structures:

Single Statement: a statement that referred to a single cell of the matrix, i.e., one person, one time and one place. For example, *"David is in the dorm in the morning."*

Parallel: a set of related statements in the same format. Parallel statements contain many inferences, for example: *"Justin went to the dorm in the morning, the bookstore at noon, the gym in the afternoon and back to the bookstore in the evening."*

Generality: any statement that involves more than one person, time, or place. For example, *"The bookstore is the most consistently visited places for the guys"* and "*David and Sammy spent more time in the Dorms that the others."* Generalities also include many plausible inferences.

Leap: any interpretation that went beyond the information given. For examples, *"David and Sammy are friends,"* "*David is probably unfit*," *"Alex manages his time well and gets everything done,"* and *"Since Justin does not return to the dorm in the* *evening I would infer that he is probably dating a student who works at the bookstore and spends evenings at her place."*

Negation: a negative statement from information given in the diagrams. For examples, *"David never goes to the gym,"* *"No one goes to the bookstore in the morning,"* and *"Students are not required to use the library or gym."*

The mean numbers of statements in each category are given in Fig. 6 (error bars indicate standard error). Generalities are the most common inference overall.

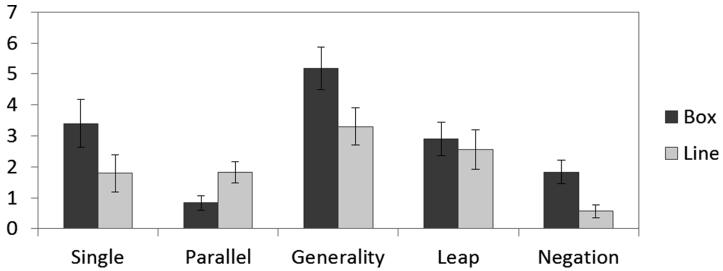


Fig. 6. Frequencies of kinds of inferences given to box and line displays.

Differences in the frequencies of the statement categories between the two displays were examined by a generalized linear model with a Poisson distribution for each dependent variable and shown in Table 1.

Table 1. Tests of differences between conditions

Inference type	Wald $\chi 2$ (1)	p-value
Single	17.828	<.001
Parallel	14.608	<.001
Generality	15.156	<.001
Leaps	0.771	.380
Negation	20.86	<.001
Word count	30.413	<.001

The analyses confirm that the box displays yielded more single statements, generalities, and negations than the line displays, and that the line displays yielded more parallel statements than the box displays. There was no significant difference in number of leaps.

4 Discussion

Diagrams of complex information use marks and place on the page to convey information effectively (e.g., [4, 6]). The marks have meanings. Lines, as in line graphs, connect like items, suggesting relations between them; they also guide the eye. Boxes, as in tables and matrices, contain like things and separate them from unlike things. Such diagrams are meant to spur a wide range of conclusions, inferences, and hypotheses. Designers of displays are faced with many decisions for portraying the data, and those choices affect the kinds of inferences that viewers make. In particular, data points can be connected by lines or enclosed in boxes. One purpose of displays of data is to encourage inferences, conclusions, and hypotheses. Different displays of the same data may encourage inferences that differ both in quality and in quantity.

Here, information about movements of people in place and time was organized with lines or boxes, corresponding to two common diagrammatic formats, line graphs and tables. Participants were asked to make as many inferences as they could from one of the displays. Overall, participants produced a large number of generalities linking information that was separated in the data, showing that they did attempt to integrate the information. In general, *People* was the dominant organizer of inferences. As predicted, the two spatial organizations of data, lines and boxes, had dramatic effects on the kinds of inferences drawn from the data,. Lines connected people over space and time. Although *People* was the dominant organizer in both cases, *People* was far more dominant when lines connected each person's movements over time, and *Time* was the dominant secondary organizer. Lines also encouraged more *parallel* inferences, inferences with the same structure and format. These are sets of inferences structured in the same way: *X went to A at time 1, to B at time 2,* etc. With boxes, people dominated as first organizer, but *Place* rather than *Time* dominated as secondary organizer. Boxes also encouraged more statements about single features of the information, more generalities involving many features, more leaps that went far beyond the information given, and more negations, that is, statements about empty cells. Importantly, both visualizations yielded large number of leaps, that is, inferences not directly supported by the data.

Displays of this information are used for exploration and understanding of the underlying phenomena driving the movements as well as conveying them to others. Visuospatial characteristics of information displays affect the kinds of inferences drawn from the information, factors like position in space, marks such as lines and boxes, and content of the dimensions. People, place, and time are three-dimensional data, and three-dimensional displays are famously difficult to comprehend, biased toward the variables on the axes (e.g., [1]). Based on previous research [3], we chose the consensus arrangement of the three variables, time on the Y axis, place on the X axis, and people as cell entries. Time and space are fixed dimensions (place was not located dimensionally here, but commonly is, in maps). Only people are movable, perhaps the reason they were selected for the cell entries.

People was by far the most popular organizer for inferences. This is most likely due to that fact that people are agents, for the most part, they decide where to go and when. *People* is also preferred to *Place* or *Time* for organizing both episodic (e.g., [5]) and autobiographical memory (e.g., [8]). In both cases, organization of memory is multiple and flexible, but organization by *People* is preferred. Location and time, like people, can be good predictors of activities, but people are agentive, and for that and a variety of other reasons, are preferred as organizers of memory.

The display format, line graph or table, affected both quantity and quality of inferences. The different patterns of inferences suggest that tables and line graphs induce different strategies for exploring the data. Those presented with tables seemed to focus on the boxes, producing more single statements that described single cells. They noticed when boxes had many entries, producing relatively more generalities, such as the crowd at the bookstore at noon. They also noted empty cells, producing negations that observed the absence of people in the bookstore in the morning or the gym at night. By contrast, those presented with line graphs used the lines to explore the data, focusing on each person's movements in turn across time and place. Lines led the eye and the mind from box to box; matrices led the eye and the mind to the boxes.

Which is better? Like almost everything, it depends. If you are tracking parcels or thieves or spies or consumers or celebrities, then lines will focus you on the important information. On the other hand, if you're entertaining many hypotheses, then use tables. Just be aware that what you choose makes a difference.

Acknowledgments. The authors gratefully acknowledge awards IIS-0725223, HHC-0905417, IIS-0855995 and IIS-0968561 from the National Science Foundation, and the Stanford Regional Visualization and Analysis Center for partial support of the research or preparation of the manuscript.

References

1. Carpenter, P.A., Shah, P.: A model of the perceptual and conceptual processes in graph comprehension. J. Exp. Psychol. Appl. **4**, 75–100 (1998)
2. Cleveland, W.S.: The Elements of Visualizing Data. Hobart Press, New York (1993)
3. Kessell, A.M., Tversky, B.: Visualizing space, time, and agents: production, performance, and preference. Cogn. Process. **12**, 43–52 (2010)
4. Nickerson, J.V., Corter, J.E., Tversky, B., Rho, Y.-J., Zahner, D., Yu, L.: Cognitive tools shape thought: diagrams in design. Cogn. Process. **14**(3), 255–272 (2013)
5. Taylor, H.A., Tversky, B.: Indexing events in memory: evidence for index preferences. Memory **5**, 509–542 (1997)
6. Tversky, B.: Visualizations of thought. Top. Cogn. Sci. **3**, 499–535 (2011)
7. Tversky, B., Zacks, J., Lee, P.U., Heiser, J.: Lines, blobs, crosses, and arrows: diagrammatic communication with schematic figures. In: Anderson, M., Cheng, P., Haarslev, V. (eds.) Theory and Application of Diagrams, pp. 221–230. Springer, Berlin (2000)
8. Wagenaar, W.A.: My memory: a study of autobiographical memory over six years. Cogn. Psychol. **18**, 225–252 (1986)
9. Zacks, J., Tversky, B.: Bars and lines: a study of graphic communication. Mem. Cogn. **27**, 1073–1079 (1999)

Exploring the Types of Messages that Pie Charts Convey in Popular Media

Richard Burns[1(✉)], Eric Balawejder[1], Wiktoria Domanowska[1], Stephanie Elzer Schwartz[2], and Sandra Carberry[3]

[1] Department of Computer Science, West Chester University, West Chester, PA 19383, USA
rburns@wcupa.edu
[2] Department of Computer Science, Millersville University, Millserville, PA 17551, USA
[3] Department of Computer Science, University of Delaware, Newark, DE 19716, USA

Abstract. In popular media, information graphics (pie charts, bar charts, line graphs) are frequently used to convey high-level intended messages. This paper focuses on the pie chart graphic type. We have collected a corpus of pie chart information graphics found in popular media, and for each chart, a team of annotators recognized its intended message. In this paper, we report on the types of intended messages that the team of annotators recognized and their inter-annotator agreement. We also briefly survey some of the communicative signals that graphic designers used which helped the annotators recognize these messages.

1 Introduction

Information graphics, such as bar charts and line graphs, are common visual devices frequently incorporated into multimodal documents to achieve a set of communicative goals [5,6]. In popular media (magazines such as *Time* and newspapers such as *USA Today*), information graphics are sometimes included in an article to convey some additional, supplemental high-level message that transcends supporting data, rather than simply providing low-level data points. For example, the grouped bar chart in Fig. 1 ostensibly conveys a high-level message that *"Women are more likely than men to delay medical treatment"*.

The idea that information graphics can be considered a form of language follows Clark [3] who noted that language is any "signal" or lack thereof, where a signal is any deliberate action that is intended to convey a message, including gestures and facial expressions. Thus, we view information graphics as a form of language, where the designer of a graphic is able to deliberately use *communicative signals* to help convey an intended message to the viewer of the graphic.

This paper presents preliminary results in our study of designing a system that can automatically reason about the most likely intended message of a pie chart, using present or absent communicative signals in the graphic as evidence.

It is non-trivial to identify the intended message of an information graphic; Carberry et al. [2] found that a graphic's message is often not contained in the

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 265–271, 2016.
DOI: 10.1007/978-3-319-42333-3_22

graphic's caption or in the article accompanying the graphic. Thus, the use of natural language processing techniques only on the graphic's caption or only on surrounding article text cannot be relied on to provide enough evidence to recognize the graphic's high-level message.

Previously, our research group has implemented intended message recognition systems for other kinds of information graphics: simple bar charts [4], line graphs [7], and grouped bar charts [1]. These three implemented systems use a Bayesian network to probabilistically capture the relationships between high-level intended messages and communicative signals that help signal the messages. Because each type of information graphic is able to convey a unique set of possible messages compared to the other information graphic types, the end-result for each of the systems has been very different. Simple bar charts, line graphs, and grouped bar charts each have a different set of message categories, and different communicative signals are utilized by graph designers to help convey the high-level intended messages.

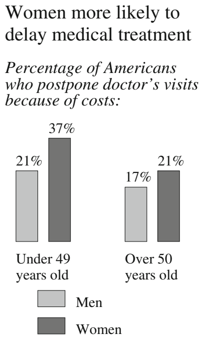


Fig. 1. From *USA Today*.

This work is the first of our knowledge that studies the problem of recognizing the intended high-level message of a pie chart when it is drawn in popular media.

We have collected a set of pie chart information graphics occurring in popular media, and examined these charts to identify (1) the types of high-level messages that graphic designers convey using pie charts, and (2) the kinds of communicative signals present in pie charts that appear likely to assist the recognition of high-level messages. Unsurprisingly, in our preliminary investigation so far, the types of recognized high-level messages and identified communicative signals are different than those in simple bar charts, line graphs, and grouped bar charts.

One application of this research is for sight-impaired individuals who cannot view information graphics. Alternative access screen readers can convert the content of a pie chart to text, but only at the level of low-level raw data: (e.g. *"the first pie chart slice is 18.5 %, the second pie chart slice is 7.3 %, ..."*). Our research aims to generate the high-level message as text for sight-impaired users.

Section 2 of the paper describes some of the messages categories that we identified and Sect. 3 presents some of the communicative signals that we found. Section 4 introduces some unexpected properties of pie charts in popular media that could be avenues for interesting future work.

2 Pie Chart Message Categories

We collected 115 pie chart information graphs from popular media.[1] Of those, we retained 90 of the charts, as the rest appeared to contain *only* data, and did

[1] The corpus of pie charts is available at: http://www.cs.wcupa.edu/rburns/piecharts.

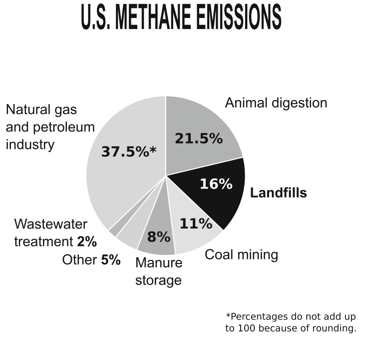


Fig. 2. From *National Geographic*

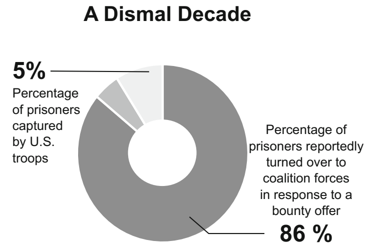


Fig. 3. Graphic from *Time Magazine.*

not appear to the annotators to convey any intended message. (Inter-annotator agreement is discussed later.) We then analyzed the corpus to generalize the kinds of high-level intended messages that we recognized into *message categories.*

There are nine pie chart message categories that we defined. Because of space constraints, we can only present graphical examples for a subset of the message categories. Below, we formally define the name of each category, the number of parameters that messages in each category take, and provide a short description.

SingleSlice($< s >$)**.** Single slice messages recognize a high-level message that involves a single, salient, pie chart slice. Generally, the pie charts that fall within this category seem to be designed so that the graph viewer compares a specific, single slice against the other slices in the pie chart. For example, consider the pie chart in Fig. 2. This pie chart ostensibly conveys that *Landfills are a significant source of U.S. methane emissions, the third highest, behind the natural gas and petroleum industry as well as animal digestion.* The parameter $< s >$ in the message category syntax is instantiated with the single pie chart slice that is to be compared against the other slices. That is, this message would be represented as: $SingleSlice(s = Landfills)$.

Versus($< s_1, s_2 >$)**.** Versus messages capture two salient slices, which are compared against each other. In contrast to single slice messages in which a salient pie chart slice is compared against the rest of the slices in the pie chart, the two salient slices in versus messages are compared with each other rather than the other slices. For example, the pie chart in Fig. 3 ostensibly conveys the message that *most prisoners were turned over to coalition forces because of bounties, rather than being captured by troops.* The versus message category is instantiated with two parameters: s_1 and s_2, the slices that should be compared with each other.

BiggestSlice( ). Biggest slice messages identify a single slice of the pie chart that is larger than all of the other slices. Because only one slice can be the largest (assuming no ties), the biggest slice message category has no parameters. For example, presumably the intended message in the pie chart in Fig. 4 is that *there* *were a greater number of male deaths than female deaths in which illicit fentanyl was detected.*

NoMajority( ). No majority messages capture that none of the slices in the pie chart are larger than 50 %. For example, the pie chart in Fig. 5 ostensibly intends to convey the high-level message that *individuals in search of work take a variable range in time in order to find a job.*

Fraction($< s >$)**.** Fraction messages represent that slice $< s >$ is a fractional percentage of the pie chart, such as the messages *juniors make up* <u>one third</u> *of the class* and <u>half</u> *of the revenue is from Philadelphia.*

AddSlices($< s_1, s_2, ..., s_n >$)**.** Add slices messages recognize the aggregation of multiple slices. Each slice that is added together is a parameter in this category.

TwoTiedForBiggest($< s_1, s_2 >$)**.** Two tied for biggest messages capture that two slices in the pie chart are approximately the same size.

SmallestSlice( ). Smallest slice messages identify a single slice of the pie chart that is smaller than all of the other slices.

NumberOfParts( ). Finally, number of parts messages capture the quantity of slices in a pie chart, for a message such as, *there are* <u>six</u> *reasons identified for not working among uninsured adults.*

2.1 Most Frequent Message Categories

The information graphic types of simple bar charts [4], grouped bar charts [1], line graphics [7], and pie charts, each have a different set of message categories though some categories do overlap. As shown in Table 1, the top two most frequent message categories for each graphic type contain around 30–50% of the collected graphics in those corpora. Notably, while the most frequent pie chart

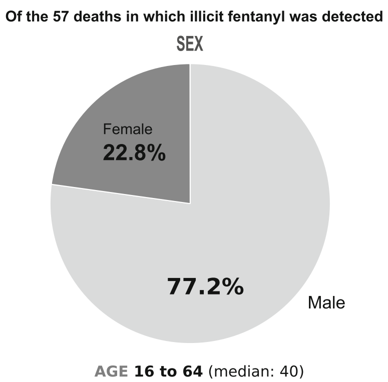


Fig. 4. From *The Philadelphia Inquirer.*

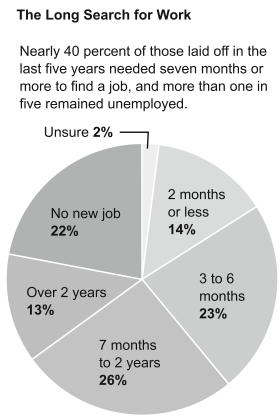


Fig. 5. From *The Philadelphia Inquirer.*

Table 1. Most frequent high-level messages by information graphic type.

Message category	Number of occurrences	Percentage
Simple bar charts		
Increasing-trend	26	23.6 %
Maximum	25	22.7 %
Grouped bar charts		
Entity-comparison	68	20.6 %
Rising-entities-all	36	10.9 %
Pie charts		
Biggest-slice	24	23.3 %
Single-slice	20	22.2 %
Line graphs		
Rising-trend	66	27.5 %
Change-trend	58	24.2 %

messages involve a *single salient slice*, the most frequent simple and grouped bar chart messages are distributed between either a *trend* message or a message that involves a *single bar entity*. The most frequently occurring messages in line graphs involve *trends*. These results highlight the importance of studying each of the information graphic types separately, and also can be used to inform the process of designing appropriate information graphics.

2.2 Annotation and Inter-Coder Agreement

The annotation of the corpus was performed with the following process: we first individually recognized the intended message for each pie chart and classified it into its appropriate intended message. Then, we conducted a consensus-based annotation by meeting as a group and discussing each of our annotations, revising any annotations if we were strongly swayed. The final annotation for each pie chart was decided by majority vote.

Three coders met and deliberated final annotations for 30 of the pie charts in the corpus. Notably, all of the individual annotators sometimes recognized exactly the same message for a pie chart before any discussion, or a majority of them agreed to exactly the same message after a discussion.[2] This level of agreement is a good result and shows that (1) the recognition of pie chart messages is not as subjective as it may initially appear, and (2) our derived and recognized set of pie chart message categories does capture the types of messages that graphic designers convey in popular media using pie charts. A summary of the inter-annotator agreement is shown in Table 2.

[2] Two annotations were only counted as matching if they had: (1) the same message category and (2) the same instantiation. For example, the two messages *SingleSlice(Landfills)* and *SingleSlice(Animal digestion)* would not be a match because their instantiations are not identical.

Table 2. Summary of the annotation agreement between coders. Table rows display *The percentage of pie charts that ...*

Percentage	Description
36.6 %	All coders recognized with exactly the same message, before any discussion
56.6 %	A majority of coders recognized with exactly the same message, before any discussion
63.3 %	All coders recognized with exactly the same message, after discussion
100 %	A majority of coders recognized with exactly the same message, after discussion

3 Communicative Signals

The presence and absence of communicative signals assist the recognition of a high-level intended message conveyed in a pie chart.

Visual Signals. One visual signal that a graphic designer may use to help communicate some intended message is prominence, by coloring a specific pie chart slice a salient coloring, or boldfacing the label of a pie chart slice. An example of this communicative signal is present in Fig. 2, which helps signal that *Landfills* should be compared against the other pie chart slices. Another example of a visual signal found in the pie chart corpus is the use of similar colors across multiple pie chart slices. For example in Fig. 3, the slices for *Bounty* and *Troops* are colored similarly (though not exactly identical), helping signal that they should be compared, while still contrasting them against the *Unlabeled 9 %* slice.[3] Another example of a visual, communicative signal is separation, when one pie chart slice is purposely drawn slightly "separated" or "exploded" away from the center of the pie, drawing additional attention to it.

Linguistic Signals. Although it does not always fully capture a graphic's intended message, the caption text of a pie chart can sometimes serve as a linguistic signal that helps convey its message. For example, in the pie chart in Fig. 6, the verb *split* helps signal the intended message that *there is no majority slice amongst the slices: "will", "will not", and "unsure"*. We have also observed instances of the article headline of a multimodal article helping to signal the intended message of a pie chart. Another linguistic clue that can serve as a communicative signal is when one pie chart slice is mentioned in the caption or article headline, while the other slices are not mentioned.

4 Conclusion

There are several avenues of future work that we are exploring: First, we are currently constructing a Bayesian network, which has a top-level node with states that enumerate all possible pie chart messages. This top-level node is linked to children leaf nodes that represent the possible communicative evidence in a graphic. Given our corpus of pie chart graphics, we will train the network to learn the probabilistic relationships between pie chart high-level intended messages and the communicative evidence that is present or absent in the charts.

[3] In the original graphic, *Bounty* is colored yellow, *Troops* is orange, and the unlabeled slice is gray.

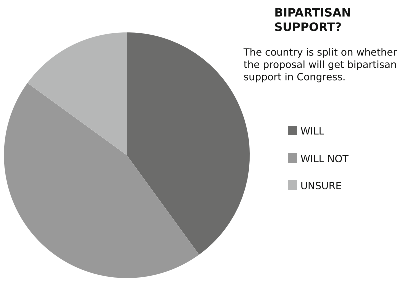


Fig. 6. Graphic from *USA Today.*

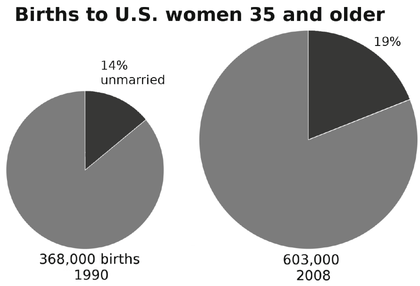


Fig. 7. From *National Geographic.*

Second, we have observed numerous instances of *multiple pie charts* drawn adjacent to one another, where the single intended message of the graphic seems to involve *both* pie charts, rather than two individual and separate intended messages. For example in the multiple pie charts shown in Fig. 7, the high-level message conveyed is that *the percentage of births to unmarried U.S. women 35 and older increased from 1990 to 2008.* This avenue of future work explores the unique types of messages and communicative signals that can be found when multiple pie charts are purposely drawn adjacent to each other.

Summary. In this paper, we have presented novel research that introduces (1) a corpus of pie charts that we have collected from popular media, (2) a sampling of the types of messages that pie charts are able to convey, and (3) examples of communicative signals that help communicate these messages. These identified messages and communicative signals are unique compared to other types of information graphics that have been previously studied.

References

1. Burns, R., Carberry, S., Elzer, S., Chester, D.: Automatically recognizing intended messages in grouped bar charts. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 8–22. Springer, Heidelberg (2012)
2. Carberry, S., Elzer, S., Demir, S.: Information graphics: an untapped resource of digital libraries. In: Proceedings of 9th International ACM SigIR Conference on Research and Development on Information Retrieval, pp. 581–588 (2006)
3. Clark, H.: Using Language. Cambridge University Press, Cambridge (1996)
4. Elzer, S., Carberry, S., Zukerman, I.: The automated understanding of simple bar charts. Artif. Intell. **175**(2), 526–555 (2011)
5. Green, N.L., Carenini, G., Kerpedjiev, S., Mattis, J., Moore, J.D., Roth, S.F.: Autobrief: an experimental system for the automatic generation of briengs in integrated text and information graphics. Int. J. Hum.-Comput. Stud. **61**(1), 32–70 (2004)
6. Iverson, G., Gergen, M.: Statistics: The Conceptual Approach. Springer, New York (1997)
7. Wu, P., Carberry, S., Elzer, S., Chester, D.: Recognizing the intended message of line graphs. In: Goel, A.K., Jamnik, M., Narayanan, N.H. (eds.) Diagrams 2010. LNCS, vol. 6170, pp. 220–234. Springer, Heidelberg (2010)

Comparison of Diagrams in Producing and Understanding Hierarchies in Three Different Application Domains

Leonie Bosveld-de Smet(✉) and Robert-Jan Verheggen

Department of Information Science, University of Groningen, Groningen, The Netherlands
l.m.bosveld@rug.nl, r.l.h.verheggen@student.rug.nl
http://www.rug.nl

Abstract. We compare diagrammatic representations of hierarchically ordered information in three different well-known application domains: organizational structure, folder and file structure, and calculation order in arithmetic expressions. Although there exists a natural link between hierarchies and trees, the diagrams conventionally associated with these domains are the tree diagram, indentation, and nested parentheses, respectively. To investigate the effects of inherent natural correspondence and convention on the production and comprehension of hierarchies, an experiment was set up. The results show that participants prefer tree representations for visualizing hierarchical structure of organizations and folders, but don't construct trees for visualizing arithmetic hierarchy. In comprehension, differences were observed between the three interfaces for both response accuracy and response time. Trees performed best on accuracy. Parentheses took most time to process. When used in their conventional domain nested parentheses performed best. For trees and indentation, the results are less clear.

Keywords: Hierarchies · Diagrams · Trees · Indentation · Nested parentheses · Construction · Comprehension · Conventions

1 Introduction

Hierarchies involve entities and connections between these entities. These connections are organized into levels that indicate status differences [6]. Hierarchies occur in many application domains, and can be visually represented in a variety of ways. The efficiency of their visual representation relies on domain knowledge, convention [2,3,5], and on the naturalness of the link between the representation and the information depicted [7]. There exists a natural link between hierarchies and tree diagrams [1,6]. This paper explores the question how tight this correspondence is, and to which extent this correspondence is overruled by conventional associations of hierarchies with diagrams other than trees. More specifically, this paper focuses on visual representations of hierarchies in three

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 272–279, 2016.
DOI: 10.1007/978-3-319-42333-3_23

generally well-known application domains: organizational structure, folder and file systems, and arithmetic expressions. For organizational hierarchy, it is common to use a tree diagram. Folder and file systems are displayed by many file browsers in an indentation representation. In arithmetic expressions nested parentheses are used to indicate the order in which operations are to be applied. Like graphs and charts, these three visual representations can be seen as schematic diagrams, as they provide a concise representation of abstract information [2]. Each of the three diagrams conveys meaning in a different way. First, the entities involved, the number and nature of their possible hierarchical connections are different. The entities are conceptually different, and in the organizational and folder hierarchy, they are less abstract than in the arithmetic one. In organizational and folder hierarchies, non terminal nodes may dominate one or more nodes at a lower level. Arithmetic expressions always involve binary branchings. Connections in organizational hierarchies are hierarchical in the sense that entities at a higher level manage and control entities at a lower level. In folder structures, the connection is one of inclusion. In arithmetic expressions, it is one of the subsequent operational orders. Second, the visualization approach differs. While the tree and indentation diagram exploit the horizontal and vertical dimension of the display space to visualize hierarchical connections, nested parentheses make use of only one dimension, and uses special characters (parentheses or similar characters) to mark levels and group entities. The tree diagram is arguably the only fully graphical diagram, as it uses a line to connect entities of different hierarchical levels. Each diagram has its variants, dependent on orientation and/or order in which objects of the same level are placed on the display. All domains involve hierarchies based upon the same logical information, but this information is represented differently in the three different domains. In order to gain insight in the relative merits of each type of diagram in producing and understanding hierarchies in the application domains selected, we have conducted an empirical experiment. In this paper, we are especially interested in effects of the interaction of natural correspondence and conventional association on use and interpretation of diagrams to communicate the hierarchical message in different application domains.

2 Method

The experiment consisted of two parts: (i) construction of a visual representation based upon a textual description of a hierarchy in the three domains selected for this study; (ii) comprehension of hierarchies in these domains via tree, indentation and nested parentheses. The diagrams were displayed with as few graphical devices as possible. Textual labels were used to denote the entities. Trees were shown top-down. Prefix notation was chosen for the indentation and parentheses representations of organizational and folder structure, but for the arithmetic expressions we decided to use infix notations. The design of the experiment was within-subject. The independent variable in both parts is the application domain with three conditions. In the comprehension part, the diagram type with three conditions is added as second independent variable.

2.1 Participants

Sixty-three participants took part in the experiment, in the period from September 7 to September 18, 2015. The participants recruited were all native Dutch speakers enrolled at the University of Groningen (33 male; 30 female; average age: 22.3 with a standard deviation of 2.2). Domain knowledge and conceptual knowledge necessary for correct interpretation of the representations were assumed to be available to all participants.

2.2 Design and Materials

In the production task, each participant was told to read a textual description (displayed on a computer screen) of an instance of a hierarchical ordering in each of the three application domains (presented in random order, and verbalized in terms suited to the domain at hand), and to draw by hand, with paper and pencil, a visual representation of the information provided. No specific instructions were given, as the goal was to elicit spontaneous drawings. Participants were allowed as much time as they needed to construct the diagrams. All drawings were categorized as belonging to a certain representation type. Categories were obtained inductively. Figure 1 illustrates the hierarchy description for folder structure (translated from Dutch).

> *Fred loves to play computer games. His Games folder is stored in the folder Applications, together with Adobe and Microsoft. The Games folder contains two subfolders: Sports and ShootingGames. Microsoft has one subfolder: Office, which contains Excel and Word. Fred has stored a folder CallOfDuty in ShootingGames. He made a folder for his game FIFA, which he stored in a subfolder Soccer of Sports.*

Fig. 1. Illustration of textual description in the folder domain for the production task.

In the comprehension task, participants were asked to respond to yes-no questions about hierarchical relations in each domain (organization, folders, arithmetics), and for each representation (tree, indentation, nested parentheses). The questions, verbalized in domain specific terms, were shown on a computer screen (display size 2560×1600) together with one of the three representations. Correct responses and response times were recorded for each question. To reduce learning effects and fatigue, only nine questions were selected from a total set of twenty-seven questions, which varied slightly in formulation for each domain. The hierarchies represented varied in complexity. A randomization algorithm for the selection of these nine questions from the whole set of twenty-seven questions was setup in such a way that all participants had to process all three visualization types in a random order for different domains. Table 1 shows examples of questions (translated from Dutch) participants were asked to respond to in the comprehension task.

Table 1. Illustration of questions for each domain in the comprehension task.

Domain	Examples of questions
Organization	Does the *Vice-Director* have a higher position in the organization than the *Sales Director*?
Folders	Starting in the folder *Pictures*, can the folder *Holiday* be opened with fewer clicks than the folder *Europe*?
Arithmetics	Must the addition of x and y be performed before the division operation can be performed?

2.3 Procedure

The experiment was conducted individually with each participant. The participant was positioned at a desk in a quiet room with a computer, screen, keyboard and mouse. Production preceded comprehension. The participant read instructions and scenarios from the computer screen and did the drawing manually. Next he/she was directed again to the computer screen to respond to the questions. The whole experiment took about twenty minutes for the construction task and five for responding to the questions.

3 Results

3.1 Production

The drawings produced were categorized as belonging to one of the following ten categories, collected inductively: tree, indentation, nested parentheses, network, treemap, table, pie diagram, iconic diagram, text, other. Figure 2 shows hand sketches that two participants constructed in response to the text given in Fig. 1. The left picture was categorized as treemap, the right one as indentation.

Table 2 shows the ten diagram types used to categorize the drawings (N = 189) produced for the three different domains by the participants (N = 63),

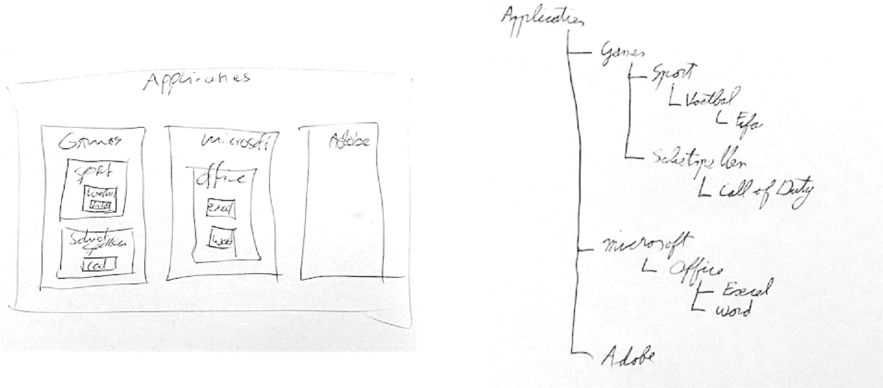


Fig. 2. Sketches drawn by two participants in response to the text in Fig. 1.

Table 2. Frequencies of representation types constructed in the production task.

Representation type	Domain type			Total
	Organization	Folders	Artithmetics	
Tree diagram	29	25	0	54
Indentation	0	10	0	10
Parentheses	0	0	7	7
Network	28	16	8	52
Treemap	0	4	0	4
Table	0	0	12	12
Iconic diagram	0	0	10	10
Pie diagram	0	0	2	2
Text	0	0	5	5
Other	6	8	19	33
Total	63	63	63	189

with their frequencies. Trees and networks were popular among the representations constructed for both organizational hierarchy and folder structure. The arithmetic scenario gave rise to the most diversity. Remarkably no tree or tree-like diagrams were drawn in response to this scenario. Table 2 also shows that, for each domain, the hierarchy representation conventionally linked to it, was produced by some of the participants. In the organization condition, nearly half of the participants draw a tree.

3.2 Comprehension

Figure 3 visualizes as box plots the results for the response times in the comprehension task. Figure 4 shows the results for accuracy as percentages of incorrect answers. Both figures give an overview of the results obtained for each representation type independent of domain, and for those for each specific domain.

The results are reported for different numbers of participants. All results of two participants were filtered out, as they were clear outliers. Additionally, outliers with respect to response times at representation and domain level were removed from the results visualized in Fig. 3. Figure 3 shows that the mean and median of the response times in the parentheses condition for all domains are higher than those in the tree and indentation condition. An ANOVA with repeated measures with a Greenhouse-Geisser correction shows that the mean scores for response times are statistically significantly different ($F(1.752, 105.121) = 11.053$, $p = 0.000$). Post-hoc testing with Bonferroni correction indicates that performance on the question task with nested parentheses has a significant impact on the time spent to complete the task (mean = 27930.61, std. dev. = 13962.096). Parentheses were slower processed than the other interfaces. A Wilcoxon Signed-Ranks test was used to test significant differences in response

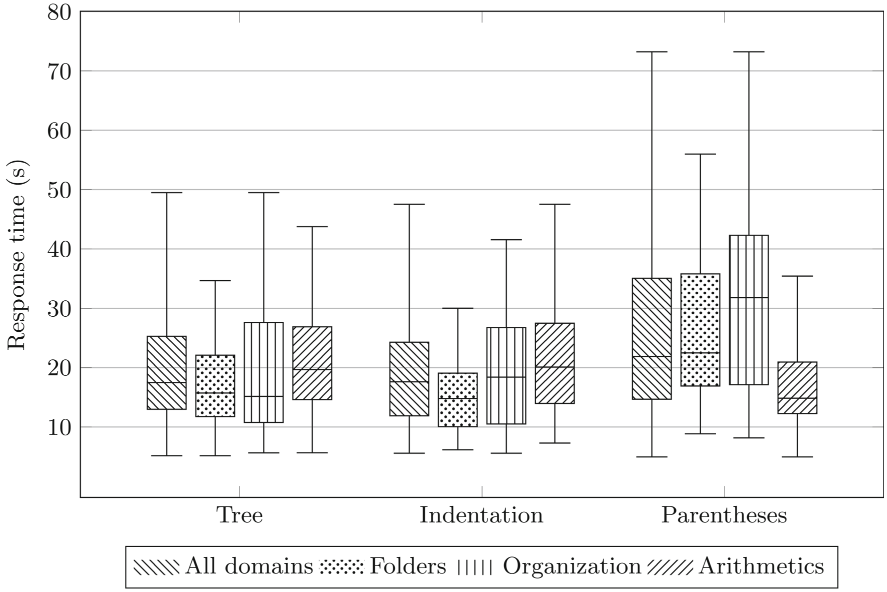


Fig. 3. Average response time for each representation type in general and for each application domain.

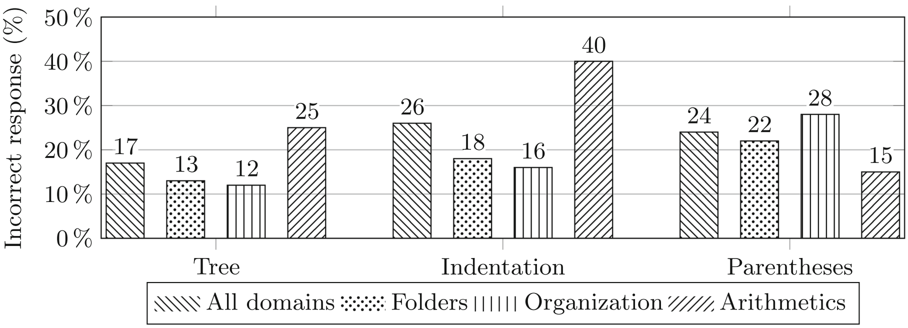


Fig. 4. Percentage of response error for each representation type, in general and for each application domain.

accuracy (see Fig. 4). This test shows that trees perform significantly better on accuracy than indentation ($Z = -3.332$, $p = 0.001$). This is especially caused by the combination of indentation and arithmetic hierarchy which turned out to yield very slow performance. Arithmetic expressions were most easily processed in a representation with nested parentheses. However, for the folder and especially the organizational domain, nested parentheses yielded worse performance than trees and indentation.

4 Discussion and Conclusion

The purpose of this paper was to get insight in the use of specific hierarchy representations and in the ease of processing them with respect to three different application domains. We were interested in the question to which extent the natural correspondence between a hierarchy and a tree would be overruled by conventional associations. Neither the results of the production task nor those of the comprehension task lead to the unequivocal conclusion that trees are the most prominent and best performing candidates for processing hierarchies in the three different domains considered in this paper, although trees turned out to yield the most accurate performance. Parentheses were clearly slowest. We discern an important difference between the tree and indentation representation on the one hand, and nested parentheses on the other. Nested parentheses are difficult to process by users in a folder or organizational hierarchy. This may be caused by the one-dimensionality of parentheses which do not exploit the levels property [4], and which look cluttered, especially in combination with (large) textual labels to denote the hierarchy entities. Interestingly, this inconvenience did not lead to participants producing fewer correct answers. In contrast, linear representations with parentheses seem to fit operational hierarchy in arithmetics far better than folder and organizational structure. The results of this study suggest that trees and especially representations with indentation perform poorly for arithmetics. This may be due to the fact that at a conceptual level arithmetic expressions are not associated with hierarchies. This observation is indeed supported by the findings of the production task, where no participant constructed a tree or tree-like diagram in response to the arithmetic scenario. The infix notation we decided to use may also have caused the slow processing of indentation representing arithmetic expressions. Finally, we assumed domain knowledge to be fully available for all domains. Future work will explore whether this is a confounding factor.

References

1. Green, T.R.: What kindergartners know about class inclusion hierarchies. J. Exp. Child Psychol. **57**, 72–88 (1994)
2. Hegarty, M., Carpenter, P.A., Just, M.A: Diagrams in the comprehension of scientific texts. In: Ban, R., Kamil, M.L., Mosenthal, P., Pearson, P.D. (eds.) Handbook of Reading Research, vol. 2, pp. 641–668 (1991)
3. Hurley, S.M., Novick, L.R.: Solving problems using matrix, network, and hierarchy diagrams: the consequences of violating construction conventions. Q. J. Exp. Psychol. **63**, 275–290 (2010)
4. Körner, C.: Concepts and misconceptions in comprehension of hierarchical graphs. Learn. Instr. **15**, 281–296 (2005)
5. Novick, L.R.: The importance of both diagrammatic conventions and domain-specific knowledge for diagram literacy in science: the hierarchy as an illustrative case. In: Barker-Plummer, D., Cox, R., Swoboda, N. (eds.) Diagrams 2006. LNCS (LNAI), vol. 4045, pp. 1–11. Springer, Heidelberg (2006)

6. Novick, L.R., Hurley, S.M.: To matrix, network, or hierarchy: that is the question. Cogn. Psychol. **42**, 158–216 (2001)
7. Zacks, J., Tversky, B.: Bars and lines: a study of graphic communication. Mem. Cogn. **27**, 1073–1079 (1999)

Posters

Evaluating the Effects of Colour in LineSets

Dominique Tranquille[1(✉)], Gem Stapleton[1], Jim Burton[1], and Peter Rodgers[2]

[1] Visual Modelling Group, University of Brighton, Brighton, UK
d.tranquille1@uni.brighton.ac.uk, {g.e.stapleton,j.burton}@brighton.ac.uk
[2] University of Kent, Canterbury, UK
p.j.rodgers@kent.ac.uk

1 Introduction

Data items often lie in overlapping sets and a number of set visualization techniques have been developed in recent years [1,4]. An example is in Fig. 1, which visualizes sets by overlaying lines on an existing visualization of data items. As LineSets are an overlay technique, they can be applied to many different data sets. Here, we apply them to network diagrams. LineSets are composed of lines that are overlaid on nodes or a node-link network; examples are given in Figs. 1 and 2. The set-lines are labelled to indicate the represented set. Nodes represent individual entities (data items). A node represents a member of a set if the set-line passes through it. Nodes that are not passed though by set-lines are not members of any of the sets. Two nodes represent related data items if an edge passes between them. Alper et al.'s initial paper on LineSets focussed on exploring the potential of their new technique [1]. Their research established that Linesets should be generated with paths that are as linear as possible as well as being smooth. However, there are a number of other graphical choices to be made when drawing LineSets; one of these choices is colour. This paper identifies how colour (hue, value, or monochrome) should be applied to LineSets drawn on networks. The study materials and collected data is available at http://www.cem.brighton.ac.uk/research/VMG/linesets-study-2015.

2 Experiment Design

The aim was to establish how different colour treatments applied to LineSets affect task performance. The three colour treatments that we used were unique hues, varying values and monochrome. For our unique hues colour treatment, we selected colours that were uniformly distinct from each other. The colour palette we created for the varying values treatment consisted of a single colour hue that was broken down into stepped levels of lightness (Fig. 2). The colours were generated from Color Brewer [3]. Black was used for all items in the monochrome treatment.

We used a between-group design with repeated measures and three groups. We recorded two dependent variables: the time taken to answer the question and whether the answer was correct. If colour treatment impacts on task performance

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 283–285, 2016.
DOI: 10.1007/978-3-319-42333-3_24

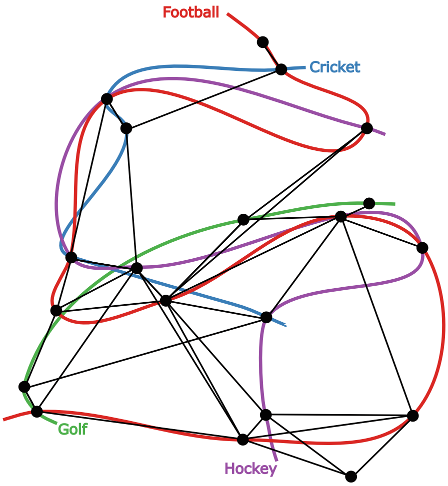


Fig. 1. Unique hues.

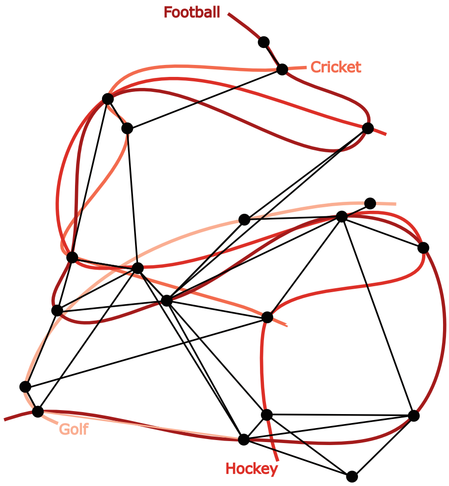


Fig. 2. Varying values.

then we expect a significant difference to exist in either accuracy or time. All of the tasks performed by participants in our study fit into Simonetto et al.'s group-level graph visualization taxonomy [5]. This taxonomy identifies four question types suited to LineSets drawn on networks: group only (about sets), group-node (about sets and data items), group-link (about sets and network connections) and group-network (about sets, data items and network connections). Covering all four task types allowed us to collect meaningful data about general task performance.

3 Statistical Analysis

We used the data collected from 60 participants together with the six pilot participants, as no changes had been made to the experiment. This gave us a total of 66 participants (55 M, 11 F, age: average 24, range 18 to 46). Each participant answered 24 questions giving us a total of 1584 observations. There were a total of 31 unanswered 'timeouts', where participants had not provided an answer within the two minutes allowed (monochrome 19, hues 6 and values 6). These were removed from our data prior to analysis, as no answer to the question was provided.

Of the 1553 questions for which answers were provided within the 2 min allowed, there were 556 errors (35.8 %). Monochrome had the highest inaccuracy rate with 258 errors (50.7 %). Varying hues and values had inaccuracy rates of 138 (26.4 %) and 160 (30.6 %) errors respectively. We conducted a chi-square test, giving a p-value of 0.000. It was established that monochrome yielded significant more errors than both varying hues and varying values. However, varying hues and varying values were not significantly different.

With regards to time taken, the overall mean completion time, for the 997 correct answers, was 38.01 s and the standard deviation was 22.14. Varying hues achieved the lowest mean completion time and standard deviation of 30.92 (18.07). Varying values and monochrome achieved mean completion times and standard deviations of 37.36 (20.91) and 49.81 (24.57) respectively. We performed an ANOVA test in order to determine the overall affect of colour treatment on time performance. However, the data were not normally distributed so we applied a log transformation which yielded a skewness of -0.01 and, thus, rendered our data suitable for analysis. The results revealed significant differences between the colour treatments, with a p-value of 0.000. A Tukey Test revealed that varying hues allowed participants to perform significantly faster than varying values. Both these treatments were significantly faster than monochrome.

4 Conclusion and Future Work

This paper set out to address the question *how does the use of colour affect the comprehension of LineSets?* We selected three colour treatments based on perceptual theories on the use of colour: monochrome, varying values and varying hues. Taking into account both our error analysis and time analysis, LineSets should be treated with varying hues.

Whilst we have focused on the use of colour in LineSets, there are other graphical properties that need to be understood. Properties such as size and shape, along with colour, can have a profound impact on perception [2]. In the context of LineSets, size corresponds to graphical properties such as set-line thickness and node size. Regarding shape, the set-lines can take many routes when being overlaid on nodes, and be either smooth or composed of a series of straight line segments. Evaluating these, and other, graphical properties will lead to a more comprehensive understanding of how to draw LineSets to aid task performance.

References

1. Alper, B., Riche, N.H., Ramos, G., Czerwinski, M.: Design study of LineSets, a novel set visualization technique. IEEE Trans. Vis. Comput. Graph. **17**(12), 2259–2267 (2011)
2. Bertin, J.: Semiology of Graphics. University of Wisconsin Press, Madison (1983)
3. Brewer, C., Harrower, M.B., Sheesley, D.H., Woodruff, A.: Color Brewer (2013). http://colorbrewer2.org/. Accessed 14 Aug 2015
4. Meulemans, W., Riche, N.H., Speckmann, B., Alper, B., Dwyer, T.: Kelpfusion: a hybrid set visualization technique. IEEE Trans. Vis. Comput. Graph. **19**(11), 1846–1859 (2013)
5. Saket, B., Simonetto, P., Kobourov, S.G.: Group-level graph visualization taxonomy. In: EuroVis - Short Papers (2014)

Negative Terms in Euler Diagrams: Peirce's Solution

Amirouche Moktefi(✉) and Ahti-Veikko Pietarinen

Tallinn University of Technology, Tallinn, Estonia
{amirouche.moktefi, ahti-veikko.pietarinen}@ttu.ee

Abstract. We commonly represent a class with a curve enclosing individuals that share an attribute. Individuals that are not predicated with that attribute are left outside. The status of this outer class has long been a matter of dispute in logic. In modern notations, negative terms are simply expressed by labeling the spaces that they cover. In this note, we discuss an unusual (and previously unpublished) method designed by Peirce in 1896 to handle negative terms: to indicate the position of the terms by the shape of the curve rather than by labeling the spaces.

Keywords: Negative term · Euler diagram · Charles S. Peirce · Syllogism

1 Introduction

Traditional Euler diagrams were first introduced to tackle syllogistic problems where only positive terms occur [1]. Hence, they hardly lend themselves to the treatment of negative terms. For instance, the outer space standing for the negation of all the terms in the argument is always shown to exist. Hence, it is not possible to express its absence without further amendments. Early logicians offered several solutions to overcome this difficulty. An obvious trick consists in replacing a negative term by a positive one during the resolution of a problem [2: 63]. For instance, a proposition "Some *x* are *not-y*" might simply be transformed into "Some *x* are *z*" (with *z* = *not-y*) and, consequently, be represented with traditional Euler diagrams. However, this method works merely when a term is not expressed twice with opposite signs in the considered problem.

Another solution would be to enhance Euler diagrams in order to represent actual relations between terms and their opposites, rather than positive terms alone. This can be achieved by representing the universe of discourse and thus devoting a finite space to the outer region of the diagram if existent, or no space at all if absent [3: 170–4]. An advantage of this solution is that it produces true Euler-type diagrams that require no additional conventions for their usage. However, this solution suffers from the complexity and the multiplicity of the figures needed for solving the problems, and thus increases the risk of misusing the diagrams.

This objection disappears in the case of Venn diagrams where all combinations of terms are first represented by compartments before syntactic signs are added to mark them and indicate their status [8]. However, such diagrams are not Euler-type since they do not represent actual information [6]. Therefore, they stand beyond the scope of the present note.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 286–288, 2016.
DOI: 10.1007/978-3-319-42333-3_25

2 Peirce's Solution

Charles S. Peirce worked on several amendments of Euler diagrams. One particular innovation from 1896 offers an unusual approach to negative terms [7]. Indeed, logicians commonly depict positive terms inside the curves [4, 5]. This usage is conventional and would not operate if we were to draw diagrams on a sphere. To indicate the term's sign, Peirce rather draws closed curves that have convex and concave sides. Then, he demands that positives are found on the concave side of the curve and negatives on the convex one. This does not infringe the common usage of Euler diagrams, since the concave side of a circle is inside it, which means that the positive term is still enclosed in the circle. However, Peirce's idea opens the way to various new shapes where the negative term is found *inside* the curve [Fig. 1a].

This new 'hyperboloid' method greatly simplifies the representation of propositions with negative terms. For instance, proposition "No *not*-A is *not*-B" which denies the existence of any outer region, is depicted with two disjoint curves [Fig. 1b].

Fig. 1. Two examples of Peirce's method: (a) "not-S"; (b) "Everything is either A or B"

Let us observe how this method applies on a syllogism whose premises are: "All S are M" and "No M is P". Since the latter premise can be transformed into "All M are *not*-P", the diagram depicts S inside M which is itself inside *not*-P [Fig. 2a]. Hence, the conclusion is "No S is P". Interestingly, syllogisms with two negative premises might be conclusive if negative terms are introduced. For instance, premises "No S is M" and "No *not*-M is *not*-P" yield the conclusion "All S are P" [Fig. 2b].

Fig. 2. Two examples of syllogisms according to Peirce's method of extending Euler diagrams

Peirce's method represents the same information in different ways, depending on the shape of the curves. This flexibility might prove convenient to represent complex propositions but it complicates their manipulation as it may not be easy to recognise diagrams' equivalences.

3 Conclusion

Modern diagrams represent negative terms along the path laid out by Euler and Venn: a curve produces two regions standing for complementary terms. The identification of the terms is conventional but is conveniently indicated by the label of the regions. Interestingly, Euler, Venn and Peirce appealed to different labeling practices. For a term S, all three would draw a circle, but Euler would put the letter 'S' *inside* the curve, Peirce *on* it and Venn *outside* it [Fig. 3]. Euler's usage is intuitive as it marks the space that stands for the class. Venn's usage is more practical since he demands a single figure for a given number of terms. Hence, the identification of the circles is unambiguous and all regions (except the outer) are kept ready to be marked. Peirce's practice is more challenging: it makes the curve stand for the *differentiæ* that disposes individuals on its both sides, depending on their predication. Hence, the curve acts as a separation line and is the object of the label. Consequently, the location of positive and negative terms is determined by the shape of the curve, not by its label.

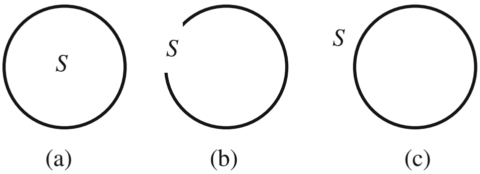


Fig. 3. The labeling conventions of Euler, Peirce and Venn

Acknowledgments. This note draws upon research supported by the Estonian Research Council PUT 267 'Diagrammatic Mind: Logical and Communicative Aspects of Iconicity' (Principal investigator Professor Ahti-Veikko Pietarinen, 2013-2016).

References

1. Euler, E.: Lettres à une princesse d'Allemagne II. Saint Petersburg: Imprimerie de l'académie impériale des sciences (1768)
2. Jevons, W.S.: The Principles of Science. Macmillan, London (1877)
3. Keynes, J.N.: Studies and Exercises in Formal Logic. Macmillan, London (1906)
4. Moktefi, A.: Is Euler's circle a symbol or an icon? Sign Syst. Stud. **43**(4), 597–615 (2015)
5. Moktefi, A., Pietarinen, A.-V.: On the diagrammatic representation of existential statements with Venn diagrams. J. Logic Lang. Inf. **24**, 361–374 (2015)
6. Moktefi, A., Shin, S.-J.: A history of logic diagrams. In: Gabbay, D.M., et al. (eds.) Logic: A History of Its Central Concepts, pp. 611–682. North-Holland, Amsterdam (2012)
7. Peirce, C.S.: On Logical Graphs. Peirce collection at the Houghton library of Harvard University, Ms 481 (1896) and Ms 1147 (1901) (1896/1901)
8. Venn, J.: Symbolic Logic. Macmillan, London (1881)

Visualising the Boolean Algebra $\mathbb{B}_4$ in 3D

Hans Smessaert[1(✉)] and Lorenz Demey[2]

[1] Department of Linguistics, KU Leuven, Leuven, Belgium
Hans.Smessaert@kuleuven.be
[2] Center for Logic and Analytic Philosophy, Postdoctoral Fellow of the Research Foundation–Flanders (FWO), KU Leuven, Leuven, Belgium
Lorenz.Demey@kuleuven.be

Abstract. This paper compares two 3D logical diagrams for the Boolean algebra $\mathbb{B}_4$, viz. the rhombic dodecahedron and the nested tetrahedron. Geometric properties such as collinearity and central symmetry are examined from a cognitive perspective, focussing on diagram design principles such as congruence/isomorphism and apprehension.

Keywords: Logical geometry · Rhombic dodecahedron · Nested tetrahedron · Congruence · Apprehension · Central symmetry · Boolean algebra

Introduction. Logical geometry systematically studies Aristotelian and related logical diagrams, focussing on both abstract-logical and visual-geometrical topics (cf. www.logicalgeometry.org) [6]. A major visual-geometrical issue is the fact that a single logical structure often gives rise to different visualisations. These diagrams are *informationally* equivalent, but they need not be *computationally* equivalent [3]: visual differences may significantly influence user comprehension. This paper presents a case study on this issue: we take one logical structure (viz. the Boolean algebra $\mathbb{B}_4$) and compare two 3D visualisations (viz. rhombic dodecahedron and nested tetrahedron) in the light of general principles of diagram design. The outcome of this comparison is a nuanced perspective: both diagrams are useful visualisations of $\mathbb{B}_4$; whichever one is ultimately adopted will depend on which logical properties of $\mathbb{B}_4$ the diagram author wants to highlight.

The Boolean algebra $\mathbb{B}_4$ will be represented by means of bitstrings of length 4, i.e. we take $\mathbb{B}_4 = \{0, 1\}^4$. A bitstring's *level (L)* is defined as the number of bit positions with value 1; e.g. 1100 and 1101 are of L2 and L3, respectively. The Aristotelian relations in $\mathbb{B}_4$ are defined as follows: b_1 and b_2 are *contradictory (CD)* iff $b_1 \wedge b_2 = 0000$ and $b_1 \vee b_2 = 1111$; they are *contrary (C)* iff $b_1 \wedge b_2 = 0000$ and $b_1 \vee b_2 \neq 1111$, they are *subcontrary (SC)* iff $b_1 \wedge b_2 \neq 0000$ and $b_1 \vee b_2 = 1111$, and they are in subalternation iff $b_1 \wedge b_2 = b_1$ and $b_1 \vee b_2 \neq b_1$.

This Boolean algebra can be visualised using a *rhombic dodecahedron (RDH)* [5]. RDH has been used both as a Hasse diagram and as an Aristotelian diagram for $\mathbb{B}_4$; cf. Fig. 1(a–b) [1]. The second visualisation of $\mathbb{B}_4$ is the *nested tetrahedron (NTH)* in Fig. 1(c–d) [2,4]. Because a tetrahedron is *self-dual*, connecting the centres of its 4 faces yields a small, 'nested' tetrahedron.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 289–292, 2016.
DOI: 10.1007/978-3-319-42333-3_26

Representing Levels. In a Hasse diagram of a Boolean algebra, the levels are visualised as (horizontal) hyperplanes that are orthogonal to the general (vertical) implication direction. This is a clear instance of the *Congruence Principle*, according to which the structure of the visualisation should correspond to the represented logical structure [7]. To what extent does this principle apply to the visualisation of levels in RDH/NTH? Fig. 1(e–j) shows the L1/L2/L3 distribution in RDH and NTH. It is clear that neither RDH nor NTH explicitly visualises the levels as (horizontal) hyperplanes orthogonal to one (vertical) implication direction. Still, NTH observes the Congruence Principle much better than RDH, albeit in a different way: levels no longer correspond to parallel hyperplanes ordered along one (vertical) dimension, but rather to the geometrical dimensions themselves. The natural geometrical ordering of 0-, 1- and 2-dimensionality for vertices, edges and faces thus corresponds to the logical ordering of the levels L1, L2 and L3.[1] By contrast, RDH is not level-preserving at all: levels do not correspond to parallel hyperplanes, and not to dimensionality either.

Representing Contradiction. Since the contradiction relation is symmetric and functional, Aristotelian diagrams often have the property of *central symmetry*: contradictory bitstrings are located at diametrically opposed vertices of the diagram and at the same distance from its centre. This is another clear instance of the Congruence Principle. To what extent does this principle also apply to the visualisation of contradictions in RDH/NTH? As to contradictory pairs of L2/L2 bitstrings, central symmetry holds in both RDH and NTH. However, as to contradictory pairs of L1/L3 bitstrings, central symmetry holds in RDH, but not in NTH: Fig. 1(c) shows that the L3 vertices are located at a much shorter distance from the centre than their contradictory L1 counterparts. Hence, NTH exhibits a lower degree of overall logico-geometrical congruence.

Contradiction is often argued to be the *strongest* opposition relation, in the sense that turning a bitstring into its contradictory involves switching the values in *all* of its bit positions. Because of the Congruence Principle, the idea of 'maximal logical distance' is sometimes visualised by means of 'maximal geometrical distance': the vertex that is farthest removed from the vertex representing a bitstring b is the vertex representing $\neg b$. Figure 1(b) shows that this property is perfectly satisfied in RDH: vertices corresponding to contradictory bitstrings are systematically located at a maximal distance removed from one another. By contrast, Fig. 1(c–d) shows that in NTH, vertices corresponding to contradictory bitstrings are systematically *not* located at a maximal distance removed from one another. Hence, NTH exhibits a much lower degree of congruence.

Collinear Vertices. NTH contains several triples of collinear vertices. For example, the bitstrings 1000, 1001 and 0001 all lie on the top edge of NTH in Fig. 1(h–i). As to the Aristotelian relations between these bitstrings, we have a contrariety between 1000 and 0001, and subalternations from 1000 and 0001 to 1001. Since the vertices of these three bitstrings are collinear, the visualisation of the contrariety overlaps/coincides with the visualisations of the two subalter-

[1] Each L2/L3 bitstring is situated at the *midpoint* of its corresponding edge/face.

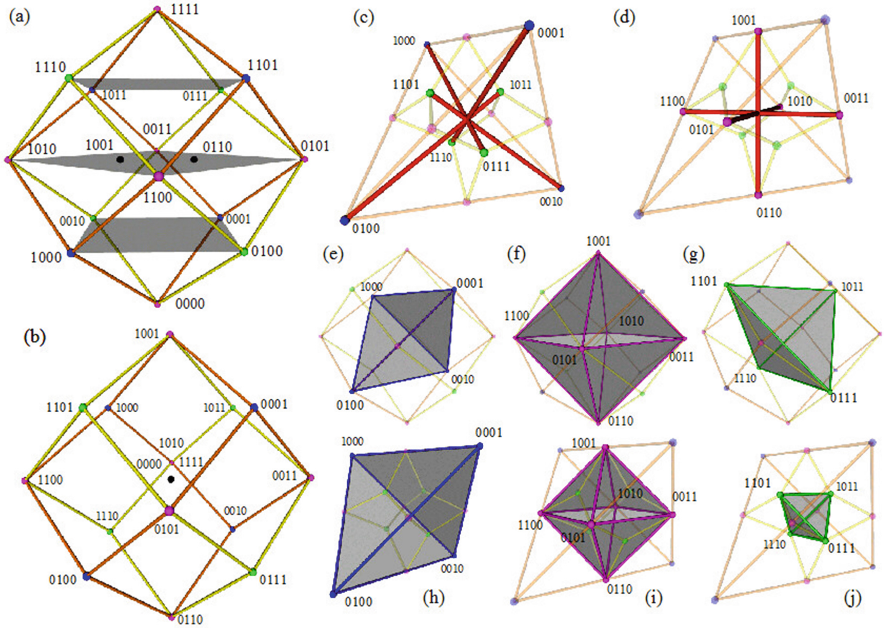


Fig. 1. The RDH and NTH visualisations of the Boolean algebra $\mathbb{B}_4$.

nations. As a result of this overlap, the user looking at NTH might have difficulties in properly distinguishing the three distinct Aristotelian relations holding between 1000, 1001 and 1001. This is a serious violation of the *Apprehension Principle*, according to which "the structure and content of the visualization should be readily and accurately perceived and comprehended" [7, p. 37]. By contrast, RDH does not have any triples of collinear vertices. Consequently, all Aristotelian relations holding between the elements of $\mathbb{B}_4$ are visualised by means of distinct (non-coinciding) lines. This systematic avoidance of overlapping visual elements means that RDH is much more in accordance with the Apprehension Principle.

References

1. Demey, L., Smessaert, H.: The relationship between Aristotelian and Hasse diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 213–227. Springer, Heidelberg (2014)
2. Dubois, D., Prade, H.: From Blanché's hexagonal organization of concepts to formal concept analysis and possibility theory. Logica Universalis **6**, 149–169 (2012)
3. Larkin, J., Simon, H.: Why a diagram is (sometimes) worth ten thousand words. Cogn. Sci. **11**, 65–99 (1987)
4. Moretti, A.: Was Lewis Carroll an amazing oppositional geometer? Hist. Philos. Log. **35**, 383–409 (2014)

5. Smessaert, H., Demey, L.: Logical and geometrical complementarities between Aristotelian diagrams. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 246–260. Springer, Heidelberg (2014)
6. Smessaert, H., Demey, L.: Logical geometries and information in the square of opposition. J. Log. Lang. Inf. **23**, 527–565 (2014)
7. Tversky, B.: Prolegomenon to scientific visualizations. In: Gilbert, J.K. (ed.) Visualization in Science Education, pp. 29–42. Springer, Netherlands (2005)

Single Feature Indicator Systems in the Openbox

Dave Barker-Plummer(✉) and Michael Murray

CSLI/Stanford University, Stanford, CA 94305, USA
{dbp,mikem05}@stanford.edu

Abstract. In this paper we describe a simple method for introducing instances of the class of Single Feature Indicator Systems into the Openbox, a framework for constructing heterogeneous proof environments. Single Feature Indicator Systems form a class of representation systems which employ a simple signalling strategy for representing information about their target domain. Using the new method allows Single Feature Indicator Systems to be implemented in the Openbox by describing the representation in a high-level language rather than by programming.

1 Introduction

In this paper we report on a project that combines theoretical work concerning the formal basis for representation systems with practical implementation of techniques to enable the development of diagrammatic reasoning systems. This is part of a generic approach to the investigation of diagrammatic representation. In this approach, we focus on patterns appearing across many specific representations, and build understanding and tools based on these patterns.

In our theoretical work we have identified and formalized a class of representation systems, called Single Feature Indicator Systems (SFIS) [2,3].

Our work on the Openbox is an approach to implementing heterogeneous proof systems by providing a common proof framework into which specific representation systems can be "plugged-in" [1]. Jamnik and Urbas have developed the MixR system with many of the same goals in mind [4].

In this paper, we describe the implementation of a new feature of the Openbox which allows users to introduce new representation systems based on SFIS into a heterogeneous proof environment by providing a high-level specification of the representation system in place of writing computer code.

2 Single Feature Indicator Systems in the Openbox

The Openbox. The Openbox is an implemented framework for the construction of heterogeneous proof environments. Openbox implements a component-based architecture into which different representations can be loaded. The Openbox supports three types of plugins: **Models** which are data structures describing the information to be represented in the system; **Editors** implementing the user interface that allows users to create and edit models, and **Engines** which implement operations that may be performed on and between models.

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 293–295, 2016.
DOI: 10.1007/978-3-319-42333-3_27

Single Feature Indicator System. Single Feature Indicator Systems (SFIS) are a formally defined class of representation systems that share a common strategy for representing information about their target domains, [2,3]. Each SFIS uses basic graphic elements of the diagram that we call *source roles*, to represent information about corresponding components of the target domain. Each source role can take on one of a finite number of *source values*.

Specifying Single Feature Indicator Systems for the Openbox. We describe a mechanism by which any SFIS can be implemented for the Openbox without programming. The mechanism consists of (1) a specification language which permits the description of an SFIS in a high-level language, (2) a specification compiler that converts these descriptions into a concrete implementation of that SFIS, using (3) a library of common editor components.

Processing the specification results in implementations of model, engine and editor components for the representation. Since all implementations of SFIS have models with the same structure and inference rules, the model and engine components are implemented abstractly within the Openbox. The specification compiler instantiates this abstract code to produce new components. Default editor components are also provided by the compiler. These default editors can be customized or replaced using the specification language.

We use an XML document to specify the desired SFIS. The central requirement of the specification language is to describe the roles and values of the SFIS. In the simplest case, these can be enumerated as lists. More interesting cases arise when the roles are specified as the product of two enumerated sets, or the product of a set with itself. Each of these idioms is expressible in the specification language. The set of values may also be enumerated, but unlike roles, the set of values need not be finite. We support the ability to specify the values by type, E.g., Integer or String, and to further restrict the type by value.

The specification compiler works in four phases[1]. (1) Parse the XML document into an internal representation, (2) Generate Java source code from the specification. Source code is emitted by the system, so that it can be refined or extended. (3) Compile the Java source code into executable code. (4) Load the executable code into the running system so that the specified components are immediately available within the system.

The specification of the roles and values of the SFIS is not sufficient to determine how best to display the information about the association between values and roles, so our specification compiler provides default editor components and the opportunity to customize or replace them.

The specification compiler chooses a default editor component to use in the implementation. These default editors are tables whose dimensions depend on the specification of set of roles. Editors can be customized with different interaction devices for selecting values, and with different ways of displaying values, such as using text, color, or icons.

[1] Videos demonstrating the end-to-end use of the system are available at https://www.youtube.com/playlist?list=PLn4GZI3tktH-mOXaMzcDmDQLYgOZXfkgw.

The specification language allows the user to override the default selection of editor components. For example, the user may wish to use a bar-graph rather than a tabular representation of values. To do this, the user provides the name of a Java class which implements the required editor type within the specification.

Our editor toolkit currently provides a number of different editor types. New editor components can be built by adhering to a specified interface for interoperating with the SFIS subsystem and require little knowledge about the Openbox architecture, and can be developed by users with basic Java programming skills.

3 Conclusion and Further Work

We have extended our generic investigation of diagrammatic reasoning by using work on the theory of diagrammatic representations to simplify the introduction of diagrammatic representations into the Openbox.

There are several future directions for this work. We would like to (1) continue to develop the editor toolkit to allow for a wider selection of more robust and full-featured representations to be used for Single Feature Indicator Systems; (2) permit the specification of homogeneous inference rules to apply within a particular SFIS, but not among SFIS in general; and (3) extend the specification language so that we can specify heterogeneous inferences both between different SFIS and between SFIS and non-SFIS representations.

As our theoretical work proceeds, we expect to identify different classes of representation systems similar to SFIS. At least some aspects of such systems might be amenable to specification for inclusion into the Openbox. We plan to extend our specification language and compiler to allow the specification of these systems.

References

1. Barker-Plummer, D., Etchemendy, J., Murray, M., Pease, E., Swoboda, N.: Learning to use the openbox: a framework for the implementation of heterogeneous reasoning. In: Cox, P., Plimmer, B., Rodgers, P. (eds.) Diagrams 2012. LNCS, vol. 7352, pp. 3–3. Springer, Heidelberg (2012)
2. Shimojima, A., Barker-Plummer, D.: The Barwise-Seligman model of representation systems: a philosophical explication. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 231–245. Springer, Heidelberg (2014)
3. Shimojima, A., Barker-Plummer, D.: A generic approach to digrammatic representation systems: the case of single feature indicator systems. In: Jamnik, M., Uesaka, Y., Elzer Schwartz, S. (eds.) Diagrams 2016. LNCS, vol. 558, pp. 83–97. Springer, Heidelberg (2016)
4. Urbas, M., Jamnik, M.: A framework for heterogeneous reasoning in formal and informal domains. In: Dwyer, T., Purchase, H., Delaney, A. (eds.) Diagrams 2014. LNCS, vol. 8578, pp. 277–292. Springer, Heidelberg (2014)

Effective Representation of Information: Generalizing Free Rides

Gem Stapleton[1(✉)], Mateja Jamnik[2], and Atsushi Shimojima[3]

[1] University of Brighton, Brighton, UK
g.e.stapleton@brighton.ac.uk
[2] University of Cambridge, Cambridge, UK
mateja.jamnik@cl.cam.ac.uk
[3] Doshisha University, Kyoto, Japan
ashimoji@mail.doshisha.ac.jp

Abstract. In order to effectively communicate information, the choice of representation is important. Ideally, a representation will aid readers in making desired inferences. In this poster, we introduce the theory of *observation*: what it means for one statement to be observable from another. Using observability, we sketch a characterization of the *observational advantages* of one representation over another. By considering observational advantages, people will be able to make better informed choices of representations. To demonstrate the benefit of observation and observational advantages, we apply these concepts to set theory and Euler diagrams. We show that Euler diagrams have significant observational advantages over set theory. This formally justifies Larkin and Simon's claim that "a diagram is (sometimes) worth ten thousand words".

Keywords: Knowledge representation · Observation · Free rides · Inference

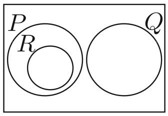
Introduction. When we want to share and understand information, we need to represent it in some notation. There is a plethora of notations available to us for this purpose. This poster, which summarizes [3], is concerned with the relative advantages of one choice of representation of information over another. Many factors can contribute to such advantages. For instance, graphical features, such as the way in which colour is used, and visual clutter (or lack thereof) can impact the ease with which information can be extracted from a representation. The particular focus of this poster is on what we call *observational advantages*.

Observation. It is *advantageous* if a representation of information allows us to simply observe other statements of interest to be true. By contrast, if we cannot observe the statement – yet it does indeed follow from the given representation – then this is a *disadvantage* of that representation. If one representation of information, r_1, has such an advantage and another, r_2, has this as a disadvantage then r_1 has an *observational advantage* over r_2. As a simple example, suppose

M. Jamnik et al. (Eds.): Diagrams 2016, LNAI 9781, pp. 296–299, 2016.
DOI: 10.1007/978-3-319-42333-3_28

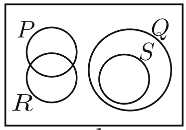
$$(P \cap Q) = \emptyset$$
$$R \subseteq P$$

(1) Set Theory



(2) Euler Diagram

Fig. 1. Multiple choices of representation.



d

Fig. 2. Free rides.

we wish to represent these two facts about three sets, P, Q and R: nothing is in both P and Q, and everything in R is also in P. There are many notations that can express this information: two examples are illustrated in Fig. 1.

Each of the sentential statements has a single *meaning-carrying relationship*. By meaning-carrying relationship, we mean a relation on the syntax of the statement that carries semantics and evaluates to either 'true' or 'false'. The first statement asserts that the intersection of the two sets is empty. The meaning-carrying relationship in (1) is that $(P \cap Q)$ and $\emptyset$ are written either side of $=$. The statement $(P \cap Q) = \emptyset$ evaluates to either true or false, depending on the interpretation of P and Q as sets. The second statement in (1) describes a subset relationship; here, the meaning carrier in (1) is that R is written on the left of $\subseteq$ and P is written on the right.

The diagrammatic representation has *many* meaning-carrying relationships. The Euler diagram uses non-overlapping curves to express the disjointness of P and Q and, similarly, curve containment to assert that R is a subset of P. Here, two meaning-carrying relationships (namely, disjointness and containment) are exploited to convey the desired information. As a consequence of the way in which Euler diagrams are formed, *additional* meaning-carrying relationships are evident. Most obviously, the non-overlapping relationship between Q and R is a meaning carrier. Thus, from the Euler diagram we can *observe* the statement that Q and R are disjoint. By contrast, this statement cannot be observed from (1) but must be *inferred* from the statements given. This is an example of an *observational advantage* of the Euler diagram over the sentential representation.

Observation can be applied to an Euler diagram to produce another statement, be it a diagram or a set-relation. It can also be applied to a set-relation to produce another set-relation or an Euler diagram. Thus, observation from a single statement, σ, is a binary relationship between σ and another statement, σ_o, denoted $\sigma \rightsquigarrow \sigma_o$, which ensures the following properties hold:

1. some of the meaning-carrying relationships holding in σ hold in σ_o, and
2. σ_o supports just enough relationships to express the meanings carried by the selected relationships in σ and nothing stronger.

Observational Advantages. The new concept of an *observational advantage* generalizes free rides introduced previously by Shimojima [2]. Our definition of an observational advantage requires three key notions to be defined: *semantic entailment*, *semantic equivalence*, and what it means for a statement to be *observable* from a set of statements. The original idea of a free ride assumes a semantics-preserving translation from one notation, N_1, into another notation, N_2, such that the translation ensures the original statements are observable from the resulting statements. We can explain free rides in detail by appealing to our chosen case study: set theory and Euler diagrams. Suppose we have a finite set of set-relations, $\mathcal{S}$, where a set-relation is a statement that asserts either set equality or a subset relationship. Further, suppose that we then identify a semantically equivalent, finite set of Euler diagrams, $\mathcal{D}$, such that each statement, s, in $\mathcal{S}$ is observable from a diagram, d, in $\mathcal{D}$; we can view $\mathcal{D}$ as being a translation of $\mathcal{S}$. Then the set-relations that are observable from the diagrams in $\mathcal{D}$ but not from $\mathcal{S}$ are *free rides* from $\mathcal{D}$ given $\mathcal{S}$.

For example, take $\mathcal{S} = \{(P \cap Q) = \emptyset, (R \cap Q) = \emptyset, S \subseteq Q\}$, which contains three set-relations, and $\mathcal{D} = \{d\}$, where d is in Fig. 2. The free rides from $\mathcal{D}$ given $\mathcal{S}$ are the set-relations that one can observe to be true from $\mathcal{D}$ but which need to be inferred, not simply observed, from $\mathcal{S}$. For instance, we can *observe* both $(P \cap S) = \emptyset$ and $(R \cap S) = \emptyset$ from d but both of these must be *inferred* from $\mathcal{S}$; in the former case, $(P \cap S) = \emptyset$ can be inferred from $(P \cap Q) = \emptyset$ and $S \subseteq Q$. By contrast, whilst the set-relation $(P \cap Q) = \emptyset$ can be observed from $\mathcal{D}$ it can also be observed from $\mathcal{S}$, so it is not a free ride from $\mathcal{D}$. Free rides are examples of what we call *observational advantages* of the Euler diagram over the original set theory representation of information. The difference between observational advantages and free rides is that observational advantages do not require the set $\mathcal{S}$ to contain only statements observable from $\mathcal{D}$.

Set Theory and Euler Diagrams. By applying our theory of observation and observational advantages to set theory and Euler diagrams, we can establish:

1. Given a finite set of set-relations, $\mathcal{S}$, no other set-relations can be observed from $\mathcal{S}$; thus, $\mathcal{S}$ is *observationally devoid*.
2. Given an Euler diagram, $d_{\mathcal{S}}$, constructed from $\mathcal{S}$, every set-relation that follows from $\mathcal{S}$ can be observed from $d_{\mathcal{S}}$; thus, $d_{\mathcal{S}}$ is *observationally complete*.

These two characterizations of what can (or cannot) be observed allow us to understand that $d_{\mathcal{S}}$ is a significantly more efficacious representation of information than $\mathcal{S}$: it has *maximal observational advantage* over $\mathcal{S}$. Thus, from a purely inferential point of view, using $d_{\mathcal{S}}$ is desirable: $d_{\mathcal{S}}$ makes informational content readily available, in the sense of observability, to end-users. As there are infinitely many set-relations that are semantically entailed by $\mathcal{S}$, the benefits of Euler diagrams over set theory are numerous.

Conclusion. In our view, the result introduced here captures the kernel in which diagrams facilitate our inference and thus excel over sentential representations. Putting this in a larger perspective, we expect to gain a fuller understanding of the relative advantages of one choice of representation over another. Linking back to the insight that a diagram is sometimes worth 10,000 words [1], our formal theory of observation and observational advantage has allowed us to prove that a diagram is sometimes worth *infinitely many* set-relations.

References

1. Larkin, J., Simon, H.: Why a diagram is (sometimes) worth ten thousand words. Cogn. Sci. **11**, 65–99 (1987)
2. Shimojima, A.: Semantic Properties of Diagrams and Their Cognitive Potentials. CSLI Publications, Stanford (2015)
3. Stapleton, G., Jamnik, M., Shimojima, A.: What makes an effective representation of information: a formal account of inferential advantages. J. Log. Lang. Inf. (2015, Submitted)

Author Index

Alharbi, Eisa 51
Alqadah, Mohanad 250

Balawejder, Eric 265
Barker-Plummer, Dave 83, 293
Blake, Andrew 32
Bosveld-de Smet, Leonie 272
Burns, Richard 265
Burton, Jim 32, 107, 283

Carberry, Sandra 265
Chakraborty, Mihir 107
Chapman, Peter 250
Cheng, Peter C-H. 17
Choudhury, Lopamudra 107
Corter, James E. 3, 258

de Vries, Erica 40
Demey, Lorenz 67, 289
Domanowska, Wiktoria 265

Fox, Amy Rae 40

Gao, Jie 258
Gerber, Aurona J. 197
Goel, Ashok 226
Grevismühl, Daniel 212

Haider, Johanna 235
Hamie, Ali 51
Howse, John 51, 250

Igarashi, Mika 172

Jamnik, Mateja 296

Kriglstein, Simone 235

Lima, Laurent 40
Linker, Sven 32
Loker, Savannah 40

Manalo, Emmanuel 157
McGreggor, Keith 226
Mineshima, Koji 123
Moktefi, Amirouche 286
Murray, Michael 293

Nickerson, Jeffrey V. 258

Pérez-Kriz, Sarah 187
Pietarinen, Ahti-Veikko 139, 286
Plöger, Christina 219
Pohl, Margit 235

Rodgers, Peter 283
Rüegg, Ulf 212

Sato, Yuri 123
Schulze, Christoph Daniel 212, 219
Schwartz, Stephanie Elzer 265
Shimojima, Atsushi 83, 296
Smessaert, Hans 67, 289
Stapleton, Gem 51, 107, 250, 283, 296
Suetsugu, Rei 172

Takemura, Ryo 98
Tanaka, Yuko 258
Thomas, Anitta 197
Tranquille, Dominique 283
Tversky, Barbara 258

Uesaka, Yuri 157, 172

van der Merwe, Alta 197
Verheggen, Robert-Jan 272
Vivancos-Pérez, Ricardo F. 187
von Hanxleden, Reinhard 212, 219

Wallner, Günter 235

Xing, Chenmu 3

Zahner, Doris 3