Cognitive Properties of Representations: A Framework

Peter C.-H. Cheng[1(✉)], Grecia Garcia Garcia[1], Daniel Raggi[2], Aaron Stockdill[2], and Mateja Jamnik[2]

[1] University of Sussex, Brighton, UK
{p.c.h.cheng,g.garcia-garcia}@sussex.ac.uk
[2] University of Cambridge, Cambridge, UK
{daniel.raggi,aaron.stockdill,mateja.jamnik}@cl.cam.ac.uk

Abstract. We present a framework for assessing the relative cognitive cost of different representational systems for problem solving. The framework consists of 13 *cognitive properties*. These properties are mapped according to two dimensions: (1) the time scale of the cognitive process, and (2) the granularity of the representational system. The work includes analyses of those processes that are relevant to the internal mental world, and those that are relevant to the external physical display too. The motivation for the construction of this framework is to support the engineering of an automated system that (a) selects representations, (b) that are suited for individual users, (c) and works on specific classes of problems. We present a prototype implementation of such an automated representation selection system, along with an evaluation.

Keywords: Representational systems · Cognitive cost · External and internal representation

1 Introduction

The motivation for (yet) another analysis of the nature of representations stems from our project that is building an automated approach to the *selection* of appropriate representations for solving problems. The motivation and goals of the project are described more fully in [9]. Representation selection must take into account: (a) the type of problem, (b) the specific representational system in which the problem may live, and (c) the users' abilities and familiarity across various representational systems. Expert teachers are able to pick alternative representations to suit each individual students' ability for specific classes of problems; thus, our aim is to design and build a system that can make similar selections. So, what kind of aspects do we need to take into account when building an automated system? In our project, we identified *formal properties* and *cognitive properties* of representational systems. Further, we are developing methods to combine those properties with information about individual users in order to suggest candidate representations for them as well as rank them according to their efficacy for each individual. In this paper, we focus on the cognitive properties. (We describe formal properties in

A. Basu et al. (Eds.): Diagrams 2021, LNAI 12909, pp. 415–430, 2021.
https://doi.org/10.1007/978-3-030-86062-2_43

detail in [18, 19, 23]). Fundamental to our approach is the assessment of the relative *cognitive cost* of alternative representations. Therefore, this requires us to state what the cognitive properties are and to formulate cost measures for them, which will be used in the calculations of an overall cost of a representational system.

Many empirical studies have been conducted on the relative benefit of selected representations for specific tasks, such as [5, 11, 27]. However, it is unclear how these findings can be applied to the assessment of the cognitive costs of representations in general: they only address particular isolated factors. In contrast, we aim to address the following research questions:

1. What are cognitive properties and where do they come from?
2. How should cognitive properties' relative importance be assessed in the context of their multitude and diversity?
3. How can cost measures of the properties be combined to give the relative order of the effectiveness of representations?

Our aim in this paper is to provide the foundational framework from which to address these questions. To be clear, we are not pursuing a general psychological theory of representational systems, but aim to engineer a system to reason about representations; in other words, we want to explore how to give computers the ability to select effective representations for humans. Give the scope of this goal, it is not possible to cover all relevant areas of the literature within this paper, so we have necessarily been selective.

The framework is presented in the next section. This is followed by the presentation of three sample solutions to one problem in three different representational systems. The five sections that then follow describe classes of cognitive properties identified by the framework. We then present an example on how the framework has been used in a prototype of an automated system for representation selection. The final discussion section reflects on the scope and limitations of our framework.

2 Analysis Framework

We use these abbreviations: R – representation; RS – representational system[1]; ER – external representation[2]; IR – internal (mental) representation; CP – cognitive property.

A *cognitive property* is a feature of a representational system that influences how information is processed, and is thus likely to affect the cognitive cost of using the representation (e.g., the number of symbols in a R can affect its cognitive cost).

By *cognitive cost*, we mean the cognitive load that a user experiences using a representational system. This might be measured empirically in terms of: the time taken to complete a problem; the number of operations or procedures used; a rating of the moment to moment subjective effort that the user perceived; the amount of unproductive

[1] Following [18], a *representational system* is an abstract entity from which many distinct individual *representations* may be created.

[2] Following [27], ERs are information and objects that exist in the external environment and can be perceived; while IRs are knowledge and structures in memory (p. 180).

effort due to errors or the pursuit of unproductive solution paths. At the level of cognitive processes, some of the factors that are known to underpin cognitive cost include (e.g., [4, 17, 20]): instantaneous working memory load; less accessible information; operators that take more effort to select or to apply; reduced ability to anticipate the consequence of applying operators; the possession of poor problem-solving heuristics; the lack of externalised memory or free-ride inferences [20]. In contrast to Sweller's [24] notation of cognitive load, our approach broadens the idea to wider temporal and granularity scales, rather than whole instructional tasks, but narrows the focus specifically on representational systems, rather than instructional interventions in general.

Table 1. Cognitive properties framework.

		Notation granularity		
		Symbol	Expression	Representational system
Cognitive type or level	0. General			Sub-RS-variety
	1. Registration	Registration-process Number-of-symbols/expressions Variety-of-symbol/expressions		–
	2. Semantic encoding	Concept-mapping ER-semantic-process IR-semantic-process		–
	3. Inference	Quantity-scale	Expression-complexity Inference-type	–
	4. Problem solution	–	–	Solution-depth Solution-branching-factor Solution-technique

A framework for cognitive properties has stringent requirements. First, it should systematically identify cognitive properties without neglecting important high impact properties. Second, the CPs should directly relate to established cognitive phenomena and accepted theoretical cognitive constructs associated with representational systems (e.g., [13]). Third, it should identify unique CPs that overlap minimally in scope.

So, to define the framework, three distinct primary cognitive dimensions have been adopted, guided by insights from [1, 16, 21]. The space is represented in Table 1. The dimensions are:

(1) The *granularity of components* of the ER: column headings in Table 1.
(2) The *type and temporal level* of cognitive processing: row headings in Table 1.
(3) Whether the component or the process is primarily *associated with the ER or IR*: see the names of some CPs in the cells of Table 1.

The framework embodies the idea that, as CPs are manifestations of interactions between cognitive and representational systems, both are conceptualised as nearly-decomposable hierarchical systems [21] that function over large ranges of spatial and temporal scales [1, 16] and are distributed between the IR and ER.

Granularity of Components. This is a dimension ranging across the size of cognitive objects that encode meaning. The *Symbol*[3] level is for elementary, non-decomposable carriers of concepts. Expressions are assemblies of elementary symbols, which occur at different hierarchical levels. The *Representational System* level concerns the complete notational system that is used in a particular case (the representation) for problem solving, which may include distinct sub-representational systems (*sub-RSs*).

Type and Temporal Level of Cognitive Processing. This dimension has two parts. The first part is composed of four temporal levels at which cognitive processes operate, ranging from 100 ms to tens of minutes (e.g., from the time to retrieve a fact from memory, to the time to develop a problem solution). These levels are: (1) *registration*, (2) *semantic encoding*, (3) *inference*, and (4) *problem solution*. Registration refers to the process of acknowledging the existence and location of objects. The encoding level considers the cost of associating symbols with concepts. The inference level considers the cost of the arguments and difficulty of inferences. The problem solution level captures the complexity of the problem state and goal structure. Relatively strong interactions occur between processes at a particular time scale, and relatively weak interactions between different time scales [1, 16]. So, for the sake of analysis, cognitive processes at scales, differing by an order of magnitude, may be treated as nearly independent. Nevertheless, short processes impact long processes cumulatively.

The second part of this dimension is a further level zero, *general*, in Table 1, which accommodates a CP that is not covered by the four temporal levels, but it is a feature that affects how information is processed too.

Association with the ER or IR. This third dimension is recognised because the nature of some processes that serve the same cognitive function may actually differ substantially between IR and ER, and so, they need to be explained in terms of different CPs.

The CP framework builds upon the taxonomy of characteristics of effective RSs compiled by [4], but diverges from that work by providing an underpinning cognitively motivated theoretical justification for the framework's structure. CPs are included in the framework on the basis that a theoretical argument can be made that the CP impacts the cost of using a representation. Inclusion makes no claim that a simple measure of cognitive cost or practical means to compute the cost is necessarily available; this issue is discussed below. As will be noted, some of our proposals need additional empirical support. Before considering the CPs named in Table 1, we present the solutions to a problem in alternative representations to provide running examples.

3 Sample Representations and Problem

We selected probability problems as one target domain for our project because they are knowledge-rich and can be solved using a large variety of alternative representations.

[3] Across disciplines, different terminology is used for *symbols* and *expressions*. From a computational perspective, [18] refers to *primitives* instead of *symbols*, and *composites* instead of *expressions*. These differences partially rise from different perspectives on what is understood by a basic/elementary unit, whether it is considered decomposable or not. As this paper focuses on cognitive aspects of RSs, we have adopted cognitive-oriented terminology.

Probability tests are a good exemplar: they are an important class of problems that have wide application in many disciplines, but are known to be challenging for problem solvers and learners. Consider this medical problem:

> *1% of the population has a disease D. There is a test, T, such that: (i) if you have the disease, the chance that T comes out positive is 98%; (ii) if you don't have the disease, the chance that T comes out positive is 3%. Suppose Alex takes the test and it comes out positive. What's the probability that Alex has the disease?*

Figure 1 shows ideal solutions in two conventional representations: algebraic *Bayesian representation* and *contingency table*; a third representation uses *Probability Space (PS) diagrams* [3]. PS diagrams exemplify how our framework can be applied to novel representations for which analysts do not have established intuitions. Also, PS diagrams provide an interesting test case as they integrate information about sets and probability relations using a coherent diagrammatic scheme that has been shown to substantially enhance problem solving and learning with little instruction [3]. The green text in Fig. 1 shows values given in the problem statement shown above.

The problem is a fairly canonical test situation, but has a complication. The test is not an independent trial, but depends on whether the disease is actually present or not. Thus, the five-line Bayesian solution (Fig. 1a) employs steps that are beyond school level probability: (1) Bayes' theorem; (2) law of total probability applied to the denominator; (3) De Finetti's axiom of conditional probability. Clearly, this solution requires a high degree of mathematical sophistication.

The contingency table solution (Fig. 1b) assumes that the user knows the arithmetic rules governing continency tables; the formulas in smaller letters at the bottom right of the cells. The solution progresses by successively entering given values of the problem statement into the cells, taking into account the arithmetic constrains. It is completed by selecting the values from the cells that correspond to the target condition probability and calculating the answer, as captured by the line below the table. Since the user must be proficient at using contingency tables, they should be able to handle the impact of lack of independence of the test and to complete only germane cells.

Students, who do not have mathematical instruction beyond 16 years of age, can solve the medical problem by drawing a diagram like Fig. 1c, after just two hours of instruction on PS diagrams [3]. A typical solution using PS diagrams might proceed by sketching the sub-diagram for a binary outcome trial first: this is the horizontal line *D* in the diagram, which consists of the slightly misaligned 'no' and 'yes' sub-segments. Then, two more sub-diagrams are drawn within line *T* (below line *D*); each one covers the two test outcomes of each state of *D*. For example, the left sub-diagram of *T* (consisting of two slightly misaligned segments on the left) covers the test outcome when the person '*does not have the disease*' (since it is under the 'no' sub-segment of *D*) and it shows a sub-segment for when '*the chance that T comes out positive is 3%,*' which is labelled with the '+' sign and the '0.03' value; thus, the sub-segment labelled with the sign '−' and value '0.97' represents the chance that T comes out negative. With kinder numbers, the diagram could be drawn to scale, nevertheless, the (green) numbers record the information given in the problem statement. Knowing the probability of each space (or sub-space), proceed to review the full diagram vertically. As required, we focus on the

positive (+) outcomes of the test (the two middle segments labelled '+' within T), which gives us a conditional sub-space that is represented with the horizontal line '*Ans.*'. Using one of the basic rules of PS diagrams, we can calculate the probability of the outcomes in that sub-space, by multiplying the values of the no_D and yes_T outcomes (0.99 * 0.03 = 0.0297), and the values of yes_D and yes_T outcomes (0.01 * 0.98 = 0.0098). Now, the probability of "*Alex has the disease*" is given by the portion of the conditional space that is yes_D within line '*Ans.*' (thicker sub-segment) which by an approximate mental calculation, is about a quarter.

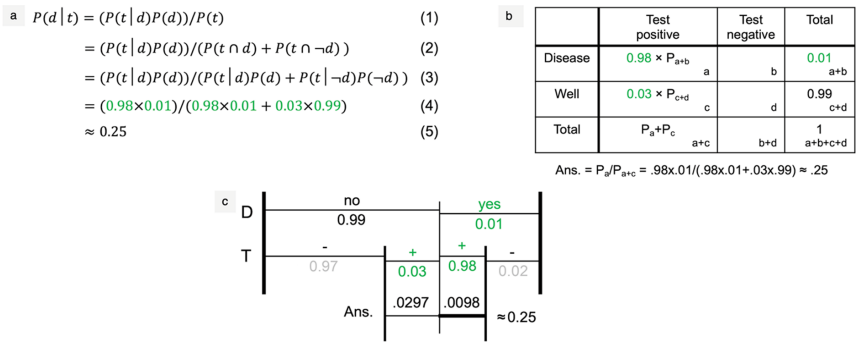

Fig. 1. (a) Bayesian representation, (b) contingency table, and (c) probability space diagram solutions to the medical problem. (Color figure online)

The comparison of these examples will informally support the claims below about cognitive cost of different CPs.

4 General Cognitive Property – Sub-RS-Variety

Much of the literature on representational systems has typically focused on RSs with a single format and made comparisons between such unitary RSs. However, all but the simplest RSs are heterogeneous mixtures comprised of *sub-RSs*. Thus, the **sub-RS-variety** is a CP, because sub-RSs are systems which must work in a coordinated fashion. This entails matching information between the sub-RSs or translating information from the format of one into another. Impacts of multiple sub-RSs include, for instance: increased frequency of attention switches between sub-RSs, with all of the attendant delays in reactivating propositions associated with each sub-RSs; greater number of inference rules to handle; more opportunity for potential errors. Thus, high heterogeneity of sub-RSs incurs a heavy cognitive cost [25].

Obviously, an RS is heterogeneous when it is composed of sub-RSs that would be independently considered as RSs in their own right. For example, in Fig. 1a, the Bayesian notation operates on the quantities of probability, P(…), separately from the set theory notation embedded within the parentheses. More formally, sub-RSs may be distinguished in four related ways. (1) A part of the RS is governed by an exclusive set of syntactic rules, likely applied to distinct operator symbols (i.e., in [12]'s terms, it

possesses a different *format* compared to the rest of the RS). (2) A part of the RS encodes a distinct set of domain concepts, so it may be a separate sub-RS: in the contingency table representation, rows and columns encode relations among sets, whereas the cell entries are formulas involving magnitudes of probabilities. (3) An RS has an indexing system that serves to coordinate between sub-RSs, but that does not directly encode domain concepts: for example, the cell labels and subscripts within the contingency table. (4) A part of the RS is a sub-RS and is spatially remote from the RS: for instance, in Fig. 1b the equation below the contingency table.

Numeration systems are in themselves RSs [28], so any RS that includes numbers has at least two sub-RSs. This is the case in our three representations in Fig. 1. However, numbers may be set aside in the count of sub-RSs because every one of our representations uses them in a similar fashion. So, the *differential* cost of their presence across the three representations will be small compared to other CPs.

The Bayesian and contingency table representations are likely to have a similar cognitive cost in terms of the number of sub-RS-variety CP. In contrast, the PS diagram does not meet many of the criteria for the existence of other sub-RSs; in fact, it may be a special case of a representation without instances of other sub-RSs, and thus, its cognitive cost is predicted to be less than the cost of the other two representations.

5 Registration Cognitive Properties

Registration is the first of the four main temporal levels of cognitive process in the framework. An RS has a vast number of possible features that might serve as symbols because any part of a feature of a graphical element could be selected arbitrarily, such as the '|' or the '–' in a '+', or even their point of intersection. Registration process establishes what particular objects, features, or groups of objects are taken to be a potential symbol (or expression), by acknowledging their existence and noting their location in the representation.

Registration occurs when we seek a symbol in the ER to match a concept (in the IR). Alternatively, we may examine an ER to find symbols in at least two ways. (1) We may use our knowledge about the RS. For example, the answer to a problem, in a problem solution, is likely to be found at the bottom of the solution – as in Fig. 1a. (2) If we are not familiar with an RS, then those features that vary with the RS are potential symbols or expressions, but constant features are not. For instance, the size of the font in the Bayesian example in Fig. 1 is fixed, so it is not meaningful, but it would be if the formulas included subscripts (as in Fig. 1b).

The **registration-process** CP concerns the various types of cognitive processes that are used to register symbols or expressions. The purpose of this CPs is to specify the relative cost arising from those processes. The processes, in order of increasing cost, are: (a) *iconic*, (b) *emergent*, (c) *spatial-index*, (d) *notational-index*, and (e) *search*. (a) The iconic registration process rapidly focuses attention upon 1 object or 1 group that is highly recognisable to the user due to its familiarity. For example, following instruction, students familiar with PS diagrams will perceive the main space (*D* and *T* lines) in Fig. 1c as a single object; or the symbol '≈' in Fig. 1a can be rapidly recognised given its location and shape. (b) Emergent registration processes occur when

a group of symbols are arranged so that they form a perceptual Gestalt (e.g., continuity, closure). For example, the numbers in parentheses in Fig. 1a, which are not part of the solution, but can be used to refer to the different algebraic statements. (c) Spatially-indexed registration processes exploit the spatial organisation in the RS, as described by [12]. (d) Notational-index registration processes exploit some alphanumeric system to organise or index objects, such as the reference letters in the contingency table of Fig. 1b. (e) Lastly, the registration process may default to mere search, perhaps using heuristics or just exhaustively, when the other processes are unavailable (e.g., find 't|¬d' in Fig. 1a). Although we consider our proposed order for these processes to be sensible, further empirical evidence is needed to confirm this order.

The other pair of CPs at the registration level address (a) the **number-of-symbols** or **expressions** and (b) the **variety-of-symbols** or **expressions**. An elementary *symbol* is a non-decomposable carrier (representation) of a concept. For example, in our three sample representations, symbols include: variables and mathematical operators, table cells, and labelled line segments, respectively. The notion of *symbols* also encompasses graphical properties of ER tokens that in themselves may encode particular concepts; for example, the thickness of a line segment in the PS diagram denoting the solution. *Expressions* are assemblies of elementary symbols, which occur at different hierarchical levels; such as algebra formulas or their parts, rows and columns of the contingency table, or the horizontal lines for a particular trial in the PS diagram. In some circumstances we may treat expressions as single objects; e.g., dividing throughout by one side of an equation to obtain a form equal to unity. So just as the number of symbols will impact the cost of using a representation, so will the number of expressions.

It is unlikely that the cognitive cost of the number-of-symbols CP will be a simple linear function of the number-of-symbols, because of the propensity of the mind to chunk information [14]. The same is likely to be true for number-of-expressions, as chunking is a hierarchical process [21]. In the Bayesian representation, the number of symbols including 'P(…)' is 14. However, the cognitive cost is more likely to be a count of the variety-of-symbols/expressions, as chunking does not operate directly on categories. For the contingency table representation, the varieties (types) include the table cells, variable names, and numbers.

6 Semantic Encoding Cognitive Properties

This set of CPs considers the cost of associating symbols and expressions with concepts, that is, the establishment of meaning (not just mere existence and location as in the registration level). Two aspects are considered. One addresses the relation between concepts and things encoding them in a representation, and the other concerns the cognitive processes.

The first CP of the first aspect is **concept-mapping**, which applies both to symbols and expressions. This CP draws upon the literature on the nature of possible matches between symbols (tokens) and expressions in the ER and concepts in the IR [7, 15]. There are five ways in which matches may occur, which are described next in likely order of cognitive cost. As our focus is cognitive, we propose a slightly different ranking to [15]. (1) *Isomorphic*: Matching occurs when each concept precisely matches one

symbol; this entails the lowest cognitive cost. (2) *Symbol-excess*: It occurs when some symbols do not represent any domain concept, they only add noise to the representation. Normally, when a user is familiar with the representation, such noise (junk) symbols can be ignored without undue effort. (3) *Symbol-redundancy*: It occurs when one concept maps to many symbols. For example, as in the Bayesian representation in Fig. 1a, the symbol '*d*' appears several times. In terms of cost, some effort is required to handle this, but since we are naturally able to deal with duplicated symbols and synonyms, the cost may not be too high. (4) *Symbol-deficit*: The cost increases in this case because there is no symbol for a concept, so the benefits of externalising memory are not available. Thus, effort must be expended to place a mental pointer to where its symbol would have appeared in the ER. (5) *Symbol-overload*: This is the worst kind of match. It occurs when multiple concepts map to one symbol. This has the grave potential of propagating error due to confusion. To avoid such errors, laborious inferences exploiting contextual information must be executed to mitigate such ambiguities. The contingency table and the PS diagram are largely isomorphic, in part because the numerical contents of cells of the *Test negative* column have been omitted from the table and the negative test values have been greyed out in the PS diagram, specifically to reduce symbol-excess for the medical problem. Finally, regarding the proposed order for these processes, we are currently working on supporting these claims with empirical evidence.

The next pair of CPs deal with cognitive processing costs. The **ER-semantic-process**, which applies both to symbols and expressions, refers to five cognitively different types of processes that associate symbols or expressions in the ER to concepts in the IR; these are listed here in our proposed rank order of cost. (1) The easiest, *known-association* encoding, depends on the familiarity of the user with the RS (e.g., people are typically familiar with numbers, such as the numbers in Fig. 1). (2) *Visual-properties* can be used to represent quantities. This generally has a low cognitive cost, but there are variations among properties that may increase the cost, such as position, length or angle for instance [5]. (3) The *linear-order* in one spatial dimension can readily encode information. For example, temporal sequencing of events *D* and *T* in the PS diagram, or placing the result of a computation to the right side (instead of the left) of an equal sign in a Bayesian solution (Fig. 1a). (4) Encoding the meaning of a symbol due to its *spatial-arrangement* in 2D is more challenging and uses devices such as: coordinate systems or arrays (e.g., the contingency table), hierarchical assemblies (e.g., the PS diagram), or networks (e.g., trees or lattices). (5) The costliest encoding is for *arbitrary* unstructured list of collections.

IR-semantic-process is the other in the pair of CPs and applies to symbols and expressions. We identify five processes within this CP, which are presented in our proposed rank order of cost (c.f., [13]). (1) The lowest are known *cases*, or prototypes, such as our understanding of the general format of a contingency table. (2) More complex and costly are *schemas*, whose slots and fillers require more processing (e.g., PS diagrams are diagrammatic configuration schemes [10]). (3) IRs based on *rules* are next, which are more costly because they have fewer constraints, so effort must be expended just to identify categories and track concepts. (4) *Mental-imagery* is more costly still, because the imagery system's limited functionality and resolution will tend to demand multiple

iterations of procedures [6]. (5) *Propositional-networks*, such as analogies, are the costliest because they are largely built on simple associations, which place little constraint on valid inferences. The form of a given RS may suggest what IR a user will likely adopt (e.g., for Fig. 1a: rules; for Fig. 1b: schema; for Fig. 1c: diagrammatic schema). So, the ordering provided by these processes provides means to estimate the relative cost of the CP.

Note that the order of our proposed processes for ER- and IR-semantic-process CPs, although sensible, is something that needs to be demonstrated empirically too.

7 Inference Cognitive Properties

This penultimate group of CPs concerns costs at the level of making inferences. One of the properties in this group is **quantity-scale**, which concerns the type of quantity or measurement scale that dominates an RS, specifically, *nominal*, *ordinal*, *interval* or *ratio* [22]. Zhang [26] considered the role of quantity scales in the design of representational systems, and the scale hierarchy is well documented [29]. Here, we claim, further, that as the more sophisticated scales have more information content, they will impose greater cognitive cost. However, it is unlikely that RSs will differ in their use of quantity scales, because this is substantially determined by the content of the problem. For example, all three of our examples in Fig. 1 involve quantities related to nominal (manipulation of sets) and ratio (manipulation of probability quantities) scales. Rather, this CP is included because users' degree of experience in reasoning with more sophisticated scales is likely to have cost implications. For this CP, we are currently conducting empirical studies about the relative costs of the scales.

The next CP in the inference group is **expression-complexity**. Obviously, the longer an expression, the more components it possesses or the more tortuous it is, the greater the costs of using it to generate new information. For instance, it is easier to understand how each part of a PS diagram constrains the size of other parts than it is to work out how the magnitudes of variables vary in relation to each other in the Bayesian representation. Expression-complexity may be decomposed into particular factors such as the depth of relations and the arity of relations. The former is the number of levels of nesting of relations. The latter is the number of arguments that relations take. The more arguments, the more information must be handled, so the greater the cost [8]. For instance, the calculation of the final answer in the Bayesian and the contingency table solutions take six numbers, whereas only two are used in the PS diagram solution.

Not all inferences have the same difficulty, so the **inference-type** CP considers various types, for which we propose this rank ordering cost: (1) *symbol-selection* (e.g., lookup a table cell entry); (2) *assign/substitute* a symbol or concept (let the top-left sub-segment line in the PS diagram in Fig. 1c stand for no_*D*); (3) *compare/match* symbols or concepts; (4) *select-expression*; (5) *substitute-expression*; (6) *calculate*; and (7) *transform-expression*, which re-arranges the structure, resulting in a new relation (e.g., writing a new line in Fig. 1a; drawing a new sub-space in Fig. 1c). The Bayesian representation in Fig. 1a is dominated by the costliest of the 7 inference-types (e.g., transform-expression), but not so for Fig. 1b and 1c. Again, some empirical evidence will be needed to support our proposed order of processes for this CP.

8 Problem Solution Cognitive Properties

To capture the impact at the overall level of problem solutions, three CPs are proposed [17]. The first two are **solution-depth** and **solution-branching-factor**, which consider the overall topology of the hierarchical problem state space that users of a representation generate when solving problems. Solution-depth is the number of steps on the most direct path between the initial state and solution. The solutions to the medical problem in Fig. 1 are ideal solutions, with no back-tracking nor branching, so the number of operations that generate the solutions is also the solution depth. The solution-branching-factor addresses the likely width of the problem space experienced by a problem solver. For example, the branching factor from step 1 to 2 in Fig. 1a is higher than in Fig. 1c: a problem solver using a Bayesian representation may need to consider several theorems to move from step 1 to 2; while a problem solver using the PS diagram just needs to draw the different events for each of those steps. A problem state space given by an RS offers the problem solver alternative paths to follow and it will increase costs in at least two ways. First, it is the simple challenge of choosing which path to follow; and second, many alternative paths may lead to impasses rather than solutions. Clearly, the heuristics possessed by a problem solver will influence the solution-depth and the solution-branching-factor.

The **solution-technique** CP considers problem solution approaches that depend on the nature of the problem, which are distinct from general heuristics, and focuses on the nature of the procedures that are used for solutions. Two problem solutions might have the same breadth and depth but may vary in the variety of operators that are used to generate expressions. For example, a solution in a PS diagram typically involves iterative applications of finding a subspace in the diagram and drawing further sub-divisions of them, whereas algebraic solutions invoke a larger range of operations that vary with the changing structure of the expressions [3]. As teachers of programming know, iterative processes are typically easier to grasp and to implement than recursive processes. Hierarchical processes also tend to be more complex than iterative processes, because they require nested sub-procedures and the management of sub-goals.

9 Example of Application

One can envisage many uses for the CP framework [9]. It may serve as a checklist of factors that instructors might consider when they develop a curriculum in order to determine the order in which to introduce different representations. More ambitiously, we are using the framework to develop an AI engine that will automatically select representations that are suited to particular problems and users with different levels of familiarity of a target pool of representations. This section of the paper summarises the role of the CPs framework in the development of our first prototype of a representation selection system called rep2rep [18] as a concrete illustration of the framework's utility. In [18], the main focus is on the formal properties and the application of our framework, whereas the underpinning cognitive rationale is the main contribution of this paper.

The general challenge is to develop computational mechanisms that formalise the CPs described by the framework in such a way that their associated cognitive costs can be accurately calculated – to enable the selection of effective RSs for problem solving.

In order to meet this challenge, we need to cover two levels of abstraction. At the lower level we have questions such as 'how do we count the number of symbols in a representation?' and 'what is the expected cost of reading any of the symbols in Fig. 1a?' – which requires a prediction of how the physical components would be chunked into discrete symbols and how much time and effort it would take. And at the higher level, we have questions such as 'how does the number of symbols affect the cost?'. For our computational formalisation, we *assume-as-given* some answers to the lower-level type of questions. We only address computationally the higher level. To be clear, this does not mean that we *have* concrete answers to the lower level-type of questions. It only means that to turn our implementation into a full computational formalisation of the framework we need to *plug in* mechanisms that yield the lower-level values.

Computationally, we encode representations abstractly as collections of *primitive terms*, *patterns*, *laws*, and *tactics*. We call these the *formal components* of a representation. Terms (or symbols) are assigned *types*, and patterns capture the idea that higher-granularity items (*composite terms*) in a representation are formed from lower-granularity items, all the way down to the primitive terms. Specifically, a pattern describes the structure of composite terms (of a certain type) which are made up from more basic terms of certain types. This abstraction – of patterns as the glue of composite terms – can capture the complexity of various grammars: from natural language, to formal mathematics, to graph-theoretic or geometric diagrams [18, 19]. Analogous to the way in which patterns describe the structure of composite terms from more basic terms, tactics encode the structure of inferences from more basic knowledge, all the way down to laws[4]. Moreover, the links between different representations (e.g., how the same problem is encoded in multiple RSs) is captured by the concept of *correspondence*. Lastly, the user's general *expertise* is captured simply as a value between 0 (novice) and 1 (expert).

Given the abstraction of representations into their formal components, the question now is how the CP framework is applied. For the work in [18], we formalised a version of each of: sub-RS variety, registration (of primitives and composite terms), concept-mapping, quantity-scale, expression-complexity, inference-type, solution-depth, and solution-branching-factor[5]. As stated above, the formalisation of these properties relies on some low-level assumed-as-givens. These take either of the following forms:

1. Given a problem-solution representation, its abstraction into formal components is assumed. This means, for instance, that the question of which terms are considered primitive (in practice, a question of chunking) must be given. Furthermore, a value of *importance* is assigned to each component, encoding its relevance with respect to the solution (e.g., a component that plays no role in the solution is considered unimportant and given a value of 0).
2. The assignment of cognitive attributes to components is assumed. This means, for instance, that whether a tactic is assigned the attribute of being a substitution or a calculation (see Sect. 7), must be given. Furthermore, the parameter values for basic

[4] In formal, sentential mathematics these would be called axioms, but we do not want to give the impression that either (i) our system only applies to axiomatic systems or that (ii) laws have to be as low level as axioms typically are.

[5] Other CPs, e.g., IR & ER-semantic-process and solution-technique, are yet to be implemented.

costs, associated with these attributes, are assumed. This means, for instance, that the cost of a single inference which is a calculation is assumed to be twice as costly as that of a simple substitution. Lacking specific and accurate empirical data, ratios such as this one were chosen arbitrarily with the simple constraint that they must preserve the rank order specified by the framework.

Given these low-level assumed values, we assign a cognitive cost for each CP using a variety of methods. For example, registration and inference-type costs are similarly computed as a sum of the basic parameter values for the given components modulated by importance and expertise (expertise is assumed to reduce the impact of noisy components, as these can be ignored). Expression-complexity and solution-branching-factor, on the other hand, are computed from the branchiness and nestiness of patterns and tactics, respectively, with a similar effect from expertise. Quantity-scales is computed via the correspondences of components to arithmetic operators, and concept-mapping is computed via the type of relation given by the correspondence map to a fixed representation. Sub-RS-variety is simply computed from the number of *modes* (a given) which are intended to capture individual formats used in the representation.

Once the cognitive cost associated to each CP is computed, they are combined in a weighted sum, with CPs in higher cognitive level and higher notation granularity being assigned greater weights. Moreover, expertise is assumed to have a stronger impact on the cost of CPs of higher notation granularity components.

Our prototype engine for representation selection can also be used to produce an informational suitability score, which estimates the likelihood that a given RS can be used to represent and solve a problem. An interesting question for future research is how the informational and cognitive computations can be used synergistically. It is clear that it depends on the application in which our framework is employed. Precise formulae for informational suitability and cognitive costs, and details of their implementation can be found in [18].

9.1 Evaluation

In [18], we presented an evaluation of the effectiveness of the implementation, which is summarised here. Since there are no other systems to compare against, the evaluation was done by comparing computed measures of informational suitability[6] (IS) and cognitive cost against data obtained from surveying expert analysts. That is, was our system producing similar rankings as expert humans? The evaluation focused on the domain of probability and the medical problem presented in Sect. 3, albeit using different values. The RS used were Natural Language (NL), Bayes, Areas, and Contingency Table. The computation of IS was done as stated at the start of this section, and the cognitive cost function was computed considering 3 user profiles, which were set through the general expertise function described above.

Eleven analysts with strong mathematical background completed an online questionnaire, which contained 2 tasks. In Task 1, participants were first shown the description

[6] Information suitability measures how well a representation encodes the informational content of a problem and is computed using the formal properties of representations.

of the medical problem. Then they were asked to give feedback on how informationally sufficient (descriptions of) RSs were using a 7-point Likert scale. In Task 2, participants were asked to rank the same RS descriptions, but for novice, expert and average users.

The mean Likert score given to different RSs in Task 1 was used to derive IS ranking, and the mean of the rank scores across different RS was used to derive the ranking of different RSs for different user profiles. In terms of IS, the rank order produced from the rep2rep system and the analysts was similar for the most and least IS RSs (Bayes and NL, respectively), but different on the Areas and Contingency Table RS. Although the correlation was not significant, it was considered that the overall ranking produced by the system was sensible. In terms of cognitive costs, the rankings given by the analysts and the rep2rep system for the expert and average profiles showed high and statistically significant correlations at $p < .05$ ($r = 0.9$), but not for the novice profile. A possible explanation of this result is that users' familiarity with the RS is not yet modelled in the system. Details can be found in [18].

Overall, the results are promising in terms of the AI system being able to recommend effective representations – although more empirical work still needs to be done.

10 Discussion

To identify cognitive properties that contribute to the cognitive cost of an RS, we formulated the analysis framework, as summarised in Table 1. We proposed 13 diverse CPs. Some relate cost to counts of instances found, some require the calculation of an average to represent some commonly occurring factor, and others propose ranking of processes as guides to relative cost. Although 13 CPs are postulated, we make no claim that they are exhaustive, and note that some are applicable at multiple levels of granularity of RSs. A key feature and potential benefit of the framework is its differentiation of CPs within a two-dimensional space of cognitive level and notation granularity. Given a particular problem-solving process, one can use the dimension to locate its position within the space and, hence, the CPs that are likely to be important factors that impact the cost of the process in different representations. Nonetheless, CPs are not perfectly orthogonal. For example, the number-of-symbols will likely increase with the number of sub-RSs. However, the distinction between these RSs is important, not just because they span very different ranges in the framework, but because we can imagine a situation where one RS *A* is comprised of two sub-RSs, and a second RS *B* without sub-RSs has an equal number-of-symbols. In that case, the RS *A* will have a higher cost because of the challenges related to multiple sub-RSs.

Whilst more extensive justification and rigorous definition could be made about the values of CPs and the rank order of the costs of particular CPs, we consider that the given notions and orders are reasonable.

Note that the three example representations in Fig. 1 encode equivalent sets of concepts. If this were not the case, then fair comparisons could not be made [2]. However, the framework does permit comparisons where the ERs of two RSs are not equivalent, as long as any difference is remedied in the IR content of the RS in deficit.

As part of our ongoing work with the framework, we are investigating how to combine the CPs into a single cost measure for whole RSs. Three critical issues will need to be addressed.

1. How can the disparate measures, with their different scales, be normalised so that they can be reasonably combined?
2. What weighting should be given to those normalised CPs, as they naturally have different levels of impact?
3. How should the weights of each CP be moderated given differences in individual's expertise with alternative RSs?

Our first prototype representation selection engine rep2rep, described in Sect. 9, provides one tentative solution to the first two issues, at least for selected CPs. More broadly and fortunately, the framework supports our analyses of the questions, because it acknowledges the range of granularity scales applicable in the use of RSs. For instance, we have some basis to examine trade-offs between changes to CPs at the lower levels (registration, semantic encoding), which have small impacts on numerous symbols and expressions, versus changes to CPs at higher levels (inference and solution), which impact just a few large-scale procedures.

Acknowledgements. We thank Gem Stapleton, from Cambridge University, for her comments and suggestions for this paper. This work was supported by the EPSRC grants EP/R030650/1, EP/T019603/1, EP/R030642/1, and EP/T019034/1.

References

1. Anderson, J.R.: Spanning seven orders of magnitude: a challenge for cognitive modeling. Cogn. Sci. **26**, 85–112 (2002)
2. Cheng, P.C.-H.: Electrifying diagrams for learning: principles for effective representational systems. Cogn. Sci. **26**(6), 685–736 (2002)
3. Cheng, P.C.-H.: Probably good diagrams for learning: representational epistemic re-codification of probability theory. Top. Cogn. Sci. **3**(3), 475–498 (2011)
4. Cheng, P.-H.: What constitutes an effective representation? In: Jamnik, M., Uesaka, Y., Elzer Schwartz, S. (eds.) Diagrams 2016. LNCS (LNAI), vol. 9781, pp. 17–31. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42333-3_2
5. Cleveland, W.S., McGill, R.: Graphical perception and graphical methods for analysing scientific data. Science **229**, 828–833 (1985)
6. Finke, R.A.: Principles of Mental Imagery. The MIT Press, Cambridge (1989)
7. 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, New York (1998). https://doi.org/10.1007/978-1-4612-1676-6_10
8. Halford, G.S., Baker, R., McCredden, J.E., Bain, J.D.: How many variables can humans process? Psychol. Sci. **16**, 70–76 (2005)
9. Jamnik, M., Cheng, P.C.-H.: Endowing machines with the expert human ability to select representations: why and how. In: Muggleton, S., Chater, N. (eds.) Human Like Machine Intelligence. Chapter 18. Oxford University Press, Oxford (2021). (in press)
10. Koedinger, K.R., Anderson, J.R.: Abstract planning and perceptual chunks: elements of expertise in geometry. Cogn. Sci. **14**, 511–550 (1990)
11. Kotovsky, K., Hayes, J.R., Simon, H.A.: Why are some problems hard? Cogn. Psychol. **17**, 248–294 (1985)

12. Larkin, J.H., Simon, H.A.: Why a diagram is (sometimes) worth ten thousand words. Cogn. Sci. **11**, 65–99 (1987)
13. Markman, A.B.: Knowledge Representation. Lawrence Erlbaum, Mahwah (1999)
14. Miller, G.A.: The magical number seven plus or minus two: some limits on our capacity for information processing. Psychol. Rev. **63**, 81–97 (1956)
15. Moody, D.L.: The "physics" of notations: toward a scientific basis for constructing visual notations in software engineering. IEEE Trans. Softw. Eng. **35**(6), 756–779 (2009)
16. Newell, A.: Unified Theories of Cognition. Harvard University Press, Cambridge (1990)
17. Newell, A., Simon, H.A.: Human Problem Solving. Prentice-Hall, NJ (1972)
18. Raggi, D., Stapleton, G., Stockdill, A., Jamnik, M., Garcia Garcia, G., Cheng, P.C.-H.: How to (re)represent it?. In: 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI), pp. 1224–1232. IEEE, Baltimore (2020)
19. Raggi, D., Stockdill, A., Jamnik, M., Garcia Garcia, G., Sutherland, H.E.A., Cheng, P.-H.: Dissecting representations. In: Pietarinen, A.-V., Chapman, P., Bosveld-de Smet, L., Giardino, V., Corter, J., Linker, S. (eds.) Diagrams 2020. LNCS (LNAI), vol. 12169, pp. 144–152. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-54249-8_11
20. Shimojima, A.: Semantic Properties of Diagrams and their Cognitive Potentials. CSLI Press, Stanford (2015)
21. Simon, H.A.: Sciences of the Artificial, 2nd edn. MIT Press, Cambridge (1981)
22. Stevens, S.S.: On the theory of scales of measurement. Science **103**(2684), 677–680 (1946)
23. Sockdill, A., et al.: Correspondence-based analogies for choosing problem representations. In: 2020 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC), pp. 1–5. IEEE, Dunedin (2020)
24. Sweller, J.: Cognitive load during problem solving: effects on learning. Cogn. Sci. **12**(2), 257–285 (1988)
25. van Someren, M.W., Reimann, P., Boshuizen, H.P.A., de Jong, T.: Learning with multiple representations. Advances in Learning and Instruction Series. ERIC (1998)
26. Zhang, J.: A representational analysis of relational information displays. Int. J. Hum. Comput. Stud. **45**, 59–74 (1996)
27. Zhang, J.: The nature of external representations in problem solving. Cogn. Sci. **21**(2), 179–217 (1997)
28. Zhang, J., Norman, D.A.: A cognitive taxonomy of numeration systems. In: Proc. of the 15th Annual Conference of the Cognitive Science Society, pp. 1098–1103. Lawrence Erlbaum, Hillsdale (1993)
29. Zhang, J., Norman, D.A.: A representational analysis of numeration systems. Cognition **57**(3), 271–295 (1995)