

Chapter 17

Mathematical Arguments and Distributed Knowledge

Patrick Allo, Jean Paul Van Bendegem, and Bart Van Kerkhove

Because the conclusion of a correct proof follows by necessity from its premises, and is thus independent of the mathematician's beliefs about that conclusion, understanding how different pieces of mathematical knowledge can be distributed within a larger community is rarely considered an issue in the epistemology of mathematical proofs. In the present chapter, we set out to question the received view expressed by the previous sentence. To that end, we study a prime example of collaborative mathematics, namely the *Polymath Project*, and propose a simple formal model based on epistemic logics to bring out some of the core features of this case-study.

17.1 Introduction

In his *Objective Knowledge*, Karl Popper famously claimed that the beliefs of the individual are irrelevant for epistemological purposes; especially when that individual is understood as the individual scientist, and the field of epistemology is restricted to scientific knowledge (Popper, 1968, 1972). Even if, in view of the present state of the philosophy of science, this account has become untenable in general,¹ one might still be tempted to believe that, at least when it comes to mathematical knowledge, Popper's conception of objective knowledge remains a sensible

¹Especially in the light of work carried out at the intersection of social epistemology and philosophy of science. See, e.g., (Kitcher, 1990) for a general defence, or (Zollman, 2007) for a specific study.

P. Allo • J.P. Van Bendegem • B. Van Kerkhove (✉)
Centre for Logic and Philosophy of Science, Vrije Universiteit Brussel, Pleinlaan 2,
Brussels, B-1050, Belgium
e-mail: pallo@vub.ac.be; jpvbende@vub.ac.be; bvkerkho@vub.ac.be

position. This, one could argue, follows from the fact that mathematical knowledge is obtained through proof. For when truth is exclusively arrived at via proof, we only need to be concerned with the correctness of that proof, and this does not depend on what any individual mathematician believes. Certainly there are many deep issues in the epistemology of mathematical proofs.² Yet, because the conclusion of a correct proof follows by necessity from its premises and is thus independent of the mathematician's beliefs about that conclusion, understanding how different pieces of mathematical knowledge can be distributed within a larger community is rarely considered an issue in the epistemology of mathematical proofs.

The above description is particularly compelling if one thinks that mathematics is just about correct proofs. By contrast, if we accept that mathematics is also about proofs that are *recognised* to be correct, or that are deemed to be *acceptable*, or widely thought to be *convincing*, knowledge and acceptance in different mathematical communities becomes relevant. This is particularly so when proofs are seen as arguments. On that conception, correctness is plainly insufficient. What arguments aim at is to be convincing, and this is something that depends on the audience that needs to be convinced of the truth (or provability) of a given theorem. This, for sure, depends on what the members of that audience know, but even that isn't enough. In addition, one would also like to know (or at least believe) that a certain argument will be convincing, which requires one to know what the members of one's audience know. These considerations not only show that what the individual mathematician knows matters for the epistemology of proofs-as-arguments, but also that an individual mathematician's beliefs about the knowledge of his audience (sometimes his peers, but equally often members of other communities or lay-people) is equally relevant. In sum: Not only do beliefs of the individual matter, but also beliefs about beliefs of others matter.

When we look at the received view about how proofs are conceived (discovered, found, designed, . . .), we find the same individualistic bias that is also present in mainstream epistemology. Proofs are the work of the individual mathematician, or at most the work of a small number of individuals. Again, one might be tempted to believe that traditional arguments for a social or interactive conception of knowledge need not apply to mathematics. Because individuals can come up with new results, and as the whole enterprise of mathematics is traditionally seen as cumulative, no real interaction is needed beyond the obvious reliance on prior results. Yet, this description no longer covers the totality of mathematical inquiry. Some proofs don't just happen to be the result of massive collaboration, but are actually results that are well beyond the reach of individuals and even smaller communities. This means that proofs cannot only be seen as arguments once they are completed (and indeed need to be used to convince the broader community), but that the process of looking for a proof can equally well be studied from an argumentation-theoretic perspective, and more generally from the perspective of social and interactive epistemology.

²What is the formal nature of proofs? Are proofs out there (in "The Book") or constructed by us? What is a surveyable proof and what not? How does one grasp (the content of) a proof? Are there more and less beautiful proofs? (etc.)

Formal models of scientific knowledge are often based on *Bayesian Models* (Bovens and Hartmann, 2003), or on *Signalling Games* (Zollman, 2007). Here, we pursue a different path and propose to use the epistemic logics that have been developed by computer scientists to reason about knowledge in *multi-agent systems* (Fagin et al., 1995). The not so obvious choice of a logical model is deliberate: We believe logic is equally relevant to the study how information flows in epistemic communities—our present aim—as it is to the study of validity—the traditional focus of mathematical proofs. This motivates the choice for a logical, as opposed to another formal framework.

Within the broader landscape of logical approaches, the choice for epistemic logics still needs further motivation. One element has already been advanced: We want to be able to reason about what different agents know and believe. But, as mentioned above, we also want to reason about knowledge in groups, and crucially also about higher order epistemic states (knowledge and belief about other knowledge or belief). This is a task that we think is best carried out within the framework of epistemic logics. Given the focus on a conception of proofs as arguments, formal models that directly deal with arguments like the logics for defeasible argumentation described in (Prakken and Vreeswijk, 2002) might be thought to be more obvious candidates. One issue with the latter type of formalism is that it primarily provides a formal model of *defeasible inference*, and this is a type of reasoning we do not readily associate with proofs in mathematics (Aberdein, 2005, 289–90). Still, as shown by Aberdein (2005) via the contrast between classical and constructive validity, Toulmin’s model of argumentation (itself a model for defeasible argumentation or inference) can be used to model mathematical proofs, and the result of doing so can be informative (see also Aberdein, 2007). The main argument in favour of the doxastic/epistemic approach then, is that it allows one to reason explicitly about how agents and communities of agents interact; a perspective that is not available to more abstract approaches to argumentation.

No one will doubt that argumentation theory has a long and respectable history within philosophy, starting with Aristotle, who, at the same time, initiated the distinction, if not opposition, between the predominantly informal realm of argumentation (to convince an audience) and the more formally oriented field of logic (to establish necessary truths). In the twentieth century this situation of course drastically changed, as new developments were initiated that aimed at a *formal* argumentation theory. Among the major contributions, we find dialogue logic, as designed by Lorenzen and Lorenz (1978), the core idea of which is that a logical system, axiomatically and/or deductively presented, can be reformulated in terms of rules for a discussion between two parties, thus showing that the distinction between logic and argumentation is partially fictitious. A similar attempt, in terms of game theory, was made by Hintikka (1985), and related models were developed by Barth and Krabbe (1982).

Of major importance in the field of *informal* argumentation theory is the work of Grootendorst and Van Eemeren (see e.g. Van Eemeren et al., 2004). The Amsterdam school, which has been the driving force behind the International Society for the Study of Argumentation (ISSA), is mainly interested in specific case studies, with

an eye on determining how participants in actual dialogues or discussions deal with arguments and their evaluation. There is moreover a direct connection with the Canadian school, with well known contributions like (Walton, 1998) and (Woods, 2003). Here too there is a strong focus on concrete situations and, above all, a careful investigation of whether fallacies are always to be rejected in a discussion or can sometimes be acceptable forms of reasoning. For example, classically, the argument *ad auctoritatem* is considered to be a fallacy, while expertise in a court of law is clearly acceptable.

The third development to be noted is that within the field of Artificial Intelligence (AI). Indeed there has been a strong interest for some time now in the modelling of argumentations (see e.g. Toulmin, 1958; Pollock, 1994; Vreeswijk, 1997). The main purpose of this research is to develop applications in decision-guided systems and intelligent agents (Reed, 1998). Recently, argumentation networks have been proposed as simple but very powerful models to represent static structures of arguments in competition with one another (Dung, 1995). However, the models remain rather abstract: arguments are viewed solely in relation to others, the basic scheme being one argument “attacking” another. Several semantics have been proposed for such networks, all taking a declarative and monological approach as a starting point (Bondarenko et al., 1997; Dung, 1995; Jakobovits and Vermeir, 1996, 1999b). In a further stage so-called dialectical semantics have been investigated (Jakobovits and Vermeir, 1999a; Prakken and Sartor, 1996), e.g. searching for applications in judicial reasoning (Verheij, 1995). Connections have already been shown between some of these semantics and the “well-founded”, “stable” ones of logic programs (see e.g. Kakas et al., 1994). More generally speaking, such argumentation networks can be applied in areas where it is important to have a *motivation* for a result or a decision, e.g., in medicine (Atkinson et al., 2005) or in bioinformatics (Jeffreys et al., 2006). As far as we are aware, applications to mathematical contexts are absent and, as stated, it is our main purpose to start filling in this gap.

As mentioned, the abstract approach to argumentation lacks a notion of players or agents; it is solely concerned with the arguments themselves. While the main logical approach to agency is to be found in the field of modal epistemic logics, further connections with argumentation remain under-explored.³ Because agency and interaction are central to our present enterprise, we privilege the framework of epistemic logics, and try to emphasise throughout this chapter to what extent it is relevant to the study of arguments in and about proofs. The upshot of the present article is to illustrate how epistemic logics that allow us to reason about the knowledge of individuals and the knowledge available within communities can be used to model proofs-as-arguments from a perspective that emphasises interaction and collaboration. This approach expands on a proposal described in Van Bendegem (1985b), but does so in a more flexible manner by modelling the structure of communities in terms of group and sub-group membership rather than with the

³For a first connection between abstract argumentation and modal logic, see (Grossi, 2010).

accessibility relation of a Kripke-model. In the following sections, we first present and analyse the Polymath Project as a potentially interesting case-study (Sect. 17.2), then on the basis of previous work by the second author (Van Bendegem, 1982, 1985a) and a number of other, more recent results develop a formalism (Sect. 17.3), and then also make it fit to capture cases plucked from contemporary mathematical practice, like the previously mentioned Polymath Project (Sect. 17.4).

17.2 Mathematics in the Cloud: The Polymath Project

In Gowers and Nielsen (2009), a mathematician and a science writer/former physicist report their endeavours in online collective problem solving. The Polymath Project, as it has been called, has drawn quite some attention lately. Some claim that it is an example of a new method or procedure of proof search in mathematics, whereas others merely see it as old methods implemented through new media, such as the internet. It is our hypothesis that the formal considerations presented here, in this chapter, can help us to clarify the matter. As it happens, we do believe that there are a sufficient number of distinctive characteristics that allow one to consider this development as indeed new and philosophically relevant.

The clue to the story is provided by the mathematician who initiated the whole process, Timothy Gowers, who wrote on his blog: “I’m interested in the question of whether it is possible for lots of people to solve one single problem rather than lots of people to solve one problem each” (Gowers, 2009a). Justin Cranshaw and Aniket Kittur made a similar comment: “In the case of finite simple group classification,⁴ the larger task being solved had hundreds if not thousands of natural and predefined subtasks whose solutions were largely independent of the solutions of the other subtasks. Gowers was instead proposing to collaboratively solve a single problem that does not naturally split up into a vast number of subtasks” (Cranshaw and Kittur, 2011). In terms of a problem-solving community, this means that a quite different structure is required and that is one of the focal points of this chapter.

Some historical background.⁵ In January 2009, Timothy Gowers opened a website, accessible to everyone, mathematicians and non-mathematicians alike, announcing that he was searching for a proof of a particular mathematical statement, and inviting them to join him in that search. Anyone could post a message about almost everything, on the obvious condition that it was somehow related to the proof search. In fact, a set of ground rules was announced to avoid the whole enterprise becoming all too chaotic. The hope was to find a proof and, if that were to occur, to publish the proof through the usual, existing channels, namely mathematical journals. Presented thus, this in fact does not really sound novel, except perhaps for the fact that a large number of people could (and did) participate. But that

⁴See, e.g., (Van Bendegem and Van Kerkhove, 2009, Sect. 4).

⁵A more complete and highly accessible version of events can be found in (Nielsen, 2011).

only seems to be a matter of scale and, although changing the scale of a process can induce important changes, it does not therefore affect the underlying standard picture of proof search. However, closer inspection does reveal some interesting elements. But let us first have a look at the problem itself.

The problem Gowers launched on the website is known as the Density Hales-Jewett Theorem (*DHJ* henceforth), for $k = 3$ at first, but later generalised to arbitrary k . This problem comes under the heading of Ramsey Theory, that is to say it involves combinatorics and so-called colouring problems, where “unavoidable” properties appear. The typical format of such problems is that, given a structure of sufficiently large size, there will always be substructures that have a particular property, the unavoidable property. More specifically, *DHJ* for $k = 3$ states the following. Let there be:

- a set $K = \{1, 2, 3\}$, with $\#K = k$
- a set $N = K^n$, i.e., the set of all words of length n , on the basis of K .

Next we need two definitions:

- A *combinatorial line* is a subset of N such that, for at least one x , the elements are of the form: $k_1 k_2 \dots k_i x k_{i+2} \dots k_n$ for $x = 1, 2, \dots, k$.
- *Example* Take $n = 6$, then: the subset $\{122132, 122232, 122332\}$ is a combinatorial line, as is the subset $\{112132, 212232, 312332\}$.
- Define the density d of a subset M of N by $d = \#M/\#N$.

DHJ for $k = 3$ says that, for every d , there exists an n such that every subset M of N with density at least d contains a combinatorial line. The “unavoidable” property here is the presence of a combinatorial line. So the theorem says that no matter how low the density of a particular subset, if the words made on the basis of the alphabet can be sufficiently long, there will always appear a combinatorial line.^{6,7}

What happened after the opening of the website? First, after Gowers, Terence Tao joined the enterprise (see below). After 6 weeks, during which 39 contributors had posted 1,228 comments (after every 100 of these, Gowers made summaries to keep an overview), not only was a proof found, but it became immediately clear how it could be generalized for arbitrary k . The proof has meanwhile been published under the pseudonym *D. H. J. Polymath* (2010), with a second paper (Polymath, 2012) being under review, reminiscent of other fictitious names covering a collective in the history of mathematics, the most famous one no doubt that

⁶This highly compact formulation of the *DHJ* can be complemented by, especially for the non-mathematician, “The gentle introduction to the Polymath project” to be found at: <http://numberwarrior.wordpress.com/2009/03/25/a-gentle-introduction-to-the-polymath-project/>.

⁷Actually, a proof already existed. It was formulated in (Furstenberg and Katznelson, 1991). However, this original proof relied on methods and techniques from domains far away from combinatorics, such as ergodic theory. So, as often happens in mathematical research, although one has a proof of the theorem, nevertheless this does not prevent mathematicians from searching for an alternative and, more importantly, an elementary proof, i.e., a proof using the concepts, proof methods and techniques of the domain itself. For more on this topic, see, e.g., (Rav, 1999).

of Nicolas Bourbaki. What can be concluded from this episode? Surely the most striking feature of the whole process is that “amateurs” could and did participate, although whether we should be as enthusiastic about this as Jacob Aron, claiming that this will “democratize the process of mathematical discovery” (see Aron, 2011), is another matter. A good reason for a restricted enthusiasm has to do with the history of amateur mathematics, probably known to all of us. Refutations of accepted results (see Hodges, 1998, for a rather depressing discussion of refutations of Cantor’s diagonal argument or Dudley, 1992, for alleged proofs of almost everything, including the parallel postulate) and easy to understand, short proofs of famous conjectures such as Fermat’s Last Theorem (even after Andrew Wiles settled the matter) or Goldbach’s conjecture are by far the most common. However, in this case the situation is different. The amateur contributions are most of the time not in the form of proofs, but in the form of suggestions, ideas, outlines of possible proofs, and so forth (see Sect. 17.2.3 below). Perhaps more importantly, the “classic” amateur is an isolated individual who sends material to mathematicians or editors of journals,⁸ preferably on a 1-1 basis, whereas here they contribute bits and pieces in an open social environment, where everything is noted and listed, for all to see. In short, your Polymath amateur is not your average “crank” amateur. Apart from that, are there any other, special features? We think there are at least three.

17.2.1 Feature 1

Combinatorial problems are rather specific mathematical problems that have two important properties: (i) the problem is usually quantified over the natural numbers, so, instead of having a direct go at the full proof, case-by-case proof search can be very revealing as to the general proof (as in the case of, e.g., Fermat’s Last Theorem), and (ii) the search for counterexamples lends itself, at least for small numbers, to an exhaustive search⁹ and, for larger numbers, to clever methods for reducing the cases to be examined. As it happened in the *DHJ* case, both research lines were pursued: next to the Gowers project, corresponding to (i), Terence Tao set up a project, complementary to the Gowers project and corresponding to (ii). More specifically, Tao asked the negative question: What is the size of the largest subset C of N that does *not* contain a combinatorial line as a function of k and n ?

The question now becomes: given k and n , what is the lower limit of d ? The reader can check for him- or herself that for $k = 3$ and $n = 2$, the size is 6. As k and n are finite, the set N will also be finite, so, in principle, all elements can be listed,

⁸One of the authors of this chapter is himself the editor of a logic journal, *Logique et Analyse*, and has a nice collection of, e.g., disproofs of Gödel’s theorems.

⁹This type of research, much loved by some of the aforementioned amateurs can indeed be seen as “outsourcing drudge work, comparable to Amazon’s Mechanical Turk” (<https://www.mturk.com/mturk/welcome>), as remarked by one of the referees.

all subsets can be listed and a step-by-step control of each subset can be checked for combinatorial lines, such that the size of C can be calculated. However, as one might expect, the concept of combinatorial explosion is definitely applicable here, as even for small k and n , the number of subsets grows exponentially and so one has to search, as mentioned, for clever tricks, techniques and methods to get sufficient insight in order to calculate the size of C .

This leads to the following interesting open question: are Polymath problems restricted to combinatorial problems (or to problems satisfying conditions (i) and (ii))? One might be tempted to answer positively, especially given the further development of the project after the success of the *DHJ* theorem (see below at the end of this section), but matters are not that simple. Note first of all that not all combinatorial (or similar) problems are good candidates. Take Fermat's Last Theorem as an example. The problem satisfies both conditions and similar to the *DHJ* theorem one could suggest the search for an elementary proof. There is however the commonly shared belief that such an elementary proof is highly unlikely, hence the risk is too high to meet failure and hence no such project will ever be launched (and the search for such a proof is definitely left to the "crank" amateurs). And it is true of course that Riemann's Hypothesis is a very unlikely candidate to be taken up as a project—in Tao's words on his blog: "I would imagine that a polymath to solve the Riemann Hypothesis will be a spectacular and frustrating fiasco; we should focus on problems that look like some progress can be made"¹⁰—but on the very same blog an important suggestion is made by Gowers, that can be seen as a generalisation of conditions (i) and (ii): "A lesson from the *DHJ* experience was that there were very different ways that people could contribute." So that seems to be the key: in a Polymath environment a methodological plurality is essential.

17.2.2 Feature 2

This feature can be described in sloganistic terms as "Lakatos visualised", in honour of Imre Lakatos, who has introduced historical and social elements in mathematics itself in his seminal study on Euler's Conjecture, entitled *Proofs and Refutations* (see Lakatos, 1976). What we mean by the "visualisation" is that, with Polymath, a register is made of all such social exchanges. Of course, we still do not know what happens in the individual minds of the participating mathematicians but, as soon as an item of information is exchanged, it is recorded. We thus obtain an extremely detailed record of the proceedings. This is obviously not only of interest for historians of mathematics, but more importantly so for the participating mathematicians. Surely this must lead to a different problem-solving practice in terms of shared knowledge in the community.

¹⁰See <http://polymathprojects.org/2009/07/27/selecting-the-next-polymath-project/>.

A first dimension is the role of authority in a problem-solving community. Anyone who had a glance at the Polymath blog will realise very quickly that Gowers and Tao are the authorities guiding the process. Not merely because both are internationally reputed mathematicians, but also because they formulated the minimal rules that any participant should respect, implying that any transgression of these rules could, in principle, lead to exclusion (although, as far as we know, this has never happened). In short they set up the framework wherein the game will be played. Their being in charge is also reflected in the number of contributions by Gowers and Tao, in comparison with other participants. Additionally, they continually made comments about the progress of the project, what suggestions and ideas were interesting and/or important to pursue, in short they both acted as filters. Another way to describe this situation is that we have here a network with two strongly connected “attractors” and no high density subnetworks, all nodes being mainly directed towards the two central ones.

A second dimension concerns the meta-level. For a similarly detailed recording of events and exchanges during the search process itself creates a “conscious” and shared meta-level, by which we mean all arguments, thoughts, ideas, etc., *about* the process itself, not so much as about the actual problem they are trying to solve. These comments are grouped together on a special blog.¹¹

17.2.3 Feature 3

What kind of contributions were made during the process? One might expect that, since the overall aim was to find a proof, everyone would have directly contributed to it. In terms of the ideal picture of a proof—a labelled list of statements, justified at each step, starting with the premisses and ending with the statement to be proved as conclusion—one could imagine Gowers writing down the first few lines, someone else contributing the next line, until it is realised that the road chosen is a dead-end, hence some backtracking is done and someone else proposes a new line to explore a different road, eventually arriving at the last line. Needless to say, that is not what happened. Actually, a stronger statement can be made: many contributions were of a quite different nature. What follows is not to be seen as an overview but as a selection of some elements that deviate from the picture sketched above:

- Looking at other domains: Very often analogies with other problems were proposed, both on the level of statements (“This problem looks very much like ...”), where the analogical problem often comes from a different domain than combinatorics, and on the level of proof procedures (“Could you make use of this or that technique that turned out to be useful for this or that problem?”), that were successful in other domains.

¹¹See <http://polymathprojects.org/category/discussion/>, for more specific comments and on a related blog, <http://polymathprojects.org/general-discussion/>, for more general comments.

- The same problem in different words: Equally often, concepts were proposed to reformulate the problem so that other problems could be related but also to get a better understanding of the problem. This is a technique that is well-known in mathematical practice: find equivalent forms of the statement you are trying to prove, hoping that the reformulation will be easier to prove. Think, e.g., about the four-colour theorem. Replace a region by a dot and, if two regions are neighbours, connect them with an edge. Now you obtain a graph and the theorem translates equivalently into: using four colours, colour the dots so that two connected dots do not get the same colour.
- Keep it simple(r): Not so much parts of possible proofs were proposed, but rather suggestions and quite vague ideas about possible routes to find the proof itself and what the proof could look like. Sometimes proofs of simpler problems, directly related to the original problem, were presented as a source of inspiration, reminding us of one of the strategies discussed in Pólya (1945, 75–85), namely, if the solution of a problem must satisfy a number of conditions, drop one of the conditions and see whether a solution can then be found, a special case of the general heuristic strategy, labeled “Decomposing and recombining”.

Consequences of this high diversity in the contributions are, at least the following:

- The community, both in terms of members and in terms of the exchanges they make, is definitely not homogeneous. Different members play quite different parts.
- Different members will come up with different ideas and concepts. What goes on in the mind of the professional high-ranked mathematician is definitely not what goes on in the mind of the high school teacher or dedicated amateur. The one is not a “light” version of the other, they do have their proper characteristics and have their own specific contributions to make.
- Room is created for chance elements to play a part. Crazy ideas get a place in such networks. The upshot seems to be that all these diverse elements create a rather special and specific problem-solving context. In this respect, Gowers has remarked: “Reading the discussion provides some kind of strange random stimulus that causes your brain to go in to fruitful places where it might not have gone otherwise” (as quoted in Aron, 2011).

What happened after *DHJ*? At the time of writing, five more problems have already been tackled. It is rather striking, although of course the specialties of Gowers and Tao have to be taken into account, that most of the problems, either directly or slightly less so, deal with combinatorics and the search for or the improvement of upper and lower bounds. Polymath 2 deals with Banach spaces and although that seems sufficiently far away from combinatorics, Gowers on the blog has remarked: “I have given definition 3 above because I think it has the potential to ‘combinatorialize’ the problem” (Gowers, 2009b). Polymath 3, the Polynomial Diameter Conjecture, states that, if G is the graph of a d -polytope with n facets, then the diameter of G is bounded above by a polynomial of d and n , and needs no further discussion. The same goes for Polymath 4, the statement that there exists

a deterministic algorithm which, given an integer k , is guaranteed to find a prime of at least k digits in length of time polynomial in k . (A proof of the theorem is under review.) Polymath 5 deals with Erdős's discrepancy problem, and Polymath 6 with improving bounds in Roth's theorem, both fitting nicely in the combinatorial domain. The question remains, as already mentioned above, whether there are problems *not* suited for this type of approach.

The case thus being laid out, it is now time to pass to the constructive part of our chapter, and propose a formalism that can capture instances of essentially collaborative mathematical research like the one met in this very example.

17.3 Formalising Shared Knowledge

Reasoning about how knowledge and belief can be shared by individuals is what epistemic logics, extended with operators for group knowledge, do best. Reasoning about how the distribution of shared knowledge and belief can be altered through interaction belongs to the domain of dynamic epistemic logic. In the present section the basic notions and insights from these fields are briefly described. We refer the reader to the Appendix for further details.

The epistemic and doxastic logics are built up in two steps. We start with the introduction of operators for single-agent knowledge (or belief) by extending the propositional language with a modal operator $[c]$ for each individual c we'd like to reason about. As is standard in the literature, we assume that these individuals are extremely powerful reasoners by stipulating that (a) they are logically and deductively omniscient, (b) they are infallible when it comes to their knowledge (they satisfy positive introspection), and (c) they are infallible when it comes to their ignorance (they satisfy negative introspection). When complemented with the standard view that knowledge is factive, this leads to a formalisation of individual knowledge that is based on the modal logic **S5**. Analogously, if we assume that beliefs can be false, but should still be consistent, the same assumptions lead to a formalisation of individual belief that is based on the modal logic **KD45**. In the remainder, we only focus on knowledge.

The operators for individual knowledge are sufficient to define two of the four main notions of group knowledge, namely *particular* and *general knowledge*. By particular knowledge in a group of individuals G , we mean knowledge by at least one of the members of G . We write $S_G\varphi$ to express that φ is particular knowledge in G . By general knowledge in a group of individuals G , we mean knowledge by all of the members of G . We write $E_G\varphi$ to express that φ is general knowledge in G . As is clear from the above, $S_G\varphi$ and $E_G\varphi$ are equivalent to, respectively, the (generalised) disjunction and conjunction of the respective knowledge claims for all individuals in G . No such straightforward reduction to claims about what the respective individuals know is available for the two remaining notions of group-knowledge. Once more, we refer the reader to the Appendix for the relevant details, and stick to an informal description of these notions. By *distributed knowledge* in G , we mean the knowledge

Table 17.1 Logical properties of group-knowledge

	$[c]\varphi$ for some $c \in G$	$[c]\varphi$ for all $c \in G$	Deductive closure	Pos & neg introspection
$D_G\varphi$			✓	✓
$S_G\varphi$	✓			(✓)
$E_G\varphi$	✓	✓	✓	
$C_G\varphi$	✓	✓	✓	✓

that can be obtained by pooling together the knowledge of all the members of G . We write $D_G\varphi$ to express that φ is distributed knowledge in G . To explain what we mean by *common knowledge* in G , we need to revert to the notion of general knowledge. Indeed, when we say that φ is common knowledge in G , we mean that φ is general knowledge in G , that it is general knowledge that it is general knowledge that φ etc. Common knowledge in a group not only amounts to general knowledge that cannot be doubted by any member of that group, but also implies that no member can doubt the fact that any other member is in such a position.

Every type of group knowledge has different logical properties, and therefore provides a model for how knowledge can be present in or available to a community with different strengths. For instance, φ being distributed knowledge within G is compatible with every member being ignorant about φ . That is, φ can be distributed knowledge without being particular knowledge. In more common terms: φ can be implicitly known without being explicitly known. Elsewhere (Allo, 2013), the first author proposed the following interpretation of each type of group knowledge: (a) distributed knowledge is implicitly available knowledge; (b) particular knowledge is explicitly available knowledge; (c) general knowledge is readily available (explicit) knowledge; and (d) common knowledge is transparently available (explicit) knowledge. This interpretation agrees with the logical properties of each type of group-knowledge as summarised in Table 17.1.

The static properties of each type of group-knowledge already give us the ability to discriminate between different ways in which something can be known or accepted within a community, but do not yet exhaust the possibilities. Consider for that matter the following principles

$$(S_G\varphi \wedge S_G(\varphi \rightarrow \psi)) \rightarrow D_G\psi$$

$$(S_G\varphi \wedge E_G(\varphi \rightarrow \psi)) \rightarrow S_G\psi$$

which express weaker types of deductive closure for particular knowledge. From the first principle it follows that the consequences of all particular knowledge are distributed knowledge. In other words, any deductive consequences of what is explicitly known are implicitly known, and to make merely implicit knowledge explicit the different agents will have to communicate.

Extensions of the standard epistemic logic known as dynamic epistemic logic (Baltag and Moss, 2004; Baltag et al., 2008; van Ditmarsch et al., 2007) bring

in the required resources to express facts about what is known after a certain announcement. Crucially, different types of announcements will alter the way in which knowledge is distributed in a different manner. The best known example is that of so-called public announcements, whose effect is modelled as the removal of all alternatives that disagree with the content of what is announced. Thus, after the public announcement of p to a group of agents G , the model will no longer contain any alternative where p is true, and hence no p -world will be a G^k -alternative (see Appendix). That is, p will be common knowledge in G .

A first alternative to public announcements that we need to consider are *fully private announcements*. These are announcements where one or more agents learn the content of what is announced, while the remaining agents are not even aware of the fact that something was announced. When, for instance, p is privately announced to an agent c while the remaining members of the group G do not notice this, the resulting model will be one where c knows that p while the remaining members will be under the false impression that the knowledge available in the group didn't change at all. Thus, for instance, if prior to the announcement it was commonly known that c didn't know whether p was the case, then after the announcement the remaining members will falsely believe that c is ignorant with respect to p .

A second alternative are the so-called *fair game announcements*, where some agents receive new information while the remaining members of the group do notice that some information is being shared with others, but remain ignorant with regard to the specific content that is being learned. That is, after such an announcement of p it could be the case that c knows that p while the remaining agents only know that c knows whether p (i.e. they do not know as much as c does, but they do know that the state of c 's knowledge has changed because they know that either c knows that p or c knows that $\neg p$).

Using these ideas, we can start to make further combinations. For instance, we can take the union of different fully private announcements to model the action where each member of a group G comes to know that p , and yet falsely believes that no other member of G comes to know that p . What is distinctive of such announcements is that after the announcement it is the case that p is generally known in G , but not commonly known in G . This type of result can be related to a well-known limitation about our means to achieve common knowledge in a group, namely the fact that common knowledge can only be attained through public announcements. That is, for p to become common knowledge in a group G , p must be publicly announced within G ; it must be transparent to all members of G that p is announced and that all members of G know about this announcement. If communication is not transparent, as is the case with the other types of announcements we've described (or as is the case with unreliable communication, as we find it in the coordinated attack problem; see, e.g., [Fagin et al., 1995](#)), then common knowledge is absolutely beyond reach.

17.4 A Model for Collaborative Mathematics?

The main reason why common knowledge is such a valuable good is that it is a prerequisite for coordinated action. This feature is traditionally associated with the coordinated attack problem: To win the battle, two generals need to ensure that they will attack the enemy at the same moment. The latter can only be achieved when it is common knowledge between them that, say, the plan is to attack at dawn. As the problem is traditionally set up, common knowledge and thus agreement can never be achieved by these generals. Since all their means of communication are unreliable (e.g. messengers that might be captured by the enemy), the generals can always doubt that the message was delivered and hence that an agreement was reached. This situation is analogous to a situation where information is exchanged with (semi-) private announcements.

Let $agree(\psi)$ be a formula that is true at states in which the players have agreed on ψ . (...) [W]e expect that if Alice and Bob agree on ψ , then each of them knows that they have agreed on ψ . This is a key property of agreement: in order for there to be agreement, every participant in the agreement must know that there is agreement. Thus, we expect $agree(\psi) \implies E(agree(\psi))$ to be valid. The Induction Rule for Common knowledge tells us that if this is the case, then $agree(\psi) \implies C(agree(\psi))$ is also valid. Hence, agreement implies common knowledge (Fagin et al., 1995, 189–90).

Assuming that collaborative mathematics requires coordinated action, one would presume that common knowledge would be as important for the success of the Polymath Project as it is for the victory of the two Byzantine generals. And hence, given the previously established connection with public announcements, one would expect that all knowledge should be shared in an entirely transparent manner. At first blush it would seem that this is exactly what is the case in the example from Sect. 17.2: The totality of what is communicated is publicly available on a website. And yet, if that were the case the whole case-study would turn out to be a rather boring example of pooling together resources (both in terms of computational capacity and in terms of mathematical knowledge). To reveal what makes the Polymath case an interesting one, we precisely need to focus on how the absence of common knowledge guarantees that the community remains sufficiently diverse, while the process of limiting this diversity ensures that the whole enterprise remains sufficiently goal-oriented to ultimately find a solution and end up with a proof.

When put in terms of announcements, the above description suggests that we need to look at the whole communication-process on the website as a mixture of both public and not so public announcements; with the public announcements enforcing a certain amount of coordination, and with the private and semi-private announcements to maintain a certain degree of diversity. Our aim in this section is to show that, first, the structure of the polymath community—with a core of two agents, and a highly diverse periphery—is indeed well-suited to maintain both diversity and a goal-oriented enterprise, and, second, to show that the structure of this type of community can be modelled with the tools described in the previous section.

Let us begin with a few methodological remarks about formal modelling. First, to design a formal model of a given phenomenon we need to decide on which features we're interested in, and thus need to be retained by the model, and which features we would like to abstract away. Here, the upshot is to focus on how information is distributed, and how that distribution can be altered by means of communication. This will mean that our model will abstract from many features of the case-study from Sect. 17.2, and in particular that a fine-grained rendering of the content of the specific mathematical problem will be absent from our model. As a result, the formal model that is further developed in this section ignores **Feature 1**, the fact that the polymath project dealt with a combinatorial problem. Such abstraction is typical of formal models of science.

A second type of abstraction follows from the fact that our formal model assumes all agents to be logically omniscient. This is a common type of idealisation in logical models. It entails that the model cannot reveal specific features that are related to the resources that are required to complete certain computations. As a result, the formal model we use will also ignore aspects that are related to the fact that by delegating the task of checking a number of limited cases computational resources can be shared (i.e. if all agents have unlimited resources, there's no point in delegating purely computational tasks).

Given these abstractions, the model we're after is primarily a model of the structure of a community understood in terms of how information can flow within that community. How can, given the formalism we've settled on, the latter be modelled? Here, we will need to make a few assumptions that strictly speaking go beyond the formalism described in Sect. 17.3. This is because the framework of dynamic epistemic logic lacks the expressive resources to make all the relevant features of this structure explicit. A *first* requirement is that we need to tie announcements to specific agents who make that announcement. To that end, we follow the standard practice of modelling the announcement of φ by some agent c as the announcement that c knows (or believes) that φ . This introduces a first type of distinction between otherwise similar announcements.

The *second* and more important requirement is that we need to account for the structure of a community. This we achieve in two ways. By picking out a number of *distinguished sub-groups* of that community, and by listing (or otherwise defining) the types of announcements that are available. By restricting the number of sub-groups we can make claims about,¹² we settle on a certain level of abstraction.¹³ In doing so, we put a limit on how finely we can discriminate between different

¹²Again, this is something that the standard formalism does not allow for since the notions of general and particular knowledge are simply definable in terms of, respectively, the conjunction and disjunction of individual knowledge claims for each member of the group. Nevertheless, such restrictions on the available groups can be introduced syntactically with, what amounts to, awareness filters. We do not pursue the details of this additional modification, but stick to a rather informal approach.

¹³An intuitively plausible way to think about such levels of abstraction is this: Sometimes we like to consider the (or a) mathematical community as a monolithic bloc (in terms of consensus in that

factions in, say, the periphery of the community in terms of what is known and/or believed in such factions. A second effect of this type of expressive limitation is that it also precludes us from indiscriminately lumping together the knowledge and belief of arbitrary agents. For instance, while it obviously makes sense to talk about the knowledge that is distributed in the community as a whole, it isn't so obvious that for any two agents c_1 and c_2 we should equally likely be able to talk about the distributed knowledge (or lack thereof) in the two-agent group $\{c_1, c_2\}$.

The latter type of limitation on the types of claims we should be able to make about different sub-groups finds a natural companion in the restrictions we impose on the available announcements. For instance, if we are not allowed to make claims about the distributed knowledge in the group $\{c_1, c_2\}$, we should also not allow these two agents to privately communicate (for this would mean that they could in principle privately aggregate their knowledge). This is just one type of limitation on the available announcements. Another very natural type of limitation is on who is allowed to make public announcements. Here too, it often makes sense to assume that not all agents can make such announcements.

Depending on the structure of a community, that is, depending on the types of announcements that are available and that can alter the distribution of information in that community, the process of communication will determine the balance between coordination and diversity. For instance, only those agents who can make public announcements will be able to enforce the agreement required for coordination (and then only relative to their own information), while the other agents will at best be able to bring in new information and thereby enhance the diversity within the community. In the last part of this section we shall illustrate these processes by looking at three broad types of interaction. This will not only illustrate the connection between available actions and the balance between agreements and diversity, but will also shed a light on some distinctive features of the Polymath community.

17.4.1 *The Inference Network*

The main purpose of an inference network, conceived as a model of distributed computing, is to make the information that is implicitly present in the network explicit. In our terms, its aim is to move from merely distributed knowledge to (at least) particular knowledge. This can be achieved in many ways. One way to do so, is to allow every agent to send his information to a unique designated agent; a so-called *wise man*.¹⁴ In our terms, this would mean that the only available actions

community), but equally often we prefer to take a more refined look that allows us to focus on more local phenomena.

¹⁴This approach trivialises the idea of *distributed computing*, but the comments we make equally apply to more refined protocols for inference networks.

are announcements that have the wise man among their recipients. Crucially, this type of network is indifferent with respect to how such announcements are made; all that matters is that the wise man should come to know every piece of information in the network. This can equally well be achieved by private announcements as it can be by public announcements. Coordinated action is, at least in principle, not required for this network to succeed in its goal; it only has an impact on how efficient the network is. For instance, if the identity of the wise man is generally known, each agent will only need to send his information to that wise man. Likewise if the communication of φ between some agent c and the wise-man c_w is public within the sub-group $\{c, c_w\}$, then c will also know that the message that φ has arrived and thus prevent it from sending it again (though this won't prevent others from sending φ as well).

17.4.2 Contributing by Solving Predefined Subtasks

While the process of solving a big problem within a group by splitting it up into smaller tasks and distributing these tasks to different agents or sub-groups shares many features with an inference network, it also gives rise to further constraints on the structure of the community. To begin with, while the presence of a wise man isn't absolutely required for an inference network, the presence of some agent or sub-community who coordinates the process by distributing the different pieces is crucial here. In our terminology, this amounts to saying that we need bi-directional information-flow: to distribute the predefined tasks, and to send back the results for aggregation. As before, some types of coordination and common knowledge can be useful. We could, for instance, require that the overall aim be commonly known, that it be publicly announced when the final goal is reached, or that the process whereby tasks are distributed amounts to reaching an agreement that a certain task is accepted. Yet, other types of common knowledge like knowledge about the different subtasks, their distribution among the community, and the completion of certain subtasks need not be present. Given the wide range of situations where problems are solved in this manner, the question of what exactly needs to be commonly known cannot and should not be settled in general (e.g. sometimes we might prefer a situation where some sub-tasks are simultaneously tackled by competing agents, who may (or may not) suspect that others might have received the same task).

17.4.3 Balancing Between Coordination and Diversity

This is the type of network or community we associate with the Polymath Project. In Sect. 17.2, this network is described as “a network with two strongly connected *attractors*”. Here, we reflect this structure by describing the types of announcements that can be made, and by listing the types of knowledge claims that can be made.

First, to capture the fact that we have a core consisting of two agents, we should be able to refer to each of these agents separately, and indeed we should require that everyone in the community be aware of their identity. Second, to reflect that these two agents are strongly connected we should be able to make claims about what they know as a group (and how they know it). Similarly, to extend these considerations to the level of available announcements, we should also stipulate that these two agents can have private interactions, and that they both (as individuals, but perhaps also as a group) should be able to make public announcements to the community as a whole. In virtue of the fact that they can interact privately, the distribution of their knowledge can be altered unbeknownst to the community as a whole; in virtue of their ability to make public announcements, they have the certainty that when they address the community as a whole, the others will listen.

Third, to capture the fact that the rest of the community is best seen as a wide periphery, we may restrict our ability to refer to each agent separately, and we almost certainly need to limit their awareness of each other. Fourth, to reflect that all members of the community can contribute, that all information is made publicly available, and yet that it is by no means certain that all members of the community actually access each and every piece of available information, we should preclude that these agents be able to make public announcements. In fact, we should model the announcements of these agents in such a way that they do not know whether they actually reach the whole community, or only some subset of that community. Finally, as already mentioned, we might also want to preclude the agents in the periphery from interacting privately.

How does such a setting contribute to the intended aim of keeping a balance between diversity and coordination? This is achieved by the following division of labour: Agents in the periphery, on the one hand, can launch new ideas, and introduce new information into the network, but do not have the means to impose these ideas. This ensures that new ideas can at all time be introduced, and that most agents cannot decisively reject these new ideas. Agents at the core, by contrast, can (as they do by introducing the problem, or by giving summaries) set and revise the agenda, and this agenda can be considered binding. This is done by picking up and promoting certain ideas that were previously introduced by agents in the community, while objecting to (or merely ignoring) others. As such, the ability of the core members to make public announcements lets them, given an already available rich supply of new information, either stimulate or constrain this diversity.

17.5 Conclusion

In the field of philosophy there is a well-entrenched research tradition on argumentation, both from an informal and a formal point of view. In the latter case, this tradition has been strongly intertwined with the development of non-classical logics. Recently, from within the field of computer science, a novel approach to argumentation has been launched, namely that of argumentation networks. However,

to date, the models have remained rather abstract: arguments are viewed solely in relation to others, the basic scheme being that of one-on-one “attacks”. Also, as far as we are aware, applications to mathematical contexts are as yet unavailable. Correspondingly, the aim of this chapter has been to demonstrate and develop the potential of this type of approach for the understanding of the construction of mathematical proofs. In particular, we have illustrated how epistemic logics that allow us to reason about the knowledge of individuals and the knowledge available within communities can be used to model proofs-as-arguments from a perspective that emphasises interaction and collaboration. In particular cases like the Polymath Project, the absence of full common knowledge seems to guarantee that the community remains sufficiently diverse, while limits to this diversity ensure that the enterprise remains sufficiently goal-oriented.

Acknowledgements The first author is a postdoctoral fellow of the Research Foundation–Flanders, which through project G.0431.09 also supported research for this chapter by the third author.

Appendix

Where \mathcal{C} is the set of individual agents, we have a modal operator $[c]$ for every c in \mathcal{C} . Thus, we say that

$[c]\varphi$ is true at a state w iff wR_cw' implies that φ is true at w' .

If wR_cw' is read as saying that w' is an epistemic alternative to w for c , and that its negation $\neg wR_cw'$ means that at w , c can exclude w' , we can say that c knows that φ at w iff

c can exclude at w all states where φ is not true,

or, equivalently, iff

φ is true at all epistemic alternatives to w for c .

Traditionally, epistemic operators are presumed to satisfy some further conditions. For present purposes we opt for the strongest possible set of conditions: We presuppose that knowledge is factive (only truths can be known), that it is positively introspective (knowing implies knowing that one knows), and that it is negatively introspective (one always knows about one’s ignorance). It is a standard result in modal logic that $[c]$ has all these properties iff R_c is a reflexive, transitive and symmetric relation (see, for instance, [van Ditmarsch et al., 2007, 2.2](#)).

In addition to being able to talk about what individual agents know, we also want to say something about what is known in communities of agents. Crucially,

knowledge can be available to communities in different guises. Where $G \subseteq C$ is a community of agents and φ is any formula, φ can be known at a state w by a community G because:

1. At w every state w' where φ is not true is excluded by at least one $c \in G$. In that case, we say that φ is *distributed knowledge* at w in G and write $w \models D_G \varphi$ to express this.
2. At w some $c \in G$ excludes every state w' where φ is not true. In that case, we say that φ is *particular knowledge* at w in G (informally, ‘someone knows’) and write $w \models S_G \varphi$ to express this.
3. At w every $c \in G$ excludes every state w' where φ is not true. In that case, we say that φ is *general knowledge* at w in G (informally, ‘everybody knows’) and write $w \models E_G \varphi$ to express this.

To explain a final way in which φ can be known in G , we first define a G^k alternative with the following inductive clauses:

- A world w is a G^1 alternative iff w is an epistemic alternative for some member of G .
- A world w is a G^{k+1} alternative iff at some G^k alternative the world w is an epistemic alternative for some member of G .

Using this notion, we can now stipulate that φ can also be known when

4. At w and for any finite k , no state w' where φ is not true is a G^k alternative. In that case, we say that φ is *common knowledge* in G and write $w \models C_G \varphi$ to express this.

References

- Aberdein, A. (2005). The uses of argument in mathematics. *Argumentation*, 19(3), 287–301.
- Aberdein, A. (2007). The informal logic of mathematical proof. In J. P. Van Bendegem & B. Van Kerkhove (Eds.), *Perspectives on mathematical practices: Bringing together philosophy of mathematics, sociology of mathematics, and mathematics education* (pp. 135–151). Dordrecht: Springer.
- Allo, P. (2013). The many faces of closure and introspection. *Journal of Philosophical Logic*, 42(1), 91–124.
- Atkinson, K., Bench-Capon T., & McBurney, P. (2005). A dialogue game protocol for multi-agent argument over proposals for action. *Journal of Autonomous Agents and Multi-Agent Systems*, 11(2), 153–171.
- Aron, J. (2011). Maths can be better together. *New Scientist*, 210(2811), 10–11.
- Baltag, A., & Moss, L. S. (2004). Logics for epistemic programs. *Synthese*, 139(2), 165–224.
- Baltag, A., van Ditmarsch, H. P., & Moss, L. S. (2008). Epistemic logic and information update. In P. Adriaans & J. van Benthem (Eds.), *Handbook on the philosophy of information* (pp. 361–456). Amsterdam: Elsevier Science Publishers.
- Barth, E. M., & Krabbe, E. C. W. (1982). *From axiom to dialogue. A philosophical study of logics and argumentation*. Berlin: Walter de Gruyter.
- Bondarenko, A., Dung, P. M., Kowalski, R. A., & Toni, F. (1997). An abstract, argumentation-theoretic approach to default reasoning. *Artificial Intelligence*, 93(1–2), 63–101.

- Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology*. Oxford: Oxford University Press.
- Cranshaw, J., & Kittur, A. (2011). The Polymath project: Lessons from a successful online collaboration in mathematics. In *Proceedings of the conference on human factors in computing systems (CHI-11)* (pp. 1–10). Vancouver, BC.
- Dudley, U. (1992). *Mathematical cranks*. Washington, DC: MAA.
- Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77(2), 321–358.
- Furstenberg, H., & Katznelson, Y. (1991). A density version of the Hales-Jewett Theorem. *Journal d'Analyse Mathématique*, 57, 64–119.
- Gowers, T. (2009a). Why this particular problem? In Gowers's Weblog. Mathematics related discussions. <http://gowers.wordpress.com/2009/02/01/why-this-particular-problem/>. Cited 11 Oct 2011.
- Gowers, T. (2009b). Must an “explicitly defined” Banach space contain c_0 or ell_p ? In Gowers's Weblog. Mathematics related discussions. http://gowers.wordpress.com/2009/02/17/must-an-explicitly-defined-banach-space-contain-c_0-or-ell_p/. Cited 11 Oct 2011.
- Gowers, T., & Nielsen, M. (2009). Massively collaborative mathematics. *Nature*, 461, 879–881.
- Grossi, D. (2010). Doing argumentation theory in modal logic. *ILLC-Preprint*, 2009–24. Amsterdam: University of Amsterdam.
- Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning about knowledge*. Cambridge, MA: MIT Press.
- Hintikka, J. (1985). A spectrum of logics of questioning. *Philosophica*, 35, 135–150.
- Hodges, W. (1998). An editor recalls some hopeless papers. *The Bulletin of Symbolic Logic*, 4(1), 1–16.
- Jakobovits, H., & Vermeir, D. (1996). Contradiction in argumentation frameworks. In *Proceedings of the IPMU conference* (pp. 821–826). Granada.
- Jakobovits, H., & Vermeir, D. (1999a). Dialectic semantics for argumentation frameworks. In *Proceedings of the seventh international conference on artificial intelligence and law* (pp. 53–62). ACM: New York.
- Jakobovits, H., & Vermeir, D. (1999b). Robust semantics for argumentation frameworks. *Journal of Logic and Computation*, 6(2), 215–261.
- Jeffreys, B. R., Kelley, L. A., Sergot, M. J., Fox, J., & Sternberg, M. J. E. (2006). Capturing expert knowledge with argumentation: A case study in bioinformatics. *Bioinformatics*, 22, 924–933.
- Kakas, A. C., Mancarella, P., & Dung, P. M. (1994). The acceptability semantics for logic programs. In P. Van Hentenriek (Ed.), *Proceedings of the 11th international conference on logic programming* (pp. 504–519). Cambridge, MA: MIT Press.
- Kitcher, P. (1990). The division of cognitive labor. *The Journal of Philosophy*, 87(1), 5–22.
- Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery* (edited by J. Worrall & E. Zahar). Cambridge: Cambridge University Press.
- Lorenzen, P., & Lorenz, K. (1978). *Dialogische logik*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Nielsen, M. (2011). *Reinventing discovery: The new era of networked science*. Princeton, NJ: Princeton University Press.
- Pollock, J. (1994). Justification and defeat. *Artificial Intelligence*, 67, 377–407.
- Pólya, G. (1945). *How to solve it*. New York: Doubleday.
- Polymath, D. H. J. (2010). Density Hales-Jewett and Moser numbers. In I. Bárány & J. Solymosi (Eds.), *An irregular mind (Szemerédi is 70): Vol. 21. Bolyai society mathematical studies* (pp. 689–753). Berlin: Springer.
- Polymath, D. H. J. (2012). A new proof of the Density Hales-Jewett Theorem. *Annals of Mathematics*, 175(3), 1283–1327.
- Popper, K. R. (1968). Epistemology without a knowing subject. In B. Van Rotselaar & J. F. Staal (Eds.), *Logic, methodology and philosophy of science III* (pp. 333–373). Amsterdam: North-Holland. (Reprinted as Chapter III of Popper 1972)
- Popper, K. R. (1972). *Objective knowledge*. Oxford: Oxford University Press.

- Prakken, H., & Sartor, G. (1996). A dialectical model of assessing conflicting arguments in legal reasoning. *Artificial Intelligence and Law*, 4, 331–368.
- Prakken, H., & Vreeswijk, G. (2002). Logics for defeasible argumentation. In D. M. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic* (2nd ed., Vol. 4, pp. 219–318). Dordrecht: Kluwer.
- Rav, Y. (1999). Why do we prove theorems? *Philosophia Mathematica (III)*, 7(3), 5–41.
- Reed, C. A. (1998). Dialogue frames in agent communication. In *Proceedings of the 3rd international conference on multi agent systems (ICMAS-98)* (pp. 246–253). Washington, DC: IEEE.
- Toulmin, S. (1958). *The uses of argument*. Cambridge: Cambridge University Press.
- Van Bendegem, J. P. (1982). Pragmatics and mathematics or how do mathematicians talk? *Philosophica*, 29, 97–118.
- Van Bendegem, J. P. (1985a). Dialogue logic and problem-solving. *Philosophica*, 35, 113–134.
- Van Bendegem, J. P. (1985b). A connection between modal logic and dynamic logic in a problem solving community. In F. Vandamme & J. Hintikka (Eds.), *Logic of discourse and logic of discovery* (pp. 249–262). New York: Plenum Press.
- Van Bendegem, J. P., & Van Kerkhove, B. (2009). Mathematical arguments in context. *Foundations of Science*, 14(1–2), 45–57.
- Van Eemeren, F. H., Grootendorst, R., & Kruiger, T. (2004). *A systematic theory of argumentation: The pragma-dialectical approach*. Cambridge: Cambridge University Press.
- Van Ditmarsch, H., van der Hoek, W., & Kooi, B. (2007). *Dynamic epistemic logic*. Dordrecht: Springer.
- Verheij, B. (1995). Two approaches to dialectical argumentation: Admissible sets and argumentation stages. In J.-J. C. Meyer & L. C. van der Gaag (Eds.), *NAIC'96: Proceedings of the eighth Dutch conference on artificial intelligence* (pp. 357–368). Utrecht: Utrecht University.
- Vreeswijk, G. A. W. (1997). Abstract argumentation systems. *Artificial Intelligence*, 90(1–2), 225–279.
- Walton, D. N. (1998). *The new dialectic: Conversational contexts of argument*. Toronto, ON: University of Toronto Press.
- Woods, J. (2003). *Paradox and paraconsistency. Conflict resolution in the abstract sciences*. Cambridge: Cambridge University Press.
- Zollman, K. J. S. (2007). The communication structure of epistemic communities. *Philosophy of Science*, 74(5), 574–587.