



# Formalization of Mathematical Proof Practice Through an Argumentation-Based Model

Sofia Almpani<sup>1</sup> · Petros Stefaneas<sup>1</sup> · Ioannis Vandoulakis<sup>2,3</sup>

Received: 1 March 2023 / Accepted: 10 May 2023 / Published online: 8 June 2023  
© The Author(s) 2023

## Abstract

Proof requires a dialogue between agents to clarify obscure inference steps, fill gaps, or reveal implicit assumptions in a purported proof. Hence, argumentation is an integral component of the discovery process for mathematical proofs. This work presents how argumentation theories can be applied to describe specific informal features in the development of proof-events. The concept of proof-event was coined by Goguen who described mathematical proof as a public social event that takes place in space and time. This new meta-methodological concept is designed to cover not only “traditional” formal proofs but all kinds of proofs and inference steps, including incomplete or purported proofs. Our approach attempts to make proof-events more comprehensive to express the complete trajectory of a mathematical proof-event until the ultimate validation of the proving outcome. Thus, we advance an extended version of proof-event calculus which is built on argumentation theories designed to capture the internal and external structure of collaborative mathematical practice and highlight the relationship between proof, human reasoning, and cognitive processes. In addition, another area in which argumentation can make a significant contribution is dealing with the defeasible knowledge of the Web which is a product of its open and ubiquitous nature. This approach seems to be sufficient for the presentation of Web-based proving processes as manifested in the case of the Mini-Polymath 4 project.

---

✉ Sofia Almpani  
salmpani@mail.ntua.gr

Petros Stefaneas  
petros@math.ntua.gr

Ioannis Vandoulakis  
i.vandoulakis@gmail.com

<sup>1</sup> School of Electrical and Computer Engineering & School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Athens, Greece

<sup>2</sup> School of Applied Mathematical and Physical Sciences, National Technical University of Athens, Athens, Greece

<sup>3</sup> School of Humanities, Hellenic Open University, Patras, Greece

**Keywords** Argumentation · Proof-events · Proving · Informal logic · Mathematical practice

## 1 Introduction

Proof is the heart of mathematics. Mathematical cognition is commonly presented as a procedure that leads to “truth” by applying logical rules of inference. However, proofs are needed everywhere, not just in mathematics. They are also needed in physical sciences, computer science, philosophy, legal argumentation, political debates, and elsewhere. Proofs cannot always be adequately captured by traditional formal mathematical proofs (Stefaneas and Vandoulakis 2012), since mathematical discovery is a more complicated process with many obstacles and dead-ends that need to be dealt with. Our goal is to present a model of mathematical discovery that depicts the connection between formal mathematics and its informal social and cognitive aspects. Our emphasis in this paper is on the exchange of arguments and counterarguments during Web proving. Web proving can restructure the way we understand mathematical practices, since it can be used to facilitate proving as a multi-agent collective activity involving people from different backgrounds, expertise, and thinking styles (Vandoulakis 2020).

In philosophy, events can be considered as objects in time or instantiations of properties in objects. Theories of events have been proposed by Donald Davidson (1967) and John Lemmon (1967) using a causal criterion and a spatiotemporal criterion, respectively. To define an event, two events are the same, if and only if they have the same cause and effect, according to the causal criterion, and if and only if they occur in the same space at the same time, according to the spatiotemporal criterion. Goguen (2001) introduced the concept of proof-events that can describe mathematical proof as a public social event that takes place in space and time. This paper claims that proving practices can be expressed as a particular type of Goguen’s proof-events as presented in (Vandoulakis and Stefaneas 2015b). Proof-events have informal social and historical components, prover-interpreter interaction, collaboration, consent, and validation (Stefaneas and Vandoulakis 2012). Proof-events can be extended in the form of a dialogue between agents that use arguments and counterarguments to check the validity of the steps of a purported proof, by linking concepts from mathematical proving and argumentation reasoning (Sect. 2).

We present a model that indicates the relationship between proof-events and logic-based argumentation theories, combining them to study informal and formal aspects of proving more adequately (Sect. 3). Argumentation is a potent reasoning tool that allows contributors in the dialogue to argue and counterargue, assert and refute, validate and invalidate steps of mathematical reasoning that aim to solve a posed problem. This leads to a deeper understanding of the often contradictory visions, perspectives, and problem-solving strategies of the contributors to a proof that ultimately concludes in their agreement and consensus (Hunter 2007). This model, named *Argumentation-based Proof-Event Calculus (APEC)*, can be used in collaborative mathematical environments where argumentation among mathematicians can be used to motivate creativity and discovery or to elucidate obscure points of a purported proof. APEC

facilitates the formalization of interactive argument schemes to describe a mathematical problem-solving activity that faces eventual contradictions and dead-ends.

Our approach is novel because we use argumentation theory techniques *to build a bridge between a formal proof and the informal social interaction aspects involved in the search for proof*. Various researchers have shown that the role of argumentation is crucial in mathematics (Aberdein 2009; Hunter 2007; Alcolea Banegas 1998; Pedemonte 2008) by adapting argumentation models, such as Toulmin's (1958) model, and comparing them with the structural components of proof. However, there has been criticism that sometimes the argument structure of Toulmin's model does not take into account the exchange of ideas between participants and thereby the justification is partial and ambiguous (Pedemonte and Balacheff 2016). Our goal in this paper is to supplement the concept of proof-events with "argument moves" of the participants that either support or attack an assumption. This is done in the wider framework of proof-events that takes into consideration not only formally validated proofs, but also informal thinking that include trials, choice of strategies, and/or possible validation or rejection of parts of a purported proof by the agents.

Pedemonte and Balacheff (2016) suggested the so-called  $ck\epsilon$ -enriched Toulmin model that captures the internal characteristics of the argument-proof structure. Nevertheless, we also wanted to express the external procedures in the practice of various participants. The APEC system can *represent the complete information and sequence of steps in the evolution of mathematical practice* which is modeled in the form of logic-based dialogues (informal external procedures) with argument moves, temporal predicates, and validation levels of argumentation (Sect. 4). At the final stage, proof may be accepted as completed, i.e. as a valid formal proof understood and recognized as true by all relevant agents. This approach enables us to examine the interplay between proof, human reasoning, cognitive processes and creativity in the mathematicians' practice more deeply.

Several studies highlight the educational aspects of argumentation and proof (Pedemonte 2008, 2007; Stylianides 2007; Krummheuer 2015; Inglis et al. 2007; Knipping and Reid 2019) and student interaction in the classroom. Even though our model can also be implemented for concept-learning and problem-solving for the sake of students, in this paper we focus on *modeling a broader perspective of the collaborative discovery process in the practice of real mathematical communities*. This context can be applied to the communication between mathematicians in a research environment where collaboration between them is essential and can lead to significant results, such as the case of mathematical practice in crowd-sourcing collaborative environments (Sect. 5). Online dialogues can be used as a rich source of argumentation repositories as this is in its purest form and provides information on how argumentation works in real-life dialogues (Vassiliades et al. 2021). To demonstrate the presented approach, this paper formalizes the Mini-Polymath 4 project, which employs Web-based interactions as part of a proving practice (see the case study in Subsection 5.1).

Thus, the objectives of this paper are:

1. to examine from a social, scientific, and cognitive perspective the common nature of arguments and proof-events and to show the relationship between the process of advancing an argument and advancing a proof;

2. to develop an APEC model to represent the “proving” procedure with argument schemes, highlighting key elements such as agent contributions (argument moves), sequences of proof-events (temporal predicates), and validation progress (levels of argumentation);
3. to show the impact of the (possibly virtual) mathematical environment on the development of arguments to attain proof; and
4. to illustrate the usability of the proposed approach in the case of the Mini-Polymath 4. We choose this problem because it is an original, real, online collaborative project employing Web-based interactions as part of proving practices.

The paper is structured into six sections. Section 2 briefly describes the theoretical background of proof-events and argumentation. Section 3 surveys the perspective leading to the development of a model for the formalization of mathematical practice as a dialogical collaborative process. Section 4 outlines the APEC formalization of proof-events in terms of argumentation theory. Section 5 highlights the APEC model’s usability and expressiveness as an abstract theoretical framework with real-life dialogues from the Mini-Polymath 4 project. Section 6 concludes with the overview of this paper. Appendix 1 includes the detailed modeling of the dialogues from Mini-Polymath 4.

## 2 Theoretical Background

Discovery, creativity, communication, and systematization are some elements that proof serves in mathematics (Villiers 1990), but often proof is perceived mainly as a method for persuasion (Weber 2010). Goguen’s (2001) broader concept of *proof-event* or *proving* (understood as a public locatable and dateable social event concerning a communicated purported proof of a posed problem) is designed to embrace any proving activity, including faulty, vague, disputed, or incomplete proofs. Vandoulakis and Stefaneas (2015a) described proof-events as activities of a multi-agent system that incorporates the history of these activities in the form of sequences of proof-events. Our purpose is to bridge the gap between formal and informal mathematical procedures by constructing a model that is closely related to the way proving actually unfolds. The comparison between proof and argumentation is based on the perception that proof (including incomplete or even false proofs, valid or invalid inference steps, ideas, etc.) can be regarded as a specific kind of argumentative discourse in mathematics (Pedemonte 2007).

The concept of proof as a *dialogue* and an activity that agents engage in is explored in (Dutilh Novaes 2020), where a dialogical account of mathematical proof is advanced to produce explanatory persuasion. The author develops a triadic conceptual scheme, consisting of the “producer” (the prover), the “receiver” (sceptic), and the “explanation itself” (the proof). In this scheme, the sceptic is mostly considered “silenced” (Dutilh Novaes 2018, p. 91), while our intention was to focus equally on the moves and counterarguments generated by the other side in order to understand the whole mathematical practice more deeply and highlight the value of the opposite side in the proving process. Another study that focuses on the agents that produce the proofs

is presented in (Hamami and Morris 2020). In the approach adopted in this work, each mathematical step corresponds to a proof activity and the formal mathematical proof is a report of the corresponding proof activities. The plan of a mathematical proof is conceived as the plan of the agents who carried out the respective proof activity. Generally, the above-mentioned studies provide similar approaches of multi-agent mathematical discourse, but from a more philosophical perspective. We also attempt to provide a formal framework through a logic-based calculus to express the informal dialogues and the steps taken in mathematical practice.

Logic-based systems for examining and assessing arguments have been broadly applied, generating various formal methods for *argumentation-based reasoning* (Arieli and Straber 2015). Argumentation theories can be used as a natural method of modeling non-monotonic reasoning, properly expressing its defeasible nature. A starting point of this paper is Pollock's (1987; 1992) approach to logical argumentation, which presented one of the first non-monotonic logics with concepts of argument and defeat. He also introduced *defeasible reasoning* where arguments are conceptualized as chains of reasoning that may lead to a conclusion, whereas additional information may destroy the chain of reasoning. The formalization developed in the present work is mainly a sequence-based realisation (Modgil and Prakken 2013) of Dung's abstract argumentation framework (1995), applying Pollock's (1987) view of defeasible reasoning with the basic structure of Toulmin's model (1958) for the representation of an argument. The present paper aims to gain from, build on, and integrate the above approaches in a way that also uses insights from other works, such as Kakas and Michael (2016), providing an abstract, theoretical exploration of logical argumentation applied principally to mathematical proving.

Argumentation is a tool of cognition that can formalize the science of common sense reasoning (Almpani et al. 2022). Many researchers tried to show that the procedure by which mathematicians evaluate reasoning is similar to argumentation, for example by adapting Toulmin's (1958) argumentation model.<sup>1</sup> to mathematical examples. Aberdein (2009; 2013; 2021) highlighted the use of arguments in mathematical conversations and practices. Pedemonte (2008; 2016; 2007) implemented the *ck*-enriched Toulmin model to indicate connections between argumentation and proof. Götz Krummheuer (2015) introduced the analysis of collective argumentation and participation using Toulmin's theory for the development of an interaction theory of mathematics learning. Christine Knipping and David Reid (2019) built on Toulmin's theory to compare and describe global argumentation structures and local argumentation aiming for a deeper understanding of proving processes in the classroom. In (Inglis et al. 2007), the full Toulmin scheme is implemented through three different warrant-types to model a wider range of argumentation. Metaxas et al. (2016) presented methodologies to study the mathematical practice in a class involved in argumentative activities by integrating Toulmin's model and argumentation schemes.

Other approaches also indicate the connection between mathematical reasoning and argumentation. Eric Krabbe (2013) presented informal mathematical proofs as arguments, applying the "pragma-dialectical" theory enriched by the theory of strate-

---

<sup>1</sup> In Toulmin's model, an argument is constituted by six interrelated components: *claim*, *data*, *warrant*, *backing*, *rebuttal*, and *qualifier*. The first three elements are considered the substantial elements of applied arguments, whereas the last three are not always necessary.

gic maneuvering to identify the four stages of critical discussion in the proving process: confrontation, opening, argumentation, concluding stage (Krabbe 2008). Aberdein (2021) highlighted the connection of mathematical reasoning with tools developed by the informal logician Douglas Walton to express argumentation schemes as a taxonomy of argumentation steps, and dialogues as a contextualisation of formality through mathematical arguments. Lakatos' "*Proofs and Refutations*" (1976) is also an enduring classic that highlights the role of dialogue between agents (a teacher and some students) at proof attempts as well as critiques of these attempts. Deep disagreement (Aberdein 2021), controversy and acceptance in mathematical practice and philosophy can be also expressed through conversation in the concept of social constructionism (Ernest 2023). Studies in the discovery of mathematics have also used the concept of "collective argumentation" to particularly examine the mathematical characteristics of dialogues, as various mathematicians/agents work together to prove a claim (Conner et al. 2014).

Related studies that analyze *original mathematical dialogues* from the perspective of argumentation are:

- a. the so-called "mixed-initiative collaborative proving" in (Pease et al. 2017), a way of formalizing social aspects of proofs by interpreting the informal logic of a Lakatos-based mathematical discovery;
- b. the analysis of Mini-Polymath 3 by Alison Pease & Ursula Martin (2012); and
- c. the modeling of mathematical dialogues with the Inference Anchoring Theory + Content (IATC) framework by Corneli et al. (2019).

The approach in (Pease et al. 2017) implements many different predicates trying to provide a well-defined formal presentation. On the other hand, the study in (Pease and Martin 2012) uses a simple typology of comments categorized as concepts, examples, conjectures, or proofs, and it can be used mainly as a description of online collaborative mathematics rather than a formal representation. The work in (Corneli et al. 2019) uses predicates that are descriptive of the procedure and can be interpreted in widely different and subjective ways (it is not easy to define concepts such as, e.g., "helpful," "beautiful," "goal," "strategy," etc.).

All these studies emphasize the relationship and continuity between reasoning in argumentation and in proving. Our approach attempts to elucidate this intrinsic relationship by modeling the argumentative dialogue between justifications and explanations offered by mathematicians during their proving activity. We choose a different approach through the more general meta-methodological framework, which involves the theory of proof-events that incorporates both proofs and arguments. We do not attempt to tag an interpretation or a description in the procedure steps, but to depict the complete proving practice and its social interactions as formally as possible. Furthermore, our approach highlights explicitly the *argument moves* that the agents implement, as well as the sequence of the steps, not only in a "temporal" manner (with the *temporal predicates*) but also in a "progressive" manner (with the *levels of argumentation*) until the ultimately validated or invalidated outcome.

### 3 Towards a Model for Formalizing Reasoning and Collaborative Proving

According to Lockhart (2008), “*Mathematics is not in the ‘truth’ but in the explanation, the argument.*” Argumentation-based proof-events can be used to advance mathematical dialogue in which all participants collaborate to critically examine posed problems and enhance thinking abilities such as problem solving, interpretation, persuasion, and creativity. Hence, the goal of mathematics interaction is no longer the cultivation of individual problem-solving skills, but the development of “collaborative problem-solving capacities” (Vandoulakis 2020).

Some of the questions that we address are:

1. *The relationship between informal proving and formal proof in real mathematical practice and communication.*
2. *The relationship between argumentative and mathematical proving activities.*
3. *The relationship between the contributions of working mathematicians and a mathematical proof as a final output.*

Cognitive science has shown (Mercier and Sperber 2011) that the dialogical nature of argumentation is similar to human reasoning in proving. Humans conducting reasoning do not necessarily follow the rules of “logic” (Kakas et al. 1992). They may change their mind concerning a previous conclusion on a matter if they are confronted with additional information. Their knowledge can be incomplete and inconsistent and, therefore, new data can invalidate any conclusions drawn (Kakas and Michael 2016). However, *it is often the case that a proof output presented in its pure form overshadows the informal and social aspects of the proving process that led to it* (Hanna 2014). Argumentation-based approaches can help their integration within wider frameworks of human reasoning — such as dialogue, debate, validation, and proving — especially in dynamic environments such as real-life mathematical environments. The model can depict the dialogues between prover(s) and interpreter(s) in a multi-agent system, expressing both the internal structure of their arguments, and therefore their cognitive thinking, as well as the external social interactions with the argumentation moves.

*We investigate mathematical proof steps as argumentation-based proof-events to elucidate the creative characteristics of argumentation that are important in proving, such as negotiation, collaboration, and fruitful mistakes.* This approach enables us to examine the kinds of reasoning agents may use in their interaction and how the dialectical activity may influence them to generate new arguments as they move from the assumptions of a problem to its proof (Conner et al. 2014). Argumentation allows the contributors to engage in dialogue in the course of their problem-solving activity to test alternative proving strategies, check a suggested argument or idea or a (part of a) purported proof until they ultimately reach agreement (Hunter 2007). This perspective can reshape mathematical discovery into an interactive, negotiable, social process.

Although a new proof is usually attributed to the problem solver, the outcome is the joint efforts of different agents each of whom has different past experiences, background knowledge, proving skills, and vision of the problem (Trninic et al. 2018). Take, for instance, Fermat’s Last Theorem which mathematicians had been attempting to prove for over three centuries, until it was finally proved by Andrew Wiles (Almpani



et al. 2017) in 1994 (after 357 years). Thus, Wiles' proof was the outcome of the creative efforts of many generations of mathematicians and their suggested proofs, which sometimes contained deficiencies and flaws (Almpani and Stefanias 2017). *We suggest a model for mathematics learning, where problem-solving is viewed as a collaborative discovery proof-event* (Almpani 2022; Stefanias et al. 2015). The system represents, in different levels of proof-events, all the history of discovery and formalizes it in the form of collaborative argumentative contributions that includes trials, conflicts, and possible validation or termination of parts of purported proofs (Almpani et al. 2019). In the final step, the formal proof is checked, understood, and confirmed by the relevant mathematical community and recognized as valid.

*Why is a calculus for an argumentative model necessary?* There is a gap in the literature about tools that can provide formal — computationally explicit — input that can manage the variety of procedures normally involved in constructing proofs, especially when they contain informal mathematical dialogues with hypotheses, arguments, counterexamples, etc. (Corneli et al. 2019).

Our contribution concerns the elaboration of an analytic framework that provides a tool to describe and assess mathematical proving based on formal structure, agent contributions, argumentation reasoning, and sequence of arguments. The above features constitute different categories of argumentation frameworks (Clark and Sampson 2008) integrated in one framework. The developed calculus bases the foundations of the justification on a core structure of *premises-warrants-conclusion*. Then, it proceeds with the number and the kind of argument moves (supporting or attacking) necessary to build the different levels of argumentation. The levels of argumentation can progress from unjustified claims (lower levels) to incontrovertibly valid proofs (higher levels). Therefore, one can track the progress in creativity, rigor, and validity of arguments offered by mathematicians (who can be either human agents or intelligent software agents) and include the informal steps in a formal analytical framework.

This type of proof-theoretical approach applied in formal argumentation frameworks can have noticeable advantages (Kakas and Michael 2016). For instance, a well-studied argument-based calculus may be implemented for analyzing or generating arguments in a semi-automated or automated way (Kakas and Moraitis 2003), or combined with Web crowd-sourcing environments for creating human-machine hybrid teams (Pease et al. 2020).

#### 4 Argumentation-Based Proof-Event Calculus (APEC)

In this section, we present the model of the described approach, the Argumentation-based Proof-Event Calculus (APEC).

Comparison of the fundamental elements of argumentation theory and proof-events shows similarities in structure, sequence, and the agents (Almpani 2022).

1. Arguments and proof-events have three common fundamental components: a set of premises for a task or problem (i.e., data in arguments and premises in proof-events), a method of reasoning (i.e., warrant in arguments and inference rules in



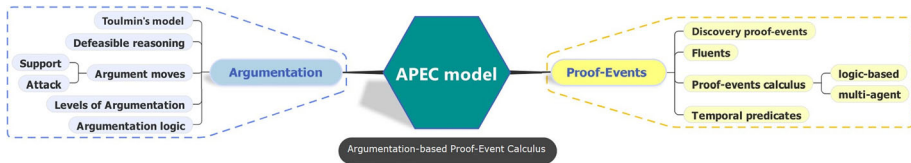


Fig. 1 APEC diagram

proof-events), and a conclusion (i.e., claim in arguments and conclusion in proof-events).

2. What is set to be proved emerges out of the history of proof-events, which is represented by sequences of proof-events (i.e., fluents (Stefaneas and Vandoulakis 2015)) or by sequences of arguments and counterarguments in the course of argumentation.
3. Argumentation involves agents or groups of agents, enacting the roles of supporter and opponent of an argument (Kakas and Moraitis 2003), enabling its adoption as a technology for multi-agent systems development. Similarly, proof-events necessitate the existence of at least two agents: a prover (the agent providing the proof) and an interpreter (the agent checking the validity of the proof) (Vandoulakis and Stefaneas 2015a).

Based on the above, our approach suggests that proving and arguing can be viewed as the functioning of a multi-agent system, the agents of which may take on different roles: the role of prover or supporter and/or the role of interpreter or opponent. These agents generate sequences of proof-events that incorporate the exchange of arguments and counterarguments between the agents participating in the proving. The concepts from argumentation and proof-event theories integrated into the APEC model are presented in Fig. 1.

The basic concepts and definitions of the APEC model are described through argumentation-based proof-events with argument moves, temporal predicates, and levels of argumentation as presented below.

**Definition 1** Argumentation-based proof-event

An argumentation-based proof-event  $e$  can be represented as a communicated argument  $\langle \Phi, c \rangle$  (Pollock 1992) designated by the pair  $e(\Phi, c)$  as  $e = \langle communicate(\Phi, c), w \rangle$ , where  $\Phi$  represents the premises of the argument based on the available *data*,  $c$  is the *claim* that refers to the conclusion of a particular problem communicated by the agent, and  $w$  are the inference rules or *warrant*<sup>2</sup> that leads from  $\Phi$  to  $c$ , so that<sup>3</sup>:

- $\Phi \not\vdash \perp$
- $\Phi \vdash c$

<sup>2</sup> Since in this work we draw an analogy between argumentation and proving, a *warrant* is an assumption that links the data to the claim, in the same way that inference rules link the premises to the conclusion in a mathematical proof. There can be different warrants leading to the same claim, as in mathematics there can be different inferences rules that lead to the same conclusion (proof), e.g., the different proofs of the Pythagorean Theorem.

<sup>3</sup> The symbol  $\vdash$  represents logical consequence (entailment), meaning that a valid logical argument is one in which the conclusion is a consequence of (entailed from) one or more premises.

– There is no  $\Phi' \subset \Phi$  such that  $\Phi' \vdash c$

where:

*claim*  $c$ : the statement/conclusion communicated by the agent,

*data*  $\Phi$ : the premises as the ground of the claim,

*warrant*  $w$ : the inference rules that connect the data to the claim.

*Counterarguments* are represented by the corresponding pair  $e^*\langle\Psi, \beta\rangle$ , where  $\Psi$  is the premises on which the claim  $\beta$  of the counterargument is based.

Argumentation may require chains or trees of reasoning, where claims are used in the assumptions to obtain further claims (Besnard and Hunter 2007), so that a proof-event could be an atomic argument or a sequence of arguments. Sequences of proof-events expressed with fluents in the calculus of proof-events (Stefaneas and Vandoulakis 2015) describe their temporal history and the interactions of the agents participating in the proof-event and, henceforth, they are useful for depicting logical arguments and counterarguments.

### Definition 2 Fluent of argumentation-based proof-events

A fluent  $f$  is a formula of the form  $e_1, e_2, \dots, e_n \rightarrow e$ ,  $n \in \mathbb{N}$ , where  $e_1\langle\Phi_1, c_1\rangle$ ,  $e_2\langle\Phi_2, c_2\rangle$ ,  $\dots$ ,  $e_n\langle\Phi_n, c_n\rangle$  is a finite, possibly empty, sequence of argumentation-based proof-events, where the conclusion of the proof-event  $e_i$  is the claim  $c_i$ , i.e.,  $concl(e_i) \equiv c_i$ , for some  $c_1, c_2, \dots, c_n \rightarrow c$  (Vreeswijk 1997). Accordingly, the meaning of the finite substantial components of the argument (Toulmin 1958) — which are abbreviated by corresponding prefixes — are defined as follows for the notion of a fluent:

claim:  $concl(e) = concl(e_1) \cap concl(e_2) \cap \dots \cap concl(e_n) \equiv c = c_1 \cap c_2 \cap \dots \cap c_n$   
 data:  $prem(e) = prem(e_1) \cup prem(e_2) \cup \dots \cup prem(e_n) \equiv \Phi_1 \cup \Phi_2 \cup \dots \cup \Phi_n$   
 warrant:  $infrul(e) = infrul(e_1) \cup infrul(e_2) \cup \dots \cup infrul(e_n) \equiv w = w_1 \cup w_2 \cup \dots \cup w_n$

A fluent contains all the necessary arguments/proving steps required to prove the desired conclusion. Therefore, the conclusion of the initial proof-event  $e$  can include the conclusions of the proof-events  $e_n$  contained in the fluent. For example, a proof may presuppose the proof of some of its subsections. Every contributing step in this procedure can be contained in a fluent. By this, we do not only mean completely correct steps, but also incomplete or faulty steps that can act as a starting point for another proving step.

However, the steps that an individual agent performs to accomplish a mathematical proof may overlap with the steps attempted or already performed by other agents. In the course of a proof, there can be various inference steps, such as attempts, impasses, confirmed or unconfirmed steps, false suggestions or implicit assumptions, intuitive ideas, intentions, etc. In this paper, the term *argument moves* is reserved for specific, active tactics or strategies among which a prover can choose to support their claim or attack an opponent's claim. Five fundamental relations are used that indicate links and conflicts in a sequence of proof-events. The argument moves — performed in a sequence of proof-events — can provide *support* (equivalent, elaboration) or *attack* (rebutting, undercutting, undermining) to the claim.

Given a claim  $c$  and an argument communicated during the proof-event  $e$ , possible argument moves, which provide support for  $c$  (Haggith 1996) or attack (Pollock 1992) are:

1. *Equivalent*( $e_1, e_2$ ): A proof-event  $e_1$  is equivalent<sup>4</sup> to a proof-event  $e_2$  for a claim  $c$  (i.e.,  $c_1 = c_2 = c$ ).
2. *Elaboration*( $e, S$ ): A proof-event  $e$  can have a set of inference rules  $S$  which elaborate or embellish upon  $c$ .
3. *Rebutting*( $e^*, e$ ): A counterargument  $e^*$  which attacks a claim  $c$  of  $e$ .
4. *Undermining*( $e^*, e$ ): A counterargument  $e^*$  which attacks a premise of  $e$ .
5. *Undercutting*( $e^*, e$ ): A counterargument  $e^*$  which attacks a warrant of  $e$ .

*Argument moves that support the claim:* A proof-event  $e\langle\Phi, c\rangle$  is equivalent to a proof-event  $e'\langle\Phi', c'\rangle$ , whenever it has the same premises and the same conclusion (although they may have different warrants). Thus, equivalent proof-events can have different ways of proving. For instance, numerous proofs have been offered for the Pythagorean Theorem, including a geometrical proof by Euclid and an algebraic proof by James Abram Garfield. Thus,

$$\textit{Equivalent}(e, e') : e\langle\Phi, c\rangle = e'\langle\Phi', c'\rangle,$$

when  $\Phi = \Phi', c = c'$  (and it might be  $w \neq w'$ ).

A proof-event  $e\langle\Phi, c\rangle$ , can have a set of inference rules  $S$  of  $e'$  which elaborate or embellish upon  $e$ , if and only if  $\Phi \cup S \vdash c$ . Thus,

$$\textit{Elaboration}(e, e') : \textit{sent}(e) \cap \textit{sent}(S_{e'}) \rightarrow \textit{concl}(e). \tag{1}$$

These moves are used for backing our claim and supporting our proof, therefore:

$$\textit{support}(e, e') \rightarrow \textit{Equivalent}(e, e') \cup \textit{Elaboration}(e, e').$$

*Counterargument moves that attack the claim:* A counterargument communicated during the proof-event  $e^*\langle\Phi, \beta\rangle$  rebuts (attacks) the conclusion of an argument communicated during the proof-event  $e\langle\Phi, c\rangle$ , if and only if  $\vdash \beta \leftrightarrow \neg c$ . Thus,

$$\textit{Rebutting}(e^*, e) : \textit{rebut}(e^*, e) \rightarrow \neg\textit{concl}(e).$$

A counterargument communicated during the proof-event  $e^*\langle\Phi, \beta\rangle$  undermines (attacks) some of the premises (defeasible inference) of the argument communicated during the proof-event  $e\langle\Phi, c\rangle$ , if and only if  $\vdash \beta \leftrightarrow \neg(\cap W_i)$ , for some  $w_1, \dots, w_n \subset W$ . Thus,

$$\textit{Undermining}(e^*, e) : \textit{undermin}(e^*, e) \rightarrow \neg\textit{prem}(e).$$

---

<sup>4</sup> In the equivalent argumentation move, both proof-events  $e_1$  and  $e_2$  can be supported, since being equivalent means that they support the same conclusion in a problem based on the same data with different warrants. Therefore, they strengthen the validity of the conclusion. However, for convenience, we consider that  $e_2$  proof-event supports  $e_1$  (since it appears “second”).

A counterargument communicated during the proof-event  $e^*(\Phi, \beta)$  undercuts (attacks) some of the inference rules (defeasible inference) of the argument communicated during the proof-event  $e(\Phi, c)$ , if and only if  $\vdash \beta \leftrightarrow \neg(\bigcap \Phi_i)$ , for some  $\Phi_1, \dots, \Phi_n \subset \Phi$ . Thus,

$$\text{Undercutting}(e^*, e) : \text{undercut}(e^*, e) \rightarrow \neg \text{infRul}(e).$$

Given an argument communicated during the proof-event  $e(\Phi, c)$ , a counterargument communicated during the proof-event  $e^*(\Phi, \beta)$  attacks the argument communicated during the proof-event  $e$ , if and only if  $e^*$  rebuts  $e$  or  $e^*$  undercuts  $e$ . Therefore:

$$\text{attack}(e^*, e) \rightarrow \text{Rebutting}(e^*, e) \cup \text{Undercutting}(e^*, e) \cup \text{Undermining}(e^*, e)$$

The above-mentioned operators are combined with *temporal predicates*: *Happens* ( $e, t$ ), *Initiates* ( $e, f, t$ ), *Terminates* ( $e, f, t$ ), *ActiveAt* ( $e, f, t$ ), *Clipped* ( $e, f, t$ ). The purpose of using the language of event calculus in describing proof-events is to express the progress of the fluents in combination with the exchange of arguments between the agents. Hence, they are formalized as in the following relations (Almpani 2022).

*Happens* ( $e, t_0$ ), which means that a proof-event  $e$  occurs at time  $t_0$ .

*Initiates* ( $e, f, t_1$ ) :  $\text{happens}(e, t_0) \rightarrow \neg \text{attack}(e^*, t_1) \cup \text{support}(e, t_1)$ , at time  $t_1$ ,

which means that, if a proof-event  $e$  occurs at time  $t_0$ , then there are no counterarguments  $e^*$  that attack the validity of the outcome of the proof-event and there is adequate support for our claim at the specific time  $t_1$ .

$$\begin{aligned} \text{Clipped}(e_1, f, t_2) : & \exists e_1, e_1^*, t_1, t_2, t [\text{Happens}(e, t_1) \cap (t_1 \leq t < t_2) \cap \text{attack}(e_1^*, t)] \\ & \cap [\nexists e_2 (\text{Happens}(e_2, t_2) \rightarrow \neg \text{attack}(e_1^*, t))], \text{ for } t_1 \leq t < t_2 \end{aligned}$$

which means that a proof-event clips when there is a terminating proof-event  $e_1^*$  between  $t_1$  and  $t_2$  and there is no proof-event  $e_2$  that attacks the counterargument  $e_1^*$  attacking the proof-event  $e_1$ .

$$\begin{aligned} \text{Terminates}(e, f, t_1) : & \exists e, e^*, t_1 ([\text{attack}(e^*, t_1) \rightarrow \neg \text{conc}(e) \cup \neg \text{prem}(e) \cup \neg \text{sent}(e)] \\ & \cap [\nexists e_2, t_2 (\text{Happens}(e_2, t_2) \rightarrow \neg \text{attack}(e^*, t_1))], \text{ with } t_1 < t_2 \end{aligned}$$

which means that a fluent terminates when there is a counterargument attacking the sequence and there is no proof-event  $e_2$  that *Happens* in time  $t_2$ , with  $t_1 < t_2$ , to defend the claim. The termination of a sequence of proof-events may be caused by the indication of the falsity of the problem (there are counterarguments that attack the conclusion of the proof-event), or the undecidability of the problem (there is a lack

of adequate warrants to prove the desideratum), or the inefficiency of the required information (there is a lack of Premises).

$$\text{ActiveAt}(e, f, t_{n+1}) : \text{Happens}(e_{n+1}, t_{n+1}) \rightarrow \neg \text{attack}(e_n^*, t_n) \cup \text{support}(e_n^*, t_n),$$

for every  $n \in \mathbb{N}, t_{n+1} > t_n$

which means that a fluent is active, if there is an argument to support the claim for every counterargument attacking the claim. This means that for every counterargument  $e^*(\Psi_i, \beta_i), i = 1, \dots, n, n \in \mathbb{N}$ , there is a proof-event  $e_{n+1}(\Phi_{n+1}, c_{n+1})$ , which  $\text{Happens}(e_{n+1}, t_{n+1})$  and defeats the attack of the counterargument  $e_n^*(\Psi_n, \beta_n)$ , for  $t_{n+1} > t_n$ .

From the above-mentioned, it concludes that:

$$\text{Happens}(e, t_1) \cap \text{Initiates}(e, f, t_1) \cap (t_1 < t_2) \cap \neg \text{attack}(e^*, t_2) \rightarrow \text{ActiveAt}(e, f, t_2),$$

which means that a fluent remains active at time  $t_2$ , if a proof-event  $e$  has taken place at time  $t_1$ , with  $t_1 < t_2$  and has not been terminated at a time point between  $t_1$  and  $t_2$ .

Consequently,

$$\forall i \leq n [\text{ActiveAt}(e, f, t_i) \cap (t_i < t_n) \cap \neg \text{Terminates}(e, f, t_i)] \rightarrow \text{Valid}(e, f, t_n),$$

at time  $t_n, i = 1, \dots, n, n \in \mathbb{N}$

which means that a fluent could be considered valid at time  $t_n$ , if it is active and there is no counterarguments to terminate it at time  $t_i$  for every  $i = 1, \dots, n, n \in \mathbb{N}$ .

To sum up, the following is a list containing all the aforementioned APEC predicates that constitute the core syntax of APEC:

– *Structural Components:*

- *prem*( $e$ ): The premises  $\Phi$  of the proof-event  $e$ .
- *concl*( $e$ ): The claim  $c$  of the proof-event  $e$ .
- *inf Rul*( $e$ ): The warrant  $w$  of the proof-event  $e$ .

– *Argumentative Moves:*

- *Elaboration*( $e, e'$ ): Statement  $S$  of  $e'$  elaborates proof-event  $e$ .
- *Equivalent*( $e, e'$ ): Proof-event  $e$  is equivalent with proof-event  $e'$ .
- *Rebutting*( $e^*, e$ ): Proof-event  $e^*$  rebuts proof-event  $e$ .
- *Undercutting*( $e^*, e$ ): Proof-event  $e^*$  undercuts proof-event  $e$ .
- *Undermining*( $e^*, e$ ): Proof-event  $e^*$  undermines proof-event  $e$ .

– *Reasoning:*

- *Support*( $e, e'$ ): Statements  $e'$  that support  $e$ .
- *Attack*( $e^*, e$ ): Statements  $e^*$  that attack  $e$ .

– *Temporal Predicates:*

- *Happens*( $e, t$ ): Proof-event  $e$  starts to happen at time  $t$ .

- *Initiates*( $e, f, t$ ): The fluent  $f$  of  $e$  initiates at time  $t$ .
- *Clipped*( $e, f, t$ ): The fluent  $f$  of  $e$  clipped at time  $t$ .
- *Terminates*( $e, f, t$ ): The fluent  $f$  of  $e$  terminates from  $e^*$ .
- *ActiveAt*( $e, f, t$ ): The fluent  $f$  of  $e$  is active at time  $t$ .
- *Valid*( $e, f, t$ ): The fluent  $f$  of  $e$  is valid at time  $t$ .

To be able to present not only the temporal process but also the validation progress of the argumentation-based proof-events more clearly and explicitly, we integrated the approach in (Kakas et al. 1992). In this paper, the argumentation framework is built in terms of logic programming rules expressing a preponderance relation among the arguments, presenting *levels of argumentation*:

*Object-level arguments*, which represent the possible decisions or actions in a specific domain.

*First-level priority arguments*, which express justifications on the object-level arguments to resolve possible conflicts.

*Higher-order priority arguments*, which are used to deal with potential conflicts between priority arguments of the previous level until all conflicts are resolved and the outcome is considered valid.

The same levels can be applied in mathematical proofs to understand the history of proof-events, starting from the statement of a problem until its validation or rejection and including all the attempts and failures (Vandoulakis and Stefaneas 2015a).

## 5 Implementation of the APEC in Online Collaborative Mathematics

According to Vandoulakis (2020), the Web as a collaborative medium may transform the way we experience proving practices, as it allows for contribution by agents with different backgrounds, knowledge, skills, and styles of thinking. Web technologies have a specific semantic structure that links opinions and arguments in a dialogue based mainly on natural linguistic models of argumentation (i.e., models that perceive argumentation as a language activity) (Bex et al. 2013a). However, Web methods do not always reflect the semantic structure of mathematical argumentative aspects explicitly enough or in depth (Rahwan 2008). They often cannot capture different types of mathematical arguments and counterarguments and are presented with difficulties in finding and evaluating arguments and their relationships (Bex et al. 2013b). There is a need for new frameworks, tools, and systems engineered into the Web to encourage dialogue, facilitate multi-agent collaboration, and promote a new online collective thinking.

Our work attempts to add to this repository of Argument Web tools by providing a semantic calculus specialized in the reasoning that takes place in mathematical practices. We believe that the reasoning that takes place in mathematical dialogues described by a machine-processable and semantically-rich argumentative structure is important to the Web vision.

One of the difficulties in the investigation of mathematical practice is that there is limited knowledge of the real process of mathematical proving and of the interaction between mathematicians during proving (Pease et al. 2019). To study mathematical

proving, we need sufficient information to capture the real-life process of mathematical discovery, not only the final product of the proof communicated in the publications. This information should provide grounds for explanations of the mathematical discovery, historical facts about the efforts undertaken by the contributors (alone and in cooperation), and data about the shaping of views and attitudes to proving outcomes (Pease and Martin 2012). Unlike traditional modes of communication, one of the Web's key features — and one that facilitates mathematical practice — is its open and ubiquitous nature, since Web-based communication enables interaction in multi-agent systems (Vandoulakis 2020). In addition, Web-based interactivity enables collaborative problem-solving, through which proof for a particular problem is achieved through spontaneously generated and exchanged arguments and counterarguments. Therefore, a source of information that can provide evidence about the mathematical proving practice presents itself in the form of Web-based crowd-sourcing projects.

Data sets of online collaborative mathematical practice can provide us with original, rich, and valuable information about the real process of mathematical discovery (Pease et al. 2020). Online blogs and forums with informal mathematical dialogues, such as Polymath, Mini-Polymaths, MathOverflow, Tricky.org, Math.Stackexchange, etc., reveal some of the hidden aspects in the evolution of mathematical proving over a period of time (Pease et al. 2019). Crowd-sourcing is a procedure similar to open sourcing where the work may be undertaken by an individual or a crowd basis, raising the number of possible contributors-provers, thus possibly gaining a deeper vision of the problem. The use of the Web as a means of crowd-sourcing and collaborative search for proof (Vandoulakis 2020) dates back to projects such as Tatami and Kumo by Goguen (Goguen 1999), and Tricky and Polymath by Timothy Gowers (Stefaneas et al. 2015). Tatami is a Web-based cooperative software system that consists of a proof assistant (Vandoulakis 2020). Kumo is a proof assistant for first-order hidden logic, which also develops websites that document its proofs (Stefaneas and Vandoulakis 2012). Tricky involves the creation of a large repository of articles useful for mathematical problem solving with the aim of assisting in mathematical proving practice (Gowers and Nielsen 2009). In Polymath, a mathematical problem was formulated, and the entire mathematical community was invited to collaborate openly to suggest ideas, approaches, comments, and pieces of proof in order to find an alternative proof (Stefaneas et al. 2015).

Polymath projects can be considered as one of the first fully documented accounts of how a mathematical problem was solved (Gowers 2009). In Polymath, contributors were encouraged to view themselves as part of a collaborative team created *ad hoc* to solve a posed problem and share their ideas even if they were “obvious,” incomplete, or faulty, as others might be able to check and correct them and discard what is useless. This form of networked brainstorming allows for tapping the full potential of various and complementary mathematical skills of the participants, thus leading to better and quicker results (Stefaneas et al. 2015).

In the next section, we discuss how the resources of online collaborative mathematics can be applied to support formulating and answering questions about mathematical proving. The data set we use are excerpts from the comments<sup>5</sup> of the *Mini-Polymath*

<sup>5</sup> <https://polymathprojects.org/2012/07/12/minipolymath4-project-imo-2012-q3>.



4 project, which allow us to integrate the arguments exchanged into dialogues represented in proof-events sequences. Some of the questions we try to highlight with this case study are:

- What knowledge can we obtain from the dialogues of online crowd-sourcing projects?
- How can the study of these projects be used to understand mathematical practice? and
- How can we present them in a systematic, illustrative, and explanatory way?

Although the dataset is not extensive, it is sufficient for our model.

### 5.1 The APEC Formalization of Mini-Polymath 4

In the Mini-Polymath 4 project, the participants contribute to the solution of a problem from the 2012 International Mathematical Olympiad, titled “*The liar’s guessing game (LGG)*” (see also Appendix 1).

The APEC model formalizes mathematical practice based on four core contexts (also indicated by the corresponding colours as follows):

**Argumentation-based proof-events** and their structural components that can be linked to the relevant sentences of the participants’ discourse.

**Argument moves and reasoning** that indicate interactions between proof-events (and their agents accordingly).

**Temporal predicates** that indicate the progress of the practice over time and whether certain proof events are active or not.

**Levels of argumentation** that indicate the progress of the proof in terms of justification.

The first two contexts connect the formal modeling of the calculus with the informal elements of the agents’ discourse and activities, and the latter two designate the progress of the proving in terms of time and validation.

We aim to present the dialogue and exchange of arguments in which the contributors were engaged through the comments functionality of the Polymath Webpage by constructing an APEC model, focusing on the proving activity of the first part of the conclusion  $c_{LGG_1}$  from the LGG problem. The second part  $c_{LGG_2}$  of the LGG problem can be modeled similarly.

The course of arguments exchange in this argumentation-based proof-event sequence is illustrated in the flow chart in Fig. 2. In this illustration the orange circles depict the argumentation-based proof-events, where the central one concerning the proof of LGG is denoted as  $e_{LGG}$  ( $e_{LGG_1}$  and  $e_{LGG_2}$  are the two conclusions of LGG), while the rest of the proof-events are denoted as  $e_{\{number\}}$ , where the number is the numbering of the related Mini-Polymath comment. The arrows depict the flow of the proof-event sequence. Labels also indicate the argument moves (green labels), the temporal predicates (blue labels) and the levels of argumentation (black labels) in the corresponding part of the sequence.

**Object-level arguments:**

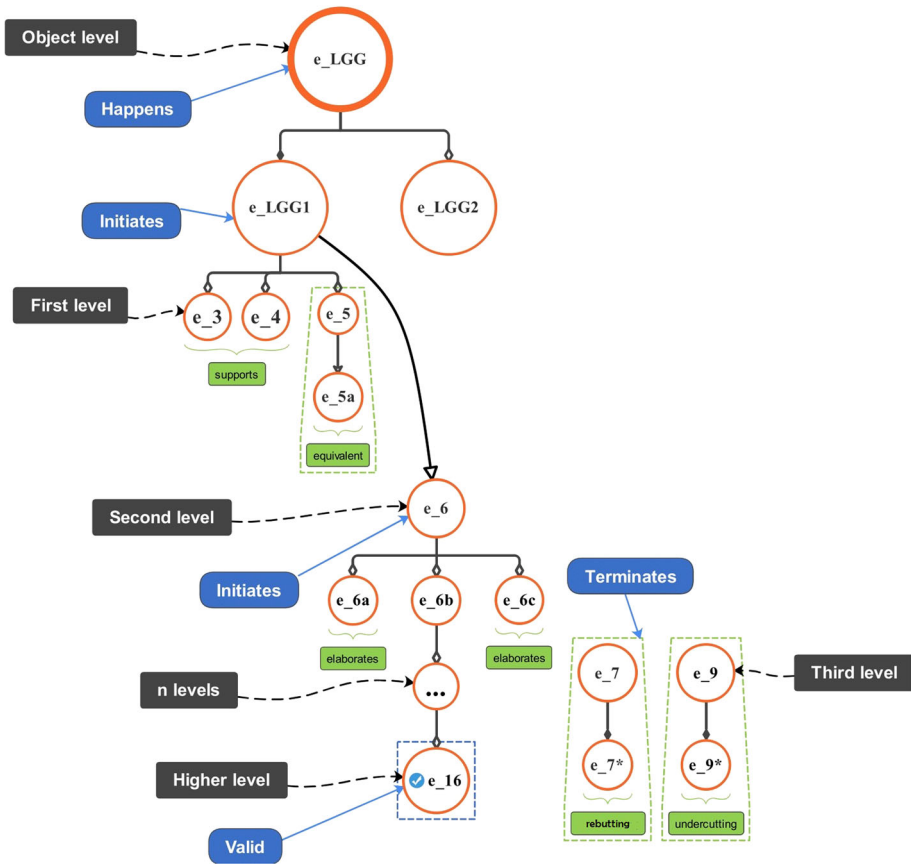


Fig. 2 Illustration of Mini-Polymath 4 through APEC model

In the object-level arguments, we have the possible initial available data and representations of arguments that can be used by the agents related to a specific domain problem that they attempt to address. Each agent may interpret and use this data differently, based on their personal perspective and background knowledge. In the use case presented, there is the LGG problem as the initial proof-event (*eLiarGuessingGame*) and two claims that need to be proved, so we have:

$$eLiarGuessingGame = eLGG_1\langle\Phi, c_1\rangle \cap eLGG_2\langle\Phi, c_2\rangle$$

where:

$\Phi = \langle$ The liar’s guessing game. $\rangle$

$c_1 = \langle$  If  $n \geq 2^k$  then B can guarantee a win. $\rangle$

$c_2 = \langle$  For all sufficiently large  $k$ , there exists an integer  $n \geq 1.99^k$  such that B cannot guarantee a win. $\rangle$

The Polymath aims to create the *warrant* of the aforementioned proof-events, as the result of collective fluents.<sup>6</sup> This initiates the proving:

$$Happens(e_{LGG}, f_0, t_1) \rightarrow Initiates(e_{LGG_1}, f_0, t_1) \cup Initiates(e_{LGG_2}, f_0, t_1)$$

**First-level and second-level priority arguments** included initial comments, attempts, and justifications of previous arguments that are not described in detail here (for the modeling of these levels see Appendix 1).

### Third-level Priority Arguments:

At this level, we have counterarguments and attacking moves on some comments and ideas of the previous levels. The proof-events are enumerated based on the numbering of the Polymath 4 comments.<sup>7</sup>

In some cases, a proof-event can be implied or assumed (correctly or faultily) from the available data, such as in the following example:

$$e_7 = \langle \Phi_7, c_7 \rangle = \langle \Phi_7: B \text{ cannot guarantee the win, } c_7: \text{ it can be "always win" for A} \rangle$$

(This proof-event is implied from the initial description of the problem.)

With counterargument  $e_7^*$ , the option that “player A can always win” was terminated.  
 $e_7^* = \langle \Phi_7^*, c_7^* \rangle = \langle \Phi_7^*: \text{ Since there is a possibility that B would win the game simply by guessing, } c_7^*: \text{ there is no "always win" for A} \rangle$

$Rebutting(e_7^*, e_7) : rebut(e_7^*, e_7) \rightarrow \neg concl(e_7)$  and

$attack(e_7^*, e_7) \rightarrow Rebutting(e_7^*, e_7)$ , where

$e_7^*$  attacks  $concl(e_7) = \langle \text{"always win" for A} \rangle$ .

$Terminates(e_7, f_3, t_{L_3}) \rightarrow attack(e_7^*, e_7)$

Argument  $e_8$  adds an observation on the warrant of  $e_{LGG_1}$ .

$$e_8 = \langle \Phi_8, c_8 \rangle = \langle \Phi_8: \text{ For the first part, proving for } n = 2^k \text{ suffices. The first approach that comes to my mind is to induct on } k, c_8: c_{LGG_1} \rangle \text{ with warrant } w_8 = \langle \text{proving for } n = 2^k \rangle.$$

With counterargument  $e_9^*$ , the related proof-event was attacked and terminated as unconstructive.

$e_9 = \langle \Phi_9, c_9 \rangle = \langle \Phi_9: B \text{ can as well ask questions in "rounds" of } k + 1 \text{ questions, } c_9: \text{ then, each round is guaranteed to have at least 1 correct answer} \rangle$

$e_9^* = \langle \Phi_9, c_9 \rangle$  with  $infRul(e_9^*) = \langle \text{While this is true, it is not very constructive [...] } \rangle$

$Undermining(e_9^*, e_9) : Undermin(e_9^*, e_9) \rightarrow \neg prem(e_9)$  and

$attack(e_9^*, e_9) \rightarrow Rebutting(e_9^*, e_9)$

Thus,

<sup>6</sup> At each level, the fluent is numbered with the corresponding level of argumentation, i.e., at first-level we have the fluent  $f_1$ .

<sup>7</sup> Another option is to number them by the agent's name (or both), depending on the information that we want to stress.

$Terminates(e_9, f_3, t_{L_3}) \rightarrow attack(e_9^*, e_9)$

**Fourth-level Priority arguments:**

At the fourth level, ideas and efforts yield some productive results thanks to fruitful (but not yet complete) cooperation, as the proving discovery progresses towards higher levels.

$e_{10} = \langle \Phi_{10}, c_{10} \rangle = \langle \Phi_{10}: \text{So for } k = 0 \text{ any version of binary search works, } c_{10}: \text{The next step should be to find the strategy for } k = 1, n = 2 \rangle$ ,  
 where  $\neg inf Rul(e_{10})$ , since the contributor claims “I first thought I have found the strategy, but it doesn’t work.”

Another prover named Mihai Nica elaborates in this proof-event with some useful lemmas that help proof-event  $e_{10}$  to progress, and finally the contribution of these comments adds a valuable input in the final proving of the first conclusion  $e_{LGG_1}$ .

$e_{10_a} = \langle \Phi_{10_a}, c_{10_a} \rangle = \langle \Phi_{10_a}: \text{I am working on this case too. Here player A can never tell two lies in a row. Here is a little observation I have made. Let } Q1, \text{ and } Q2 \text{ be questions that player B can ask, and I will use the notation like } [\dots], c_{10_a}: \text{Here is a cute little lemma: If B asks } Q1 \ Q2 \ Q1, \text{ then A must give the same answer for } Q1 \text{ both times it is asked, or else tell the truth for } Q2 \rangle$ .

$e_{10_b} = \langle \Phi_{10_b}, c_{10_b} \rangle = \langle \Phi_{10_b}: \text{Let } Q1, Q2 \text{ be questions. If player B asks the sequence of questions } Q1 \ Q2 \ Q1 \ Q1 \text{ and gets answers } A1 \ A2 \ A3 \ A4 \text{ (each } A_i \text{ is either an L (lie) or a T (truth)), } c_{10_b}: \text{By the last lemma for the sequence of questions } Q1 \ Q2 \ Q1, \text{ player B knows that either } A2=T \text{ or the answers to the first three questions are } LTL, TLT, \text{ or } TTT [\dots]. \text{ I think the second lemma can be used to make a binary search by making } Q1 = \text{half the numbers, } Q2 = \text{the other half of the numbers} \rangle$ .

$support(e_{10}, e_{10_a}) \rightarrow Elaborate(e_{10}, S_{e_{10_a}})$ ,

$support(e_{10}, e_{10_b}) \rightarrow Elaborate(e_{10}, S_{e_{10_b}})$

where:

$prem(S_{e_{10_a}}) = \langle \text{If B asks } Q1 \ Q2 \ Q1, \text{ then A must give the same answer for } Q1 \text{ both times it is asked } [\dots] \rangle$

$prem(S_{e_{10_b}}) = \langle \text{If player B asks the sequence of questions } [\dots] \text{ then player A is forced to reveal } [\dots] \rangle$

$Initiates(e_{10}, f_4, t_{L_4}) \rightarrow support(e_{10}, e_{10_a}) \cup support(e_{10}, e_{10_b})$

(continues for comments 11–15)

**Higher-order Priority Arguments:**

At the higher level, we have the justification and the proof of LGG’s first conclusion  $e_{LGG_1}$ , as the outcome of collective argumentation-based proof-events.

$e_{16} = \langle \Phi_{16}, c_{16} \rangle$ , where  $\Phi_{16} = prem(e_{16}) = \langle \text{We can assume } N = 2^k + 1, n = 2^k [\dots]. \text{ Then we can keep asking if } b_1 \text{ is 1, there are two possibilities.} \rangle$

$c_{16} = conc(e_{16}) = conc(e_{16_a}) \cup conc(e_{16_b})$ , where:

$conc(e_{16_a}) = \langle k + 1 \text{ times we get the answer NO, then we exclude the number } 10\dots 0 \rangle$

$conc(e_{16_b}) = \langle \text{There is a YES answer. Then we stop asking about } b_1 \text{ and ask } b_2 = 1, b_3 = 1, \dots, b_{k+1} = 1. \text{ After we are done we can exclude the number } [\dots] \rangle$

We have several proof-events in comment 16 that add in the proving discourse, either by supplementing claims of previous agents or by questioning some incomplete claims. We can see that it is a live procedure where each comment comes to fill a piece of the “proving puzzle” until its ultimate completion. In this proof puzzle, it often happens that even the attempt to add a “wrong piece” can contribute to the process, since something that does not work was tried, and it can now be safely excluded as an option.

$support(e_{16}, e_{16_a}) \rightarrow equivalent(e_{16}, e_{16_a})$ , where

$w_{16_a} = \langle \text{Another way (which seems to solve the first question). We ask the sequence of question } Q_i: \text{ “Does } b_i = 1? \text{” in a row. That makes } k + 1 \text{ questions } [\dots]. \text{ We have excluded a possibility, which by the reduction of comment 15 is enough.} \rangle$

$Rebutting(e_{16_b}^*, e_{16_a}) : rebut(e_{16_b}^*, e_{16_a}) \rightarrow \neg concl(e_{16_a})$ , where

$e_{16_b}^* = \langle \text{Which number will you exclude in that case? (It might not be in the range)} \rangle$ , and

$attack(e_{16_b}^*, e_{16_a}) \rightarrow Rebutting(e_{16_b}^*, e_{16_a})$

$ActiveAt(e_{16_a}, f_n, t_{L_n}) \rightarrow support(e_{16_a}, e_{16_c}) \cup support(e_{16_a}, e_{16_d})$ , with

$support(e_{16_a}, e_{16_c}) \rightarrow Elaborate(e_{16_a}, S_{e_{16_c}})$ , and

$support(e_{16_a}, e_{16_d}) \rightarrow Elaborate(e_{16_a}, S_{e_{16_d}})$ , where

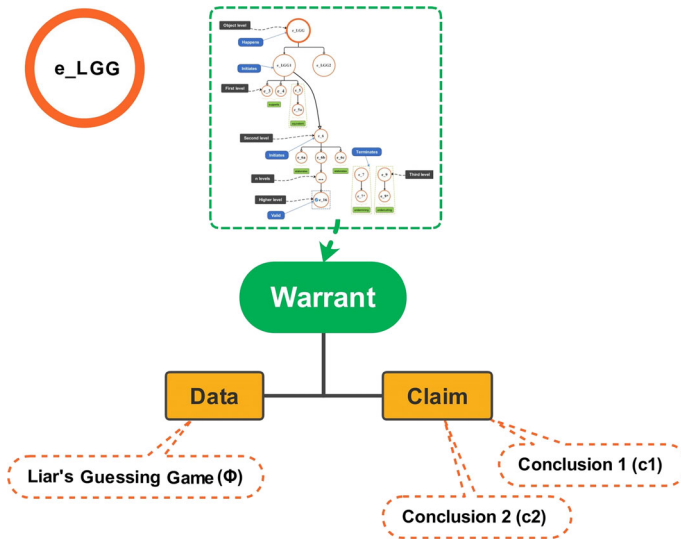
$e_{16_c} = \langle \Phi_{16_c}, c_{16_c} \rangle = \langle \Phi_{16_c}: \text{When } c_1 = 1, c_{16_c}: \text{then the number might be out of the range.} \rangle$

$e_{16_d} = \langle \Phi_{16_d}, c_{16_d} \rangle = \langle \Phi_{16_d}: \text{I’m not sure I totally understand your argument, but your argument lead me towards the following: Let } B_i \text{ be the subset of } \{0, \dots, N - 1\} \text{ with } 0 \text{ as the } i^{th} \text{ digit in their binary expansion } [\dots], c_{16_d}: \text{So } x \text{ cannot be } s_i \text{ and we have the required win. On the other hand, if } A \text{ always says that } x = s_i \text{ for any } i, [\dots] \text{ and } B \text{ wins.} \rangle$

$Valid(e_{LGG_1}, f_n, t_{L_n}) \rightarrow support(e_{16}, e_{16_c}) \cup support(e_{16}, e_{16_d}) \cap \neg attack(e_{16_b}^*, e_{16})$

The Mini-Polymath example illustrates the agents’ contribution in the process of proving. *The information we obtain from this type of project indicates that the characteristics and quality of dialogues can affect mathematical thinking and practice.* Firstly, the central aim of proving itself is to convince the rest of the community about the justification and the validity of one’s approach and outcomes. Moreover, all agents contributed significantly to the procedure, since various people had to participate in reaching their common goal, which was the proof of the LGG problem (in Fig. 3, the warrant is justified based on the contributions of all participants).

*Crowd-sourced mathematics is valuable in the study of mathematical practice,* revealing the way that mathematicians think and debate. Proving, at least at its inception phase, can be understood as an inquiry implemented by exchange of ideas: a collaborative dialogue between mathematicians with the common aim of solving an open problem, which none of the participants in the conversation has specifically pre-determined (Aberdein 2021). Such exchange of arguments can definitely be found in mathematics, at least in the context of mathematical discovery. At the end of the



**Fig. 3** Illustration of argumentation-based proof-event  $e_{LGG}$

sequence of proof-events, the agent who takes on the role of the administrator has an overview of the entire participation “history” of each prover-agent to the sequence of proof-events, so that he/she can analyse the overall contribution of each agent and integrate them into the final proof.

Additionally, in these types of collaborating environments, the participants might have less fear of committing mistakes, and therefore different solutions can be tried out and corrected. Argumentation is more efficient in interactive contexts as it permits counterarguments to be addressed and stronger arguments to surface, and tools such as the APEC model can provide considerable aid in this procedure. *It can be applied by provers and interpreters to identify and distinguish arguable elements on others' positions, but also on their own thinking.* The design and implementation of such learning environments can enhance the development of meta-cognitive activity and creativity in mathematics (Nadjafikhah et al. 2012).

To sum up, the historical road-map of proving in Mini-Polymath 4 can be experienced as a cooperative activity, connecting people with different backgrounds, perspectives, and interests. At each point of the proving trajectory our APEC framework illustrates the current state of the formal and informal reasoning in proving. This creates a link between:

- the informal and social aspect in the natural mathematical dialogue during the discovery of a proof; and
- the formal and computational aspect of abstract argumentation reasoning and semantics in a proof.

*The APEC framework adds an additional dimension and performs a significant role in making these connections sufficiently detailed in a systematic and explainable way,*

*demonstrating the applicability of argumentation techniques to mathematical proving and thinking.*

## 6 Conclusions

In this paper, we introduced the Argumentation-based Proof-Event Calculus (APEC) based on Toulmin's, Pollock's, Dung's, and Kakas' argumentation theories that extend the proof-event calculus with the integration of arguments. Our main aim was to develop a new approach in which theories and techniques from argumentation theory can be applied to bridge the gap between formal proving and human reasoning. We illustrated how proving and arguing, as processes, have many common aspects from a social and a cognitive point of view. We presented a calculus defining argumentation-based proof-events, argument moves, and temporal predicates and we analysed them in terms of levels of argumentation. This enables us to model conflicting arguments or unresolved moves, similarities and contradictions in multi-agent dialogues, the social collaboration between provers and interpreters, the controversy of previously accepted proofs, and so on. These aspects are often unseen or ignored in traditional mathematical models.

The original contribution of the present paper is that this calculus is formal, practical, and has the expressive power to represent real mathematical proving, as shown in the case of Mini-Polymath 4. We suggested a model for multi-agent proving, where problem-solving is implemented as a collaborative discovery proof-event. The model provided the analysis of the step-by-step components (argumentation-based proof-events) of mathematical practice, distinguishing the process of searching for proof (informal proving) from the final product of this process (formal proof). The combination of proof-events-based theory and logic-based argumentation makes it possible to dive into the micro-structure of the proving process, which allows us to also track the informal aspects of conveying mathematical information at all steps of proving, as well as in the external structure of proving, highlighting the social roles and the interactions of the contributors. Thus, Web-based mathematical problem-solving is strengthened by structured and semantically rich argumentative dialogues and can serve the development of collective mathematical knowledge through online multi-agent practices (as in the presented case of Mini-Polymath 4), transforming the way we understand mathematical proving.

**Funding** Open access funding provided by HEAL-Link Greece. This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included



in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## A The detailed dialogue and formalization of Mini-Polymath

“The liar’s guessing game is a game played between two players A and B. The rules of the game depend on two positive integers  $k$  and  $n$  which are known to both players. At the start of the game, A chooses two integers  $x$  and  $N$  with  $1 \leq x \leq N$ . Player A keeps  $x$  secret, and truthfully tells  $N$  to player B. Player B now tries to obtain information about  $x$  by asking player A questions as follows. Each question consists of B specifying an arbitrary set  $S$  of positive integers (possibly one specified in a previous question), and asking A whether  $x$  belongs to  $S$ . Player B may ask as many such questions as he wishes. After each question, player A must immediately answer it with yes or no, but is allowed to lie as many times as she wishes; the only restriction is that, among any  $k + 1$  consecutive answers, at least one answer must be truthful. After B has asked as many questions as he wants, he must specify a set  $X$  of at most  $n$  positive integers. If  $x$  belongs to  $X$ , then B wins; otherwise, he loses.”

### Object level arguments (the statement of the problem):

$e_{LiarGuessingGame} = e_{LGG1}(\Phi, c_1) \cap e_{LGG2}(\Phi, c_2)$ , where:

$\Phi = \langle \text{The liar's guessing game.} \rangle$

$c_1 = \langle \text{If } n \geq 2^k \text{ then B can guarantee a win.} \rangle$

$c_2 = \langle \text{For all sufficiently large } k, \text{ there exists an integer } n \geq 1.99^k \text{ such that B cannot guarantee a win.} \rangle$

$Happens(e_{LGG}, f_0, t_1) \rightarrow Initiates(e_{LGG}, f_0, t_1)$

### First-level priority arguments (First attempts):

$Support(e_{LGG1}, e_3) \rightarrow Elaboration(e_{LGG1}, e_3)$

$Support(e_{LGG1}, e_4) \rightarrow Elaboration(e_{LGG1}, e_4)$

$e_3 = \langle \Phi, c_1 \rangle$ , with

$infRul(e_3) = \langle \text{The fact that player A has to choose the number } N \text{ at the beginning of the game is intriguing. The number of possibilities for } x \text{ is originally } N, \text{ so it would seem like large } N \text{ would make the game harder for B. I suspect that B can counteract the difficulty by asking many more questions for large } N \text{ than small } N. \rangle$

$e_4 = \langle \Phi, c_1 \rangle$ , with

$infRul(e_4) = \langle \text{Ramsey Theory} \rangle$ .

### Second-level Arguments (Induction):

$e_6 = \langle \Phi, c_1 \rangle$ , with

$infRul(e_6) = \langle \text{Induction with respect to } N. \rangle$

$Initiates(e_6, f_2, t_6) \rightarrow support(e_6, t_{L_2})$

$support(e_6, t_{L_2}) \rightarrow Equiv(e_6, e_{12}) \cup Elab(e_6, Se_{6_a}) \cup Elab(e_6, Se_{6_c})$

$Elab(e_6, Se_{6_a}) = \langle \text{It seems to me that if we could ask a series of questions to guarantee that } x \text{ falls inside, say, } [0, N/2], \text{ then we could reduce to a previous case, but once we find such a series of questions we more or less have solved the problem.} \rangle$

$Elab(e_6, Se_{6c}) = \langle \text{It suffices to prove it for } N = n + 1. \text{ See comment 12 (i.e. } e_{12}). \rangle$

**Third-level Arguments (Guessing answers of B):**

$e_7 = \langle \Phi_7, c_7 \rangle = \langle \Phi_7: B \text{ cannot guarantee the win, } c_7: \text{ it can be “always win” for A} \rangle$   
 (this proof-event can be implied from the problem.)

$e_7^* = \langle \Phi_7^*, c_7^* \rangle = \langle \Phi_7^*: \text{ Since there is a possibility that B would win the game simply by guessing, } c_7^*: \text{ there is no “always win” for A} \rangle$

$Rebutting(e_7^*, e_7) : rebut(e_7^*, e_7) \rightarrow \neg concl(e_7)$  and

$attack(e_7^*, e_7) \rightarrow Rebutting(e_7^*, e_7)$ , where

$concl(e_7) = \langle \text{“always win” for A} \rangle$ .

$Terminates(e_7, f_3, t_{L3}) \rightarrow attack(e_7^*, e_7)$

Argument  $e_8$  adds an observation on the warrant of  $e_{LGG1}$ .

$e_8 = \langle \Phi_8, c_8 \rangle = \langle \Phi_8: \text{ For the first part, proving for } n = 2^k \text{ suffices. The first approach that comes to my mind is to induct on } k, c_8: c_{LGG1} \rangle$  with warrant  $w_8 = \langle \text{proving for } n = 2^k \rangle$

$e_9 = \langle \Phi_9, c_9 \rangle = \langle \Phi_9: B \text{ can as well ask questions in “rounds” of } k + 1 \text{ questions, } c_9: \text{ then, each round is guaranteed to have at least 1 correct answer} \rangle$

$e_9^* = \langle \Phi_9, c_9 \rangle$ , with  $infRul(e_9^*) = \langle \text{While this is true, it is not very constructive. Player A can just answer about half truth and half lies, making this strategy hard to implement.} \rangle$

$Undermining(e_9^*, e_9) : Undermin(e_9^*, e_9) \rightarrow \neg prem(e_9)$  and

$attack(e_9^*, e_9) \rightarrow rebut(e_9^*, e_9)$

Thus,  $Terminates(e_9, f_3, t_{L3}) \rightarrow attack(e_7^*, e_7)$

**Fourth-level arguments (proof for  $k = 1$ ):**

$e_{10} = \langle \Phi_{10}, c_{10} \rangle = \langle \Phi_{10}: \text{ So for } k = 0 \text{ any version of binary search works, } c_{10}: \text{ The next step should be to find the strategy for } k = 1, n = 2 \rangle$ , where

$\neg infRul(e_{10})$ , since the contributor claims “I first thought I have found the strategy, but it doesn’t work.”

$e_{10a} = \langle \Phi_{10a}, c_{10a} \rangle = \langle \Phi_{10a}: \text{ I am working on this case too. Here player A can never tell two lies in a row. Here is a little observation I have made. Let } Q_1, \text{ and } Q_2 \text{ be questions that player B can ask, and I will use the notation like:}$

Q’s:  $Q_1 Q_2 \dots$  A’s:  $L T \dots$  To denote that we asked  $Q_1$ , then  $Q_2$  and we received a lie and a truth respectively (of course, B doesn’t know which),

$c_{10a}$ : Here is a cute little lemma: If B asks  $Q_1 Q_2 Q_1$ , then A must give the same answer for  $Q_1$  both times it is asked, or else tell the truth for  $Q_2$ . Proof: There are 5 possible ways A can answer. LTL, LTT, TLT, TTL, TTT. From here we see that if the answers to  $Q_1$  are different, then the only possibilities are LTT and TTL, in either case the answer to  $Q_2$  must be true.)

$e_{10b} = \langle \Phi_{10b}, c_{10b} \rangle = \langle \Phi_{10b}: \text{ Let } Q_1, Q_2 \text{ be questions. If player B asks the sequence of questions } Q_1 Q_2 Q_1 \text{ and gets answers } A_1 A_2 A_3 A_4 \text{ (each } A_i \text{ is either an L (lie) or a T (truth)) } [ \dots ], c_{10b}: \text{ Then player A is forced to reveal one of the following pieces of information to player B. (i.e. player B will know which of them is true.): i) } A_2 = T, \text{ ii) } A_3 = A_4 = T, \text{ iii) } A_2 = A_4. \text{ By the last lemma for the sequence of questions } Q_1 Q_2 Q_1, \text{ player B knows that either } A_2 = T \text{ or the answers to the first three questions are LTL, TLT, or TTT } [ \dots ]. \text{ I think the second lemma can be used to make a binary search by making } Q_1 = \text{ half the numbers, } Q_2 = \text{ the other half of the numbers} \rangle$ .

$support(e_{10}, e_{10_a}) \rightarrow Elaborate(e_{10}, S_{e_{10_a}}),$

$support(e_{10}, e_{10_b}) \rightarrow Elaborate(e_{10}, S_{e_{10_b}})$

$prem(S_{e_{10_a}})=\langle$ If Player B asks the same question twice in a row and the answer is the same both times, then it must have been true both times)

$prem(S_{e_{10_b}})=\langle$ “Let Q1,Q2 be questions. If player B asks the sequence of questions Q1 Q2 Q1 Q1 and gets answers  $A_1 A_2 A_3 A_4$  (each  $A_i$  is either an L (lie) or a T (truth)). Then player A is forced to reveal one of the following pieces of information to player B. (i.e., player B will know which of them is true.”)

$Initiates(e_{10}, f_4, t_{L_4}) \rightarrow support(e_{10}, e_{10_a}) \cup support(e_{10}, e_{10_b})$

(continues for comments 11–15)

**Higher-lever Arguments (proof of LGG1):**

$e_{16} = \langle \Phi_{16}, c_{16} \rangle$ , where

$\Phi_{16} = prem(e_{16})=\langle$  We can assume  $N = 2^k + 1, n = 2^k$ . It means that  $x$  has at most  $k + 1$  binary digits ( $k + 1$  digits only for  $n = 2^k$ ):  $x = b_1 b_2 \dots b_{k+1}$ . Then we can keep asking if  $b_1$  is 1, there are two possibilities.

$c_{16} = conc(e_{16}) = conc(e_{16_a}) \cup conc(e_{16_b})$ , where

$conc(e_{16_a}) = \langle k + 1$  times we get the answer NO, then we exclude the number 10 ...0 )

$conc(e_{16_b}) = \langle$ There is a YES answer. Then we stop asking about  $b_1$  and ask  $b_2 = 1, b_3 = 1 \dots b_{k+1} = 1.$  ) After we are done we can exclude the number for which all the last  $k + 1$  answers would have been lies whose first digit is 0 (because of the YES answer.)

$support(e_{16}, e_{16_a}) \rightarrow equivalent(e_{16}, e_{16_a})$ , where

$e_{16_a} = \langle$ Another way (which seems to solve the first question). We ask the sequence of question  $Q_i$ : “Does  $b_i = 1$ ?” in a row. That makes  $k + 1$  questions. Then we must have at least one of the digits right. In particular, let  $y = c_1 \dots c_{k+1}$  be such that  $c_i = 0$  if the answer to  $A_i$  is Yes, and  $c_i = 1$  if the answer to  $A_i$  is No. Then  $x \neq y$ . We have excluded a possibility, which by the reduction of comment 15 is enough.)

$Rebutting(e_{16_b}^*, e_{16_a}) : rebut(e_{16_b}^*, e_{16_a}) \rightarrow \neg concl(e_{16_a})$ , where

$e_{16_b}^* = \langle$ Which number will you exclude in that case? (It might not be in the range)), and

$attack(e_{16_b}^*, e_{16_a}) \rightarrow Rebutting(e_{16_b}^*, e_{16_a})$

$ActiveAt(e_{16_a}, f_n, t_{L_n}) \rightarrow support(e_{16_a}, e_{16_c}) \cup support(e_{16_a}, e_{16_d})$ , with

$support(e_{16_a}, e_{16_c}) \rightarrow Elaborate(e_{16_a}, S_{e_{16_c}})$ , and

$support(e_{16_a}, e_{16_d}) \rightarrow Elaborate(e_{16_a}, S_{e_{10_d}})$ , where

$e_{16_c} = \langle$ When  $c_1 = 1$ , then the number might be out of the range.)

$e_{16_d} = \langle$ I’m not sure I totally understand your argument, but your argument lead me towards the following:

Let  $B_i$  be the subset of  $\{0, \dots, N - 1\}$  with 0 as the  $i^{th}$  digit in their binary expansion (note we’re leaving out one member).

Let B ask  $B_1, \dots, B_k$  in that order, and let  $b_i$  be 0 if A says yes to  $B_i$  and 1 else. Then let  $s_i$  be the number with binary expansion  $a_0 a_1 \dots a_i a'_{i+1} \dots a'_k$  where  $a'_j = 1 - a_j$ . Now ask  $\{s_0\}, \dots, \{s_k\}$  in order.

Suppose A answers at least once that  $x \neq s_i$ , and pick the first such instance of this. Then if  $x = s_i$ , A will have lied for the last  $k + 1$  questions, i.e.  $B_i, B_{i+1}, \dots, B_k, \{s_0\}, \dots, \{s_i\}$ . So  $x$  cannot be  $s_i$  and **we have the required win**. On the other hand, if A always says that  $x = s_i$  for any  $i$ , then if  $x$  was the one member we didn't manipulate, A lied  $k + 1$  times (all  $\{s_i\}$  questions). So if A says that  $x = s_i$  for all  $i$ , then the one member we didn't manipulate is actually not  $x$ , so we've discarded one member, and *B wins*.)

$Valid(e_{LGG_1}, f_n, t_{L_n}) \rightarrow support(e_{16}, e_{16_c}) \cup support(e_{16}, e_{16_d}) \cap \neg attack(e_{16}^*, e_{16})$

## References

- Aberdein A (2009) Mathematics and argumentation. *Found Sci* 14:1–8. <https://doi.org/10.1007/s10699-008-9158-3>
- Aberdein A (2021) Dialogue types, argumentation schemes, and mathematical practice: Douglas Walton and mathematics. *J Appl Logics* 8(1):159–182
- Aberdein A, Dove IJ (2013) The argument of Mathematics. *Logic, Epistemology, and the Unity of Science* 30. <https://doi.org/10.1007/978-94-007-6534-4>
- Alcolea Banegas J (1998) L'argumentació en matemàtiques. In: E. Casaban i Moya (ed.), XIIè Congrés Valencià de Filosofia. València, pp 135–147
- Almpani S (2022) Argumentation and rule-based logic in mathematical proving and legal artificial intelligence applications. [Doctoral dissertation, National Technical University of Athens, Greece, ND: 51572]. National Archive of PhD Theses, <https://doi.org/10.12681/eadd/51572>
- Almpani S, Stefaneas P (2017) On proving and argumentation. AIC 2017, 5th International Workshop on Artificial Intelligence and Cognition, Larnaka (2017)
- Almpani S, Stefaneas P, Vandoulakis I (2017) On the role of argumentation in discovery proof-events. C3GI 2017, 6th International Workshop on Computational Creativity, Concept Invention, and General Intelligence, Madrid
- Almpani S, Stefaneas P, Vandoulakis I (2019) On the significance of argumentation in discovery proof-events. In: 16th International Congress on Logic, Methodology and Philosophy of Science and Technology: Bridging Across Academic Cultures, Prague, 5-10 August 2019, Book of Abstracts
- Almpani S, Stefaneas P, Frangos P (2022) Argumentation-based logic for ethical decision making. *Studia Humana* 11:46–52. <https://doi.org/10.2478/sh-2022-0015>
- Arieli O, Straber C (2015) Sequent-based logical argumentation. *Argument Comput* 6:73–99. <https://doi.org/10.1080/19462166.2014.1002536>
- Besnard P, Hunter A (2007) Elements of Argumentation. The MIT Press. [https://doi.org/10.1007/978-3-540-75256-1\\_3](https://doi.org/10.1007/978-3-540-75256-1_3)
- Bex F, Lawrence J, Snaith M, Reed C (2013) Implementing the argument web. *Commun ACM* 56:66–73. <https://doi.org/10.1145/2500891>
- Bex F, Lawrence J, Snaith M, Reed C (2013) Implementing the argument web. *Commun ACM* 56:66–73. <https://doi.org/10.1145/2500891>
- Clark D, Sampson V (2008) Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *J Res Sci Teach* 45:293–321. <https://doi.org/10.1002/tea.20216>
- Conner A, Singletary LM, Smith RC, Wagner PA, Francisco RT (2014) Identifying kinds of reasoning in collective argumentation. *Math Think Learn* 16(3):181–200. <https://doi.org/10.1080/10986065.2014.921131>
- Corneli J, Martin U, Murray-Rust D, Nesin G, Pease A (2019) Argumentation theory for mathematical argument. *Argument J* 33:173–214. <https://doi.org/10.1007/s10503-018-9474-x>
- Davidson D (1967) The logical form of action sentences. In: Rescher N (ed) *The Logic of Decision and Action*, University of Pittsburgh Press, pp 81–95
- Dung PM (1995) On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artif Intell* 77:321–358
- Dutilh Novaes C (2018) A dialogical conception of explanation in mathematical proofs. Springer International Publishing, Cham, pp 81–98

- Dutilh Novaes C (2020) *The dialogical roots of deduction: Historical, cognitive, and philosophical perspectives on reasoning*. Cambridge University Press
- Ernest P (2023) Rejection, disagreement, controversy and acceptance in mathematical practice: episodes in the social construction of infinity. *Global Philosophy* 33(1):1–22. <https://doi.org/10.1007/s10516-023-09652-8>
- Goguen J (1999) Social and semiotic analyses for theorem prover user interface design. *Formal Aspects Comput* 11:272–301
- Goguen J (2001) What is a proof, (accessed 3 march 2022). <http://cseweb.ucsd.edu/~goguen/papers/proof.html>
- Gowers T (2009) Polymath1 and open collaborative mathematics. <http://gowers.wordpress.com/2009/03/10/>
- Gowers T, Nielsen M (2009) Massively collaborative mathematics. *Nature* 461:879–81. <https://doi.org/10.1038/461879a>
- Haggith M (1996) *A meta-level argumentation framework for representing and reasoning about disagreement*. PhD thesis, Dept. of Artificial Intelligence, University of Edinburgh
- Hamami Y, Morris RL (2020) Plans and planning in mathematical proofs. *Review of Symbolic Logic* pp 1–40
- Hanna G (2014) *Mathematical proof, argumentation, and reasoning*. Springer, Netherlands, Dordrecht, pp 404–408. [https://doi.org/10.1007/978-94-007-4978-8\\_102](https://doi.org/10.1007/978-94-007-4978-8_102)
- Hunter R (2007) Can you convince me: Learning to use mathematical argumentation. In: Woo J, Lew H, Park K, Seo D (Eds.), *In: Proceedings of the 31st conference of the International Group for the Psychology of Mathematics Education, PME, Seoul, vol 3*, pp 81–88
- Inglis M, Mejia Ramos J, Simpson A (2007) Modelling mathematical argumentation: the importance of qualification. *Educ Stud Math* 66:3–21. <https://doi.org/10.1007/s10649-006-9059-8>
- Kakas A, Michael L (2016) Cognitive systems: argument and cognition. *IEEE Intell Inform Bulletin* 17:14–20
- Kakas A, Moraitis P (2003) Argumentation based decision making for autonomous agents. In: *Proc. Second International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS, Melbourne, Australia*, pp 883–890
- Kakas A, Kowalski R, Toni F (1992) Abductive logic programming. *J Log Comput* 2:719–770. <https://doi.org/10.1093/logcom/2.6.719>
- Knipping C, Reid D (2019) Argumentation analysis for early career researchers. Kaiser G, Presmeg N (Eds), *Compendium for Early Career Researchers in Mathematics Education* pp 3–31, [https://doi.org/10.1007/978-3-030-15636-7\\_1](https://doi.org/10.1007/978-3-030-15636-7_1)
- Krabbe E (2008) Strategic maneuvering in mathematical proofs. *Argumentation* 22:453–468. <https://doi.org/10.1007/s10503-008-9098-7>
- Krabbe E (2013) *Arguments, Proofs, and Dialogues*. Springer, Dordrecht, pp 31–45
- Krummheuer G (2015) Methods for reconstructing processes of argumentation and participation in primary mathematics classroom interaction. In: Bikner-Ahsbahs A, Knipping C, Presmeg N (Eds), *Approaches to Qualitative Research in Mathematics Education* pp 51–74., [https://doi.org/10.1007/978-94-017-9181-6\\_3](https://doi.org/10.1007/978-94-017-9181-6_3)
- Lakatos I (1976) *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge University Press, Cambridge
- Lemmon EJ (1967) Comments on D. Davidson’s “The Logical Form of Action Sentences”, Pittsburgh: Pittsburgh University Press, pp 96–103
- Lockhart P (2008) A mathematician’s lament. [https://www.maa.org/external\\_archive/devlin/LockhartsLament.pdf](https://www.maa.org/external_archive/devlin/LockhartsLament.pdf), retrieved MAA Online from on 21 November, 2020
- Mercier H, Sperber D (2011) Why Do Humans Reason? arguments for an argumentative theory. *Behavioral and Brain Sciences* 34:57–74. <https://doi.org/10.1017/S0140525X10000968>
- Metaxas N, Potari D, Zachariades T (2016) Analysis of a teacher’s pedagogical arguments using Toulmin’s model and argumentation schemes. *Edu Stud Math*. <https://doi.org/10.1007/s10649-016-9701-z>
- Modgil S, Prakken H (2013) A general account of argumentation with preferences. *Artif Intell* 195:361–397. <https://doi.org/10.1016/j.artint.2012.10.008>
- Nadjařikhah M, Yafitian N, Bakhshalizadeh S (2012) Mathematical creativity: some definitions and characteristics. *Procedia-Soc Behavioral Sci* 31:285–291. <https://doi.org/10.1016/j.sbspro.2011.12.056>

- Pease A, Martin U (2012) Seventy four minutes of mathematics: an analysis of the third mini-polymath project. AISB/IACAP World Congress 2012: Symposium on Mathematical Practice and Cognition II. Part of Alan Turing Year 2012:19–29
- Pease A, Lawrence J, Budzynska K, Corneli J, Reed C (2017) Lakatos-style collaborative mathematics through dialectical, structured and abstract argumentation. *Artif Intell* 246:181–219. <https://doi.org/10.1016/j.artint.2017.02.006>
- Pease A, Aberdein A, Martin U (2019) Explanation in mathematical conversations: an empirical investigation. *Philosophical Transactions of the Royal Society A* 377:n.pag
- Pease A, Martin U, Tanswell F, Aberdein A (2020) Using crowdsourced mathematics to understand mathematical practice. *ZDM* 52:1087–1098. <https://doi.org/10.1007/s11858-020-01181-7>
- Pedemonte B (2007) How can the relationship between argumentation and proof be analysed? *Educ Stud Math* 66:23–41. <https://doi.org/10.1007/s10649-006-9057-x>
- Pedemonte B (2008) Argumentation and algebraic proof. *ZDM* 40:385–400
- Pedemonte B, Balacheff N (2016) Establishing links between conceptions, argumentation and proof through the ckg-enriched Toulmin model. *J Math Behav* 41:104–122. <https://doi.org/10.1016/j.jmathb.2015.10.008>
- Pollock JL (1987) Defeasible reasoning. *Cogn Sci* 11:481–518
- Pollock JL (1992) How to reason defeasibly. *Artif Intell* 57:1–42
- Rahwan I (2008) Mass argumentation and the semantic web. *Semantic Web* 6(1):29–37. <https://doi.org/10.1016/j.websem.2007.11.007>
- Stefaneas P (2012) Vandoulakis I (2012) The web as a tool for proving, *Metaphilosophy* 43(4):480–498. <https://doi.org/10.1111/j.1467-9973.2012.01758.x/abstract>
- Stefaneas P, Vandoulakis I (2015) On mathematical proving Computational Creativity Concept Invention and General Intelligence Issue. *J Artif General Intell* 6:130–149
- Stefaneas P, Vandoulakis I, Martinez M, Foundalis H (2015) Collective Discovery Events: Web-Based Mathematical Problem-Solving with Codelets. Atlantis Press, Paris, pp 371–389
- Stylianides A (2007) Introducing young children to the role of assumptions in proving. *Math Think Learn* 9:361–385. <https://doi.org/10.1080/10986060701533805>
- Toulmin SE (1958) *The uses of arguments*. Cambridge University Press
- Trninc D, Wagner R, Kapur M (2018) Rethinking failure in mathematics education: A historical appeal. *Thinking Skills and Creativity* 30:76–89. <https://doi.org/10.1016/j.tsc.2018.03.008>, the Role of Failure in Promoting Thinking Skills and Creativity
- Vandoulakis I (2020) Web-based collaboration: A prospective paradigm of mathematical learning. *Humanistic futures of learning: Perspectives from UNESCO Chairs and UNITWIN Networks* pp 161–163
- Vandoulakis I, Stefaneas P (2015a) Mathematical style as expression of the art of proving. In: *Handbook of the 5th World Congress and School on Universal Logic, Istanbul, Turkey, UniLog*
- Vandoulakis I, Stefaneas P (2015b) Proofs as spatio-temporal processes. Abstracts of the 14th Congress of Logic, Methodology and Philosophy of Science, Nancy, July 19–26, 2011, Volume: Full version will appear in: Pierre Edouard Bour, Gerhard Heinzmann, Wilfrid Hodges and Peter Schroeder-Heister (Eds) 14th CLMPS 2011 Proceedings, *Philosophia Scientiae*, 19(1) pp 111–125. <https://doi.org/10.4000/philosophiascientiae.1010>
- Vassiliades A, Bassiliades N, Patkos T (2021) Argumentation and explainable artificial intelligence: a survey. *The Knowledge Engineering Review* 36:n.pag., <https://doi.org/10.1017/S0269888921000011>
- Villiers M (1990) The role and function of proof in mathematics. *Pythagoras* 24:17–24
- Vreeswijk GA (1997) Abstract argumentation systems. *Artif Intell* 90:225–279
- Weber K (2010) Mathematics majors' perceptions of conviction, validity, and proof. *Math Think Learn* 12(4):306–336. <https://doi.org/10.1080/10986065.2010.495468>