Argumentation Frameworks for Constructing and Evaluating Deductive Mathematical Proofs

Nadira Boudjani, Abdelkader Gouaich and Souhila Kaci

Laboratoire de l'Informatique, de Robotique et de Microélectronique de Montpellier

{boudjani, gouaich, kaci}@lirmm.fr

Problem statement

Argumentation in the didactics of mathematics:

- Enhances critical thinking and meta-cognitive skills
- Increases the student's motivation through open interactions
- Allows mutual learning amongst the students

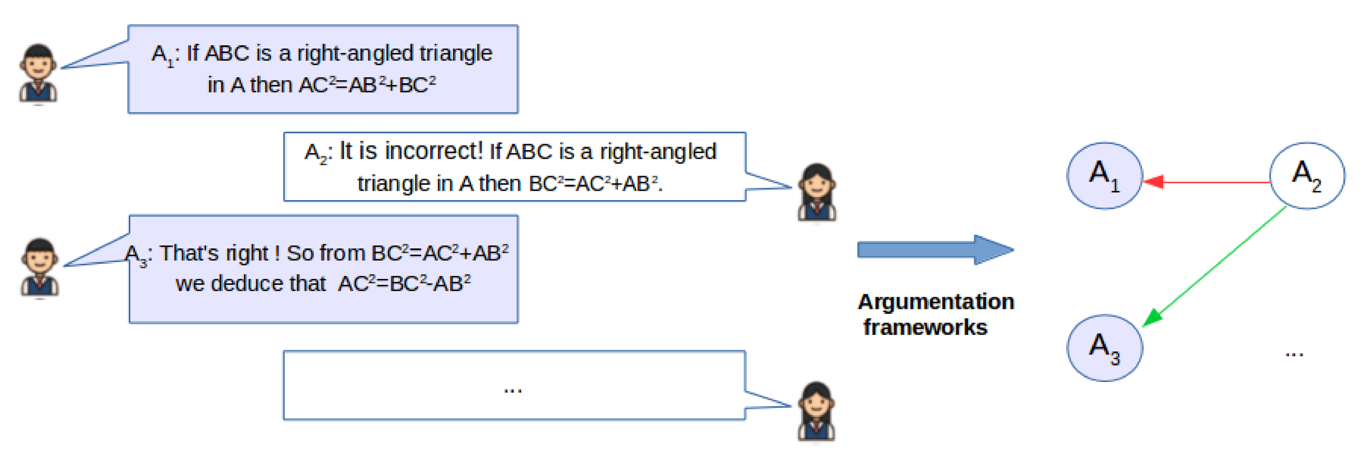
Yet, the following shortcomings are pointed out:

- The language to outline the proof differs from the language of the final proof ⇒ Difficulties to write the final proof
- Evaluating the informal debate ⇒ difficulty of evaluation and to provide constructive feedback
- Evaluating the final proof ⇒ Loss of information to identify misunderstandings

Proposition The CLEAR system (lirmm.fr/mathgame/tmp) has a twofold objective:

- Allow students to build deductive proofs using **structured argumentative** debates
- Help the instructors evaluate these proofs in order to identify misconceptions and provide a relevant feedback

Argumentation in AI

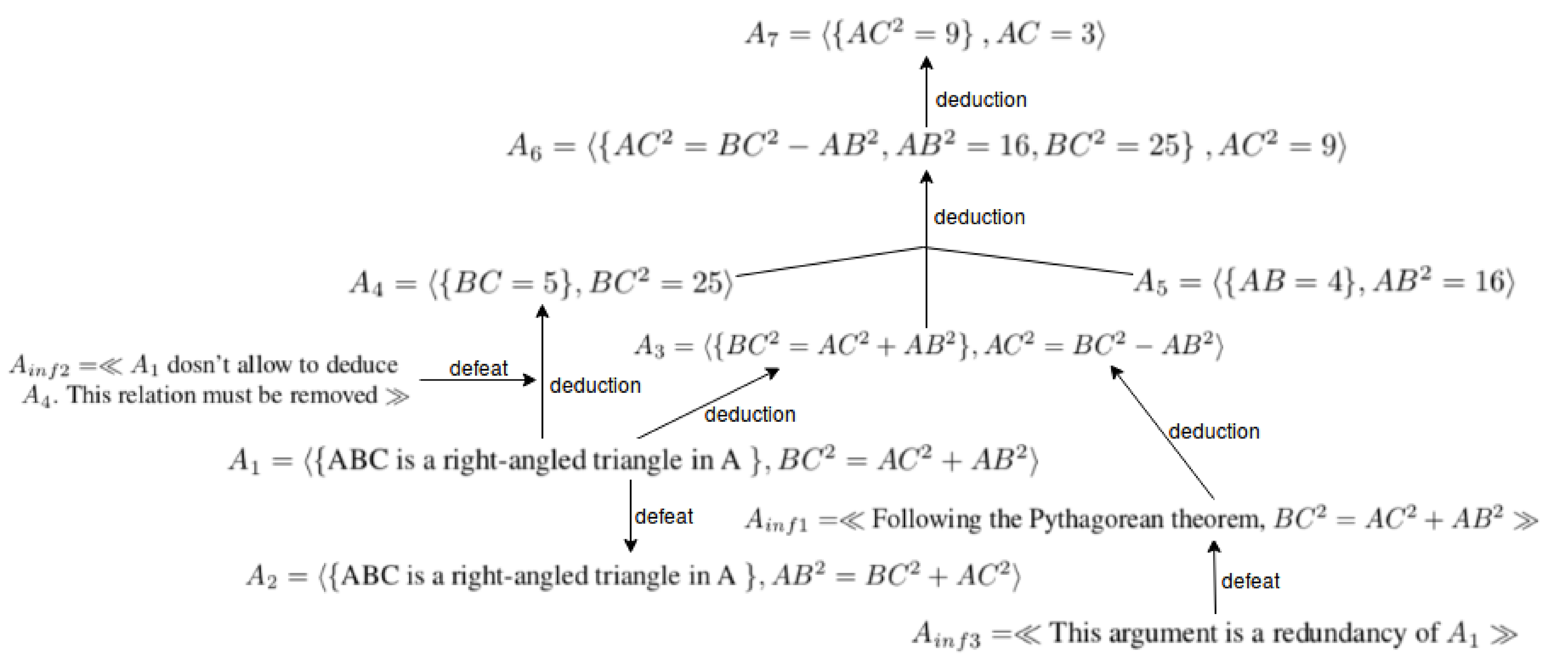


The Defeat-Support argumentation framework is a tuple $\langle \mathcal{A}, \mathit{Def}, \mathit{Supp} \rangle$, where $\mathcal{A}$ is a set of arguments, $\mathit{Def} \subseteq \mathcal{A} \times (\mathcal{A} \cup \mathit{Def} \cup \mathit{Supp})$ is a defeat relation, and $\mathit{Supp} \subseteq \mathcal{A} \times (\mathcal{A} \cup \mathit{Def} \cup \mathit{Supp})$ is a necessary support relation.

Semantics for acceptability of arguments

- $\mathcal{S} \subseteq \mathcal{A}$ is *admissible* iff $\mathcal{S}$ is conflict-free and defends all its elements.
- $\mathcal{S} \subseteq \mathcal{A}$ is a *preferred extension* iff $\mathcal{S}$ is the largest (for set inclusion) admissible set.

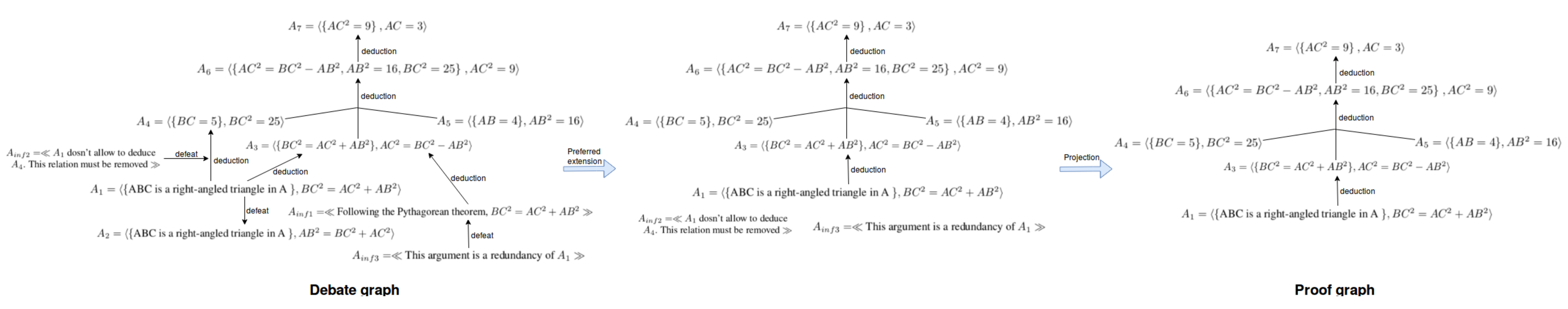
Construction of deductive proofs



The students build a debate graph with:

- **Formal arguments:** $\langle \{P_i\}, C \rangle$, where $\{P_i\}$ are the premises and C the conclusion
- **Informal arguments:** free text
- **Deduction relations:** deduction between arguments
- **Defeat relations:** conflict on arguments and relations

Analysis of the debate graph



The proof graph is obtained by:

1. Computing the acceptable arguments using the preferred semantics
2. Projecting the acceptable arguments on deduction relation

Experiment #1: Students

Goal: Are formal argumentation frameworks suitable for building deductive proofs?

Experimental protocol: Undergraduate students from the department of mathematics ($n = 8$) and the department of computer science ($n = 16$) solved 3 exercises: linear algebra, probability, and analysis.

Results

(a) The average and standard deviation of the number of arguments and relations.

	Formal argument	Informal argument	Defeat	Simple deduc.	Multiple deduc.
Exercise 1	10(2.2)	0(0)	0.25(0.7)	9.37(5.57)	4.25(3.24)
Exercise 2	7.37(3.11)	0.5(0.92)	0.5(0.75)	6.75(2.76)	2.37(2.26)
Exercise 3	4.85(2.54)	0.14(0.37)	0(0)	3.57(2.5)	0.85(1.06)

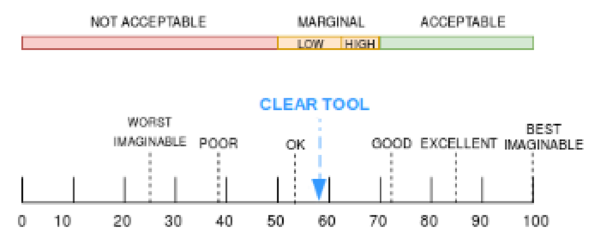
(b) The average and standard deviation of actions.

	Delete relation	Edit argument	Pass turn
Exercise 1	3(3.81)	1.87(2.35)	0.12(0.35)
Exercise 2	0.5(1.41)	2(2.67)	0.12(0.34)
Exercise 3	0.71(1.49)	1.14(0.89)	0.28(0.75)

The average and standard deviation of the number of arguments, relations and actions per exercise (CS)

Item	average rating /5
C1: Comprehension of formal argument and argumentation theory	3,46
C2: Representation of proof by graph and its visualization: Arguments and relations	3,34
C3: Building formal argument by selecting premise(s) and conclusion	3,37
C4: Adding relations: Deduction and defeat	2,94
C5: Importance of having multiple deduction relation	3,63
C6: Importance of having edit argument action	4,43
C7: Importance of having delete relation action	4,75
C8: Importance of having pass action	1,78
C9: Importance of having informal arguments	2,56
C10: Relevance of building proofs in pair	3,78

The average rating of a self-reported questionnaire on the different aspects of CLEAR (CS)



Average score of CLEAR tool (CS)

Experiment #2: Instructors

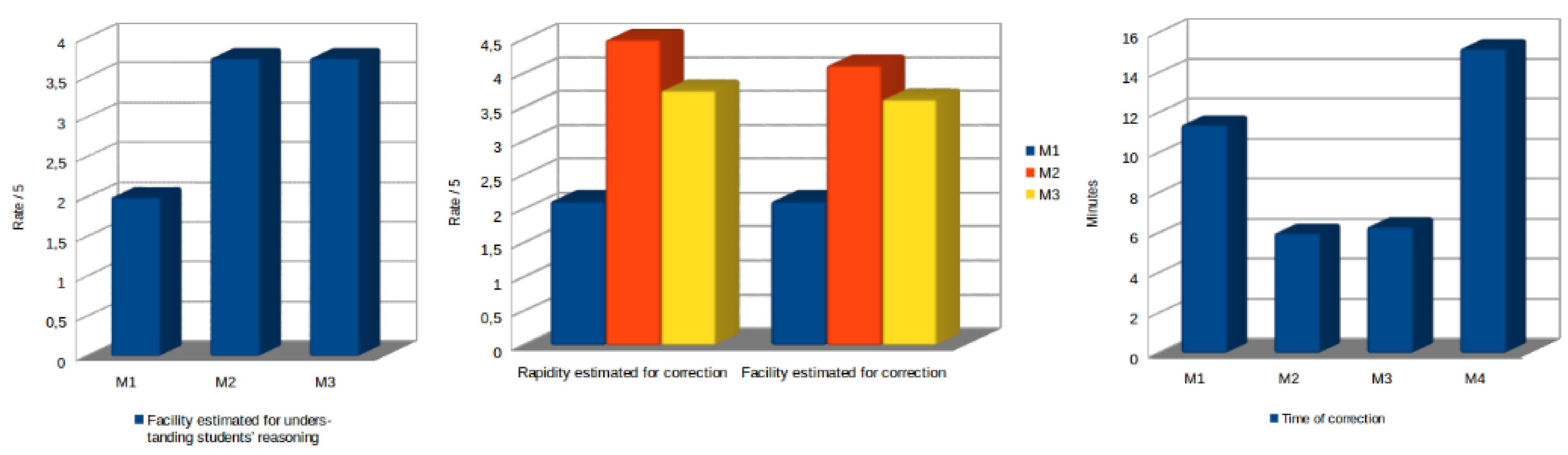
Goals:

(i) What are the representations that allow instructors to easily correct deductive proofs?

(ii) What are the representations that allow instructors to provide a relevant feedback?

Experimental protocol: ($n = 8$) teachers (high school and lecturers) evaluate proofs by four methods: debate journal ($M1$); proof graph ($M2$); classical proof ($M3$); and having all the three representations ($M4$)

Results



Conclusion

The results of the experimental studies suggest that:

- AI argumentation frameworks are suitable for building deductive proofs without any prior knowledge on argumentation theory
- The usability of CLEAR has been rated as 'OK' on the Standard Usability Scale
- The instructors have considered that the proof graph representation facilitates the assessment of proofs and helps understand the students' reasoning.