Edinburgh Research Explorer

An Experimental Comparison of Diagrammatic and Algebraic Logics

An Experimental Comparison of Diagrammatic and Algebraic Logics

Daniel Winterstein, Alan Bundy, Corin Gurr & Mateja Jamnik

Edinburgh University & Cambridge University

Abstract. We have developed a diagrammatic logic for theorem proving, focusing on the domain of metric-space analysis (a geometric domain, but traditionally taught using a dry algebraic formalism). To evaluate its pragmatic value, pilot experiments were conducted using this logic - implemented in an interactive theorem prover - to teach undergraduate students (and comparing performance against an equivalent algebraic logic). Our results show significantly better performance for students using diagrammatic reasoning. We conclude that diagrams are a useful tool for reasoning in such domains.

1 Introduction

Euclidean plane geometry has always been taught using diagrammatic reasoning. Traditionally though, only algebraic proofs are allowed in the slippery realms of more abstract geometries. We have investigated using diagrams in such a domain, that of *metric-space analysis*. It is a hard domain, and even great mathematicians such as Cauchy have made mistakes in this subject.[1] Students typically find it daunting, and we conjecture that the dry algebraic formalism used in the domain is partially responsible for these difficulties. Currently our logic only covers a fraction of the domain, but this was sufficient to run some short tutorials on the concept of *open sets*. This allowed us to experimentally compare our diagram logic with an equivalent algebraic logic.

2 The Logics

We can only give a brief overview of the logics here.[2] The algebraic logic uses natural-deduction rewrite rules. The diagram logic is specified using *redraw rules*, which are a visual adaptation of rewrite rules. Redraw rules are defined by an example diagram transformation; figure 1 shows an example. The diagrams in our logic are made up of example objects with relation statements. These statements can be represented in three different ways: *implicitly* (where the relation is true for the objects drawn, e.g. $a \in B$ for $a = \frac{1}{2}$, $B = [0, 1]$), *graphically* (using

---

[1] See E.Maxwell "Fallacies in Mathematics" Cambridge University Press, 1959.

[2] For more information, please refer to the paper by the same authors in *Diagrammatic Representation and Inference* Springer-Verlag, 2002.

conventions such as 'a dotted border indicates an open set') or *algebraically*. For both reasoning styles, students were restricted to constructing valid forward reasoning proofs using equivalent rule-sets. The algebraic proofs produced could be described as 'typical text-book proofs'.

Both logics were implemented in a user-friendly interactive theorem prover, which we call *Dr.Doodle* for its drawing mode. Figure 2 shows a sample screenshot. This system was designed to minimise the potential effect of differences in the interface/presentation methods, so that logics could be compared without other factors unduly affecting the results.
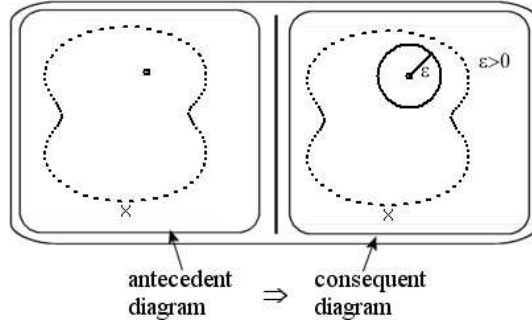

Fig. 1. A redraw rule for "X an open set, $x \in X \Rightarrow \exists \epsilon > 0$ s.t. $\{x' : |x' - x| < \epsilon\} \subset X$"

3 Experimental Design & Results

Two experiments were conducted: the first used 10 1st year students (for whom the material covered was new), and the second 10 2nd-3rd year students (who had seen the material before). For each experiment, the students were randomly split into 2 groups of 5. One group worked using diagrammatic reasoning, the other used algebraic reasoning. Over two 45-minute lessons, the students were taught to use the Dr.Doodle system, then tested on a set of exercises. Lessons were conducted entirely on computer without human interaction. Results from the 1st years showed that the test-exercises were too ambitious, producing a very coarse-grained measure of ability. The exercises were therefore modified for the 2nd experiment, which gave a better spread of results (it is probably this which is responsible for the difference in the results between experiments). For each student, we calculated two measures: 'score' (correct exercise questions), and 'inefficiency' (wasted actions, based on analysis of user-logs). Informal feedback was also gathered via a questionnaire.

Experiment	Reasoning Style	Score	Inefficiency
1st years	Diagrams	30% σ=22%	1.49, σ=0.98
1st years	Algebra	30% σ=14%	2.74, σ=1.99
2nd/3rd years	Diagrams	63%, σ=6.4%	1.27, σ=0.03
2nd/3rd years	Algebra	42%, σ=20%	1.63, σ=0.19
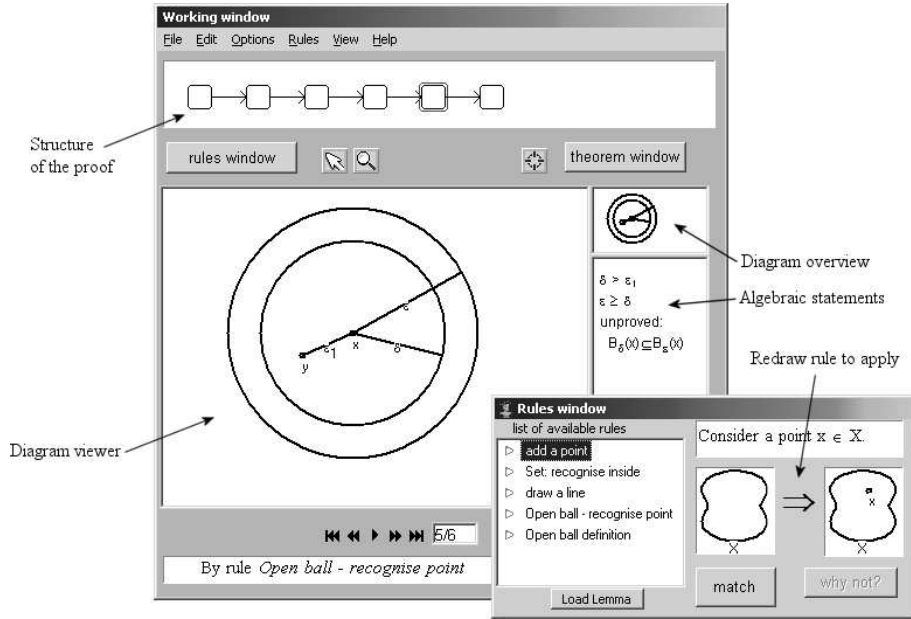
Fig. 2. Screenshot of Dr.Doodle in diagrammatic reasoning mode.

Both experiments show students working more efficiently when using diagrams. The second experiment also shows the diagrams group scoring higher. With such a small sample, we would not expect statistically strong results. However the results from the second experiment *do* show statistically significant support for the conjecture that diagrammatic reasoning is better at this task[3] (both measures are significantly better at 95% confidence using a one-tailed t test). Informal feedback gave comments such as: "The pictures were useful for helping understand what was going on. Better than written explanations a lot of the time."

4   Conclusion

These positive results are not surprising. As the domain is a geometric one, we would expect visual representations to be useful. We conclude that diagrammatic reasoning is a useful tool in this field. However further experiments are desirable, especially as these experiments did not look at the interesting questions of *how* and *why* diagrams are useful here (and hence how general these findings are). This work is described in more detail in the first author's forthcoming Ph.D. thesis. Hopefully this project will be extended to produce a tutorial system for this.

---

[3] Although it is also possible that the difference comes from having mental access to multiple representations, rather than from diagrammatic reasoning *per se*.