Dr.Doodle: A Diagrammatic Theorem Prover

Daniel Winterstein, Alan Bundy & Corin Gurr

Edinburgh University

Abstract. This paper presents the Dr.Doodle system, an interactive theorem prover that uses diagrammatic representations. The assumption underlying this project is that, for some domains (principally geometry), diagrammatic reasoning is easier to understand than conventional algebraic approaches – at least for a significant number of people. The Dr.Doodle system was developed for the domain of metric-space analysis (a geometric domain, but traditionally taught using a dry algebraic formalism). Pilot experiments were conducted to evaluate its potential as the basis of an educational tool, with encouraging results.

1 Introduction

Diagrams are commonly used in virtually all areas of representation and reasoning. In particular – although current theorem-provers make very little use of diagrams – they are invaluable in mathematics texts. They are used in a variety of ways, including to give examples showing why a theorem is true, to give counter-examples, to explain the structure of a proof, and to prove a theorem outright. Insight is often more clearly perceived in these diagrammatic proofs than in the corresponding algebraic proofs. We have developed a system for producing such proofs in the domain of *metric-space analysis*, based upon a new diagrammatic logic.

As well as contributing to the development of more accessible theorem provers, this work also opens the exciting possibility of developing computer-based diagrams in new directions. If we consider the very real differences between text and hypertext, we see that diagrammatic reasoning on computers need not be just a straight conversion of diagrammatic reasoning on paper. Our work has led to the development of animated diagrams with a formal semantics as a meaningful representation for quantifiers (see [8]). The rigour required for doing mathematics forces a thorough investigation of the mechanics of such reasoning.

1.1 Our Domain: Metric-Space Analysis

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*. This is a hard domain, and even great mathematicians such as Cauchy have made mistakes in this subject [5]. Students typically find it daunting, and we conjecture that the dry algebraic formalism

used in the domain is partially responsible for these difficulties. Currently the system only covers a fraction of the domain, but this was sufficient to run some tutorials on the concept of *open sets*. This allowed us to experimentally compare the use of diagrams with an equivalent algebraic approach.

2   The Dr.Doodle System

The Dr.Doodle system is an interactive theorem prover for non-inductive reasoning, with diagrammatic representations for metric-space and real-line analysis concepts (involving objects such as functions, sets and lengths, and properties such as open, closed and continuous). The user selects which rule to apply at each step. In places this involves drawing, and the interaction sometimes resembles using a graphics program.
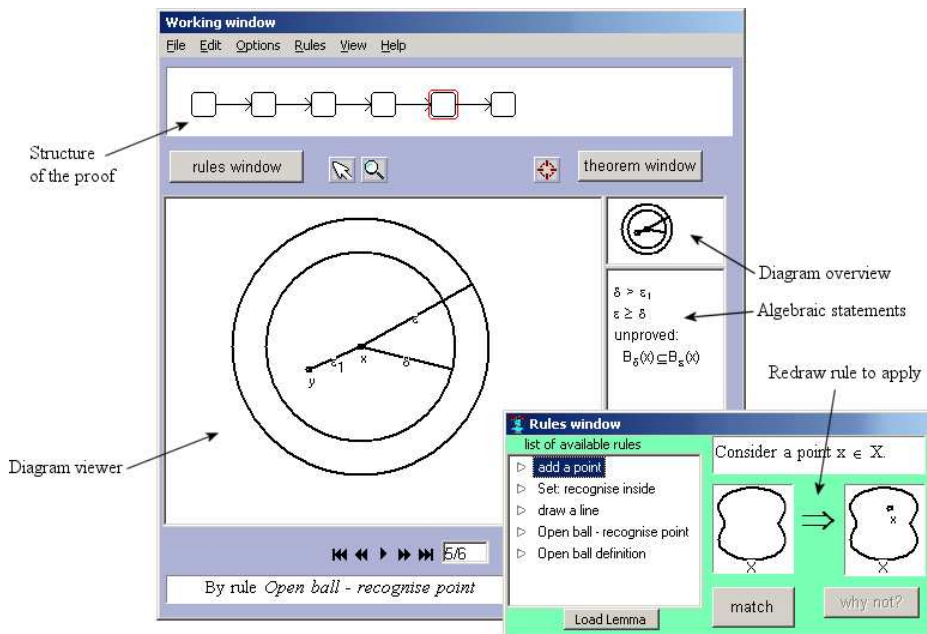

Fig. 1. Screenshot of Dr.Doodle.

2.1   Dr.Doodle Diagrams

The diagrams in our system are so-called *heterogenous representations*, combining graphical and textual elements. They consist of example objects with conjunctions of relation statements. Where possible, the example objects should constitute a model for the statements; if no model can be found (e.g. because the statements are inconsistent) then a *near* model is used. Disjunctions are represented using multiple diagrams (an intuitive representation, but with scaling

issues). 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 conventions such as 'a dotted border indicates an open set') or *algebraically*. These methods are used as appropriate (e.g. relations such as $A \subset B$ are usually implicit, but can also be stated algebraically if necessary).

The system produces diagrams from an internal proof state as follows:

1. The example objects are drawn.
2. The drawing is tested to see which relations are implicitly represented (note that these tests take into account drawing resolution).
3. Often, a diagram will represent relations that are not specified in the internal proof state. Sometimes this is desirable. For example, if we represent $A \subset B$ by drawing A inside B, then a diagram for $A \subset B$, $B \subset C$ will inevitably also represent $A \subset C$ – which is, of course, true. Such 'free rides' are one advantage of using diagrams (c.f. [7]).

   Any extra relations found in the previous step are tested to see if they follow from the specified relations as a result of free rides. Unwanted extra relations are then explicitly removed by the system.
4. Relations which are not implicitly represented by drawing the objects (e.g. $y = f(x)$) are explicitly added to the diagram.

This process involves important design choices at several stages:

- How to choose and draw the example objects.
- Which relations can be represented implicitly.
- How to represent the explicit relations (e.g. arrows for function relations, but predicates for properties such as 'surjective').

There is a trade off here between simplicity, flexibility, and intuitive appeal, and our choices are tailored to the domain considered.

2.2 Dynamic Diagram Logic

We can only give a brief overview here of the logic implemented in Dr.Doodle. For more information, please refer to [8]. Often diagrammatic reasoning is presented as a question of interpreting static diagrams. Here we consider *dynamic* diagrammatic reasoning, where the process of drawing is important, as opposed to just the finished diagram.

The inference rules are specified using *redraw rules*, which are a visual adaptation of rewrite rules. Redraw rules are defined by an example diagram transformation; figure 2 shows an example. When a rule specifies creating new objects, these are typically drawn-in by the user (automated drawing, whilst interesting, requires both model-generation and aesthetic judgement to select 'good' examples). Theorems are stated in the same manner. A proof consists of a demonstration that the theorem antecedent can always be redrawn to give the consequent diagram using an accepted set of rules.This paradigm - whereby inference rules are mainly presented as drawing acts - restricts users to constructing forward reasoning proofs; future work includes handling other proof structures.
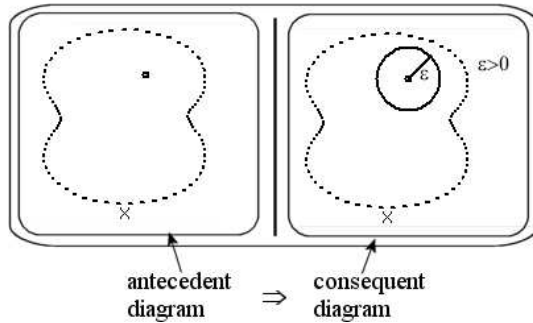
Fig. 2. A redraw rule for "X an open set, $x \in X \Rightarrow \exists \epsilon > 0$ s.t. $\{x' : |x' - x| < \epsilon\} \subset X$"

An Equivalent Algebraic Logic Dr.Doodle also features an algebraic reasoning mode, where redraw rules are converted into natural-deduction rewrite rules. Although the two types of proof appear very different, the reasoning used is equivalent, and the Dr.Doodle system can convert a diagrammatic proof into an algebraic one (though not vice versa at present).

3 Experimental Evaluation

Pilot experiments were conducted to test the use of Dr.Doodle in a practical setting (teaching mathematics undergraduates). These experiments compared performance using Dr.Doodle in algebraic and diagrammatic modes (hence evaluating the potential of the diagrammatic logic – since this is the innovative aspect of the system – independently of the interface), and are described in [9]. The results show statistically significant improvements – both in exercise scores and 'efficiency' – when using diagrammatic reasoning. 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."

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).

4 Related Work

Diagrammatic reasoning is a relatively unexplored research area, and there are few comparable systems. There are several powerful geometric theorem provers (e.g. Cinderella [4]), but these use proof techniques that are very hard for humans to follow. The work by Howse *et al* on spider diagrams is more relevant. These diagrams are used to represent set-theoretic statements, and can also support reasoning [2].

Barwise, Etchemendy *et al*'s HyperProof arguably sets the standard for educational applications of theorem provers [1]. It is a much more developed system than Dr.Doodle, but handles a very different domain. Aimed at philosophy students learning logic, it uses the blocksworld domain to give concrete visual meaning to predicate logic statements. Diagrammatic inferences in HyperProof involve reading information from the diagram, or testing propositions against the diagram. Psychometric studies by Stenning *et al* show that this is beneficial for some students [6]. Moreover, which students will benefit may be predictable using simple aptitude tests for spatial reasoning.

The most closely related system is Jamnik's Diamond, which uses geometric reasoning about area to prove natural number arithmetic theorems [3]. The user supplies proofs for example cases of a theorem, from which a general proof is extracted and checked. This project was initially conceived as extending Jamnik's work to a continuous domain whilst exploring practical applications. However the differences between countable domain and continuous domain reasoning have led to a very different system.

5   Conclusion

We have described the Dr.Doodle system and given an overview of the logic it uses based on diagrammatic representations. The aim of such work is to produce theorem provers whose proofs can be readily understood, based on the idea that diagrammatic representations are easier and more intuitive for some domains. Our pilot experiments support this idea. We now intend to develop Dr.Doodle further to produce a tutorial system for mathematical analysis.

This work is described in more detail in the first author's forthcoming Ph.D. thesis. A demonstration version of the Dr.Doodle system is available from the first author.

References

1. J.Barwise & J.Etchemendy "Heterogeneous Logic" in *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, AAAI Press/MIT Press, 1995.
2. J.Flower & G.Stapleton "Automated Theorem Proving with Spider Diagrams" in *Computing: the Australasian Theory Symposium (CATS)*, 2004.
3. M.Jamnik "Mathematical Reasoning with Diagrams"CSLI Press, 2001.
4. U.Kortenkamp & J.Richter-Gebert "The Next Version of Cinderella" in *First International Congress Of Mathematical Software*, World Scientific, 2002.
5. E.Maxwell "Fallacies in Mathematics" Cambridge University Press, 1959.
6. J.Oberlander, R.Cox & K.Stenning "Proof styles in multimodal reasoning" in *Logic, Language and Computation*, CSLI Press, 1996.
7. A.Shimojima "Operational Constraints in Diagrammatic Reasoning" in *Logical Reasoning with Diagrams*, OUP, 1996.
8. D.Winterstein, A.Bundy, C.Gurr & M.Jamnik "Using Animation in Diagrammatic Theorem Proving" in *Diagrams 2002*, Springer-Verlag, 2002.
9. D.Winterstein, A.Bundy, C.Gurr & M.Jamnik "An Experimental Comparison of Diagrammatic and Algebraic Logics" in *Diagrams 2004*, Springer-Verlag, 2004.