1997 Workshop on Automated Reasoning

Louise Dennis, Alan Bundy and Helen Lowe

April, 1997

This report is actually a collaborative effort. The individual contributions are credited to their authors by initials at the end of each section: Louise Dennis [LD], Alan Bundy [AB] and Helen Lowe [HL]. Alan Frisch contributed some editorial improvements.

1 Introduction

The 1997 Automated Reasoning Workshop, like the previous Automated Reasoning Workshops, was run as part of the AISB 1997 Workshop series. This allowed the workshop to share it's first invited talk, by Jeff Paris, with the entire workshop series. The other invited speaker, Ray Reiter, also gave a talk to the Practical Reasoning and Rationality workshop. However, it was felt that the simultaneous occurrence of both the Practical Reasoning workshop and the AR workshop didn't help the drift of the AR workshop towards the "CADE Community" (more on this anon).

The Automated Reasoning Workshop, itself, was organised by a committee chaired by Michael Fisher. It was run as a number of informal panel sessions with a couple of invited talks. The overall tone was very introspective with a lot of questioning of what direction the field should be taking and seemed to reveal a certain amount of disappointment with progress made so far. [LD]

2 Interworkshop Plenary Talk: "Representation Dependence: Not Seeing Is Not Believing" – Jeff Paris

Jeff Paris (Manchester) presented a talk to all the workshops being held at Manchester. It was founded on criticisms of his work on modelling belief in Intelligent Agents. He had shown that under certain "common sense" constraints belief (as a function of the agent's knowledge) was the maximum entropy function (ME). The particular criticism he focused on was that ME was "representation dependent", that is the belief that fish live in a river varies depending on whether fish are represented simply as "fish" or as "fresh water or sea water fish". He showed that this objection was flawed since it either failed to account for the fact that the agent's knowledge had changed, or it failed to allow the agent to reason that although some category of objects might be reducible to sub-categories, it also might not.

He added that there were other, more significant, criticisms of ME. In particular he said he didn't claim that this was the way in which humans form beliefs. [LD]

3 Invited Talk: "Modelling Dynamical Systems in the Situation Calculus: Some Representational and Computational Issues for Automated Reasoning" – Ray Reiter

Ray Reiter (Toronto) advocated the use of the situation calculus for robotic reasoning in direct opposition to the behavioural robotics school of Brooks. Ray called his approach "Cognitive Robotics" and his reasoning system is called "Golog". Among the interesting points made in his talk were the following. Negation as failure is not enough in robotics — you need sensors to update your knowledge of the world. Second order dynamic logic is used for initial modelling, but this is translated into first order logic for execution for efficiency reasons. Reasoning can be kept tractable if it is restricted to rewriting with definitional equivalences. It is necessary to use induction to form recursive plans. [AB]

4 Panel A: Bridging the Gap between the AI Reasoning and the Automated Deduction Communities

This panel was chaired by Alan Frisch (York). The panelists were Alan Bundy (Edinburgh), Lincoln Wallen (Oxford), Ray Reiter (Toronto), Hans-Jürgen Ohlbach (Imperial) and Ian Gent (Strathclyde). Alan Frisch opened the panel session by pointing out that the workshop appeared to be drifting towards what he termed "the CADE community" and away from "the KR community". He submitted a plea for the workshop to maintain its diversity.

Alan Bundy continued this theme, splitting the field into maths, formal methods and common sense reasoning and discussed the different foci of these groups, in particular the fact that knowledge representation was the major area of research in common sense reasoning and was often regarded as "given" for maths and formal methods, though he suggested that it returned as a problem at the meta-level when representing search control knowledge.

Lincoln Wallen emphasised that Automated Theorem Proving (ATP) was a subset of Common Sense Reasoning (CSR). He also pointed to Ray Reiter's GOLOG system saying that CSR systems needed to produce real-time solutions and this meant using something simple and reserving the complicated formalisms for analysis of those simple systems.

Ray Reiter enlarged this and talked about "feasible reasoning". He suggested that CSR should exploit database methods and monotonic logics that are equivalent to nonmonotonic ones to produce systems with trivial reasoning processes once the relevant analysis of these methods and equivalences was done.

Hans-Jürgen Ohlbach discussed an experiment to build a program that could generate a dialogue between a car salesman and a customer. Once again the reasoning part was a database with a hierarchy of definitions.

Ian Gent noted that both areas were concerned with satisfiability and discussed conflict directed back-jumping as a search method.

The open discussion centred around two areas. The first was the various differences suggested between (broadly) theorem proving and common sense. The contention that knowledge representation was given was challenged since it could affect search and even the solutions (e.g Cartesian versus Euclidean geometry). The general feeling was that search was the wrong paradigm for real-time agents and CSR should focus on producing systems that made mistakes and took the consequences, whereas ATP required specific problem solving behaviour.

The second area was a discussion of Otter and whether there was anything to be learnt from its approach. It was felt that it was limited to certain areas and it was likened to a calculator rather than a theorem prover. People were divided over whether speed of search was a misguided effort or not, similarly whether the twitching of settings counted as cheating or not. [LD]

5 Panel B: The Future of Automated Reasoning Research

Alan Bundy (Edinburgh) chaired this panel and Lincoln Wallen (Oxford), Rob Arthan (ex-ICL, now Lemma 1) and Dominic Semple (EPSRC) were the panelists. Lincoln emphasised the growing importance of computer science practioneers being skilled professionals. Rob discussed the role of automated reasoning systems in formal development of high-assurance systems. Although there has been a lot of progress in the last decade or so, there is still a great deal of work left to do before the potential of mechanised proof is realised in practice. Much of what is needed is consolidation — not so much building bridges as filling in the ravines that the pioneers have already bridged. Dominic discussed the new structure of EPSRC and the continued support within it for basic and theoretical research. However, grant applicants must ensure they address the needs of the "users" of their research. These users could be fellow academics. Automated reasoning research is well supported by EPSRC. There are no figures for AR per se, but Theoretical Computer Science takes about 30% of Dominic's budget and AI about 25%. [AB]

6 Compulog Session

Manfred Kerber (Birmingham) is the organiser of the Automated Deduction Systems Area of the Compulog Network of Excellence. He took the opportunity of the AR workshop to hold some Compulog events. Compulog money was used to sponsor Ray Reiter's talk. Uli Furbach was invited to give a talk on "Theorem Proving and Logic Programming", but unfortunately this had to be cancelled because Uli had an accident.[1] His place was taken by Laurent Vigneron who spoke on "Automated Deduction and Rough Information". There was also a business meeting which discussed what role the Compulog AD area might play. It was decided that we needed to find a distinctive niche and not just to organise another workshop for the sake of it. Three possibilities were discussed in the resulting brainstorming session: try to bridge the gap between AD and either logic programming or knowledge representation; run a summer school for students; focus on technology transfer and industrial contacts. These topics are up for further discussion. [AB]

You can find out more about this area of Compulog at:

`http://www.cs.bham.ac.uk/~mmk/compulog/compulog.html.`

7 Automated Reasoning Web Page

Alan Bundy gave a short report on the work of the Research Strategy Group (RSG) of the Conference of Professors and Heads of Computing (CPHC). The role of RSG is to represent UK CS research to bodies like Foresight and EPSRC. To do this it was building a map of the field which it was mounting on the WWW. This map was multi-level ranging from products and services via "generic questions" and "key issues" to actual research areas. An example generic questions is: "how can computer systems be made more reliable?", and a key issue that needs to be addressed to solve this question is the "provision of machine assistance with proof obligations". Automated reasoning is working on this key issue, among others. The mappings between each level are many-to-many.

Each research area is to be asked to prepare the web page for its area. This should not only list the major UK research groups with links to their pages, but also link back to the key issues and say how they are being addressed. Automated reasoning has been selected as a pilot area for this.

There was general enthusiasm to take on this task and on its importance for the field, *e.g.* in promoting collaboration, drawing our work to the attention of users, funding agencies, *etc.* Alan Frisch agreed to edit the page, but needs technical help on WWW technology. [AB]

[1] Manfred organised us to send him a get well soon card.

8 Panel C: Can Formal Methods Flourish without Automated Reasoning

This panel was chaired by Ursula Martin (St. Andrews) and the panelists were Harrie de Swart (Tilburg), Rob Arthan (ex-ICL, now freelance), Konrad Slind (Cambridge), and Helen Lowe (Napier).

Harrie began the session. His position was that there are a number of problems with current automated reasoning tools.

1. None of them support Dijkstra-Hoare style proof development.
2. They are built in layers and are cumbersome; they will not run on average-sized computers.
3. At the level of the user interface, they are constructing proofs rather than programs; the latter is what the user is interested in.

Another aim for a system is that is should not force formal support on the user where it is not welcome. He concluded by saying that what Formal Methods needed was an ATP as a special tactic which could be invoked where it was wanted.

Rob Arthan of ProofPower fame spoke next. He replaced Cliff Jones at short notice. He said that the Z method would not survive, but that Z itself would — for specifying safety critical and low level code. He said that if the question the panel was addressing was instantiated with Formal Methods = Z and Automated Reasoning = mechanised theorem provers, then he believed that Formal Methods itself could not survive without AR. Real problems are too complicated for practitioners to maintain true rigour. Z was most effective if used early on in the specification process. He pointed out that people are not in the habit of doing proof during specification.

Helen Lowe spoke next. She pointed out that so-called automated theorem provers are not really that automatic as the user often has to persuade the system to prove the theorem by introducing lemmas, etc. There is a steep learning curve with this. Entering the mindset of a formal prover is hard because people are more used to rigorous proof styles. Developers should consider theorem proving in the context of supporting the whole development process, and design for a sensible division of labour (humans to do the interesting bits and machines to do the tricky and tedious).

Then it was Konrad Slind, who said that most successful proofs were interactive (really). The work lay in taking a real property, and manipulating it so the proof of it was automatable. He talked of the trade off between simple and more expressive logics: ease of formulation versus finding proofs. The real resource cost is user time. He gave us a thought experiment involving two users. The first is good at proof but has only a simple proof checker. The second is dumb but has a state of the art theorem prover. They are both given a problem slightly out of range of the systems. Konrad's experience leads him to believe the former would succeed first. Whereas lack of the right background can be alleviated by tools, when the tools fail you are left to draw on your personal resources.

A discussion followed at which some of the points raised were as follows.

Is ATP the only bottleneck for FM? — No, also cultural factors.

Compared with building bridges, software development tends to get by with testing rather than rigour. The public don't expect software to work, whereas they do expect bridges to stay up.

An example was given where the person would not have finished a proof they did by hand if they had used a tool.

CaDiZ was mentioned, it is good for part of inductive proofs — unfortunately, although the inductions themselves are simple, the side conditions spiral out of control.

Some talk about mistakes in hand proofs — anecdotal evidence says they are always correctable (i.e. people don't end up proving things that aren't true).

Interface issues: proof is about understanding and explanation, the aim is to know what is going on not formal proof per se.

Helen Lowe asked what FM ATP system builders used. ProofPower of course does; the latest Nqthm variant rather incestuously proves itself, apparently. (This question went down like a visitors' goal in the home end at Ibrox.)

Finally, do we need one big project or special purpose projects? Hitherto we have had LCF/ML, MKRP, etc. [HL]

9 Panel D: How Intelligent Should Reasoning Systems Be?

Michael Fisher (Manchester Metropolitan) chaired this panel and posed the problem by saying that brute force provers like Otter had been successful, *cf.* the recent proof of the Robbin's conjecture, so was there really a need for proof planning and other intelligent control techniques? Andrew Adams (St. Andrews) did not directly address the question but argued the role of automatic reasoning systems as "intelligence amplifiers". Andrew Ireland (Heriot Watt) listed 4 reasons for needing proof planning, namely: where huge search spaces needed to be controlled; where open systems were needed to support interoperability; where lemmas needed to be speculated, *etc*; and where there was a need for user interaction. Manfred Kerber (Birmingham) argued that the wider problem of reasoning was "AI-complete", *i.e.* required "real" artificial intelligence for solution. He advocated the study of Polya to suggest what kind of intelligence needed to be added to theorem provers. Judith Underwood (Glasgow) said that certain applications only required "fast but dim" provers, but that others required a deeper understanding from the prover.

In the question period Alan Bundy claimed that the Otter proof of the Robbin's conjecture had required something like proof planning, but that this was provided by the human users in breaking down the problem into lemmas and setting the many switches and dials on Otter. [AB]

10 The Next Workshop on Automated Reasoning

The organising committee for this workshop series are committed to a continuing annual series. To ensure this, the committee is trying to reach a long-term agreement with AISB that would couple the workshop to the annual AISB workshop series and would give the Automated Reasoning Workshop some financial independence.

However, in 1998 AISB will be hosting ECAI-98 and therefore will not be organising a Spring workshop series. The workshop series associated with ECAI-98 would not be an appropriate venue for us because ECAI would require attendees to register for the full conference. This would not suit the many AR workshop attendees who are not AI researchers (*e.g.* those from formal methods). Therefore, in 1998 we will associate the AR workshop with the British Colloquium for Theoretical Computer Science (BCTCS), which is being held in St. Andrews in the week of 30 March. Ursula Martin (St. Andrews) will serve as programme chair and local arrangements chair for both events. [AB]

You can find out more about the AR workshop series at:

`http://www.cs.york.ac.uk/~frisch/AutoReason/`