Agent-Based Theorem Proving

Christoph Benzmüller
Fachbereich Informatik
Universität des Saarlandes
chris@ags.uni-sb.de

Volker Sorge
School of Computer Science
University of Birmingham
V.Sorge@cs.bham.ac.uk

1 Introduction

We report on a commencing project that is concerned with the transfer of techniques from multi-agent systems research to the domain of automated reasoning. The goal is to achieve more flexible organisation within a theorem-prover, not only to provide a better approximation to the human reasoning approach, but also to solve more complex theorems.

In previous work we have already experimented to some extend with the transfer of multi-agent techniques to theorem proving. We have developed a distributed architecture to support interactive theorem proving [3] within the ΩMEGA theorem-proving environment [1]. Its automation has led to a prototype distributed reasoning system [2], which comprises several hundred independent reasoning processes of which some encapsulate full reasoning agents such as independent automated theorem provers, model generators, and computer algebra systems. The cooperation of the single components is achieved via a central hierarchical blackboard architecture and the overall proof is constructed within a centralised proof object. A significant bottleneck is the problem of communicating partial proofs between the different systems involved.

Many of our problems are typical for distributed problem solving environments and have been tackled with some success in field of multi-agent systems [6; 4]. Here single agents are comprised of autonomous computational entities that have both self perception and a perception of their environment containing other agents. This enables them to independently pursue their goals as well as to flexibly form societies with other agents in order to cooperatively achieve goals. Some of the techniques we try to incorporate from by multi-agent systems research are, for instance, ways to negotiate between agents, efficient communication languages, distribution of limited resources and heuristics to evaluate performance.

The main goal of our project is to build a flexible automated theorem proving system based on the agent paradigm that enables us to tackle hard problems by cooperation within a society of heterogeneous reasoning agents. The single agents are independent from each other but can communicate and cooperate in order to produce a proof. Thereby we will exploit the considerable insights we have gained in our previous work when we experimented with an agent-based approach in ΩMEGA. Currently, we extend our approach, bringing together work on distributed theorem-proving, the development of societies of agents, and the modelling and evolution of cooperation, in order to investigate how an *agent-oriented* approach may be used to improve theorem-proving. We will, in particular, tackle the shortcomings of the predecessor system, such as the communication bottleneck, the limited distribution, and the lack of autonomy of the reasoning agents.

In the following we present some of the major research tasks of the project.

2 Distributed proof search

The main aim of the agent-based system is to enable the distribution of the proof search among groups of reasoning agents. There are two possible ways to distribute this search: (1) to tackle one subproblem using several reasoning agents concurrently or cooperatively, and (2) to work on different subproblems in parallel. This essentially corresponds to *and-or-parallelism* within the search.

From the theorem proving point of view, this type of search poses several challenging problems. (1) The identification of possible *dependencies between parallel proof attempts*: Which agents might interfere with each other and how? What is the relevant information that needs to be exchanged while tackling different subgoals in parallel? When should the parallel proof attempts by resynchronised? (2) The problem of *backtracking*: How can backtracking be organised in parallel proof attempts with heterogeneous reasoning agents? Which are the proof states the system can backtrack to and what kind of information, such as logical dependencies between different proof branches, can be exploited? What are the criteria for backtracking? (3) The construction of a joint *uniform proof object*: If a proof attempt has been successful we are still interested in the proof that was constructed, i.e. it is necessary to reconstruct a uniform proof object from the results of the distributed search. For this the agents involved have to retain some information about their own search. Therefore, we have to investigate appropriate granularity of this information and how should it be represented, communicated, and maintained? A final uniform proof object is also particularly important in order to *guarantee correctness* of the proof. Since very heterogeneous systems can be involved (for instance, provers for classical and constructive logic) assembly of a distributed proof might not lead to a correct proof.

3 Flexible cooperation

The major means to achieve a flexible behaviour of the system is to enable the *dynamic formation of clusters* of reasoning agents. These clusters can be of various types. For example, agents that complement each other can attempt to cooperatively solve a problem (for instance, the cooperation of a first order and higher order theorem prover). Agents with similar abilities can form clusters that have a certain reasoning expertise and can work concurrently on given problems (for instance, a collection of first order theorem provers). This increases the likelihood that a problem at hand will be solved by one of the agents. Clusters can also be comprised of agents with contrasting abilities (for instance, an ATP and a model generator), which would enable the classification of some problems, for example, an assertion is a theorem or has a counter model. In order to form clusters it is necessary that the agents have some knowledge of their own and each others abilities.

In our prototype system, a significant bottleneck was the problem of *communicating partial proofs* between the different systems involved — in some cases, the communication time outweighed the actual time needed for proof search. Therefore, one major concern will be to find a concise representation for both problems and proofs in order to communicate them more efficiently. Currently, all communication is routed via a central structure in a uniform format, a higher-order natural deduction calculus. However, this is not necessarily desirable as certain agents may be able to communicate more concisely between each other, for instance, by exchanging sets of clauses. Therefore, we aim at communication languages in which relevant elements of the proof process can be effectively transferred.

4 Resources

We have to *identify the resources* that are limited and that have to be managed in the agent-based system and to develop resource-guided heuristics for spawning new threads within the theorem-prover. Thereby we will consider as resources not only the classical resources such as computation time or memory, but also those special to the domain of automated reasoning. These additional resources have generally to do with logical dependencies between parallel proof attempts. For example, we can consider the instantiation of a variable in the proof as a resource. If one agent wants to instantiate this variable with a particular term it can affect the proof search of other agents. Therefore, the resource would be consumed by the agent which will need to be broadcasted to the other agents, or the agent might even have to negotiate first to obtain the right to instantiate this variable. Another sensible resource may consist in the *size* of the representation of the partial proofs exchanged between the agents.

5 Architecture for Reasoning Agents

With the background of the above requirements for our system we have to investigate what understanding our reasoning agents need to have of themselves and of other agents. The former is necessary for an agent to gain sufficient self-recognition to know its strengths and weaknesses and to judge its abilities to contribute to the solution of a given problem. The latter is generally important for the exchange of information between agents and in particular helpful in negotiating with other agents about forming societies that cooperatively pursue a task. For both cases our agents need to have heuristics to judge themselves and others. These heuristics can be influenced by a performance measure for past runs and which can vary for different application domains.

Therefore, will investigate SimAgent [5], a very flexible toolkit for exploring agent architectures. This will be studied in order to decide which of its features may be useful for the purposes of an agent-based theorem prover. Examples might include mechanisms for controlling resource allocation, mechanisms for interfacing symbolic condition-action rules running within an agent's cognitive system with "lower level" mechanisms treated as "black boxes", and mechanisms for linking agent behaviours with graphical displays. We also want to investigate to which extend already existing networks of mathematical reasoning systems can be integrated or exploited in our framework.

References

[1] Christoph Benzmüller, Lassaad Cheikhrouhou, Detlef Fehrer, Armin Fiedler, Xiaorong Huang, Manfred Kerber, Michael Kohlhase, Karsten Konrad, Erica Melis, Andreas Meier, Wolf Schaarschmidt, Jörg Siekmann, and Volker Sorge. ΩMega: Towards a Mathematical Assistant. In William McCune, editor, *Proceedings of CADE–14*, volume 1249 of *LNAI*, pages 252–255, Townsville, Australia, 1997. Springer Verlag, Germany.

[2] Christoph Benzmüller, Mateja Jamnik, Manfred Kerber, and Volker Sorge. Experiments with an Agent-Oriented Reasoning System. In Franz Baader, Gerhard Grewka, and Thomas Eiter, editors, *KI 2001: Advances in artificial intelligence : Joint German/Austrian Conference on AI*, volume 2174 of *LNAI*, pages 409–424, Vienna, Austria, 2001. Springer Verlag, Germany.

[3] Christoph Benzmüller and Volker Sorge. Ω-Ants – An open approach at combining Interactive and Automated Theorem Proving. In Manfred Kerber and Michael Kohlhase, editors, *Symbolic Computation and Automated Reasoning – The CALCULEMUS-2000 Symposium*, pages 81–97, St. Andrews, United Kingdom, 2001. AK Peters, Natick, MA, USA.

[4] Nicholas R. Jennings and Michael J. Wooldridge, editors. *Agent Technology: Foundations, Applications, and Markets.* Springer Verlag, , Germany, 1998.

[5] A. Sloman and B.S. Logan. Building cognitively rich agents using the sim_agent toolkit. *Communications of the Association for Computing Machinery*, 43(2):71–77, 1999.

[6] Gerhard Weiss, editor. *Multiagent Systems: a Modern Approach to Distributed Artificial Intelligence.* MIT Press, Cambridge, MA, USA, 1999.