ANL-90/28

ANL--90/28

DE91 001097

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue
Argonne, Illinois 60439-4801

THEORY INSTITUTE IN AUTOMATED REASONING

held at Argonne National Laboratory
August 6-10, 1990

Larry Vos

Mathematics and Computer Science Division

August 1990

This work was supported by the Applied Mathematical Sciences subprogram of the Office of Energy Research, U.S. Department of Energy, under Contract W-31-109-Eng-38. Special funding was provided by the Physical Research Program Administration at Argonne National Laboratory.

MASTER

EB

Contents

1	Introduction	1
2	Background and Objectives	2
3	Details of the Institute	3
3.1	Participants	3
3.2	Organization	4
4	Evaluation and Future Work	7

Theory Institute in Automated Reasoning

held at Argonne National Laboratory,

August 6-10, 1990

Abstract

On August 6-10, 1990, Argonne National Laboratory hosted a Theory Institute in Automated Reasoning. The institute was organized by the Mathematics and Computer Science Division and was supported by special funding from Argonne's Physical Research Program Administration. The focus of the Institute was on the obstacles confronting the effective automation of reasoning. The objective was to lay the groundwork for formulating a theory governing the interrelationship of representation, inference rule, and strategy. Here we summarize the activities that took place during the week-long Institute. We also present an evaluation of the progress achieved—progress that includes the solution of challenge questions, the increasing use of both our database of problems and our automated reasoning program OTTER, and the discovery of new uses for OTTER.

1 Introduction

On August 6-10, Argonne National Laboratory hosted a Theory Institute in Automated Reasoning. The Institute was organized by the Mathematics and Computer Science Division and supported by a special grant from Argonne's Physical Research Program Administration.

The Institute gathered some of the world's best researchers in automated reasoning, logic, and mathematics to attack various obstacles confronting the automation of reasoning. The focus of the Institute was on the obstacles confronting the effective automation of reasoning. The objective was to lay the groundwork for formulating a theory governing the interrelationship of representation, inference rule, and strategy. The choices for each materially affect the performance of an automated reasoning program. A deep understanding of this interrelationship would measurably increase the attractiveness and usefulness of automated reasoning programs and contribute markedly toward reaching the goal of producing a self-analytical program.

The structure of the Institute was intentionally kept informal; participants were encouraged to work in small groups to attack a specific problem or to test a new approach. Short presentations were arranged when a researcher wished to summarize his progress, or put before the entire group questions to solicit additional suggestions about an attack, or promote general discussion for an exchange of ideas. This combination of working groups with informal presentations proved extremely successful.

Several important challenge problems in mathematics and logic were answered. New insights were gained about basing the choice of inference rule on the properties of the problem under consideration. Most important, many of the participants have begun experimenting with our automated reasoning program OTTER and are using our database of axioms, lemmas, and test problems to experiment with new ideas and to compare the effectiveness of various automated reasoning programs.

In the remaining sections of this report, we summarize the activities of the Theory Institute in Automated Reasoning. First, we review the objectives of the Institute. We then list the participants, present summaries of the informal presentations, and briefly discuss pertinent issues raised. We conclude with an evaluation of the Institute and suggestions for new avenues of research.

2 Background and Objectives

Over the past decade, the field of automated reasoning has made tremendous strides. Automated reasoning programs have provided valuable assistance in solving a wide variety of problems that include answering open questions from mathematics, designing and validating logic circuits, and proving that computer programs meet their specifications. Nevertheless, significant obstacles remain—deep and challenging theoretical problems that, if unsolved, limit the automation of reasoning [see Ref. 6 for a full discussion of these obstacles].

In particular, researchers do not yet know how to determine which inference rule, strategy, or notation is most effective for attacking a given problem. The difficulty is compounded by the fact that all three elements are tightly coupled: if, for example, the chosen inference rule does not mesh well with the chosen representation, then the likelihood for success may be reduced to zero. Unfortunately, too little is known at this time about the precise nature of this coupling. A principal objective of the Institute, then, was to lay the groundwork for formulating a theory governing the interrelationship of representation, inference rule, and strategy.

Such a theory would provide *metarules*—guidelines for selecting the appropriate representation, inference rule(s), strategy or strategies, transformations for canonicalization (demodulators), and type of information-discarding procedure (subsumption) to be employed. The theory would, moreover, help reduce the severity of at least some of the following obstacles to the automation of reasoning:

- *Clause Retention.* A reasoning program keeps too many deduced clauses (too many conclusions) in its database of information.
- *Inadequate Focus.* The reasoning program gets lost too easily.
- *Redundant Information.* The reasoning program generates the same clause (or proper instances of clauses already retained) over and over again.
- *Clause Generation.* The program draws far too many conclusions, many of which are redundant and many of which are irrelevant even though they are not redundant.
- *Size of the Deduction Steps.* The inference rules do not take deduction steps (steps of reasoning) of the appropriate size.

- *Choice of Transformations for Canonicalization (Demodulators)*. The program does not always use the most effective transformations to rewrite information into a canonical form.
- *Database Indexing*. The program requires too much time to find the appropriate information in its database.

Several existing programs do address individual obstacles. PTP, for example, uses a Prolog-like approach to avoid information retention—but is limited to solving simple problems [Ref. 4]; and RRL makes heavy use of canonicalization to minimize clause retention and redundant information—but is limited principally to problems that involve complete sets of reduction [Ref. 2]. Our objective in the Institute, however, was not to identify which programs work well on a specific class of problem, but to understand *why* certain constructs or strategies work better than others. Because Argonne's program OTTER offers a more general testbed than most automated reasoning systems, many of the participants used the Institute as an opportunity to experiment with OTTER. Their questions and feedback helped pave the way for the development of a uniform theory of reasoning and the eventual design of a powerful and versatile reasoning system.

3 Details of the Institute

3.1 Participants

To retain a workshop atmosphere, we limited the number of participants to 20. The participants were selected from a variety of areas related to automated reasoning and mathematics—in particular, programming languages, set theory, program verification, and databases. The list of participants is given below.

Alan Bundy	University of Edinburgh
Edmund Clarke	Carnegie Mellon University
Tammi Henry	University of Tennessee
Larry Hines	University of Texas
Deepak Kapur	State University of New York at Albany
Matt Kaufmann	Computational Logic
Ken Kunen	University of Wisconsin
Vladimir Lifschitz	Stanford University
Ewing Lusk	Argonne
William McCune	Argonne
Ross Overbeek	Argonne
Dana Scott	Carnegie Mellon University
Mark Stickel	SRI
Rick Stevens	Argonne
Robert Veroff	University of New Mexico
Richard Waldinger	SRI
Steve Winker	Argonne
Larry Wos	Argonne
Hantao Zhang	The University of Iowa

3.2 Organization

With the exception of an introductory presentation by the organizer, L. Wos, no formal presentations were scheduled. Nevertheless, the participants were encouraged to give informal presentations and demonstrations. Highlights of these presentations are given below.

We note that, because our reasoning program OTTER provides a more general testbed than the other special-purpose systems, many of the Institute participants used this week to experiment with OTTER.

Monday, August 6

Larry Wos (Argonne National Laboratory) opened the Institute with an overview of the obstacles to the automation of reasoning, emphasizing the interrelationship between representation, inference rule, and strategy. Examples from Boolean algebra, group theory, and Robbins algebra were used to illustrate the effect on program performance. Wos then presented a series of questions and topics that the Institute participants might consider. He concluded with a discussion of the kernel strategy, a systematic approach for searching for fixed point combinators.

The individual participants then introduced themselves, sketching their goals and research interests.

During the afternoon, Ross Overbeek (Argonne) and Robert Veroff (Univ. of New Mexico) introduced small groups of researchers to the use of Argonne's automated reasoning program OTTER. Veroff handed out a set of examples to be used in conjunction with the *OTTER 2.0 Users Guide* [Ref. 3]. The remainder of the afternoon was spent experimenting with Veroff's test problems. We note that such experimentation is essential to making progress in automated reasoning, and we found it refreshing that the researchers were so willing to try a new system. Indeed, some of the participants returned in the evening to experiment.

Tuesday, August 7

Much of Tuesday morning was devoted to an enthusiastic presentation by Dana Scott (Carnegie Mellon) on the use of the symbolic manipulation program Mathematica. Scott demonstrated how Mathematica (on a Macintosh) successfully handles problems involving functions and their periodicity. His goal is to have researchers combine the use of automated theorem-proving programs with Mathematica.

During the afternoon, Robert Veroff (Univ. of New Mexico) summarized research he and Wos have been conducting during the past year on linked inference rules in which equality plays a small role. (This research was funded by Laboratory Exploratory Research Funds during the past year.) Linked inference rules directly address the obstacle of information retention; these rules mesh well with a strategy that permits the user to define which information is significant. Each participant was given a copy of the preprint "The Linked Inference Principle, I: The Formal Treatment," by R. Veroff and L. Wos, which presents a detailed, formal treatment of the linked inference principle [Ref. 5]. The response to the presentation was excellent; numerous questions were raised that merit further investigation—for example, whether chaining and hyperchaining (to cope with transitivity) can be captured with linked inference, and whether one can control or classify links.

The Institute participants then continued experimenting with OTTER and comparing its performance with other reasoning programs, including RRL (Kapur and Zhang's program, accessed remotely in Texas) and the Boyer-Moore theorem prover. So enthusiastic were the researchers that

several again returned to the Laboratory in the evening to continue the experimentation.

Wednesday, August 8

Alan Bundy (University of Edinburgh) chaired a session on controlling general-purpose reasoning programs, such as OTTER. Discussion focused on advice for assigning weights, choosing inference rules, subdividing clauses, and deciding which axioms to use. Since OTTER is sensitive to the choice of inference rules, weights, and other parameters, this advice is especially useful. Our intention is to use the discussion as a basis for future research focusing on effective choices, which is directly germane to the obstacle of formulating needed metarules. In particular, we intend to follow up on several ideas about rewrite rules, which are directly linked to the formulation of metarules.

Toward the end of the session, Ross Overbeek (Argonne) and Larry Hines (University of Texas) initiated a lively debate about the difficulty of choosing parameters for the reasoning program. They believe they can adapt a sequencing algorithm (currently used to attack problems in structural biology) to have OTTER select the best parameters for attacking simple problems. Other researchers claim that this is a naive view of the obstacles faced with an automated reasoning program. Both groups did agree on two points: (1) to date, no one has *systematically* carried out experiments in which each change of parameter has been carefully recorded and its effects analyzed, and (2) different programs appear better suited for particular classes of problem.

To address the first point, we plan to begin producing "folios" recording such experiments; these would be distributed to the automated reasoning community, perhaps via electronic mail or perhaps as Argonne technical reports. To address the second point, we plan to ask researchers to collect ten problems for each of the leading automated theorem-proving programs (OTTER, RRL, CLAM [1], and others); the problems will include those that run easily, those that stress the program, and those that are still out of reach. Our hope is that the collection of problems will provide insight into the relationship between properties of problems and a program's effectiveness.

Richard Waldinger (SRI) initiated a discussion early Wednesday afternoon on the value of paramodulation and negative paramodulation. A program using paramodulation treats equality as "understood"; with negative paramodulation, the program reasons from inequalities [see Ref. 7]. Waldinger has generalized negative paramodulation and other inference rules, by replacing certain monotonicity axioms. We naturally found his comments pleasing, since we formulated these rules. His approach also provides a justification for the type of generalized rewriting that researchers are doing with the theorem prover CLAM. A. Bundy, after the completion of Waldinger's talk, led a discussion on inductive inference. Among the questions discussed were the following: (1) How can we automatically suggest lemmas needed for proofs? (2) How can we automatically suggest generalizations needed for proofs? (3) Is it better to use explicit induction or inductive completion? (4) Can we extend our techniques to existential theorems? Bundy explained how meta-variables can be used, and discussed the benefits of "rippling." Kapur suggested that inductive completion could be used to consistency checking, but he has found explicit induction to be better suited for proving interesting theorems.

There was also a special session during the afternoon, led by Veroff, to introduce Dana Scott and Edmund Clarke (both of Carnegie Mellon) to the use of OTTER. We found that small sessions like this one were more effective than group lessons.

Thursday, August 9

Ewing Lusk (Argonne) led a session on the use of parallel-processing computers for theorem

proving. He and Mark Stickel (SRI) are working on parallelizing a small automated theorem-proving program called *PTTP*. The basic procedures of automated reasoning—drawing a conclusion, rewriting each conclusion into canonical form, testing its significance, and comparing it with other information—appear well suited for parallel treatment.

A new tool for theorem proving was reported by William McCune (Argonne). He and a graduate student (T. Henry) have designed a graphical interface for theorem provers that allows a set of formulas to be created, updated, edited, and transformed in a X window environment. The interface, called *FORMED*, allows one not only to see what occurs at each step, but to try different orderings of the transformations, observing any changes in the clause formation. McCune and Henry demonstrated the use of *FORMED* in conjunction with *OTTER*.

A highlight of the day was the successful solution of a challenge problem involving disjunction, issued by Dana Scott Thursday morning. Steve Winker (Argonne) first solved the problem using certain “tricks”; Ken Kunen (Univ. of Wisconsin) with Winker then tried the problem using paramodulation and obtained a proof in 20 seconds with *OTTER* on a Sun workstation.

Friday, August 10

Continuing the successful attack on specific problems, several of the Institute participants began work on a problem that J. Kalman has been investigating. The problem involves determining the nature and size of a particular formula in formal logic. Kalman has obtained preliminary results—but only by *manually* manipulating the results in conjunction with *OTTER*; his objective was, of course, to obtain a completely automated proof. By Thursday evening, Deepak Kapur found a way to “order the equalities” such that *RRL* could obtain a proof in two runs, in a total of 40 CPU minutes. Early Friday, McCune and Lusk adapted Kapur’s approach and obtained a completely automated proof requiring only 1-1/2 minutes (still in two runs).

Edmund Clarke (Carnegie Mellon Univ.) presented his work with *Mathematica*. He has developed a program called *Analytica*, which combines theorem proving with *Mathematica*. Like his colleague Scott, Clarke feels strongly that algebraic systems like *Mathematica* work well with problems involving induction.

Wos closed the Theory Institute by highlighting the different issues raised and challenges remaining. He noted that a variety of programs now exist for attacking problems involving reasoning, and that new programs such as the Argonne graphics editor promise to make the use of such programs easier for the mathematician and logician. The most encouraging note is the willingness of researchers to experiment with different approaches and existing programs.

4 Evaluation and Future Work

The success of the Theory Institute can be measured by the following:

1. New attention to metarules. Bundy, for example, has provided a counterexample to Overbeek’s suggestion about parameter selection using *OTTER*’s default values.
2. Increased use of our database of axioms, puzzles, and test problems. Zhang and Kaufmann, for example, are using the database for experimentation.

3. Increased accessibility of our reasoning program. Kaufmann has written preliminary notes about how to use OTTER—what to do and what not to do. We plan to develop these notes into a tutorial, which will complement the OTTER Users Guide.
4. New applications of our automated reasoning systems. Scott prepared an ordered list of 68 theorems in sentential calculus and asked us to try these with our automated reasoning programs. Results are encouraging: using hyperresolution and OTTER, McCune obtained proofs of 33 of the theorems; using a parallel automated reasoning program, McCune then proved 48; and using weighting and OTTER, Wos obtained proofs for all 68, in 16 minutes on a Sun workstation.
5. Active experimentation. Bundy is close to completing a proof, using OTTER, for Scott's challenge problem. Scott has prepared additional challenge problems from a variety of fields.

We also drew several conclusions about future research directions:

- Systematic experimentation is needed to determine which parameters are most effective (e.g., whether a maximum weight or non-standard weight is of much value, which demodulators to include in the input, or whether to use back subsumption).
- A standard set of problems is needed to compare schemes. Specifically, we need to evaluate which general-purpose schemes do best and which domain-specific schemes can be applied to other domains.

References

1. Bundy, A., van Harmelen, Frank, Horn, Christian, and Smaill, A., "The OYSTER-CLAM System," *Proc. 10th International Conference on Automated Deduction, in Lecture Notes in Artificial Intelligence, Vol. 449*, ed. M. E. Stickel, Springer-Verlag, Berlin, 1990, pp. 647-648
2. Kapur, D., and Zhang, H., "RRL: Rewrite Rule Laboratory User's Manual," Department of Computer Science, University of Iowa, Report 89-03, 1989
3. McCune, W. W., *OTTER 2.0 Users Guide*, ANL-90/9 (March 1990)
4. Stickel, Mark E., "A Prolog Technology Theorem Prover: Implementation by an Extended Prolog Compiler," *J. Automated Reasoning* 4 (4), 1988, 353-380
5. Veroff, R., and Wos, L., "The Linked Inference Principle, I: The Formal Treatment," MCS-P160-0690
6. Wos, L., *Automated Reasoning: 33 Basic Research Problems*, Prentice-Hall, Englewood Cliffs, NJ, 1988
7. Wos, L., Overbeek, R., Lusk, E., and Boyle, J., *Automated Reasoning: Introduction and Applications*, Prentice-Hall, Englewood Cliffs, NJ, 1984

- END -

DATE FILMED

11 / 16 / 90