Guiding Linear Deductions with Semantics

Thesis submitted by
Marianne Elizabeth Brown BSc(Hons) *JCU*
December 2003

for the Degree of Master of Science
in the School of Information Technology
James Cook University

Statement on Access

I, the under-signed, the author of this work, understand that James Cook University will make this thesis available for use within the University Library and, via the Australian Digital Theses network, for use elsewhere.

I understand that, as an unpublished work, a thesis has significant protection under the Copyright Act and I do not wish to place any restriction on access to this thesis.

_____ _____

Signature Date

Declaration

I declare that this thesis is my own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

_____ _____

Signature Date

Electronic Copy

I, the undersigned, the author of this work, declare that the electronic copy of this thesis provided to the James Cook University Library is an accurate copy of the print thesis submitted, within the limits of the technology available.

_____ _____

Signature Date

Acknowledgements

I would like to thank Dr. Geoff Sutcliffe, who was my supervisor for the first three years, and Dr. Bruce Litow, who was my supervisor for the last three years. This thesis would not have been finished without their support. I would like to also acknowledge the support and motivation provided by the staff at the School of Information Technology including Dr. Greg Allen and Mrs. Beverley Frangos. To my fellow research students Alan McCabe, Chris Christensen and Jarrod Trevathan, thanks for the coffee and conversations. And finally to family and friends, who for the most part, tried to avoid the inevitable question, “So, how’s the thesis going?”

Abstract

Guidance is a central issue in Automatic Theorem Proving systems due to the enormity of the search space that these systems navigate. Semantic guidance uses semantic information to direct the path an ATP system takes through the search space. The use of semantic information is potentially more powerful than syntactic information for guidance. This research aimed to discover a method for incorporating semantic guidance into linear deduction systems, in particular model elimination based linear systems. This has been achieved. The GLiDeS pruning strategy is a simple strategy of restricting the model elimination deduction to one where all A-literals are false in the guiding model. This can be easily incorporated into any model elimination based prover. Evaluation of the GLiDeS strategy has shown that when "good guidance" has been achieved, the benefit of this guidance is significant. However attempts to develop a heuristic for predicting which model will provide "good guidance" has been largely unsuccessful.

Original Contributions

1. Developed novel strategy (GLiDeS) for applying semantic guidance to full linear deduction systems.
2. Shown that the new GLiDeS strategy is sound but incomplete.
3. Shown that GLiDeS is complete for a small group of problems termed *Semantic Horn* and that this result is essentially equivalent to renaming [Slagle, 1967].
4. Implemented system to demonstrate ease of including GLiDeS into an existing linear theorem proving system, PTTP.
5. Evaluated performance of the GLiDeS semantic guidance strategy and concluded that overall the GLiDeS strategy does not provide significant improvement to PTTP's performance.
6. It has been shown that when good guidance is achieved the improvement in performance is significant. GLiDeS dramatically reduces the amount of search space covered before a proof is found (as reflected by the number of inferences made). In the best case, PTTP covered on average 8 times the search space that GLiDeS covered (See NHN_SEQ Table 6.7).

Material from this thesis has appeared in the following publications:

- M. Brown. Selecting Semantics for Use with Semantic Pruning of Linear Deductions, In McKay,Bob and Slaney, J. editor, *AI 2002: Advances in Artifical Intelligence. 15th Australian Joint Intelligence Canberra, Australia, December 2002 Proceedings*, number 2557 in LNAI. Springer-Verlag, Berlin Heidelberg, Germany, 2002.
- M. Brown and G. Sutcliffe. PTTP+GLiDeS - Semantically Guided PTTP, In D. McAllester, editor, *Automated Deduction - CADE-17: 17th International Conference on Automated Deduction, Pittsburgh, PA, USA, June 17-20, 2000 Proceedings*, p 719, number 1831 in LNAI. Springer-Verlag, New York, USA, 2000.

- M. Brown and G. Sutcliffe. PTTP+GLiDeS: Guiding Linear Deductions with Semantics. In N. Foo, editor, *Advanced Topics in Artifical Intelligence: 12th Australian Joint Conference on Artificial Intelligence, AI'99*, number 1747 in LNAI, pages 244–254. Springer-Verlag, 1999.

Contents

1 Introduction and Technical Notation **1**

1.1 Background 2

1.2 The Need for Semantic Guidance 3

1.3 Research Objectives 3

1.4 Notation and Terminology 4

1.4.1 Basic Terminology 4

1.4.2 Notation 5

1.5 Thesis Structure 7

2 Linear Deduction Systems **8**

2.1 Ancient History 8

2.2 Prolog Technology Theorem Provers (PTTP) 15

2.3 Guidance Strategies Employed by Linear Deduction Systems 20

2.4 Summary 21

3 Semantic Guidance Strategies **24**

3.1 Early Resolution Systems 24
3.2 Linear Systems 26
3.3 Modern ATP systems 27
3.4 Model Generation 29
3.5 Summary 30

4 Guiding Linear Deductions with Semantics 31

4.1 Theory 31
4.1.1 Formal Notation 32
4.1.2 Completeness and Soundness 35
4.2 GLiDeS System 40
4.3 Summary 40

5 Model Generation and Selection 42

5.1 Model Generation 42
5.2 Model Selection Heuristic 43
5.3 Implementation and Performance 45
5.3.1 Initial Implementation 46
5.3.2 Final Implementation 50
5.4 Some Examples 55
5.4.1 PUZ014-1 and PUZ013-1 55
5.5 Summary 61

6 PTTP+GLiDeS Implementation and Performance 62

6.1 Initial Implementation 62

6.1.1 Performance 64

6.2 Ordering of Clauses 67

6.2.1 Performance of GLiDeS+Model selection heuristic 71

6.3 Performance of GLiDeS + Model selection heuristic across TPTP 72

7 Conclusion 77

7.1 Summary of objectives 77

7.2 Future Work 79

A Semantic Checking 85

List of Figures

1.1 Two different visual representations of the same Model Elimination deductions: (a) traditional vertical representation, (b) tableau style representation. 6

2.1 Resolution proof for Example 2 12

2.2 Search tree for linear resolution to depth 6 13

2.3 Search tree for SL-resolution to depth 6 14

2.4 Search tree for Linear input resolution to depth 7 15

2.5 Model Elimination deduction for $S = \{\mathtt{p} \vee \mathtt{q}, \sim\mathtt{p}, \sim\mathtt{q} \vee \mathtt{r}, \sim\mathtt{r} \vee \mathtt{s}, \mathtt{p} \vee \sim\mathtt{t}, \sim\mathtt{r} \vee \sim\mathtt{s} \vee \mathtt{t}\}$ 16

2.6 Search tree for Model Elimination to depth 6 16

2.7 Two ME linear deductions (a) using identical ancestor pruning and (b) without identical ancestor pruning 22

2.8 Two ME tableau deductions showing (a) a regular tableau and (b) a tableau without regularity 23

4.1 Tableau ME proof for the clause set $\{\mathtt{p} \vee \mathtt{q}, \mathtt{p} \vee \sim\mathtt{q}, \sim\mathtt{p} \vee \mathtt{q}, \sim\mathtt{p} \vee \sim\mathtt{q}\}$. 32

4.2 non-GLiDeS proof with regularity 37

4.3 GLiDeS proof without regularity. Internal nodes that violate regularity are highlighted 37

4.4 A GLiDeS deduction for MSC006-1. 38

4.5 An ME deduction for MSC006-1. 39

4.6 Architecture of the PTTP+GLiDeS system. 41

5.1 Tableau ME proof for PUZ013-1 60

5.2 Tableau GLiDeS ME proof for PUZ013-1 60

List of Tables

5.1 Results of the best and worst models in the 14 cases where there was a significant difference in the outcomes. 47

5.2 Results of the best and worst models in the 18 cases where there was a significant difference in the outcomes. 51

5.3 Model clause for PUZ013-1 and PUZ014-1 58

5.4 Four models generated by MACE for PUZ014-1 59

6.1 Summary of experimental data. 65

6.2 Results for problems where semantic guidance rejected some inferences. Non-Horn problems are marked with "$*$". 66

6.3 Summary of experimental data. 68

6.4 Comparison of results using no ordering, ascending ordering, and descending ordering based on the "trueness" rating of each clause. 69

6.5 Summary of experimental data. 72

6.6 Results for problems solved by either PTTP or PTTP+GLiDeS where a model was generated and GLiDeS had some effect 73

6.7 Summary of results for PTTP and PTTP+GLiDeS across all 7 SPCs for unsatisfiable CNF problems from the TPTP Library v2.4.1 75