Automatic Heuristic Selection, on a Problem by Problem Basis, Using an Analytical Model and *In Situ* Sampling

Author: Santiago Franco

Supervisor: Dr. Mike Barley

The Department of Computer Science

The University of Auckland

New Zealand

March 2012

A thesis submitted in fulfillment of the requirements of Doctor of Philosophy in Computer Science

Abstract

One of the main drives in automated planning is to automate the translation from the knowledge level description of the problem to the symbol level implementation of a solver for that problem. The field is at the point where there are techniques that can automatically synthesize very large numbers of heuristics that can be used by heuristic search based problem solvers [Holte *et al.*, 2005], [Edelkamp, 2006], [Culberson and Schaeffer, 1994], [Haslum *et al.*, 2007], etc. However, currently there is no *a priori* method available to automatically tailor them to a specific problem instance. Nor are there any automated techniques to automatically determine which heuristics are best for a problem. There are some automated techniques to generate problem specific heuristics based on *in situ* gathered information but they are not competitive with state-of-the-art manual heuristic selection.

Currently, the state-of-the-art is for the user to determine this *a priori*. This is usually done by running each of the heuristics on a wide range of problems in the chosen domain. The heuristic that is best on average for the domain is then used for each problem in the domain. The current approach is empirical, this thesis explores an analytical approach; which uses a simplified run-time formula that represents the important components of the problem-solver. This formula predicts the impact of the selected heuristic subsets upon run-time. Some of the formula's parameters can be determined *a priori* (e.g., how long it takes to check for a goal state) while others can only be determined *a posteriori*. Unfortunately, finding out those values after the problem has been solved is not usually very useful.

This thesis explores the possibility of using the early parts of the problem-solving task(*in situ* sampling) to determine approximate values for those parameters. A system called RIDA* was created to do this. The danger is that the cost of determining these approximate values may outweigh the benefits those values bring. The ideal situation is to only do as much extra work as is needed to obtain the approximate values that allow us to make a decision that saves enough work to compensate for the time invested.

We found that RIDA* was competitive, and in some cases significantly faster, than manual heuristic selection. However, there are two caveats: RIDA* has a scalability issue for large heuristic sets due to the combinatorial explosion in the heuristics search space. Also RIDA*'s *in situ* sampling effort is currently determined *a priori* for the domain, it does not have an *in situ* mechanism to adjust its sampling effort. Both of these caveats are the subject of future directions for research. This thesis work(RIDA*) significantly advances the automation of *in situ* heuristic selection.

Finally, RIDA* uses new data structures which significantly reduce the costs associated with *in situ* sampling. Significantly reducing *in situ* sampling costs is critical for any technique using problem-specific heuristics to be competitive with state-of-the-art manual heuristic selection.

To my mother, Maria Reyes Aixela Barasona, wife, Susan Elizabeth Quantrell and my two beautiful children, Rodrigo and Megan.

Acknowledgement

It is with great pleasure that I am able to take this opportunity to express my sincere appreciation to my supervisor, Dr. Mike Barley, to whom I will always be indebted for his guidance, wisdom, dedication to my work, and, most of all, endless support.

I must also express my gratitude to the Computer Science Department at the University of Auckland. Whenever I needed any help, all the administrative and IT support staff were helpful and friendly. I would also like to thank my colleagues in the postgraduate office, always ready for a good nature chat whenever the thesis proved to much, or a bit of LaTeX advice that saved me hours of online search (Thanks Al!).

I would like to mention that I am very thankful to the University of Auckland for fully financing my doctoral studies through a Doctoral Scholarship.

Also thanks to my dear friend, Dr. Clara Palleja-Lopez, whose friendship, advice and support helped me finishing this thesis.

Finally, I must express my deepest gratitude to my wife, Susan Elizabeth Quantrell, whose love and support helped me to never give up and finish this thesis.

Contents

List of Figures

List of Tables