Progress Report Towards PhD Degree, Year Two

Research Topic: Rational Metareasoning in Problem Solving Search

David Tolpin

Department of Computer Science

Ben-Gurion University of the Negev

Under the supervision of Professor Solomon Eyal Shimony

Author's Signature: _____

Advisor's Signature: _____

Chairman's Signature: _____

September 20, 2011

1 Research Progress

During the second year of my Ph.D. studies I worked on rational deployment of search heuristics (Section 1.1), as well as on improvements to Monte Carlo sampling algorithms for tree search (Section 1.2). I also prepared revisions of two journal papers [4], [1] for publication later in 2001.

1.1 Rational Deployment of CSP Heuristics

We proposed a rational metareasoning approach to decide when and how to deploy heuristics, using CSP backtracking search as a case study. The heuristics examined are various *solution count estimate* heuristics for value ordering, which are expensive to compute, but can significantly decrease the number of backtracks. These heuristics make a good case study, as their overall utility, taking computational overhead into account, is sometimes detrimental; and yet, by employing these heuristics adaptively, it may still be possible to achieve an overall runtime improvement, even in these pathological cases. Following the metareasoning approach, the value of information (VOI) of a heuristic is defined in terms of total search time saved, and the heuristic is computed such that the expected net VOI is positive.

As a result of the research [3], we suggested a model for adaptive deployment of value ordering heuristics in algorithms for constraint satisfaction problems. As a case study, the model was applied to a value-ordering heuristic based on solution count estimates, and a steady improvement in the overall algorithm performance was achieved compared to *always* computing the estimates, as well as to other simple deployment tactics. The experiments showed that for many problem instances the optimum performance is achieved when solution counts are estimated only in a relatively small number of search states.

The methods introduced in this paper can be extended in numerous ways. First, generalization of the VOI to deploy different types of heuristics for CSP, such as variable ordering heuristics, as well as reasoning about deployment of more than one heuristic at a time, are natural non-trivial extensions. Second, an explicit evaluation of the quality of the distribution model is an interesting issue, coupled with a better candidate model of the distribution. Such distribution models can also employ more disciplined statistical learning methods in tandem, as suggested above. Finally, applying the methods suggested in this paper to search in other domains can be attempted, especially to heuristics for planning. In particular, examining whether the meta-reasoning scheme can improve reasoning over deployment of heuristics based solely on learning methods is an interesting future research issue.

The paper was presented at the 11th International Joint Conference on Artificial Intelligence, Barcelona, Spain.

1.2 Rational Monte Carlo Search

UCT, a state-of-the art algorithm for Monte Carlo tree sampling (MCTS), is based on UCB, a sampling policy for the Multi-armed Bandit Problem (MAB) that minimizes the accumulated regret. However, MCTS differs from MAB in that only the final choice, rather than all arm pulls, brings a reward, that is, the simple regret, as opposite to the cumulative regret, must be minimized. This ongoing work [2] aims at applying meta-reasoning techniques to MCTS, which is non-trivial. We introduced policies for multi-armed bandits with lower simple regret than UCB, and an algorithm for MCTS which combines cumulative and simple regret minimization and outperforms UCT. We also developed a sampling scheme loosely based on a myopic version of perfect value of information. Finite-time and asymptotic analysis of the policies was provided, and the algorithms were compared empirically. [2]

2 Publications

References

[1] David Tolpin and Solomon Eyal Shimony. Rational value of information estimation for measurement selection. In *URPDM*, 2010.

[2] David Tolpin and Solomon Eyal Shimony. Doing better than uct: Rational monte carlo sampling in trees. *CoRR*, abs/1108.3711, 2011.

[3] David Tolpin and Solomon Eyal Shimony. Rational deployment of csp heuristics. In *IJCAI*, pages 680–686, 2011.

[4] David Tolpin and Solomon Eyal Shimony. Semi-myopic measurement selection for optimization under uncertainty. *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, 2011, to appear.