Comparisons and Evaluations of Planners

Signposts for the AAAI-94 Workshop

Austin Tate
Artificial Intelligence Applications Institute
The University of Edinburgh
80 South Bridge, Edinburgh EH1 1HN, UK

There are a number of research activities which are contributing to improved communication, classification, comparison and evaluation of work on AI planners. There is a closing of the gap between formal approaches to characterising and analysing planning algorithms and practically motivated planners (many of which use detailed knowledge of the application domain in their problem solving). Drummond [4] notes that, until recently, the formal community have concentrated on goal achievement methods whereas the practical planning community have used a variety of other methods many of which are worthy of more detailed study.

We are therefore moving out from the narrower study of the truth criterion for goal achievement in planning to consider other matters which have been important in practical planning systems for many years:

- hierarchical and abstraction level models of planning domains.
- domain knowledge and constraints modelling, pre-compilation and use.
- Search control and opportunism.

The papers in this workshop represent a number of the most important trends in comparing and evaluating planning systems. Here are some themes and signposts to begin the process of integrating many of the ideas now emerging.

Generic Frameworks to Describe Actual Planners

1. The reviews in [1] and the tutorial in [17] give categories of planning systems and provide definitions of the technical features within them.
2. The description of "standard" planner components and the Plan Modification Operator cyclic processing abstraction in [14].
3. The emerging KADS model of planning as a generic task [18] using Plan and World Description features, Function Structure, Control Structure, etc.
4. The framework for planning application analysis in [6].

Simplifying Frameworks for Reconstruction of Planning Algorithms

1. The Refine-Plan framework (termination criteria, Plan Modification Operator selection method, bookkeeping method and tractability refinement methods) in [7].
2. Extensions to this framework for Hierarchical Task Network (HTN) planning [2],[8],[9],[19].
3. Extensions to this framework to allow an explicit agenda of outstanding issues in the plan [15].
4. The terminology of the <I-N-OVA> constraint model of plans [15].
5. The approach to planning as propositional satisfiability described in [10].

Empirical Comparative Studies

1. Comparative studies in which variations of approach and technique are made within the same planner implementation, e.g. [7],[11].
2. Recent empirical evaluations of aspects of practical planning systems, e.g. those conducted by Pollack and her colleagues:
    - Joslin and Pollack [5] – study of the merits of using an opportunistic estimator to prioritise the order in which issues (flaws) are processed in a planner.
    - Young and Pollack [20] – study of the value of delaying certain types of condition satisfaction during planning.
3. Studies of the value of using forward checking constraint satisfaction to prune search in scheduling problems [3].

Common Plan Representations and Ontologies

1. The ARPA/Rome laboratory Planning Initiative KRSL Plan Ontology [12].
2. The terminology of the <I-N-OVA> constraint model of plans [15].

Integrated Systems

There is a greater level of interest in planners and schedulers which are situated within a command, planning and execution environment. Planners are only as good as the plans they generate. With this in mind, Zilberstein [13] calls for evaluations of planners to be related to the actual execution or use of the plans generated. Crawford and Baker begin to relate empirical studies of scheduler performance to the characteristics of "real" scheduling problems.

References

[1] Readings in Planning (eds. J.Allen, J.Hendler and A.Tate), Morgan-Kaufmann, 1990.

[2] Barrett, A. and Weld, D.S., Task-Decomposition via Plan Parsing, Proceedings of the Twelfth National Conference on Artificial Intelligence (AAAI-94), Seattle, USA, 1994.

[3] Crawford, J.M. and Baker, A.B., Experimental Results on the Application of satisfiability Algorithms to Scheduling Problems, Proceedings of the Workshop on Comparative Analysis of AI Planning Systems, AAAI-94, Seattle, USA, 1994.

[4] Drummond, M.E., On Precondition Achievement and the Computational Economics of Automatic Planning, in Current Trends in AI Planning (eds. C.Backstrom and E.Sandewall) IOS Press, Sweden, 1993.

[5] Joslin, D. and Pollack, M.E., Least-Cost Flaw Repair: A Plan Refinement Strategy for Partial-Order Planning, Proceedings of the National Conference on Artificial Intelligence (AAAI-94), Seattle, WA, USA, 1994.

[6] Gil, Y. and Linster, M., On Analyzing Planning Applications, in preparation, USC/ISI and Digital Equipment Corporation, 1994.

[7] Kambhampati, S., Design Tradeoffs in Partial Order Planning, Proceedings of the Second International Conference on AI Planning Systems (AIPS-94), Chicago, IL., USA, 1994.

[8] Kambhampati, S., Comparing Partial Order Planning and Task Reduction Planning: a Preliminary Report, Proceedings of the Workshop on Comparative Analysis of AI Planning Systems, AAAI-94, Seattle, USA, 1994.

[9] Kambhampati, S. and Hendler, J., A Validation-Structure-Based Theory of Plan Modification and Reuse, Artificial Intelligence, May, 1992.

[10] Kautz, H. and Selman, B., An Empirical Evaluation of Knowledge Compilation, National Conference on AI, Seattle, USA, August 1994.

[11] Knoblock, C.A. and Yang, Q. Evaluating the Tradeoffs in Partial-Order Planning Algorithms, Paper presented to the AAAI-94 Workshop on Comparative Analysis of AI Planning Systems, Seattle, USA, July 1994.

[12] Lehrer, N.B. (ed.) Knowledge Representation Specification Language (KRSL), Version 2.0.2, February 1993, USAF Rome Laboratory.

[13] , Zilberstein, S., The Utility of Planning: How to Evaluate and Compare Planning Systems, Paper presented to the AAAI-94 Workshop on Comparative Analysis of AI Planning Systems, Seattle, USA, July 1994.

[14] Tate, A., The Emergence of “Standard” Planning and Scheduling System Components, in Current Trends in AI Planning (eds. C.Backstrom and E.Sandewall) IOS Press, Sweden, 1993.

[15] Tate, A., Representing Plans as a Set of Constraints - the <I-N-OVA> Model, to appear in ACM SIGART Notes, January 1995. Available by ftp from aiai.ed.ac.uk in file bat/inova.ps.

[16] Tate, A., Drabble, B. and Kirby, R., O-Plan2: an Open Architecture for Command, Planning and Control, in Intelligent Scheduling, (eds, M.Zweben and M.S.Fox), Morgan Kaufmann Publishers, Palo Alto, CA., USA, 1994.

[17] Tate, A., Hendler J.A. and M.E.Drummond, AI Planning: Systems and Techniques, AI Magazine Vol. 11 No. 2, pp. 61-77, Summer 1990, American Association for Artificial Intelligence.

[18] Valente, A., Knowledge-Level Analysis of Planning Systems, Proceedings of the Workshop on Comparative Analysis of AI Planning Systems, AAAI-94, Seattle, USA, 1994.

[19] Yang, Q., Formalizing Planning Knowledge for Hierarchical Planning, Computational Intelligence, Vol. 6, No. 1, pp12-24, 1990.

[20] Young, R.M., Pollack, M.E. and Moore, J.D., Decomposition and Causality in Partial-Order Planning, Proceedings of the Second International Conference on AI Planning Systems (AIPS-94), Chicago, IL, USA, 1994.