The Third International Planning Competition: Temporal and Metric Planning

Maria Fox and Derek Long
University of Durham, UK
maria.fox@durham.ac.uk   d.p.long@durham.ac.uk

Introduction

The International Planning Competitions, run by Drew McDermott in 1998 and Fahiem Bacchus in 2000, have provided an important spur to the planning community, encouraging the development of planning technology and of a wide selection of planning bench mark domains. One of the most important contributions has been the introduction of a widely accepted standard for domain description, Drew McDermott's PDDL language (McDermott 2000), leading to the sharing of domains and planning systems. An equally important outcome, which has been extremely beneficial to the community, has been that planning research has made rapid progress in the four years since the competitions began and has risen to many interesting challenges.

A complaint that has been often levelled at competitions in various research communities (theorem proving, natural language understanding, etc.,) is that they have tended to encourage a focus on winning for its own sake, rather than on tackling real problems that might have longer term interest for application. Competitions can have the negative effect of turning developers' attention inwards so that systems are honed on artificial benchmark problems at which they can excel, rather than extended to meet real challenges. A consequence of this is that potential participants can be put off taking part in the competition if their technology is not tailored for efficient solution of such problems. The community is at risk of being deprived of seeing exciting and adventurous new developments whilst simple and unscaleable approaches occupy the limelight. Although this has not yet become a serious issue for the planning competitions it is a real danger and it is important to anticipate the danger and try to ensure that the competition remains relevant to the wider planning community.

With these points in mind we decided to focus the 2002 competition on planning with temporal and metric domains. It is generally agreed that temporal modelling and reasoning is essential for the application of planning to practical problems, and certain high-profile application areas, such as space and aerospace applications, have encouraged developers to turn their attention towards these issues. Although some developers have already considered the management of temporal constraints in their planning systems (Laborie & Ghallab 1995; Muscettola 1994; Currie & Tate 1991) this has not been widespread and there has been little agreement over the modelling of time. Certainly there have been no commonly accepted standards for modelling temporal planning domains and no temporal bench mark problems.

This document briefly introduces the objectives of the 2002 competition and the extensions we have made to PDDL to support temporal modelling. Finally, we outline the structure of the competition, in terms of the tracks being run and the domains being considered in those tracks.

Competition Objectives

There has always been a wide gap between basic research and application in the planning research community. Whilst basic researchers have developed elegant algorithms, and heuristics, for intelligent navigation of the combinatorial search spaces underlying (for the most part) propositional planning problems, the applications-led part of the community has confronted the need to model out much of the search involved in solving complex realistic problems involving the integration of many heterogenous constraints. Although both sides of the community have made important contributions in planning their contributions have diverged because they are concerned with solving such different problems. This has been evident in the fact that PDDL has been seen as a language for basic research and not for application. Application-led planning systems tend to have their own custom-built languages, with features known only to their developers, making it difficult, or impossible, to compare their performance on bench mark problems with that of PDDL-based planners.

A particular challenge we faced was to make PDDL expressive enough to begin to bridge this gap. For a domain description language to be a successful standard it has to meet the needs of all parts of the community. Having such a standard will make it possible for more application-driven problems to make it into the bench mark domain collections, and having challenging bench mark problems drives research forward and avoids the complacency that arises from concentrating on successfully solving easy problems. In extending PDDL we have focussed on temporal and metric modelling – there are many other challenges that must be addressed before it can be said that the language meets the needs of all parts of the community but we think that the extensions we have made constitute an important first step.

We hope that the competition will precipitate a burst of activity in temporal and metric planning which will help to close the gap between the theory and application of planning and to raise the expectations and ambitions of the planning research community.

PDDL2.1

In order to support the objectives outlined above we developed a series of extensions to the domain description standard PDDL. These mainly comprise temporal extensions – in particular the ability to express *durative actions* supporting both discrete and continuous temporal modelling. The most adventurous temporal extension we made models actions with delayed effects in terms of the instantaneous initiation and termination of *processes*, active on numeric state over continuous intervals of time. This extension constitutes such a radical departure from the commonly adopted temporal modelling approaches in which actions, rather than state, have duration that there has not been time for it to be sufficiently widely adopted to justify its use in the competition. However, this extension, which we refer to in the following paragraphs as level 5, we consider to be a philosophically accurate model of how state changes over time. It has a direct relationship with timed hybrid automata theory (Henzinger 1996), which we used to provide the formal semantics, making it capable of modelling complex domains with discrete (logical) and continuous (numeric) elements.

In addition to providing several temporal extensions we tidied up some of the syntax already present in PDDL for modelling metric quantities such as consumable resources. The extended language is called PDDL+.

As mentioned above, PDDL+ comprises five levels, the first four of which constitute the officially agreed extensions to PDDL for use in the 2002 competition. We refer to these four levels as PDDL2.1. The full details of PDDL2.1 and PDDL+ are given in (Fox & Long 2001b). Level 5 and its formal semantics are described in (Fox & Long 2001a).

The five levels of PDDL+ are organised in the following way. Level 1 corresponds to the propositional and ADL levels of McDermott's PDDL. Nothing new is added or changed, ensuring backward compatibility with the previous language standard. Level 2 supplements level 1 with numeric variables and the ability to test and update their values instantaneously, and represents a minor revision of McDermott's proposals (McDermott 2000). Levels 3 and 4 contain *durative actions* – actions the effects of which occur some time after the instant of their execution. Level 3 durative actions discretize time, whilst level 4 durative actions can have continuous effects. Because durative actions allow only restricted access to the states of the variables they affect, level 4 simplifies the modelling of real time domains. Level 5 is a natural extension of level 4 and presents a way forward for the design of an expressive planning domain description language, capable of representing arbitrary real time, mixed discrete/continuous domains. All levels constitute completed languages with formal semantics and provide a strong foundation for further development of the PDDL sequence.

One of our objectives in extending PDDL has been to supply a means by which more interesting plan metrics than plan length can be used to judge the value of a plan. The use of numbers in a domain provides a platform for measuring consumption of critical resources and other parameters. An example of a metric we can now model is that fuel use must be minimized or that overall execution time must be minimized, or both (and other such mixed-objective examples).

Durative Action Semantics

A very important extension introduced in the durative actions modelling at levels 3 and 4 is the ability to associate conditions and effects with the end points of durative actions. A common approach (Smith & Weld 1999; Haslum & Geffner 2001; Bacchus & Ady 2001) to durative action modelling has been to treat them as indivisible units with preconditions that must remain true throughout their intervals of execution and effects that are undefined until these intervals are complete. This modelling unnecessarily prohibits much of the interesting concurrency that might be exploited in a domain. When conditions and effects are associated with the two end-points there is more scope for interleaving of actions and thereby exploiting concurrency. Conditions that must in fact remain true throughout the interval of execution of an action can be handled separately as invariant conditions, providing greater flexibility in the modelling process.

The view of a durative action having end points derives from our philosopical position on whether actions or states have duration. In our view, logical change must be instantaneous and only numeric values can be updated continuously. When a durative action is applied it changes logical state and initiates a period of time within a state that might be subject to numeric update. When that period has elapsed a new logical state is entered. Our durative action syntax at levels 3 and 4 are simplified versions of level 5 in which time is discretized in various ways. We see levels 3 and 4 as progressive stages in the development towards level 5, not as alternatives to the modelling approaches used in level 5. Of course, the implementation of these semantics is a matter for the planners handling levels 3 and 4.

Associated with the semantics of the actions is the question of plan validity. When the view is taken that actions have two instantaneous end points, decarmating a period of time in a state, the question arises under what circumstances two actions can overlap. In a direct implementation of our semantics two actions with mutually inconsistent end-points can overlap if their end points do not occur simultaneously. A plan that allows two conflicting end points to co-occur is therefore invalid, whilst one that allows a tiny fraction of time to elapse between the occurrences of these end points is valid. The way we determine the size of this fraction of time is to provide it as part of the domain specification and to validate plans by ensuring that at least this fraction is used to separate mutually exclusive action end points. This complication, which is a natural consequence of the semantics of concurrent activity, modifies the construction of valid plans in a way that must be observed by all competing planners.

Tracks and Domains

As in previous competitions we have operated two tracks in the competition: the *fully automated* and *hand-coded* tracks. Within each of these tracks we have operated five different levels of complexity from STRIPS through to complex domains involving numeric quantities and continuous time. Whilst some of the domains are similar in style to the classic bench marks (the DriverLog, Depots and Zeno Travel domains), others have been deliberately designed to appeal to the more applications-led part of the community (the Satellite, Mars Rovers and Relief Mission domains). Still other domains were designed to combine the features of classic bench marks in a way that makes them much more challenging. For example, DriverLog combines a version of the classic Logistics domain with the need to allocate drivers to the trucks. This simple combination results in a very challenging domain.

At the time of writing, interesting challenges await us in interpreting and presenting the results of the competition. Planners in the hand-coded track might well display a performance advantage over those in the fully automated track, but this has to be traded off against the effort investment involved in defining the control rules. Fully automated planners have the advantage of being domain-independent but this has to be traded off against their lack of competitiveness in highly structured and complex domains. Both approaches have strengths and weaknesses, and it is hoped that the competition will give both an opportunity to display their best features.

Conclusions

The 2002 competition sets new challenges for the community in temporal and metric planning. We have established a new collection of bench mark domains which begin to bridge the gap between the theory and application of planning and we have extended the PDDL standard to enable the modelling of complex real time planning domains. Whilst many problems remain to be solved before industrial applications of planning become the norm rather than the exception, these are important steps in the right direction.

However, the real contributions to the community are made by the competitors themselves – those individuals and teams brave enough to put their planners to the test in a public, and somewhat unforgiving, forum. It is always difficult to make a system robust enough to operate reliably under competition conditions, but the additional challenges of coping with an extended language, complex multi-objective functions and a modified validity criterion have considerably exacerbated the usual difficulties. We hope that much will be learned from their efforts and that the results of the competition will motivate the community to address the open challenges in planning with time and resources.

References

Bacchus, F., and Ady, M. 2001. Planning with resources and concurrency: a forward chaining approach. In *Proc. of 17th IJCAI*.

Currie, K., and Tate, A. 1991. O-plan: the open planning architecture. *Artificial Intelligence* 52(1):49–86.

Fox, M., and Long, D. 2001a. PDDL+ : Planning with time and metric resources. Technical report, University of Durham, UK.

Fox, M., and Long, D. 2001b. PDDL2.1 : An extension to PDDL for expressing temporal planning domains. Technical report, University of Durham, UK.

Haslum, P., and Geffner, H. 2001. Heuristic planning with time and resources. In *Proc. of European Conf. on Planning, Toledo*.

Henzinger, T. 1996. The theory of hybrid automata. In *Proceedings of the 11th Annual Symposium on ogic in Computer Science. Invited tutorial.*, 278–292. IEEE Computer Society Press.

Laborie, P., and Ghallab, M. 1995. Planning with sharable resource constraints. In *Proc. 14th Int. Joint Conf. on AI*.

McDermott, D. 2000. The 1998 AI planning systems competition. *AI Magazine* 21(2).

Muscettola, N. 1994. HSTS: Integrating planning and scheduling. In Zweben, M., and Fox, M., eds., *Intelligent Scheduling*. San Mateo, CA: Morgan Kaufmann. 169–212.

Smith, D., and Weld, D. 1999. Temporal Graphplan with mutual exclusion reasoning. In *Proceedings of IJCAI-99, Stockholm*.