Science, Engineering, and Community in the Planning Competition

Manuela M. Veloso
Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213 U.S.A.
veloso@cmu.edu

Abstract

I argue that the planning competition makes sense if it targets and serves three main goals: (i) the advancement of the scientific state of the art of the planning area; (ii) the creation of a concrete effective planning system for relevant real world planning problems; and (iii) the forum for the research community to incrementally move towards the achievement of goals (i) and (ii). In this brief paper, I discuss these three main goals and present suggestions for the continued implementation of the planning competition.

Introduction

Others may remember it differently, but as I recall it, in 1998, I felt very strongly about a planning competition because of two main reasons. The first reason was that traditionally the planning research (similarly to the empirical machine learning research) had a very strong component on the topic of "my planner is faster than yours." Given the difficulty of comparing different planners at the conceptual level, researchers had resorted to empirically comparing their planners on the set of available domains. The results have always shown, and as scientifically expected, that no planner universally dominates all the other planners. However, the community had a true desire to understand the differences among different planners and we expected the planning competition to help us reach this goal in a structured manner. The second reason, probably more personal, was that the RoboCup competition had just started with great success in 1996–1997. The remarkable advances experienced in just two years of the RoboCup competitions showed that the events provided a great way to advance the state of the art of the field. And researchers were enthusiastic about competitions!

The RoboCup competition started in 1996–1997, the planning competition in 1998, and many other competitions have followed, including the Trading-Agent competition. Speech and language had traditionally had earlier competitions, AAAI had also its robot competitions, and there may have been several others.

I view that these competitions, and the planning competition in particular, serve three main goals : (i) the scientific; (ii) the engineering; and (iii) the community. I discuss each one of these goals and conclude with some summary thoughts.

Throughout the discussion, I try not to separate the classical, temporal, or probabilistic planning problems and try to make the discussion inclusive of these different planning frameworks. I view that the overall planning competition may be organized as different *"leagues"* to address such different problems separately. I could have organized the discussion along such leagues, but most of the points presented can be applied across problems. This paper contributes ideas for brainstorming (most of them based on my own experience of ten years of the RoboCup competition), and does not intend to present a "theory" of planning competitions.

The Scientific Goal

We all know that planning is very complex in its most general statement, namely model states of the world, model agent's actions, and find a set of actions (totally or partially ordered) that transforms some initial state into some final desired goal configuration. Stated in that general way, planning is a "once-in-a-lifetime" and "never-executed" single activity: the *generation* of a complete plan for each new initial state and goal configuration. I cannot decouple planning from two inevitable extensions to that planning generation activity: plans need to be *executed* and planning needs *to learn* from past planning and execution experience.

I believe that the planning competition can be organized to see how we advance the planning sciences, namely: *generation*, *execution*, and *learning*. These three scientific challenges have different instantiations in different types of domains – e.g., deterministic, non-deterministic, probabilistic, temporal, repetitive – but they still remain as core problems of the general planning problem.

The overall planning problem is therefore quite overwhelming and one of the roles of the planning competitions, while serving its scientific goal, is to help understand the multiple components of the *complete* planning problem. The planning competitions have included some efforts to address the core challenges. I discuss three main directions that could be part of future planning competitions, namely the understanding of specific problems as "targeted challenges," a continued approximation of "planning and execution in the real world," and an increased exploration of the ultimately inevitably necessary "planning and learning."

Targeted Challenges

Targeted challenges are competing setups especially designed to understand specific aspects of the planning problem. Examples follow.

The impact of different domain representations The aim is to let the competitors explore different representations of a domain just described in English. Domain representation is completely open in terms of the predicates and operators used. Different representations combined with different planning algorithms, whether domain independent or domain dependent ones, certainly impact planning efficiency. For example, in the logistics domain, we can represent the location of an object as `(at ?obj ?loc)` or `(at-obj ?obj ?loc)` or `(at-obj-loc ?obj ?loc)`. And, as another example, we can represent the blocksworld domain with multiple different operators, and even a single `MOVE-BLOCK-FROM-TO` operator moving a block directly from a location to another with a set of conditional effects to handle the different situations for state updates.

The complexity of specific goal/state combinations The aim is to understand subproblems of complex domains and try to empirically contribute to the understanding of the precise sources of the complexities of planning in particular domains.

Partial goal satisfaction The aim is to develop planning approaches capable of returning approximate or partial solutions to given classical planning problems. The approaches could explore removing or adding goals, or removing or adding state. Examples in a navigational path finding domain would be to add to the state a missing key to a door, or to try to reach a different location. The concept to explore would be make to minimally change the problem so that the problem is still partially, if not completely, solved.

"Open" challenge The aim is to challenge every competitor to demonstrate a compelling aspect of their planning approach. The character of any general competition includes devising testing problems and evaluation metrics to compare the competitors and to declare the winners. A good testing problem set tries to fairly sample the problem space, as best understood. However it can be that the invented testing problems may not let specific planners shine at their best. This targeted challenge allows competitors to demonstrate the situations that show their planner at its best. The qualification of the open challenge is qualitative as a jury will have to balance the significance of the chosen demonstration and its results with respect to the general planning problem.

The concept of these targeted challenges has a dynamic character in the sense that they may be revised and adapted for each new competition to address new specific problems that may arise. Ideally, the targeted challenges will eventually contribute to resolve very concrete subproblems shared by the planning research community. The targeted challenges can be viewed as exercises to address open questions in the community.

Planning and Execution in the Real World

The ultimate goal of planning is to create plans that are executed in the real world. The planning competition can help the field to approach the complexities of the real world. The "real world" for planning can be viewed of two rather different types: the web/data and the physical spaces. The planning competition can create a track-based framework for making progress in these two spaces of domains.

The web/data track The web and its complexities in data interpretation, management, and services has proven to be an ideal space to foster AI research. It would be interesting for the planning competition to include a special track for this challenging domain.

The robot/physical world track Planning algorithms typically make many assumptions stating the planning problem, which are not trivial to verify for mobile robots acting in the physical world. These assumptions include that the state of the world is given, the effects of actions, even if probabilistic, are known, and that universal plans (MDP/POMDP policies) to handle a realistic level of uncertainty are feasible to generate. These are strong assumptions when it comes to the physical world. The planning competition can put to test these multiple assumptions and planning approaches using domains of the physical world.

These two proposed real-world tracks share several technical and implementation aspects.

Multiagent/multi-robot challenge When thinking of planning in the real world, many such tasks involve multiagent or multi-robot approaches: Trade, negotiations, rescue, adversarial, construction, soccer, assistance to mention just a few examples of potentially many such multiagent/multi-robot planning tasks. The planning competition can intensively provide a framework to make progress in this area, addressing in particular issues of teamwork, communication, and heterogeneous agents in sensing and planning capabilities.

IPC@... The overhead of creating two "real-world" tracks as planning competitions at ICAPS could be overwhelming. However this should not draw us back from pursuing such tracks and we should discuss novel and creative solutions to still challenge and compare our planning approaches in real-world domains. To that effect, we can investigate the possibility of connecting such tracks of our IPC (International Planning Competition) with existing web and robot-based competitions. Examples would be IPC@Trading-Agent and IPC@RoboCup which would allow us to evaluate planning-based solutions within the Trading-Agent and the RoboCup competitions.

Planning and Learning

Learning from planning and/or execution experience is clearly the ideal approach to address the problems of accurately creating an effective substrate for planning. Research in planning and learning has traditionally progressed both within classical and nondeterministic/probabilistic domains.

The probabilistic domains in their MDP or POMDP representations inherently handle the learning problem also. So the existing probabilistic planning competition can address or be extended to address both the effective model solution in terms of policy generation, and the model or policy learning.

Planning and learning can be a major component of the planning competition serving our scientific goal, creating tests to advance the state of the art of multiple facets of the overall planning and learning problem.

Model learning track Models in general can be viewed as the input to the plan generators. Modeling a domain is a well recognized difficult problem, and models are usually assumed to be given. The planning test is usually set only at the plan generation level. A model learning track within the planning competition could target the learning of actions from experience. Furthermore, and related to the previously discussed targeted challenge to understand domain representations, this track could include the learning of different domain models in terms of improvement of planning performance.

Control learning track Planning includes a respectable search problem. Domain-specific heuristic and control knowledge are now well known to be effective to help reduce the planning search. The planning competition is an ideal framework for the development and comparison of control learners.

Planner learning track The introduction of domain-specific planners opened an opportunity for learning of such planners from observation of example plans. Planner learning differs from model and control learning as it is not about learning knowledge for a domain-independent planner. The interesting distinguishing aspect of this track would be that there is no planner before learning, i.e., there is no algorithm that solves planning problems. Given example plans, the competitor learner in this track acquires the ability to solve new problems.

These multiple planning and learning tracks can be evaluated in terms of an equivalent set of metrics, including number of examples needed, learning time, solvability horizon, and quality of the plans generated.

The Engineering Goal

The planning competition has a fantastic side effect in terms of actual and effective algorithm implementations. Although the participants in the competitions are competitors at the event level, they are researchers who share the goal of advancing the state of the art of the planning area, also in terms of reaching good implementations. Even if one specific competing entry wins an overall competition, we know that other competitors probably excelled at subproblems of the complete problem. Because of our underlying shared research goal, we would like to build upon all the distributed partial expertise. Several mechanisms can serve this purpose.

Sharing of code Participants release their implementations for use of others. This practice should be continued and probably even enforced for qualifying participation.

Modular implementations Modular implementations should be encouraged for several reasons. First, to make visible well bounded expertise in a subproblem. Second, to allow for an effective sharing of subparts of code in a plug-and-play manner. Newcomers with new ideas of their own to address a specific problem, could still participate in the overall planning competition by building a planner from shared modules to which they can add just their focused new ideas. Third, modular code sharing allows for an excellent way for researchers to give credit to each other in a fair way.

Building the *"ideal" planner*, capable of strategically switching among approaches to effectively handle different domains, and capable of learning from its own practicing and real execution experience, is a theoretically unachievable dream. But even approximating such an ideal planner is clearly not the business of a single researcher. Through code sharing in general and modular code sharing in particular, the hope is that the planning competition will enable the dynamic and joint creation of good approximations of the ultimate ideal planning system.

The Community Goal

Finally the planning competition serves a tremendous goal to the community. It provides exciting problems for researchers, and students in particular, and potential entertainment to all participants and spectators.

The competition relies on the community in terms of a variety of aspects, in particular for organizing and technical committees. This community goal could be more extensively discussed but I consider three main aspects of major importance.

Setup of rules The rules (domains, metrics, training, testing, etc.) of the competition are in principle dynamic and should be revised every year. This revision can be a controversial issue as different competitors may favor changes along different aspects. The technical committee for the competition has therefore an important role in terms of: (i) proposing changes, (ii) allowing extensive time for discussion (by email or other remote method) by the community, and finally (iii) settling on the final set of changes well ontime ahead of the competition dates.

Promotion of the competition Without effective promotion, the competition can converge to always have the same participants. Furthermore, potential new participants may be "intimidated" by veteran participants. The organizing committee has therefore the continued important role of: (i) making shared modular code well available to the community, (ii) maintaining the web sites updated with the rules of the future competition and the results of the past ones, and (iii) promoting interactions among the community of participants.

Community build up of the competition Finally, setting up a planning competition is not an easy task if we want to achieve a large set of ambitious goals. The community needs to collaborate with the organizing committee and contribute to the concrete setting of the competition. Of particular difficulty is the creation of new challenging planning domains related to real-world problems. One very concrete way for the involvement of the community is *to require* each participant to contribute a complete planning *domain* (classical, temporal, probabilistic, etc.) to enter the competition. Other examples would be to provide example plans or problem generators. Such concrete requirements would serve as the “registration cost” for the competition.

The community goal is as important as the scientific and engineering goals, as the planning competition is a vacuum without participants. There is no progress without researchers interested in pursuing a problem. The planning competition includes the goal of maintaining the community enthusiastic about the overall planning problem.

Conclusion

The paper discusses three types of goals for the planning competition, namely a *scientific goal* to serve the advancement of the scientific state of the art of the area, an *engineering goal* to contribute to the creation of an actual planning product, and a *community goal* to assert the research community an integral part of the competition.

As stated in terms of goals, the planning competition itself can be viewed as a planning problem with specific steps that can be taken towards the complete, partial, or incremental achievement of its goals. To that extent, as an additional proposal, it would be beneficial to engage on a regular “roadmap” discussion as a monitoring of the execution of the planning competitions. This workshop should be a valuable step to this monitoring.

Although in the paper I have presented a positive view for the future of the planning competition (and of every competition for the matter), competitions can lead to a divergence with respect to ultimate goals. I.e., “winning” a competition at some time could be the result of factors that may not be the real long-term relevant factors. So, I am not completely sure that competitions, and the planning competition in particular, are actually “needed.” (History clearly shows many advances of science without competitions.) But they are certainly an aspect of the current days and they have shown to be of relevance to advancing the state of the art of several fields. The challenge is to continue to understand how to continuously analyze and improve the implementation of any specific competition.