# A Call for Participation:
# Supporting Large Scale Studies of Planning Systems

**Mark Roberts and Adele Howe**
Computer Science Dept., Colorado State University
Fort Collins, Colorado 80523
mroberts, howe@cs.colostate.edu
http://www.cs.colostate.edu/meps

## Introduction

The biennial planning competitions have significantly influenced research. Planners are standardized to PDDL and often made publicly available. Thanks in large part to this standardization and availability, we recently began analyzing runtime data for a large set of classical planners and problems in order to discover when and why one planner (planning technology) is amenable to solving a problem instance (problem set) (Roberts, Howe, & Flom 2007).

The data collection for our project involved running individual planners on each problem instance. We have collected 86 planner versions and almost 5000 problems to date; we have around 30 running. Collecting the planners and problems was arduous because of the separate sources; running some planners is challenging because they rely on different compiler platforms. While there is no clear way to resolve the platform issue, we think that creating a repository may help in obtaining, tracking, and storing the growing collection of systems, problems, and supporting tools.

We outline below the details for a repository of planners, problems, and planning tools. As a start, we hope to emulate the UCI Machine Learning Repository (Asuncion & Newman 2007), which is a leading source of datasets in the Machine Learning community[1].

But we also see the potential for incorporating two additional levels of support into the repository. First, we think it might be straightforward to include the progress and results of the continuing International Planning Competitions. Second, we see the potential for automating the repository to generate runtime data from the systems and benchmark problems as new problems and planners are added.

## The Repository

The basic functionality we envision is as a repository with a wiki (or wiki-like) interface. The interface would allow researchers to post material independently of a central authority. In this way, the repository acts as consolidating hub from which published and unpublished material is referenced. Already there exist a number of information hubs with content we hope to consolidate.

---

[1]We thank to David Aha for insight into the success of the UCI Machine Learning Repository.

The repository need not be limited to only competition planners and problems. Not all planners are competition planners and there frequently exist post-competition versions that fix bugs and/or further analyze the planner. There are also a variety of tools that hold value for the community; for example, there are tools perform domain analysis, visualization, domain modeling, or solution validation. We envision a broader portal that tracks and manages systems as entities distinct from competition participation.

## Competition Portals

If the repository were generally accepted, then the repository could also act as a unifying framework for hosting the information from future planning competitions. For example, one way the repository could be used is that the author(s) of each planning system could:

- start a repository page that describes the competition version and provides a link to the abstract plus any links to outside (published or unpublished) papers on the system,
- upload one or more executables for the planning system as well as a collection of the performance data as run on the current problems,
- link the system to existing problem and competition pages already in the repository, and
- optionally upload the source code for the system.

These pages would be marked as private during the competition and opened to the public after the competition. Note that the initial description of each system happens as *part* of the already existing process for producing the IPC booklet. The competition portal would then provide a link to the specific planner version page that was used in the competition.

Over time, other researchers could incorporate the previous competitions, systems, and tools into the repository. Since the IPC websites are spread across different servers; a clear first step is to bring together these disparate pages.

## Supporting Empirical Studies

On the most grand scale, the repository could include performance data of planners to support large scale studies of planner performance in manner previously done for seven early planners (Howe & Dahlman 2002). There could be several ways to read "supporting empirical studies."

1. As a start, this could mean linking to historical data from papers already written or in progress. In this way, a researcher interested in studying a particular problem set or planner could quickly locate existing results.

2. A long-term goal would be to provide actual runtime data of planners on problems. For example, all STRIPS-capable planners could be run on the entire set of STRIPS problems. Each planner version could link to runtime data that shows its performance when run on all problems.

3. Another (almost too good to be true) goal might be to also provide problem and planner features.

Collecting the performance and feature data sounds daunting, but it can be automated by scripts so as to minimize the maintenance effort; in fact, our scripts already perform much of this automation. Of course, the performance data would also need to be run from a homogeneous set of computing resources.

We think it is worthwhile to consider adding empirical support to the repository in support of multiple purposes: 1) that historical comparison between planning systems becomes an integral part of the competitions; 2) that performance baselines are more easily constructed to guage the direction of the field; and 3) that providing the runtime data accelerates analysis of the planners and problems.

A fair criticism of supplying performance data is that one should not compare on the basis of "most problems solved" or "fastest solver" since it is not clear that doing so is in the best interest of the field. We offer in response that this has been a central concern in the IPCs. Also, the community can shape the presentation of the data in whatever way seems appropriate.

## Who will build it? Will they come?

A natural question is how this system will actually get built. Fortunately, part of the answer to this question is that some of the framework already exists. In terms of content, it may be possible to leverage existing catalogs as a starting point.

Our effort catalogs just the classical planners (Howe & CSU MEPS Group 2007) and STRIPS problems from around 1998 to IPC5. Another effort has a similar focus on planners (St. Amant 2003). The next most complete, but much broader, effort catalogs both planning and scheduling systems (Liu & Artform 2004). Both these latter catalogs are over three years old. There are other sites as well (such as the individual competition pages), but these seem to be the most complete. Almost every planner has its own web-page run by the author or group of the planner.

In terms of setup and maintenance, a modest amount of funding could provide enough to allow for a one student to design the initial system; though we could also enlist a volunteer startup group. The content would be managed by the community. So the only ongoing maintenance costs associated with the project would be for website storage and delivery bandwidth, for which we may be able to leverage an existing website.

## Summary

We hope we have convincingly argued the value of cataloging and compiling the past and future efforts of the planning competitions into a repository of the planning systems, tools, and problems. We feel that contributions should come from from many sources while still providing some community-based quality control over the entire process. We see the key benefits as:

- linking planners to specific publications,
- linking domains and problems to specific publications that highlight analysis of the domains,
- easing the addition of new problems and refinement of existing problems,
- providing a central store for tracking and posting recent advances in understanding the computational complexity of competition problems,
- retrospectively identifying the current benchmark problems and when and where they have been used.

In closing, we would like to propose a way that the hub could start small – as part of the existing competition framework – then expand as needed:

- create the initial wiki-like repository,
- solicit (through funding, volunteer effort, or both) a small group of individuals to post the initial content on the site,
- host the results of the next planning competition,
- as part of the competition results, publicly announce and release the repository for public consumption.

The order allows time to shape the initial design, content, and scope of the project. It seems natural to consider that the repository can be built in conjunction with and publicized within the context of the IPC; though it will likely contain non-competition material. Upon public release, it then allows the hub to expand by community effort.

## References

Asuncion, A., and Newman, D. 2007. UCI machine learning repository. URL: http://www.ics.uci.edu/~mlearn/MLRepository.html. University of California, Irvine, School of Information and Computer Sciences.

Howe, A., and CSU MEPS Group. 2007. Modeling domain independent planning to advance application: Data. URL: http://www.cs.colostate.edu/meps/nsf-data.html. Colorado State University, Department of Computer Science.

Howe, A., and Dahlman, E. 2002. A critical assessment of benchmark comparison in planning. *Journal of Artificial Intelligence Research* 17:1–33.

Liu, D., and Artform. 2004. Planning database. URL: http://scom.hud.ac.uk/planet/repository/. University of Huddersfield, Artform group.

Roberts, M.; Howe, A.; and Flom, L. 2007. Learned models of performance for many planners. In *ICAPS 2007, Workshop AI Planning and Learning*, to appear.

St. Amant, R. 2003. AI planning resources. URL: http://www4.ncsu.edu/ stamant/planning-resources.html. North Carolina State University.