Suggestions for the Knowledge Engineering Competition

Tiago Stegun Vaquero[1] and **Flavio Tonidandel**[2] and **José Reinaldo Silva**[1]
[1]Escola Politécnica - Universidade de São Paulo
Design Lab, Mechatronics and Mechanical Systems Department - São Paulo, Brazil
[2]Centro Universitário da FEI
IAAA - Artificial Intelligence Applied in Automation Lab - São Bernardo do Campo, Brazil
tiago.vaquero@poli.usp.br, flaviot@fei.edu.br, reinaldo@usp.br

Abstract

The challenge of a competition has attracted attention in the academic world, first by the fact that it deals with motivation and second because it can play an important role in the research improvement as a testbed for techniques, models and applications. In this article we raise a discussion on ICKEPS competition which had two editions so far. We review the evolution and history of the competition, highlight some points that should be considered in ICKEPS, and finally elaborate a format that try to fulfill all the expectations.

Introduction

The Knowledge Engineering for Planning and Scheduling has becoming a important research area for the AI planning community, mainly because of the recent efficiency improvement and the rising demand for practical planning systems applied in real life applications. With the improvements of planning engines seen in the literature, two main facts could be detached: the application of planning in real life problems become more and more reliable; and a sound design life cycle is required to apply planning achievements in real domains.

When dealing with real world and complex problems, Requirements Engineering (RE) and Knowledge Engineering (KE) processes become fundamental factors to the success of applications. Some of these processes are requirements gathering and analysis, knowledge acquisition, requirements specification, validation, verification, modeling, planning technique selection, plan (output) analysis, algorithm optimization and knowledge and model maintenance (Vaquero et al. 2007; Bartak and McCluskey 2006). By applying these processes while building an application, the designer is actually structuring its design life cycle and rising the possibilities of making the right application and reducing possible costs of fixing errors in further design stages. Since real planning application deals usually with high levels of complexity, RE and KE aspects must be considered and investigated for such challenging Artificial Intelligence area. One of the main initiatives taken by the planning community to investigate KE concepts applied for Planning & Scheduling (P&S) is the International Competition on Knowledge Engineering for Planning and Scheduling, so called ICKEPS. The ICKEPS is a bi-annual event that aimed to accelerate knowledge engineering research in AI P&S and to encourage the development and sharing of prototype tools or software platforms that promise more rapid, accessible, and effective ways to construct reliable and efficient planning and scheduling systems (Bartak and McCluskey 2006).

In fact, creating a competition as a way to improve a particular research area is very common. Some examples are SAT (The International SATisfability Competition), CASC (The Conference on Automated Deduction ATP System Competition), TAC (Trading Agent Competition) and many others. This strategy has been used by the planning community to develop innovative planning techniques since 1998 with International Planning Competition (IPC) (McDermott 2000; Bacchus 2001; Long and Fox 2003; Hoffmann and Edelkamp 2005). The same strategy is supposed to be applied to ICKEPS, but the results were not an entire success. Even with all the substantial work made by the ICKEPS organizers, the competition seems to have some issues and improvements to enhance ICKEPS. Indeed, the competition had few editions, only two (ICKEP-1 was in 2005 and ICKEPS-2 in 2007), which are very similar in the concept but different in structure. In the first edition of the event, only seven competitors participated. In the second event, there were four competitors and no competition (because of the low number of competitors), but a workshop instead.

This paper aims to discuss some ideas and suggestions as an attempt to make not only the competition ICKEPS more visible and efficient, but also let the KE competition plays an important role to announce novelties to P&S research area. We focus on the entire life-cycle process by keeping in mind that individual and simple tools are very welcome in ICKEPS as well as complete systems. It is expected that with the issues raised in this paper, new and refined strategies could be used toward a clear contribution of the ICKEPS for planning community and for a coordinated relation with IPC (International Planning Competition).

The remainder of this paper is organized as follows. In the first section, we briefly describe the past ICKEPS competitions emphasizing their main objectives and structure. In the next section, we go into a discussion about important issues concerning the competitions. Then, we give some suggestions of competition format for the next ICKEPS competition. Lastly, the concluding remarks are provided.

ICKEPS Competitions

As described by (Bartak and McCluskey 2006), ICKEPS combines high academic objectives with the capacity to expose recent application tools and its achievements to solve new classes of planning problems. In fact, the creation of such competition was a great initiative towards the application of planning and scheduling systems in real life problems and also to encourage the development of tools that comprise the whole KE area.

The competition was created in 2005 and organized by Roman Barták and Lee McCluskey during the 15^{th} ICAPS, when the first edition of the event took place (ICKEPS-1). The organizers decided to launch the first competition with a very simple format and to assemble a panel of judges to evaluate the competing application systems. The event ran in two main stages. In the first one (pre-conference stage), the competitors were asked to submit a short paper describing their tools. The competition committee analyzed the papers by evaluating the relevance of each tool and then sent feedback to the participants. In the second stage (during the conference) the competitors were asked to make general presentations to the audience by explaining their KE tools and also individual presentations to the judges who would give the final decision. Due to the high diversity of the participant tools, some basic criteria were formulated that the judges could use to select the winner (Bartak and McCluskey 2006). The main criteria were: *potential, scope of the system, usability, interoperability, innovation, wider comparison with knowledge engineering tools in AI, building quality, and relevance to the scope* (Bartak and McCluskey 2006).

In the ICKEPS-1 there were seven competitors presenting interesting tools. Because of the diversity of the tools, they were grouped into two categories in order to facilitate the judgment: the 'general tools' which joined the tools ModPlan (Germany), the winner GIPO (UK), and itSIMPLE (Brazil); and the second category, called 'specific tools', with Hamlet (Spain), the winner ARMS (Hong Kong), Tailor (USA), and PlanWorks (USA).

It was clear that the first edition established a high interest of KE aspects in planning community, showing the potential of available tools. However, the competition focused on qualitative attributes given subjective results as mentioned in (Edelkamp, Frank, and Kellershoff 2007). Indeed, the tools were evaluated solely on the short papers, presentations and demonstrations to the judges. No quantitative aspect (metric) was used to rank the competitors at that time.

Analyzing the weak points of a pure qualitative evaluation of the 1st edition, the organizers of the 2nd competition turned its focus into quantitative aspects of the tools. Stefan Edelkamp, Jeremy Frank and Mark Kellershoff created and organized the 2nd edition of the International Competition on Knowledge Engineering for Planning and Scheduling (ICKEPS-2) that was held in the 17^{th} ICAPS in Providence. ICKEPS-2 *was intended to provide a continuation of knowledge-based and domain modeling as a bi-annual event, in synergy to the bi-annual International Planning Competition IPC* (Edelkamp, Frank, and Kellershoff 2007). The main objectives of the competition were to continue to accelerate knowledge engineering research in AI and encourage the development of KE tools for planning and scheduling.

The structure and format of ICKEPS-2 had a different approach compared to the previous event. The organizers provided a client-server environment showing a more realistic planning situation, which was a good initiative. The competition was now structured in three stages. In the first stage ("Pre-conference"), competitors were encouraged to submit a short paper describing the tool to the organizers. The competitors were also asked to make their tools available for download prior to the competition to let the judges evaluate the tools. The short papers were reviewed and feedback was provided to all candidates. During second stage ("Evaluation through Simulation"), which happens prior to the conference, the organizers made available a planning and scheduling simulation framework that competitors could use to test their tools. The interaction between the tools and the simulator was recorded in a log file in order to be evaluated by the judges. Following, competitors received a short text description of five planning and scheduling domains, described in (Edelkamp, Frank, and Kellershoff 2007). The domain models had to be tested with problem instances, resulting in plans together with feedback from the simulator about the problem solving process. An important point to be mentioned is that the choice of planning techniques and design tools, including domain description language, was up to the competitor. The only requirement to be satisfied was the simulator interface. For the third stage ("performed during the conference") the competitors had to make a short presentation to discuss how their tools helped them to solve the simulated domain, and also to discuss other features of their tools not tested during the simulations (Edelkamp, Frank, and Kellershoff 2007).

The criteria used to evaluate the participants included both qualitative and quantitative aspects. Some of the criteria used were: user assistance in potential, scope, usability, interoperability, innovation, building quality, relevance, domain simulation applicability (criteria related to the interaction with the simulator, the number of domains modeled, and others).

Unfortunately, the competition received only four competitors: itSIMPLE$_{2.0}$, GIPO IV, Source Control Server, and Mini Zink. The small number of competitors led to the elimination of the simulation stage of the competition which forced the organizers to turn the ICKEPS-2 into a workshop. Thus, this workshop intended to illustrate the state-of-the-art in knowledge engineering tools, the capacity of the tools to solve new classes of domains, and also raise discussion on how far current technology is from planning practice.

Competition Issues

Analyzing the last two ICKEPS, we enumerated some issues that we must be aware in order to make the event reaches its main goals and receives the deserved attention from the AI community. For each issue we provide a brief discussion of how it could improve ICKEPS.

Competition preparation. Any competition requires an initial preparation and for ICKEPS this preparation stands

for deciding the suitable domains to be used. In fact, the organizers must be aware that competition preparation must benefit both single tools and complete systems, and that all selected domains must have a reason to be proposed. Indeed, all domains must comprise real applications, but they also have to be suitable and simple enough to let single tools to participate. Such domain selection would demand a consult to companies and institutions that are interested in particular classes of problems (e.g. NASA for space mission planning domains). In our point of view, one of the critical problem of ICKEPS-2 that lead the competition to have a small number of competitors is the lack of suitable domains that can let single and simple KE tools to participate of the second phase (domain modeling and simulators). For example, the organizers could release more than just complete and complex real domains descriptions, like partial modeled domains as well as naive models (not well modeled) of a single real application. From that range of domain types, tools that improve the model of a existing domain (e.g. ARMS, Tailor and PlanWorks), tools that extract domain knowledge (e.g. ModPlan) and tools that can convert one model into another would probably feel more comfortable to get in the competition.

Not only proposing different type of domains is really suitable but it is also important to highlight which feature is being tested in each problem. In fact, any selected domain must have a reason to be used as a benchmark for competition. Therefore clear criteria to select domain for the competition could lead us to define qualitative and quantitative evaluation suitably since a good balance from design life cycle, knowledge engineering, and AI problem solving techniques could be achieved. In addition, all selected domain must be feasible to be used in IPC, since the synergy between IPC and ICKEPS is required. This synergy is the topic of the next issue.

Synergy between ICKEPS and IPC. One of the main points that can really contribute to the success and improvement of ICKEPS is the synergy and synchronization between ICKEPS and IPC (and also the possible Scheduling Competition). All the potential and expectation of IPC events could be used to give a boost to ICKEPS which can also help to improve IPC. In fact they are supposed to work in a close loop, feeding each other with new challenges. In (Edelkamp, Frank, and Kellershoff 2007) such synergy is mentioned but so far it is not implemented.

In fact, KE tools must be connected to the current planners as well as planning systems must take into account new features of arisen domains and challenges from ICKEPS. For instance, a good synergy could arise from a selection of outstanding planners from previous IPC to simulate domains modeled by the KE environments during ICKEPS. Such planners could be selected by their performance (as a general planner or in different categories of problems) in the IPC one year before the ICKEPS. Conversely, some new modeled domains (bringing new features and challeges) raised in ICKEPS could be used in the next IPC providing heuristics, constraints or any piece of knowledge that could take the process of planning more efficient or result in best plans. Thus, besides the new tools, approaches, methods and roadmaps generated by both competitions, the IPC would provide selected planners to the ICKEPS, whereas ICKEPS would provide new planning domain models and new challenges for the future planners. Figure 1 illustrates the envisioned synergy between these competitions.

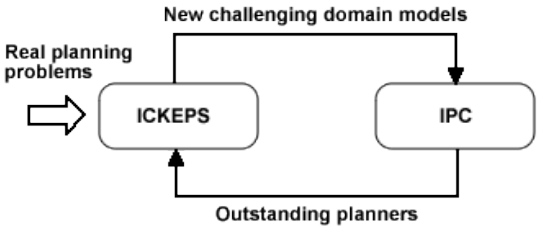

Figure 1: Synergy between ICKEPS and IPC

In addition, supposing that all KE tools need to use PDDL (Planning Domain Definition Language) (Gerevini and Long 2006) to communicate with planners in the competition, it is possible to see that PDDL development could profit from a clear and necessary direction on how it should be enhanced toward a specification language that could be fully interpreted by planners. Therefore, extensions of PDDL could be inspired by both new IPC challenges and demands raised by real applications in the ICKEPS results. This synergy is important for the ICKEPS, IPC and planning researches at all.

Qualitative and Quantitative Evaluation. The first ICKEPS focus basically on qualitative aspects of applied tools while the second edition included quantitative measures. The quantitative measure included by the ICKEPS-2 is based on the number of modeled domains that ran suitably in the provided simulators. Obviously, the mixing between quantitative and qualitative evaluation is welcome and should be maintained. As well-seen by the ICKEPS-2 organizers, the competition seems weak without a clear way of ranking the systems, so quantitative aspects are really necessary.

The quantitative results must be captured automatically, without any influence of judge or human factor. We can clearly see this in the IPC that qualify, automatically, a planner by its results over a plenty of problems. We must follow the same idea in ICKEPS. However, we must keep in mind that it is difficult to compare whether a domain is better modeled than other since the result of a planning problem (the plan) is strongly influenced by the planner selected to solve such problem. Sometimes, the complexity of a modeled domain can lead all planners to fail, making any kind of comparison impossible. It would be an interesting idea to introduce these complex models as a new challenge for planners in further IPC competitions.

One category with distinguished awards. When established many quantitative criteria, it is natural to think whether the competition should be divided in several categories with different winners or not. This is the reality of

IPC competition but it can not be fit well in the ICKEPS scope. The purpose of ICKEPS is to investigate the entire modeling life cycle process where the result of a KE tool is a modeled domain while the purpose of IPC is to investigate the results of a planner applied to solve specific domain problems. On one hand, the result of a planner in IPC can vary from domain types and planning techniques, therefore, it is suitable that planners are classified in different categories of domain types in order to evaluate planning techniques. On the other hand, since the ICKEPS results show how real and good the model is and what the life-cycle process is, it is unclear how to create categories. Therefore, we believe that ICKEPS must have only one single category but with distinguished awards in different parts of a design life cycle.

In the entire life cycle process there are many topics/stages that can be evaluated separately. Each evaluation could lead to a distinguished award and the composition of all these specific evaluation can configure the ICKEPS winner. Distinguished award can clarify and present to all planning community the best improvement in each part of the life cycle process. In addition, distinguished awards for topics of the entire process can let single tools to compete.

What has being analyzed? Both events defined clear criteria to be applied during judgment. However, it seems that the criteria and metrics are more (maybe solely) related to the tool (interoperability, innovation, relevance, usability, support, scope, and others). In fact, as IPC, the result of ICKEPS has been treated solely as a final result, which is the final domain model. We believe that the life cycle process evaluation by itself must be also taken into account. Processes such as structuring the problem, specification, modeling, analysis, plan analysis, simplification methods, maintenance, new KE approaches and RE methods that goes beyond the use of tools must also be analyzed. Recently, more and more papers bring the application of planning into real and complex problems and they are probably using some kind of discipline or method to create suitable models in PDDL. These methods are very important and they must be shared and evaluated.

Real life domains. The second ICKEPS brought challenging planning domains and realistic problems. That was an excellent initiative from the organizers. The authors believe that the ICKEPS could become one of the main entrance to bring new planning domains into the planning community. In fact, it would be very interesting if both industry and academic research contributed to the formulation of future ICKEPS.

Nowadays, when a industry or a new academic research interest want to propose an application for planning systems, they generally apply this application to IPC, where someone will probably model this application in PDDL and then released as a new domain for IPC competitors. However, it is important to note that the resulting model from the PDDL description is sometimes restricted (e.g. PDDL limitations). With the synergy between IPC and ICKEPS, new applications could be firstly proposed to ICKEPS that will release domain models to the IPC competitors or improve PDDL language in the appropriate time, taking planning community and researches to solve real planning applications for real.

Inputs and Outputs. In the ICKEPS-1 there was no input to competitors in order to evaluate their tools. In the ICKEPS-2, the input was a set of text descriptions of given planning domains (Edelkamp, Frank, and Kellershoff 2007). The existence of such description as an input to the competition was a great initiative. However, it would be very interesting if the domain text description had a standard format or structure (off course, that depends on the abstraction level of the description). Now, concerning the output, there was no standard output format at all. This fact made the process of understanding each output interface a little bit time consuming. Therefore, the choice of a good requirements representation would be important to the community as well as to the competition. Probably that will be a long discussion that could survive for several editions of ICKEPS, but unify the interpretation of problems is a key point.

Criteria and Metrics. In addition to the criteria defined for ICKEPS-1 and ICKEPS-2, we suggest the evaluation of the entire KE processes and use the life cycle approach performed for ranking and classifying the systems qualitatively. This evaluation is a result analysis from the judges. For the quantitative aspects we would like to see the evaluation of the domain model. Since this evaluation depends on the planner used to solve domain problems, it is suitable to use a plenty of good planners (from previous IPC) and count for instance how much problems ran faster than the others using all available planning systems. This metric is a suitably measure of how appropriate is the domain model. Counting solely the number of solved problem (ignoring the quality) does not say anything, since a failure to solve a problem is not necessarily caused by the domain but also by the planner.

The interesting thing here is that it would be possible to see different results from different KE approaches and, finally, study the impact of knowledge-based process in the planning world.

Expected results. After each competition we hope to see: a clear link between planning competitions; challenging domains being analyzed and solved by planners with higher quality, motivating even more the application of planning and scheduling techniques outside the academic context; roadmaps for representation languages and techniques; refined roadmaps for KE to Planning & Scheduling; improvements on KE tools, processes and algorithms; a clear vision of the main existing and remaining gaps in applying our tools and softwares into real-world problems; a closer relation with industry and key institutions since their problems could be used as benchmark; and a higher number of competitors.

Suggestions for Future Competitions

We propose in this section a competition of KE based mainly on ICKEPS-2. Although ICKEPS-2 format was not fully tested since the small number of competitor turn it into a

workshop, the second competition was built on a concept which has the same root that also driven us to propose a new format for ICKEPS. In the following topics we show the proposed format for ICKEPS.

Preparation. The first stage is the preparation of the competition. Before any other stage, the definition of which planning systems will be selected to evaluate domain models and problems is crucial. Also, the domains and problems of the competition will be carefully selected based mainly on the community interest and roadmaps. Indeed, the selected planners will influence the choice of the domains and problems. The selected planners would be those that outperform during previous IPC in any track.

Paper submission. In this stage a short paper is required. The paper must highlight methods, disciplines and supportive KE approaches for each tool or system, including a briefing of the implementation. The tool or KE System will only be eligible to compete if this short paper is accepted.

Warm-up stage. In this stage competitors would receive a set of domain descriptions and models to submit their tools. This phase will give a feedback to the competitors about: the minimum requirements for the competition; how their environments and tools perform; what would be expected from them; and what should be improved. This phase is not public and will not be judged. These domains and models would be classified in different classes (or topics):

- naive models: This class of domain models would provide some domain models that are not entire optimal modeled, i.e., the model of the domain is more complex or simple than necessary. These set of models would give a chance for tools that intend to improve models by knowledge extraction or interaction with the domain designer.

- partial modeled domains: This class of domain models would be given to the competitor partially modeled together with the description. For example, these domains could be described as set of plans or pre-established processes (workflow). These domains would give a chance for individuals or simple tools to extract domain knowledge, complete the domain model by inference or give a chance to show how a tool can improve a domain model.

- applications descriptions: This class of domains will provide domain descriptions of real applications. The description must be textual and can provide some applications details and/or examples of desired plans in such domain. This set of domain descriptions should give a chance to tools to reach a desirable model for such applications.

- complex applications descriptions: This final class is similar to previous one, except for the fact that the applications are far difficult from the previous one. This difficulty is related to domains that have great potential to avoid the available planners to succeed. Once modeled, these domains make an interesting set of domains to be proposed on the next IPC competition.

Therefore, in this stage, the competitors will manipulate, create and improve domains descriptions. A set of problems for each domain must be given, and all competitors must test their models by using the selected planners. It is important to say that, in order to test models by planners, all domains must be described in PDDL.

Competition. The third stage is the competition itself. Here the novelty is that new domains in each topic will be provided and new problems of the released domains in warm-up stage will be presented. All the competitors will test and evaluate their tools on such new domains and problems. All the results of the selected planners from modeled domains will compute points for the competitors (quantitative analysis). In fact, for each problem modeled by an specific tool and solved faster than the others, one point is counted to such tool. Similarly, one point is counted for each best quality result for each problem of an specific domain. Judges will also evaluate tools by considering the entire life-cycle process and to analyze models that planners can not run, e.g., domains from complex application descriptions (qualitative analysis). These judges analysis will be composed by a report to be filled with points (like paper review process) by observing the entire tool working on selected domains and over the results of the modeling process.

Presentation. A final presentation of the papers of the competitors that are in the competition should be made to open discussion about methods and features addressed by the tools. Papers accepted to this workshop must bring inspiring novelties to environments, specific knowledge concerning particular problems, issues for life cycle managements, etc. Pure implementation would be not enough to justify the presentation.

Final Report. After the event and the competition, a final report should be prepared by the judges and steering committee where the feedbacks to the community and to other competitions are transformed in a paper to be published and made available. This feedback could contain for example the domain models for next IPC, roadmaps, some PDDL improvements, the winner and distinguished awarded tools.

Results. All theses stages intend to evaluate quantity and qualitative results, all life-cycle requirements and all details of the tool in the final stage. The judges must give grades for each considered topic and these grades will classify a winner of ICKEPS. The winner of ICKEPS encapsulates the entire desire of KE for P&S concept. However, we believe that we should give four distinguished award. Since the life-cycle has many topics/stages that should be evaluated, we hope that some tools or techniques can show some distinguished performance in some topics but not necessarily in all. Declare only one single winner can obstruct some good and real contribution in each stage of the entire life-cycle. In addition, distinguished awards will stimulate individual and simple systems to participate in the competition and to continue to improve their tools. An individual or complete system can win one or more distinguished awards as well as the competition.

The distinguished awards could be provided in four items:

- Gathering & Structuring;
- Analysis & Validation;
- Specification;
- Modeling & Solution Evaluation.

The first item, Gathering & Structuring, will take into account the points received in the evaluation of partial modeled domains topic by planners and some grades provided by judges. This item, for example, could comprise tools like ARMS and Hamlet (which participate in ICKEPS-1) besides complete tools like itSIMPLE, GIPO and ModPlan. This first item also motivates tools developed for learning phase of IPC - Learning Track to participate and to compete.

The second item, Analysis & Validation, comprises tools like Tailor and PlanWorks (from ICKEPS-1) and any other tools that can analyze plans and actions by performing verification and validation of the entire model and giving feedback for the end user (designer). This item can be graded by using the performance of tools during the test of naive domain models in the planner. Judges will also consider validation aspects of the KE systems during the judgment.

The third item, called Specification, focus on domain specifications and how the KE tool can help end users to achieve their goals, i.e., the designers to reach a complete model for a specific application. This item can be influenced by the result of planners by considering applications descriptions, (complex or not) classes of domains and by qualitative analysis from judges.

The last item, Modeling & Solution Evaluation, is also influenced by the results from the class of domains with applications descriptions and by qualitatively analysis of judges. Tools like PlanWorks could have a good performance in this item, for instance.

Following the format described above, we believe that we can motivate new tools to participate, not only tools developed for IPC Learning Track, but also individual tools that can make an important contribution for a small part of the entire life-cycle as well as tools that can hold the life-cycle process at all.

Conclusion

From the experience of these two editions of ICKEPS, we can list the names of at least ten different KE tools and systems, some that could be classified as general environments and others that are addressed to specific features or problems of the entire life-cycle. Naturally there are others that will show up soon (like tools that will be used during the learning phase of learning track of IPC). Thus, the first conclusion is that a good number of competitors exists, and that it is worthwhile to encourage them to come to the competition. After all, the main objective of the competition is first to bring KE and AI Planning and Scheduling techniques to real problems, associating different expertise that comes from Computer Science and Engineering Design.

The synergy between IPC and ICKEPS is discussed and it could be used as another point to motivate participants since both IPC and ICKEPS competition has as the main goal the improvement of the state-of-art tools for planning and scheduling. In the ICKEPS we suggest different topics/classes that allow the evaluation of different kind of KE tools for showing a winner and distinguished tools. By grouping quantitative and qualitative analysis it is possible to highlight the best approaches in some aspects of the life cycle of modeling process.

Thus, part of the effort is to share experience, to give suggestions, professional evaluation, and even open code to those interested in producing good tools. Finally, this suggested format of the ICEKSPS competition would bring a minimum standard to all tools and systems and can also broadcast the achievements of planning and scheduling techniques.

References

Bacchus, F. 2001. The aips00 planning competition. *AI Magazine* 20(3):4756.

Bartak, R., and McCluskey, L. 2006. The first competition on knowledge engineering for planning and scheduling. *AI Magazine* 26(1):97–98.

Edelkamp, S.; Frank, J.; and Kellershoff, M. 2007. Knowledge engineering through simulation. *Proceedings of the International Knowledge Engineering Competition.* International Conference on Automated Planning and Scheduling. Providence.

Gerevini, A., and Long, D. 2006. Preferences and soft constraints in pddl3. In Gerevini, A., and Long, D., eds., *Proceedings of ICAPS workshop on Planning with Preferences and Soft Constraints*, 46–53.

Hoffmann, J., and Edelkamp, S. 2005. The deterministic part of ipc-4: An overview. *Journal of Artificial Intelligence Research* 24:519579.

Long, D., and Fox, M. 2003. The 3rd international planning competition: Overview and results. *Journal of Artificial Intelligence Research 20. Special issue on the 3rd International Planning Competition.*

McDermott, D. 2000. The 1998 ai planning competition. *AI Magazine* 20(2).

Vaquero, T. S.; Romero, V.; Tonidandel, F.; and Silva, J. R. 2007. itsimple2.0: An integrated tool for designing planning environments. In *Proceedings of the 17th International Conference on Automated Planning and Scheduling (ICAPS). Providence, Rhode Island, USA.*