The Fifth International Competition on Knowledge Engineering for Planning and Scheduling: Summary and Trends

Lukáš Chrpa, Thomas L. McCluskey, Mauro Vallati, Tiago Vaquero

■ *We review the 2016 International Competition on Knowledge Engineering for Planning and Scheduling (ICKEPS), the fifth in a series of competitions started in 2005. The ICKEPS series focuses on promoting the importance of knowledge engineering methods and tools for automated planning and scheduling systems.*

The International Competition on Knowledge Engineering for Planning and Scheduling has been running since 2005 as a biennial event promoting the development and importance of the use of knowledge engineering methods and techniques within this area. The aim of the competition series is to foster developments in the knowledge-based and domain modeling aspects of automated planning, to accelerate knowledge engineering research, and to encourage the creation and sharing of prototype tools and software platforms that promise more rapid, accessible, and effective ways to construct reliable and efficient automated planning systems.

ICKEPS 2016 aimed specifically (1) to provide an interesting opportunity for researchers and students to experience the challenges of knowledge engineering; (2) to motivate the

  ISSN 0738-4602

planning community to create and improve tools and techniques for supporting the main design phases of a planning domain model; and (3) to provide new interesting and challenging models that can be used for testing the performance of state-of-the-art planning engines. In order to achieve the mentioned aims, ICKEPS 2016 focused on on-site modeling of challenging scenarios, performed by small teams.

This article summarizes the ICKEPS held in 2016. More information about the competition, including complete scenario descriptions, can be found on the ICKEPS 2016 website.[1]

Format and Participants

ICKEPS 2016 format included two main stages: On-site modeling and demonstration.

During the on-site modeling stage, each team received a set of scenarios description and had to exploit the available time for generating the corresponding models. Four scenarios were provided. Two of them — Star Trek, Rescue of Levaq, and Roundabout — required temporal constraints, while the other two — RPG and Match-Three, Harry! — only required classical reasoning. Participants were free to select the scenarios to tackle and had no restrictions on the number and type of tools that can be used. The only constraints were on the available time (six hours were given) and on the maximum size of teams (at most four members).

The day after the on-site modeling, each team had 10 minutes to present and demonstrate the aspects of the knowledge engineering process they exploited for encoding the scenarios. Specifically, teams were expected to discuss the division of work among team members, the tools used, key decisions taken during the encoding, and the issues they faced.

Teams were then ranked by a board of judges, which included Minh Do (NASA, USA), Simone Fratini (ESA, Germany), Ron Petrick (Heriot-Watt University, UK), Patricia Riddle (University of Auckland, New Zealand), and David Smith (NASA, USA). The evaluation process will be described in the corresponding section below. Noteworthy, judges were presented during the demonstrations session and had the opportunity to ask questions and discuss relevant aspects of the knowledge engineering process the teams followed.

The competition had two tracks: the PDDL track, where teams had to generate PDDL models using PDDL features up to those introduced in version 3.1, and the Open track, where teams could encode models in any other language. However, for the open track, participants were also required to provide a planner able to deal with the selected language. Sixteen people, divided into six teams, took part in the competition. One team entered the Open track, while the remaining five decided to participate in the PDDL track.

Participants came from institutions in Australia, Brazil, Canada, USA, Japan, and the United Kingdom. The level of expertise of participants covered various academic ranks, that is, Ph.D. students, lecturers, research fellows, and professors. One team was composed only of industry experts.

Evaluation

The board of judges evaluated each team by considering two main aspects: the exploited knowledge engineering process and the quality of the generated models.

The knowledge engineering process was assessed once for each team, regardless of the number of scenarios the team was able to encode. Three main criteria were taken into account: teamwork, method, and tools. Teamwork focused on the degree of cooperation and effective collaboration among team members. In terms of the method, effectiveness and systematicity of the knowledge engineering process were assessed. Finally, the innovation and originality of exploited tools, and their actual usefulness (that is, the support their use provided to the process) were evaluated.

To assess the quality of the generated models, the organizers provided the judges with the models the teams had submitted along with quantitative and qualitative information about these models. Qualitative information included evaluations about correctness, (that is, whether all the requirements were correctly handled); readability (how easy it was to read and understand the model); generality (if the domain model could be reused on different problem instances); and originality, where the use of innovative ways for modeling element or interactions was evaluated. Quantitative information included statistics on the number of types, number of predicates, number of operators, total number of lines, and the average (maximum) number of parameters, effects, and preconditions per operator. Moreover, in the PDDL track, the run time and quality of solutions generated by 10 well-known planners (5 classical and 5 temporal) were provided to judges. For teams participating in the Open track, the corresponding performance of the planner(s) submitted by the participants were provided to judges.

In accordance with the aims of the competition, emphasis was given to good practice in knowledge engineering, with particular regard to the degree of cooperation between the members of each team. For this reason the judges used a 0–100 scale, where up to 45 points could be awarded for the knowledge engineering process, and the remaining 55 points could be assigned according to the number and quality of generated models, as follows: Star Trek, the Rescue of Levaq (up to 20 points); Roundabout (up to 15 points); Match-Three, Harry! (up to 10 points); and RPG (up to 10 points).

Interesting in Hosting ICWSM-19?

AAAI, in cooperation with the ICWSM Steering Committee, is currently seeking proposals for a host city for the Thirteenth International AAAI Conference on Web and Social Media (ICWSM-19). The conference is typically held Monday – Thursday during the timeframe of mid-May through mid-June. Final selection of a site will be made by August 2017. For more information about proposal requirements, please write to icwsm19@aaai.org.

Note: ICWSM-18 will be held at Stanford University in Palo Alto, California USA.

Results

The board of judges acknowledged the efforts of all the competitors. Honorable mentions were then awarded in two categories:

The Innovative Methodology Award was presented to Emre Savas and Michael Cashmore. This team generated a complete domain transition graph for the RPG scenario by hand, analyzed the graph to remove bad states and transitions, and then created a compact and elegant model for the domain.

The Dilithium Crystal Award was presented to Sara Bernardini, Maria Fox, and Chiara Piacentini. This team was the only one to produce a working model that correctly captures most of the requirements of the Star Trek Scenario, which was the most difficult domain in the competition.

The Overall Winner Award was presented to the team composed of Nir Lipovetzky and Christian Muise. This team demonstrated a great ability to develop high-quality models quickly in multiple scenarios, while utilizing, and at the same time enhancing, model development tools for PDDL.

Given the positive feedback from competitors and judges, we believe that ICKEPS 2016 was a success. It is therefore envisaged that future ICKEPS will exploit a similar format.

Reflections

We observed that the generated models showed significant differences, even on easier scenarios, where, for instance, the number of operators ranged from two to seven, with remarkable impact on readability and generality. The impact on different planning approaches has to be assessed, in order to advance the state-of-the-art of knowledge engineering.

Two items were of concern at ICKEPS. First, most teams did not use any tools (except text editors), and thus relied only on their expertise. Second, existing tools do not effectively support cooperation: to cope with the growing complexity of planning applications, planning experts have to cooperate and coordinate the knowledge engineering process. In addition, the number of participants of ICKEPS is still not very large, especially when compared with the latest edition of the International Planning Competition (Vallati et al. 2015). This suggests that the planning community underestimates the importance of knowledge engineering, despite of its enormous impact on applicability of domain-independent planning in real-world scenarios.

Reference

Vallati, M.; Chrpa, L.; Grzes, M.; McCluskey, T. L.; Roberts, M.; and Sanner, S. 2015. The 2014 International Planning Competition: Progress and Trends. *AI Magazine* 36(3): 90–98.

Lukáš Chrpa is a research fellow at University of Huddersfield. His main research interests are in the area of AI planning, machine learning, and knowledge engineering. He was a co-organizer of the 2014 International Planning Competition (IPC).

Thomas L. McCluskey is a coleader of the PARK research group of the University of Huddersfield, UK. His research interests include automated plan generation, domain modeling, information extraction, knowledge engineering, machine learning, data mining, requirements validation, and formal specification. He cofounded the Knowledge Engineering for Planning and Scheduling competitions. He also co-organized IPC-2014.

Mauro Vallati is a senior lecturer at University of Huddersfield, UK. His main research interest is in AI planning. He was a co-organizer of the 2014 International Planning Competition.

Tiago Vaquero is a postdoctoral fellow at MIT CSAIL and Caltech. His research interests include autonomous systems, automated planning and scheduling, knowledge engineering, robotic space exploration, artificial intelligence, and robotics in general. He was a co-organizer of the 2012 edition of the International Competition on Knowledge Engineering for Planning and Scheduling.