Artificial Intelligence 229 (2015) 258–281

Contents lists available at ScienceDirect

Artificial Intelligence

www.elsevier.com/locate/artint

RoboCup@Home: Analysis and results of evolving competitions for domestic and service robots ☆

Luca Iocchi [a,*], Dirk Holz [b], Javier Ruiz-del-Solar [c], Komei Sugiura [d], Tijn van der Zant [e]

[a] Department of Computer, Control, and Management Engineering, Sapienza University of Rome, Italy
[b] Autonomous Intelligent Systems Group, University of Bonn, Germany
[c] Department of Electrical Engineering & Advanced Mining Technology Center, Universidad de Chile, Chile
[d] Information Services Platform Laboratory, National Institute of Information and Communications Technology, Japan
[e] Cognitive Robotics Laboratory, University of Groningen, The Netherlands

ARTICLE INFO

Article history:
Received 17 March 2014
Received in revised form 15 July 2015
Accepted 8 August 2015
Available online 11 August 2015

Keywords:
Robotic competitions
Artificial intelligence and robotics
Benchmarking

ABSTRACT

Scientific competitions are becoming more common in many research areas of artificial intelligence and robotics, since they provide a shared testbed for comparing different solutions and enable the exchange of research results. Moreover, they are interesting for general audiences and industries. Currently, many major research areas in artificial intelligence and robotics are organizing multiple-year competitions that are typically associated with scientific conferences.

One important aspect of such competitions is that they are organized for many years. This introduces a temporal evolution that is interesting to analyze. However, the problem of evaluating a competition over many years remains unaddressed. We believe that this issue is critical to properly fuel changes over the years and measure the results of these decisions. Therefore, this article focuses on the analysis and the results of evolving competitions.

In this article, we present the RoboCup@Home competition, which is the largest worldwide competition for domestic service robots, and evaluate its progress over the past seven years. We show how the definition of a proper scoring system allows for desired functionalities to be related to tasks and how the resulting analysis fuels subsequent changes to achieve general and robust solutions implemented by the teams. Our results show not only the steadily increasing complexity of the tasks that RoboCup@Home robots can solve but also the increased performance for all of the functionalities addressed in the competition.

We believe that the methodology used in RoboCup@Home for evaluating competition advances and for stimulating changes can be applied and extended to other robotic competitions as well as to multi-year research projects involving Artificial Intelligence and Robotics.

☆ This paper was submitted to the Competition Section of the journal.

* Corresponding author.
E-mail addresses: iocchi@dis.uniroma1.it (L. Iocchi), holz@ais.uni-bonn.de (D. Holz), jruizd@ing.uchile.cl (J. Ruiz-del-Solar), komei.sugiura@nict.go.jp (K. Sugiura), tijn@ai.rug.nl (T. van der Zant).

http://dx.doi.org/10.1016/j.artint.2015.08.002
0004-3702/

1. Introduction

Artificial intelligence and robotics competitions have significantly increased their scope and visibility in recent years and today most research areas have some type of competition. Such competitions provide frameworks where research groups can compare the results of developed methods and give opportunities to define standard benchmarks for solving specific problems, comparing different solutions, and disseminating the best solutions available to the community. Moreover, tasks and results are often used in scientific papers to compare new approaches with existing ones.

Competitions go beyond typical experiments. Analogies and differences between robotic experiments and competitions have been investigated [1,2]. Experiments (and benchmarking) in robotics are characterized by the following: i) a common testbed (usually provided through detailed specifications), ii) specific performance metrics, iii) reproducibility, and iv) repeatability. On the other hand, scientific robotic competitions involve: i) specifications of competition environment, ii) specification of robot requirements and constraints, iii) specific performance metrics to rank the participants, iv) information about how they are organized.

The main difference is of course found in the scope of these two activities; the former aims at demonstrating and measuring the performance of a system or component to solve a particular problem, while the latter aims at directly comparing different solutions in a predefined testbed. A common element is the environment specifications that can be more or less detailed in both cases, depending on the design of the experiment/competition. For example, RoboCup@Home gives no specifications about the shape and the size of the apartment, the kind of furniture, and the objects used in the tests.

In competitions, since performance metrics are defined (in terms of a scoring system) to measure a system's overall ability, they are generally not intended to measure specific functionalities or internal features. This is not in contrast with robotics experiments, but the latter are more often used to measure the performance of a system's components to solve a specific problem rather than the entire system.

The following are some important differences between experiments and competitions: 1. in competitions, perfect reproducibility is not possible, since recreating the same scenario (including the same level of background noise) is typically not possible, and 2. repeatability is usually not considered within the competition (each test is normally conducted only once) for organization reasons.

Competitions have thus both similarities and differences with respect to robotic experiments, and the definition of a proper relation between them is an on-going effort. For example, a proposal has been presented for improving the scientific aspects of experiments within robotic competitions and challenges [1]. The proposed framework is based on a modular integrated system and the interoperability of components, where participants can focus on particular modules. Even though this idea has been partially implemented in some competitions, it has not been fully exploited yet. The RoCKIn project also aims for better integration between competitions and scientific benchmarking [3] by collecting benchmarking data during tests that combine the logged internal states of robots and the ground truth acquired through an external system. RoCKIn competitions will thus be an important contribution to this research area.

Another important aspect of a competition is its evolution. Many existing competitions are held periodically (e.g., annually), changing tests over time to address more difficult problems, to enlarge the variety of problems, to provide different experimental conditions, etc. When a competition is run for many years, it is important to evaluate how its design and organization affect the solutions implemented by the participating research groups. More specifically, the choices in defining the tasks to be addressed during a competition and the score that determines the rank of the participants impact how the participants develop their solutions to the specified problems. A good test and scoring system design allows for suitable development of corresponding solutions from participants.

In this article, we are mainly interested in analyzing the results of competitions over the years. In particular, we focus on how this analysis has been carried out in the last years within the RoboCup@Home competition, which is the largest competition for domestic and service robots. The evaluation of a competition's effectiveness based on the scores of the teams over the years is important for better fueling the competition and the novel approach introduced in RoboCup@Home for addressing this issue is the main contribution of this article. Although our analysis reported in this article is specific to the RoboCup@Home competition, we believe that its principles can be applied both to other existing competitions and to new ones, and it could even be used for evaluating multi-year projects or challenges in AI and robotics where some tasks are repeated and changed over the years.

RoboCup@Home is a competition where domestic and service robots perform several tasks in a home environment, interacting with people and with the environment in a natural way. Natural interaction means that a robot is expected to interact with the environment and with other people, as any person would do. So natural forms of human–robot interaction include speech and gestures, but not joysticks or keyboards.

During the competition, the teams are required to perform several tests. Since their total score is the sum of the scores obtained in each test, teams are motivated to perform well in every test in order to gain a high rank. Each test requires a combination of different functionalities (including navigation, object perception and manipulation, person detection, and tracking, etc.) and the score is related to the accomplishment of the task.

RoboCup@Home started in 2006, and its main characteristic is that it changes tests every year while maintaining the same basic functionalities. By changing the difficulty and the combinations of the functionalities to be integrated, we aim