Picking the right robotics challenge

Juxi Leitner recounts how he and his team took part in — and won — the 2017 Amazon Robotics Challenge and reflects on the importance of solving big picture problems in robotics.

Challenges are essential to push applied research forward. For example, the long-running RoboCup has stimulated advances in multi-robot systems and swarm intelligence research for more than two decades. Challenges can also bootstrap new markets: the DARPA Grand Challenge for autonomous vehicles created an initial spark for the billion-dollar investments in self-driving car companies. Amazon saw the potential of robotic technologies and started the Amazon Robotics Challenge in 2015, then called the Amazon Picking Challenge. A company like Amazon has over 100,000 robots in its warehouses, but they mainly move stock around. Creating robots that can not only navigate around environments, but also interact with a range of items within them — for instance, moving objects from shelves into boxes — turns out to be a thorny problem.

Robotic manipulation of objects is complicated due to the interaction of physics, perception and control. Although humans make dexterous movements of the hand all day — such as finding the right key for a car and turning the ignition — these movements are problematic from a scientific and engineering perspective. It turns out that the human hand is an evolutionary marvel of hardware, consisting of bones, joints, ligaments, muscles, arteries and nerves controlled by the brain. At present, designing a robotic hand that emulates biology is so difficult that most researchers simply use a gripper or a suction cup.

The goal of the Amazon Robotics Challenge was to bring industrial and academic robotics researchers together to solve, or partially solve, key problems in how robots grasp and handle objects robustly. Progress over the past decade in machine learning, vision science and hardware design has enabled robots to handle an expanding set of objects and to do so more robustly, so the timing seemed right for such a challenge. Yet the results of the 2015 Amazon Robotics Challenge were sobering. Out of 25 teams, only a few robots managed to successfully pick up objects.

In 2015, my colleagues and I at the Australian Centre for Robotic Vision (ACRV) started our first foray into grasping and manipulation. After early work on a Baxter robot, we felt prepared to enter the Amazon Robotics Challenge in 2016. Using mainly custom, off-the-shelf robots and parts, we quickly developed a system. But just as quickly, we realized that a versatile and functional

robot could not be assembled 'out of the box'. A robot consists of many subparts, leading to complex system interactions. In the 2016 challenge, we built a baseline system that had the right parts to perform grasping. But there were a lot of moving pieces and the rigid structure of our design fell short. Our project failed but provided an important takeaway: flexible integration was what we needed, with components and modules working together.

An important personal realization was that bringing together a team motivated to solve the problem was essential. Moving forward, we had 22 members on 'Team ACRV', mainly undergrad and PhD students. Managing the team was a challenge in itself, especially since we were based thousands of kilometres apart in Adelaide, Brisbane and Canberra. The rules changed and allowed more flexibility in design for the competition in 2017, so we built our robot from scratch, both software and hardware. We conducted weekly full system tests, enabling the comparison of updates and improvements from a holistic viewpoint. This process kept us from merely improving subsystems and losing focus on our end goal. This iterative and flexible approach meant that when something went wrong, we had a pretty good guess what it was. In the end, our solution was the only one that did not use an industrial or humanoid arm. Instead, we designed a Cartesian coordinate robot with a claw and a suction gripper for a 'hand' and a sliding mechanism that picked up objects from above (pictured). We nicknamed our Cartesian manipulation robot Cartman[1].

The 2017 challenge had three stages. First, robots picked specified objects from an assortment of items and placed them in boxes — the 'pick' task. Second, robots selected target items out of a container and placed them in storage — the 'stow' task. And third, robots put items into storage and then lifted a selection of them and put them into boxes — a combined pick and stow task. Compared to

previous years, robots had less space to work in, forcing them to deal with objects next to or on top of each other. Another change was that half of the objects in a task were only revealed 45 minutes before the competition started, so teams could not prepare in advance by programming their robots to manipulate specific objects. To tackle this last added difficulty, we created a computer vision system that could be trained on photos of new objects taken from different angles, which were fed into a deep neural network, the latest in machine learning. Although we didn't place in the top three teams in the first two tasks, our robot performed so well in the finals (the third task) that we took home first place, including an US$80,000 prize.

Robotics needs challenges that sit between current challenges and a big unifying grand challenge, such as the DARPA Robotics Challenge. We recently proposed a Tidy Up My Room Challenge, a teaser of which occurred at the International Conference on Robotics and Automation meeting in 2018. This challenge asks, 'How do you know that an object is out of place?' Visually, a book may look the same on the floor or on the coffee table, yet one place is considered 'tidier'. The challenge is multi-tiered, with increasing complexity in perception, reasoning and manipulation. It provides a way of benchmarking and comparing robotic systems on a task level, instead of focusing on sub-problems. This framework allows researchers to explore a wider design space, including robotic systems that are soft, flexible and deformable, while being less reliant on high-precision object detection and localization. Fundamentally, such challenges bring researchers together to solve outstanding problems, getting us closer to the robots of the future. ❐

Jürgen Leitner[1,2]

[1]Australian Centre for Robotic Vision, Brisbane, Queensland, Australia. [2]LYRO Robotics, Brisbane, Queensland, Australia.
e-mail: j.leitner@qut.edu.au

References
1. Morrison, D. et al. in IEEE International Conference on Robotics and Automation 7757–7764 (IEEE, 2018).