A gripping problem

Designing machines that can grasp and manipulate objects with anything approaching human levels of dexterity is first on the to-do list for robotics.

BY RICHARD HODSON

In the 1973 science-fiction classic *Westworld*, the robots that staff the film's fictitious holiday resort are indistinguishable from their human guests, except for one small clue: the engineers, we are told, haven't perfected the hands yet.

The capabilities of real-world robots fall a long way short of those of *Westworld*'s murderous hosts, but on this small point, reality and fiction are in agreement: hands, and the manipulation of objects, are particularly challenging aspects of robotics. "Grasping is the critical grand challenge right now," says Ken Goldberg, an engineer at the University of California, Berkeley.

In the past 50 years, robots have become very good at working in tightly controlled conditions, such as on car-assembly lines. "You can build a robotic system for one specific task — picking up a car part, for example," says Juxi Leitner, a robotics researcher at the Australian Centre for Robotic Vision (ACRV), based at the Queensland University of Technology in Brisbane. "You know exactly where the part is going to be and where the arm needs to be," he says, because the robot picked up the same thing from the same place "the last million times".

But the world is not a predictable assembly line. Although humans might find interacting with the countless objects and environments found beyond the factory gates a trivial task, it is tremendously difficult for robots. ▶

▶ These unstructured environments represent the next frontier for researchers across the field of robotics, but such environments are particularly problematic for robots that grip. Any robot hoping to interact physically with the outside world faces an inherent uncertainty in how objects will react to touch. "We can predict the motion of an asteroid a million miles away far better than we can predict the motion of a simple object being pushed across the table," Goldberg says.

Some researchers are using machine learning to empower robots to independently identify and work out how to grab objects. Others are improving the hardware, with grippers ranging from pincer-like appendages to human-like hands. And roboticists are also gearing up to tackle the challenge of manipulating objects gripped in the hand.

Advances in robots' ability to handle objects could have enormous societal impact. Commercial entities, particularly those involved in the movement of varied goods, are following developments closely. "There's a big demand. Industry really wants to address this because of how fast e-commerce is growing," says Goldberg. With interest greater now than ever, "It's an opportunity for the research to really be put into practice."

LEARNING TO LEARN

The heightened industry interest is exemplified by an annual competition organized by e-commerce giant Amazon for the past three years. The Amazon Robotics Challenge asks teams of researchers to design and build a robot that can sort the items for a customer's order from containers and place them together in boxes. The items are varied, ranging from bottles and bowls to soft toys and sponges, and are initially jumbled together, which makes it a difficult task in terms of both object identification and mechanical grasping.

In July 2017, Leitner's ACRV team claimed victory with a robot called Cartman, which resembles a fairground 'claw' game. An aluminium frame supports the claw assembly. The robot has two tools for picking up objects, known as end effectors: a gripper with two parallel plates, and a suction cup backed by a vacuum pump. For each object the robot encounters, the researchers specify which effector it should try first. If that doesn't work, the robot switches tools.

First, however, the robot must find the item it's looking for. The team tackled this challenge by using machine learning. The main input for the robot is an RGB-D camera, a technology that is popular among roboticists and that can assess both colour and depth. The camera looks down from the effector into the boxes below. From this vantage point, Leitner explains, Cartman labels each pixel according to the object it belongs to — a form of deep learning known as semantic segmentation. Once a cluster of pixels representing the desired object is found,

the camera's depth-sensing capability helps the robot to work out how to grab the item. "In simple terms, we attach to the bit that pokes out the most," says Leitner.

Rapid advances in machine learning underpin a lot of recent progress in grasping. "Software has been the bottleneck for ages, but it's becoming more advanced thanks to deep learning," says Pieter Abbeel, a deep-learning specialist at the University of California, Berkeley. These developments have, he says, opened up "whole new avenues of robotics applications".

Abbeel is co-founder and chief scientist of covariant.ai, a start-up in Emeryville, California, that uses deep learning to train robots. Rather than program a robot to perform a specific action, humans provide demonstrations that the robot can then adapt to deal with variations of the same problem.

"GRASPING IS THE CRITICAL GRAND CHALLENGE RIGHT NOW."

The human trainer views the camera feed of the robot arm through a headset, and uses motion controls to guide the robot arm to pick up objects. The process feeds a neural network with data on the approach taken. "With just a few hundred demonstrations done in this particular way, you can train a deep neural network to acquire a skill," says Abbeel. "And I don't mean acquire a specific motion that it's going to repeatedly execute, but acquire the ability to adapt the motion to whatever it's seeing in its camera feed."

Goldberg uses machine learning to teach robots to grasp, too. But rather than gather data from real-world attempts, his Dex-Net software is trained virtually. "We can simulate millions of grasps very quickly," he says. The software lets an industrial robot pick objects from a pile with a success rate of more than 90%, even if it hasn't seen those objects before. It can also decide for itself whether to use a parallel-jaw gripper or suction tool for a particular object.

Dex-Net's fourth incarnation will be presented in 2018. According to a metric being developed by Goldberg and roboticists around the world to aid reproducibility, known as mean picks per hour, the Dex-Net system is now among the fastest pickers around. It can achieve over 200 picks an hour — still behind human capacity, estimated at 400–600 picks an hour, but far ahead of the numbers achieved by the teams at the most recent Amazon Robotics Challenge (see 'A measure of success').

But Dex-Net's simulated world is imperfect. The model assumes that objects are rigid, for instance, and does not account for objects that contain liquid. Simulation, Abbeel says, might not always be the easiest way to learn. "The real world provides a simulator for free," he says. "It's great to leverage both ways of doing it."

SOFTLY DOES IT

The combination of a parallel-jaw gripper and suction employed by both Goldberg and Leitner is a popular choice: most teams at the 2017 Amazon Robotics Challenge took this hybrid approach. But how these tools grasp objects — planning contact points before moving into position as precisely as possible — is very different to how we humans use our hands.

"When you pick something like a pen up off a table, the first thing you touch is the table," says Oliver Brock, a roboticist at the Technical University of Berlin. We do not think about where we need to place our fingers. The softness of human hands allows for something called compliant contact — the fingers mould against the surface of the object. "Because you get a lot of surface contact, you can much more intuitively reach out and grab," says Daniela Rus, a roboticist at the Massachusetts Institute of Technology in Cambridge. "With soft fingers, we change the paradigm of grasping."

Many are seeking to exploit the benefits of compliance in grasping by building softness into robot grippers. Brock's lab has developed the RBO Hand 2, a human-like hand with five silicone fingers. The fingers are controlled by the movement of pressurized air, which allows them to curl and straighten as required.
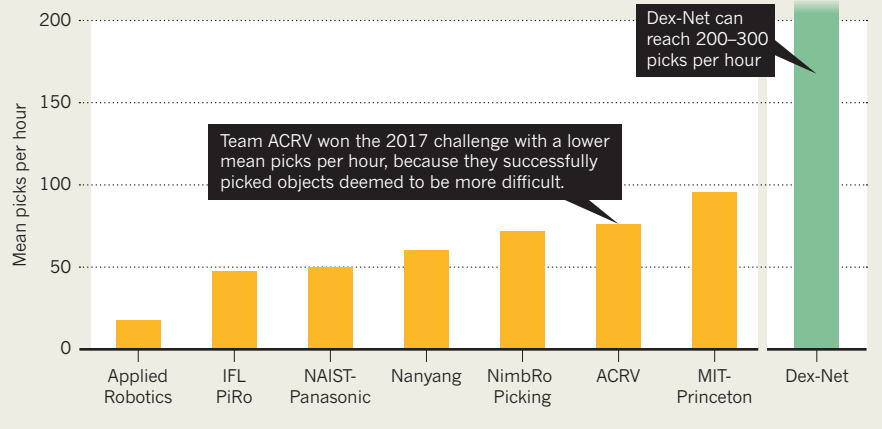
Although the human-like arrangement of the fingers might not be suited to every task, it is ideal for interacting with the world we inhabit. "The world is designed for anthropomorphic hands," says Brock. But there's also a romantic element to the anthropomorphic design of his robotic hand. "It's embarrassing," he says. "People, even roboticists, are more fascinated by things that look human."

The benefits of softness are already drawing commercial interest. Soft Robotics, in Cambridge, Massachusetts, produces air-actuated grippers that have a more claw-like design than Brock's research model. The robots are already being trialled in a factory setting, handling delicate produce without damaging it.

Another start-up, Righthand Robotics in Somerville, Massachusetts, is adding softness to the claw and suction set-up so popular with roboticists. Its claws have three flexible fingers, arranged around a central suction cup that can be extended to draw in objects. The design takes inspiration from birds of prey, says the company's co-founder Lael Odhner. In these birds, most of the forearm musculature is attached to a single group of tendons that reaches to the tip of the claw, he says. Similarly, all the power of the motors in Odhner's robotic claws is put into a single closing motion. This

A MEASURE OF SUCCESS

To make it easier to compare experimental results in robotic grasping, Ken Goldberg, an engineer at the University of California, Berkeley, and others are preparing a metric known as mean picks per hour that could be used as a standard measure across the field. By this metric, the Dex-Net system seems to be considerably more capable than the robots that competed in the 2017 Amazon Robotics Challenge. They are all, however, still far behind humans, who can manage 400–600 picks per hour.

Dex-Net can reach 200–300 picks per hour

Team ACRV won the 2017 challenge with a lower mean picks per hour, because they successfully picked objects deemed to be more difficult.

simple action improves reliability — a crucial consideration for grippers intended for commercial use — at the expense of the ability to perform delicate motions. The extendable suction cup makes up for this shortfall. "It replaces potentially dozens of fine actuators that you would otherwise have to place in the hand," says Odhner.

IN OUR HANDS

Softness is still relatively new to robotics. "It's an extraordinarily powerful concept, but people are just beginning to figure it out," says Rus.

A common criticism is that it's hard for a soft robotic hand to perform a useful action with a grasped object. "You grab it very well," explains Rus, but "you don't know exactly the orientation of the object inside the hand". This makes it tricky to manipulate the object. Goldberg echoes this: "If you just have a soft, enveloping type of hand, then you've really reduced your visibility of that object," he says.

The way people solve this is simple: we touch. But few robots have been granted this capability. "People agree that it is important," says Brock — it's just very difficult to do. He is pursuing two approaches to giving his soft robot hands a sense of touch. The more-mature method involves embedding tubes of liquid metal in a silicone sheet wrapped around the finger. His team can then monitor applied forces using the electrical resistance along the tubes. "It's measuring strain all over the finger and inferring from that, through machine learning, what actually happens to the finger." The team is currently assessing how many of these strain sensors are required on each finger to measure various forces.

Brock's other approach to incorporating touch centres on acoustics. In a proof-of-concept test, a microphone was placed inside the air chamber of a soft finger. The sound that was recorded enabled researchers to identify which part of the finger was touching something, the force of the touch and the material of the object. The ability to tuck the microphone deep inside the finger — and therefore avoid reducing compliance — sidesteps a key issue with sensorizing soft fingers. Brock says that the details of this work will be published soon, and that he plans to work with acoustics specialists to improve his design.

Many roboticists think that there is unlikely to be a universal solution to grasping. Even if robots could achieve human levels of dexterity, Rus points out, "there are lots of things that we cannot pick up with a human hand." But as robots become increasingly adept at handling variability, more tasks currently performed by humans will become automatable.

A report published last year estimates that most occupations could be at least partially automated (go.nature.com/2rjtnud). But Goldberg stresses that robotics does not have to put people out of work. "My goal is not to replace humans," he says. "What I want to do is assist humans."

Whatever the outcome, a lot of development needs to be done before any robot revolution can take place. Leitner's team might have won Amazon's challenge last year, but its robotic arm fell to bits on the first day of competition, and most teams experienced technical issues of some kind or another. "These systems are not super robust yet," Leitner says. "If you were to take that system and put it into an Amazon warehouse, I'm not sure how long it would actually work for."

"I'm someone who is exploring a new world," says Brock, "and I'm not ready yet to spring in to the limelight and make industrial applications left and right." But, as Goldberg says, "there's no barrier to putting these into practice. It's already starting to happen." ■

Richard Hodson is supplements editor for Nature.