



robots are getting smarter all the time.

Developments in hardware and processing power have made it possible to automate tasks much more complex than spot welding or pick-and-place.

Indeed, we live in an age when society may soon have to decide how to integrate autonomous aircraft with conventional air traffic. Personal service robots may one day set the table for us or pour our drinks.

In industry, machines and their artificial intelligence can be productive at ever-more-complex tasks, but the cost of programming them can be prohibitive, especially to small and mid-size enterprises. It is not just a one-time cost of developing the large computer code. Smaller manufacturers are frequently producing short runs and so face a need for frequent reprogramming.

Current manual programming approaches rely on formulating a simple version of the problem and applying a search-based planning algorithm to discover a solution. To generate a problem within the conceptual grasp of a human programmer, the simplifying assumptions might include considering all objects being manipulated to be rigid and ignoring dynamics.

The challenge and cost of programming can increase significantly when the process involves the management of deformable or fluid materials, typically encountered in many routine tasks in industry as well as in our daily lives. But imagine an alternative that can eliminate the need to write lengthy programs that try to foresee the numerous variations possible in a complex task. Suppose you could program a robot by showing it how to do the job.

Robots at EPFL learn to feel the difference between skilled and unskilled welders so they can provide assistance when needed.

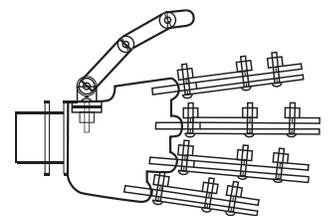
RESEARCH IS MAKING GREAT STRIDES IN IMITATION LEARNING—THE PROGRAMMING

robot see



OF AUTONOMOUS MACHINES BY INDIVIDUAL AND CROWD DEMONSTRATIONS.

robot do



BY KRISHNANAND N. KAIPA, JOSHUA D. LANGSFELD, AND SATYANDRA K. GUPTA

The core idea is that humans cannot spontaneously perform challenging tasks; instead, they gain experience and improve their performance over time as they execute repeated trials.

This is often called “learning from demonstrations” or “imitation learning.” And it isn’t a far-fetched idea. Labs at several institutions—for example, the Swiss Federal Institute of Technology at Lausanne, the University of Maryland, Massachusetts Institute of Technology, and Worcester Polytechnic Institute—are experimenting with technology that may one day make imitation learning common for machines.

The underlying idea of this approach is to allow an agent to acquire the necessary details of how to perform a task by observing another agent (who already has the relevant expertise) perform the same task. Usually, the learning agent is a robot and the teaching agent is a human. Often, the goal of imitation learning approaches is to extract some high-level details about how to perform the task from recorded demonstrations.

Research into imitation learning has achieved some impressive results ranging from training unmanned helicopters to perform complex maneuvers to teaching robots general-purpose manipulation tasks.

One early implementation reported in 2004 was focused on teaching a helicopter to hover in place and perform a few maneuvers.



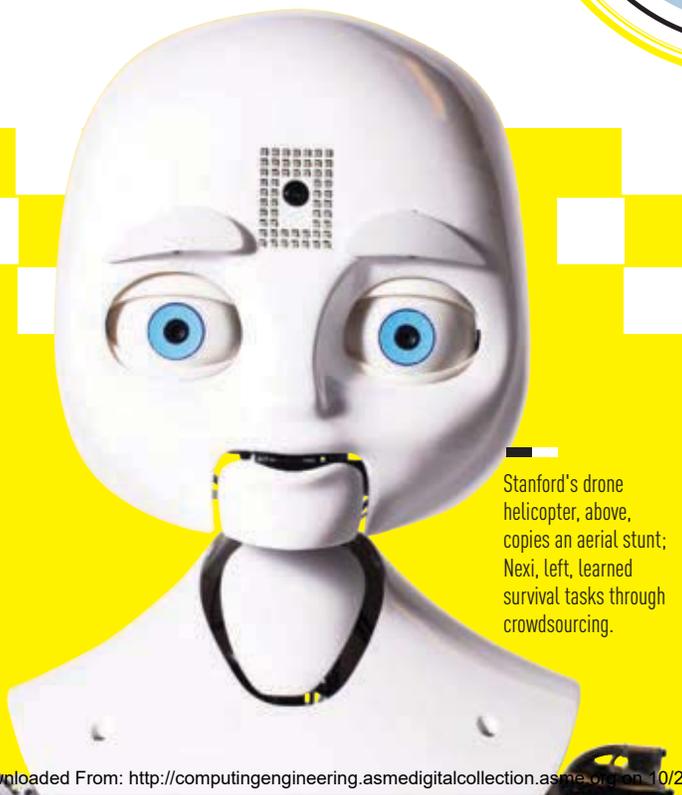
Learning of more complex maneuvers like in-place flips from human demonstrations was reported in 2010. Researchers at the Stanford AI Lab led by director Andrew Ng were able to train the small helicopter to perform complex stunts by observing the behavior of an expert pilot performing them.

The robot could not simply copy the inputs given by the pilot since no two runs of a stunt are exactly equal. The pilot is continually compensating for disturbances while performing the task. Successful replication of a stunt by an autonomous robot requires a learning algorithm to extract the desired characteristics of the task from one or more demonstrations and to develop a policy to reproduce those characteristics at an acceptable level of performance. The apprenticeship learning approach has the advantage of not requiring any input from a human programmer to define the stunt motions. This is difficult to do, as helicopter dynamics are complicated and often only known implicitly, assuring that any hand-coded control algorithm is likely to fail on maneuvers that require capabilities of the helicopter close to its limits. Learning from expert demonstrations serves as a feasible option in such cases.

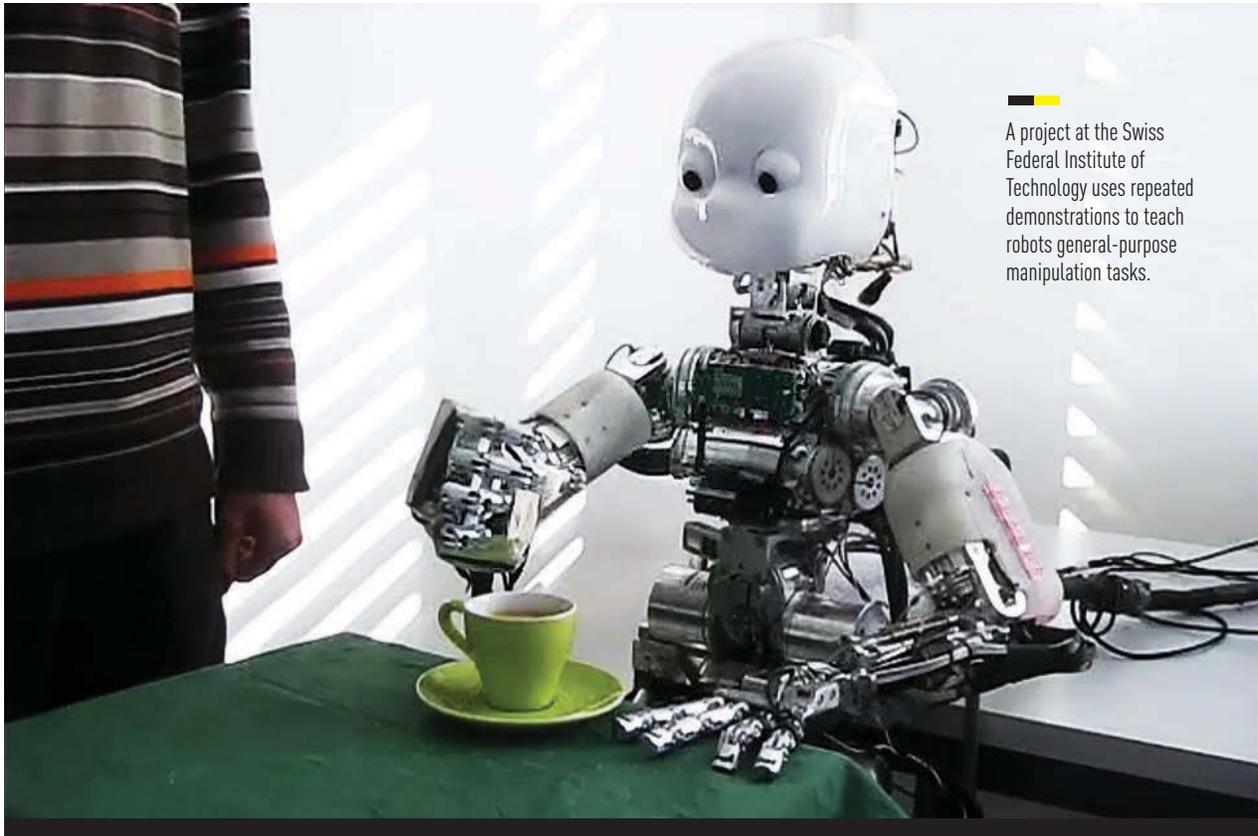
The general approach used by the Stanford researchers involves the assumption that the pilot is demonstrating a noisy version of the desired trajectory to execute the stunt. Multiple demonstrations of the same stunt have differing amounts of noise in different sections, and so a good averaging algorithm is able to extract an appropriate trajectory. Using an expectation maximization algorithm, the system is able to produce the average trajectory by simultaneously aligning the demonstrations temporally and finding trajectory elements that can be achieved by the helicopter, given its dynamics.

The net result is that the helicopter can perform the stunts with a consistency comparable to that of the expert pilot. The helicopter was able to transition between learned stunts without returning to neutral flight state between stunts.

Researchers at the Learning Algorithms and Systems Laboratory of the École Polytechnique Fédérale de Lausanne, led by Aude Billard, have been incorporating learning from demonstrations for a variety of robot tasks. They are primarily concerned with teaching humanoid robots various general-



Stanford's drone helicopter, above, copies an aerial stunt; Nexi, left, learned survival tasks through crowdsourcing.



A project at the Swiss Federal Institute of Technology uses repeated demonstrations to teach robots general-purpose manipulation tasks.

purpose manipulation tasks, from setting a table with cutlery and flatware to putting objects into a container. These tasks may involve learning when to grasp, how to grasp, and what trajectories to follow.

The researchers have developed algorithms that are able to extract the desired characteristics of a task from repeated demonstrations. The approach enables different demonstrators to illustrate different aspects of a manipulation task and generalizes the demonstrations into a single cohesive model for achieving the task goals. They use a probabilistic framework to encode the demonstration data and extract important constraints for achieving the task.

most traditional approaches to imitation learning in the robotics area only utilize a small number of successful human demonstrations. These demonstrations are used to construct a model that identifies parameters to be used by the robot in doing the same task. If the robot is unable to do the task using the parameters prescribed by the model, then the approach fails.

The reasons for failures are often insufficient number of demonstrations or the subtle differences between the robot and the human that are not modeled. This phenomenon is generally referred to as the correspondence problem in the imitation learning and cognitive science communities.

Relying on demonstrations or a model that captures all the differences between the human and the robot is impractical. We need a robust approach to imitation learning that antici-

pates failures in the transfer of skills from the human to the robot and has built-in features to recover from it.

Human operators often need to perform challenging tasks multiple times in order to be able to reach an acceptable level of performance. Typically, humans make a lot of errors during early phases of learning.

They learn the appropriate coordination by using the motor error to adjust their neural command over repeated trials. This provides a different approach to imitation learning: In addition to learning from successful demonstrations, robots can also learn from errors made by human operators and how they recovered from these errors in subsequent trials.

We and our collaborators in the Maryland Robotics Center at the University of Maryland in College Park are working on developing imitation learning algorithms, with a particular emphasis on learning from failures. The core idea is that humans cannot spontaneously perform challenging tasks; instead, they gain experience and improve their performance over time as they execute repeated trials.

Accordingly, all demonstrations, whether successful or not, can be recorded and learned from. The robot learns a model of the human's behavior which can provide a means to act in a novel variation of the task and a strategy to adjust its behavior when failure occurs. It imitates how the demonstrator converged to a successful trial.

We are currently focused on the problem of pouring liquid into a moving container placed on a rotating platform. This scenario takes inspiration from an assembly line at a

small or medium-size manufacturing firm, where task requirements might change rapidly and purchasing specialized automation hardware for each variation would be too expensive.

The goal is to use general-purpose robotic manipulators that can be easily trained. The task of pouring a liquid into a container while it is moving is a challenging problem for current autonomous planners because it is difficult to accurately model the fluid dynamics in real time. The task is therefore relatively more amenable to directly learning from human demonstrations.

During experiments, we observed that humans approach this task with no experience and rapidly converge to a successful behavior. Initial results indicate that there is valuable information in a trial where the demonstrator failed at performing the task, primarily in terms of how that experience affected their behavior in the next trial.

Our approach involves extracting this adjustment strategy as a function of current performance in order to imitate not just the task itself, but how to improve and succeed in cases of failed attempts. The algorithm developed on these ideas was illustrated using a robotic arm that failed initially, learnt from its failures, and eventually succeeded at performing the pouring task.

p eople often wonder how well the learned components of autonomy will perform in situations not encountered during demonstrations. Extensive physical experiments to predict reliability would be costly in time and money. Clearly a demonstration that might pose a threat to the human or the robot has to be avoided. So conducting demonstrations in the virtual world is emerging as an attractive alternative.

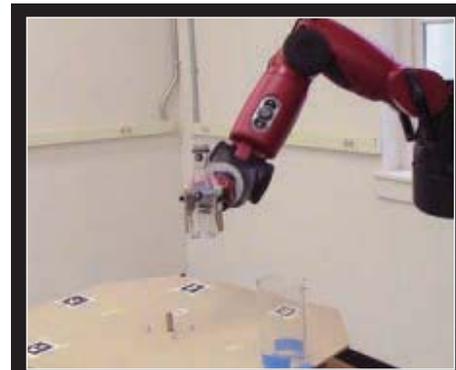
Over the last few years, tremendous progress has been made in the area

A demonstrator at the Maryland Robotics Center pours liquid into a container on a revolving platform. The arrangement is intended to simulate some of the challenges of a factory assembly line.



of physics-based robot simulators. For example, the DARPA Robotics Challenge is making an extensive use of simulation technology to test autonomy components.

By combining advances in multi-player online games and accurate robot simulations, new games can be developed in which humans can compete and collaborate with each other by teleoperating virtual robots. This advancement means that demonstrations need not be confined to a few experts. Instead, anyone with an Internet connection can participate in the training of a new robot. That is how



The robotic arm is taught by successful and unsuccessful

DARPA used a publicly distributed Anti-Submarine Warfare game to learn how to track quiet submarines.

Integrating virtual world demonstrations with advances in crowdsourcing takes imitation learning to a new level.

At the MIT Media Lab, in conjunction with Worcester Polytechnic Institute, Cynthia Brazeal and Sonia Chernova are working on enabling large-scale robot learning from crowdsourcing of demonstrations. Their work is inspired by the Restaurant Game project, where thousands of human players were able to interact with each other in the form of virtual avatars inside a restaurant, generating example behaviors of how customers and waiters behave.

This large amount of data could then be mined to produce generalized behavior models that respond appropriately to previously unseen contexts. The concept was extended to a virtual survival game on Mars, where a

human and a robot must collaborate on a physical task of salvaging items and escaping to a spaceship.

In this scenario, physical constraints on the robot—such as an inability to traverse stairs when only wheeled locomotion is available—are modeled. Both the robot and human virtual agents would be controlled by human players, producing a database of samples that illustrate how the two agents should collaborate given the motion and task constraints.

The results from this virtual interaction

TO LEARN MORE

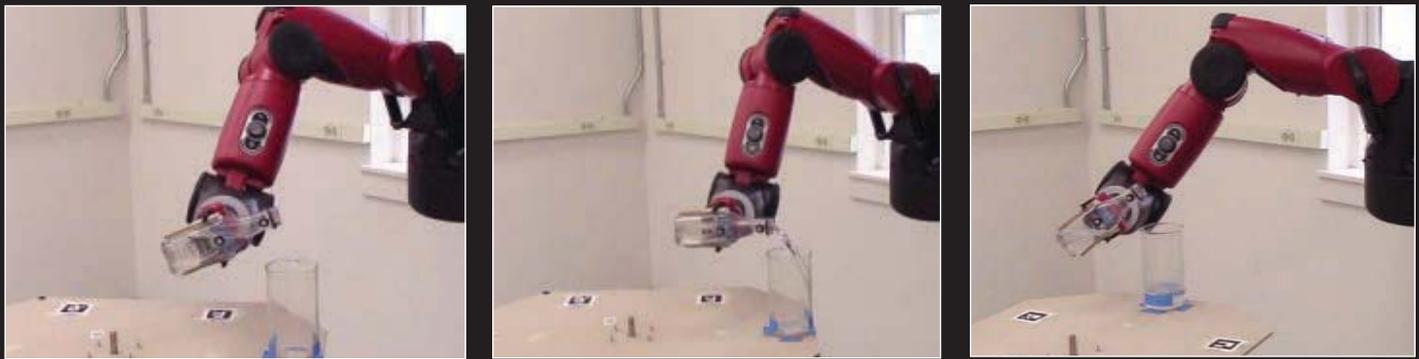
"A Survey of Robot Learning From Demonstration," B. D. Argall, S. Chernova, M. Veloso, and B. Browning. (2009). *Robotics and Autonomous Systems*, 57(5): 469–483.

"Autonomous Helicopter Aerobatics Through Apprenticeship Learning," P. Abbeel, A. Coates, and A.Y. Ng. (2010). *The International Journal of Robotics Research*, 29(13): 1608–1639.

given that the two scenarios have substantial differences that change human behavior.

Crowdsourcing provides a rich diversity in demonstrations and hence enhances the probability of generalization. Some of the participants are likely to exhibit unconventional thinking and demonstrate a highly creative or innovative way of doing a task. This is great news for optimizing robot performance.

For some people, this way of training robots might serve as a means to earn money by performing demonstrations (basically acting as robot tutors). Playing games that involve robots is likely to be entertaining for at least a segment of the population.



demonstrations. It learns from its own failures, much as humans do, and adjusts its behavior to pour the liquid with reasonable accuracy into the moving container.

were extracted and applied to a physical version of the survival scenario where Nexi, a mobile, dexterous, and social robot assisted a human participant in performing the survival tasks of gathering needed items. Their results indicated that the robot was able to incorporate crowd-sourced knowledge in its behavior to perform the collaborative task at a comparable level of ability as a predefined behavioral script, executed by a teleoperator.

The researchers report seeing spontaneous collaborative behavior between the human and robot solely based on the crowdsourced data. They conclude that using the ability of crowds to generate the behavior policy of the robot is promising, but still requires significant work with several open challenges. For example, they must determine how to assure that the player of the virtual game will act in the same manner as with a physical robot,

"Catching Objects in Flight," S. Kim, A. Shukla, and A. Billard. (2014). *IEEE Transactions on Robotics*, 30(5). In press.

"Crowdsourcing Human-Robot Interaction: New Methods and System Evaluation in a Public Environment," C. Breazeal, N. DePalma, J. Orkin, S. Chernova, and M. Jung. (2013). *Journal of Human-Robot Interaction*, 2(1): 82–111.

"Towards Imitation Learning of Dynamic Manipulation Tasks: A Framework to Learn From Failures," J. Langsfeld, K.N. Kaipa, R. J. Gentili, and S.K. Gupta. (2014). *Simulation Based System Design Lab Technical Report*.

"Keyframe-Based Learning From Demonstration," B. Akgun, M. Cakmak, K. Jiang, and A.L. Thomaz. (2012). *International Journal of Robotics Research*, 4(4): 343–355.

"Active Learning From Demonstration for Robust Autonomous Navigation," S. David, J.A. Bagnell, and A. Stentz. (2012). *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 200–207.

This paradigm can also be used in situations where a robot is stuck during a difficult task and needs a creative solution to get out of the bind.

We expect demonstrators recruited via crowdsourcing to be non-experts and therefore to fail, but robots can still learn from those failures, just as humans do. Imitation learning methods that make use of both successful and failed demonstrations are suited to exploiting the benefits of crowdsourced demonstrations.

Automatically learning reasoning rules and skills from a vast amount of demonstration data is an interesting challenge and will keep the research community busy for many years to come. But this seems to be the much-needed crucial advancement to reduce the cost of autonomous robots. **ME**

KRISHNANAND N. KAIPA is a research assistant professor and **JOSHUA D. LANGSFELD** is a graduate student in the Department of Mechanical Engineering at the University of Maryland. **SATYANDRA K. GUPTA** is a professor in the Department of Mechanical Engineering and the Institute for Systems Research at UMD.