ARTIFICIAL INTELLIGENCE

Task-agnostic self-modeling machines

Robert Kwiatkowski[1]* and Hod Lipson[2,3]

A robot modeled itself without prior knowledge of physics or its shape and used the self-model to perform tasks and detect self-damage.

We humans are masters of self-reflection. We form a mental picture of ourselves by revisiting past experiences and use that mental image to contemplate future scenarios. Our mental self-image contains information about our body configuration in physical space. Our self-image also gives us the ability to link future actions with likely sensations. Your mental self-image allows you to imagine yourself walking on the beach tomorrow, smelling the ocean and feeling the sand under your feet.

An accurate self-image will be key to allowing robots to learn and plan internally without resorting to costly training in physical reality for each new task. The ability to self-simulate can create an illusion of one-shot learning, whereas in actuality, adaptation involves incremental learning or planning inside an internal self-image. A self-image can also be used to identify and track damage, wear, or growth.

Humans likely acquire their self-image early in life and adapt it continuously (*1*). However, most robots today cannot generate their own self-image. Although recent advances in machine learning have allowed robots to become increasingly adept at understanding the world around them, when it comes to understanding themselves, most robots today still rely on a hard-coded simulator (*2*,*3*). These designer-provided simulators are laborious to construct and invariably become outdated.

As an alternative to self-modeling, many robotic systems do without a model altogether by using end-to-end training for a specific task, applying techniques such as model-free reinforcement learning (*4*). Such task-specific learning may be good for narrow artificial intelligence (AI) but lacks the generality and transferability required for robots capable of continuously learning new tasks throughout their lives.

Here, we suggest that, because the robot itself is persistent across multiple tasks, there is strong incentive to extract a self-model and then reuse that self-model repeatedly to learn new tasks. Moreover, by separating the self from the task, every future experience can be used to refine a common self-model, leading to continuous self-monitoring.

Early adaptive control methods also attempted to tune parameters of a fixed analytical self-model (*5*). We previously used evolutionary algorithms to find the morphology most consistent with the robot's recorded action-sensation pairs (*6*), but both approaches make many assumptions. A key question remained: Can a robot create a self-model with no prior knowledge?

First, we chose a physical robot with four coupled degrees of freedom. The robot recorded action-sensation pairs by moving through 1000 random trajectories (Fig. 1, step 1). Actions correspond to four motor angle commands and sensations correspond to the absolute coordinate of the end effector. This step is not unlike a babbling baby observing its hands. The entire captured dataset is provided in (*7*).

Importantly, when a robot's motors are commanded to achieve some target angles, they do not necessarily reach those angles due to hysteresis, self-collision, structural flexibility, and other effects. Therefore, a high-fidelity self-model must capture not just the direct geometric transformations from the robot's base to the end effector, but an implicit relationship between current positions, new motor commands, past positions, and past motor commands.

We used deep learning to train a self-model (Fig. 1, step 2). Using the acquired self-model, the robot could apply a simple planning strategy to accomplish a variety of tasks. We tested the performance of the robot on two separate tasks: a pick-and-place task and a handwriting task (Fig. 1, step 3), both in open and in closed loop.

Closed-loop control allows the robot to recalibrate its actual position between each step along the trajectory by using positional sensor feedback. In contrast, open-loop control involves carrying out a task based entirely on the internal self-model, without any external feedback, like reaching for your nose with eyes closed. This test is also frequently used to test human cerebellar dysfunction such as dysmetria (*8*).

We ran multiple tests with both explicit and implicit representations. Explicit models capture the relationship between motor commands and end-effector position but cannot handle self-collision. Implicit models capture the sequential relationship between state-action pairs and thus are more general. In open-loop tests where planning was completed successfully, the median distance between the physical effector and the target was 4.3 cm. In closed loop, the self-model achieved a median physical error of 0.6 cm, an error lower than the accuracy between the analytical model and the physical robot. The explicit model achieved a median accuracy of 0.65 cm. These results suggest that acquired self-models would be able to successfully execute internal planning and learning on par with a conventional simulator.

The second test involved a combination of subtasks and gripper actuations. The robot used its self-model to plan how to pick and deposit nine red balls, each 20 mm in diameter. In closed loop, the self-models had sufficient accuracy and consistency to execute a pick-and-place task with precision similar to analytical forward kinematics. The open-loop pick rate was 44%, whereas the place rate was 100% of successful picks, and most failures were a result of the planner, not of the self-model.

To demonstrate that the same self-model could be used for other tasks without additional task-specific retraining, we performed a second task involving simple handwriting with a marker. This task was used for qualitative assessment only.

We concluded by replacing one of the robot arms with a longer and slightly deformed

[1]Department of Computer Science, Columbia University, New York, NY 10027, USA. [2]Department of Mechanical Engineering, Columbia University, New York, NY 10027, USA. [3]Data Science Institute, Columbia University, New York, NY 10027, USA.
*Corresponding author. Email: robert.kwiatkowski@columbia.edu

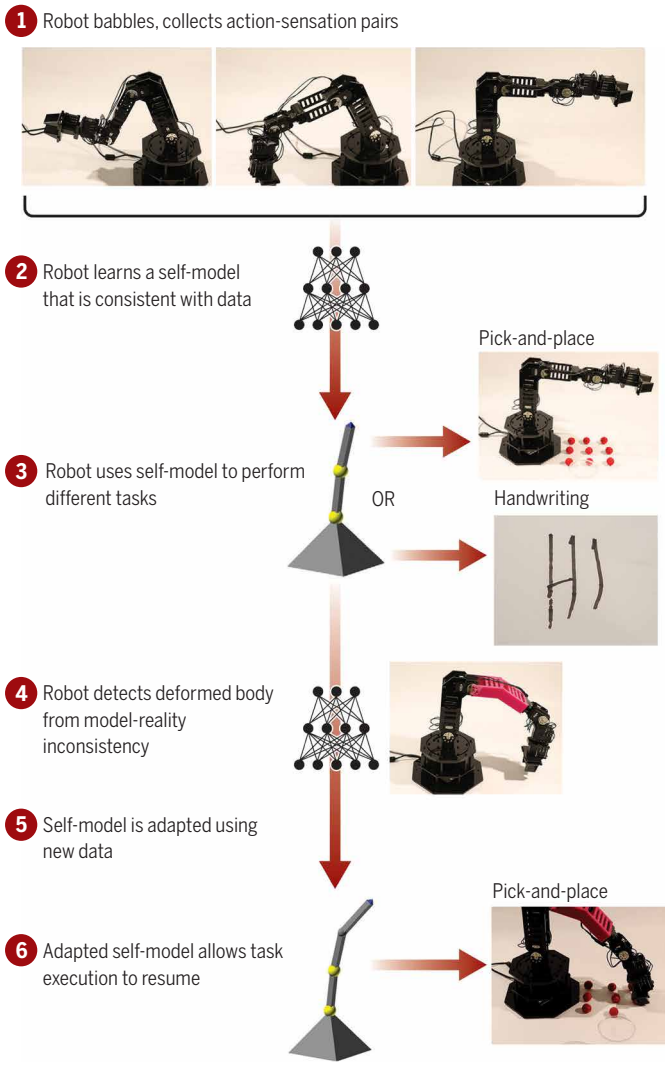

Fig. 1. Self-model generation, usage, and adaptation. An outline of the self-modeling process from data collection to task planning. (Step 1) The robot recorded action-sensation pairs. (Step 2) The robot used deep learning to create a self-model consistent with the data. (Step 3) The self-model could be used for internal planning of two separate tasks without any further physical experimentation. (Step 4) The robot morphology was abruptly changed to emulate damage. (Step 5) The robot adapted the self-model using new data. (Step 6) Task execution resumed.

part as a proxy for unanticipated morphological damage (Fig. 1, step 4). The robot was able to detect the change and retrain the self-model using 10% additional data (Fig. 1, step 5). The new self-model allowed the robot to resume its original pick-and-place task with little loss of performance (Fig. 1, step 6).

Robotics research has historically split between two camps: model-predictive control and model-free learning. We propose a hybrid where machine learning acquires a self-model that is then reused to perform planning or learning internally. This way, data collected in the course of any task can help refine the self-model and thus transfer to other tasks.

Self-imaging will be key to allowing robots to move away from the confinements of so-called narrow AI toward more general abilities. We conjecture that this separation of self and task may have also been the evolutionary origin of self-awareness in humans.

CREDIT: A. KITTERMAN/*SCIENCE ROBOTICS*

SUPPLEMENTARY MATERIALS

robotics.sciencemag.org/cgi/content/full/4/26/eaau9354/DC1
Section S1. Related work
Section S2. Experimental robot platform
Section S3. Training the self-model
Section S4. Results
Section S5. Self-model adaptation to damage
Section S6. Limitations
Fig. S1. The robot and data used for this study.
Fig. S2. Self-modeling process.
Fig. S3. Self-model architecture and training.
Fig. S4. Two tasks.
Fig. S5. Accuracy degradation.
Fig. S6. Diagram of the reach of the WidowX robot arm taken from Trossen Robotics.
Fig. S7. Rendering of the CAD model used to 3D-print the deformed arm length.
Fig. S8. Analytical model position versus physical robot position.
Fig. S9. The distribution of accuracies when using the forward kinematics model.
Fig. S10. The distribution of accuracies visualized onto the reachable space.
Fig. S11. The forward kinematics learner architecture.
Table S1. Summary of results of trajectory planning and pick-and-place tests.
Table S2. Joint position.
Table S3. Table of model parameters in the model used to conduct all self-modeling tests.
Table S4. Table of positions the arm was instructed to go to in the pick-and-place test.
References (*9*–*16*).
Movie S1. Overview video.
Data S1. 3D printed part model .stl file.

REFERENCES AND NOTES

1. P. Rochat, Five levels of self-awareness as they unfold early in life. *Conscious. Cogn.* **12**, 717–731 (2003).
2. N. Koenig, A. Howard, Design and use paradigms for Gazebo, an open-source multi-robot simulator, in *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2004), pp. 2149–2154.
3. E. Todorov, T. Erez, Y. Tassa, MuJoCo: A physics engine for model-based control, in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2012), pp. 5026–5033.
4. J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, Proximal policy optimization algorithms. arXiv:1707.06347 [cs.LG] (20 July 2017).
5. I. Doré Landau, R. Lozano, M. M'Saad, A. Karimil, *Adaptive Control: Algorithms, Analysis and Applications* (Springer, 2011).
6. J. Bongard, V. Zykov, H. Lipson, Resilient machines through continuous self-modeling. *Science* **314**, 1118–1121 (2006).
7. R. Kwiatkowski, Source code for Task-Agnostic Self-Modeling Machines, GitHub; https://github.com/rjk2147/Task-Agnostic_Self-Modeling_Machines.
8. J. D. Schmahmann, Disorders of the cerebellum: ataxia, dysmetria of thought, and the cerebellar cognitive affective syndrome. *J. Neuropsychiatry Clin. Neurosci.* **16** (3), 367–378 (2004).
9. B. Damas, J. Santos-Victor, An online algorithm for simultaneously learning forward and inverse kinematics, in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2012), pp. 1499–1506.
10. J. Sturm, C. Stachniss, W. Burgard, A probabilistic framework for learning kinematic models of articulated objects. arXiv:1405.7705 [cs.RO] (16 January 2014).
11. N. Haber, D. Mrowca, L. Fei-Fei, D. L. K. Yamins, Emergence of structured behaviors from curiosity-based intrinsic motivation. arXiv:1802.07461 [cs.LG] (21 February 2018).
12. B. Amos, L. Dinh, S. Cabi, T. Rothörl, S. Gómez Colmenarejo, A. Muldal, T. Erez, Y. Tassa, N. de Freitas, M. Denil, Learning awareness models. arXiv:1804.06318 [cs.AI] (17 April 2018).
13. J. Chung, C. Gulcehre, K. Cho, Y. Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv:1412.3555 [cs.NE] (11 December 2014).
14. A. Krizhevsky, I. Sutskever, G. E. Hinton, ImageNet classification with deep convolutional neural networks, in *Advances in Neural Information Processing Systems 25* (NIPS 2012), pp. 1097–1105.
15. S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167 [cs.LG] (11 February 2015).
16. D. P. Kingma, J. Ba, Adam: A method for stochastic optimization. arXiv:1412.6980 [cs.LG] (22 December 2014).

Acknowledgments: Supported by DARPA MTO grant HR0011-18-2-0020.

10.1126/scirobotics.aau9354

Citation: R. Kwiatkowski, H. Lipson, Task-agnostic self-modeling machines. *Sci. Robot.* **4**, eaau9354 (2019).

Task-agnostic self-modeling machines

Robert Kwiatkowski and Hod Lipson

Sci. Robotics **4**, eaau9354.
DOI: 10.1126/scirobotics.aau9354