

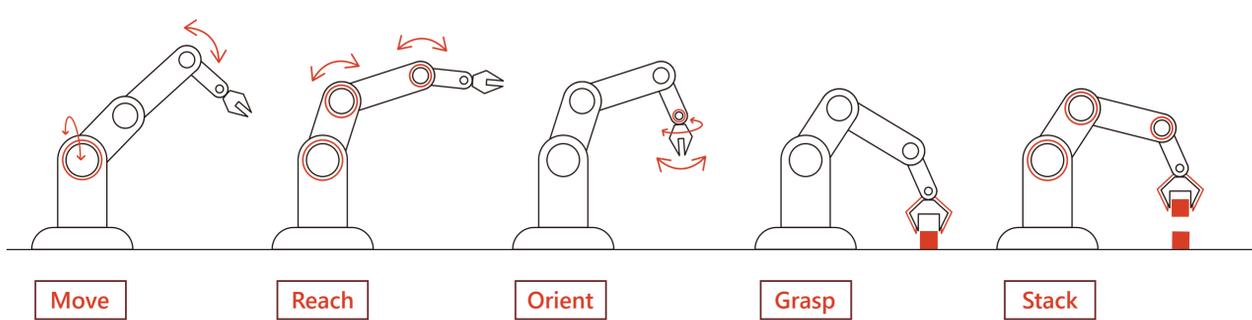
How AI learns through reinforced practice

Some tasks are simple and consistent.
But many others are fuzzy and variable.



Programming an AI may be the right choice for simple, consistent tasks. But deep reinforcement learning (DRL) is often the best way to prepare an AI for success with those fuzzy, variable tasks. DRL lets the AI practice in a simulated environment, adapting to changing conditions and figuring out how to achieve its goals. Here's an example of an autonomous AI that is powered by both programming and DRL.

A team at Microsoft created a simulation in which an AI brain using a Kinova Jaco robotic arm learned to place a rectangular block precisely on top of a cube.

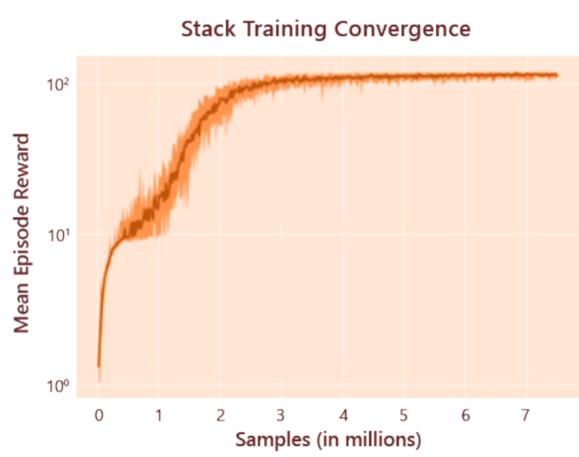
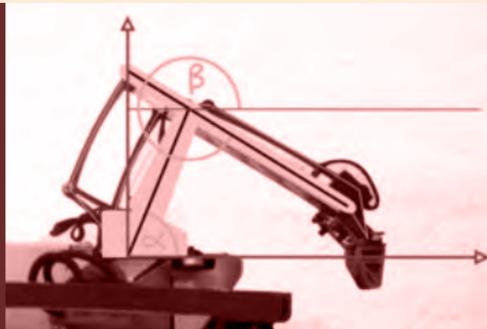


The robotic arm AI had to learn specific skills:



- Move**
Moving the arm back and forth laterally from the shoulder.
- Reach**
Extending the arm in a straightening motion using elbow and wrist.
- Orient**
Orienting the end effector hand palm-down around a block using the wrist.
- Grasp**
Grasping a block by closing the fingers.
- Stack**
Placing a block on another block using a combination of shoulder, elbow, wrist, and fingers.

The move and reach skills are fairly simple. The team taught those skills by programming them using inverse kinematics, calculating how to adjust the robot arm axes to end up in a certain position.

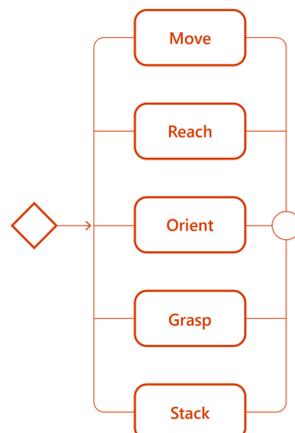


The orient, grasp, and stack skills are fuzzier, so the AI learned them using DRL modules. The research team set up practice sessions with goals and rewards for each individual skill.

For example, the AI was rewarded for keeping the robot wrist farther from the shoulder to stop it from folding in on itself. Through trial and error, the AI figured out how to maximize rewards and minimize penalties when performing each skill. This diagram shows how the AI gained competence while practicing the stacking skill.

Finally, the AI learned the supervisor skill, which involved selecting the right skill to use at the right time in order to stack the block correctly.

To teach the AI efficiently, the research team chose learning methods that were best suited to each skill. Programming worked well for the simple skills. DRL was ideal for the complex skills because the AI could run through millions of iterations in the simulation to determine what worked best. In particular, the AI was very quick to learn the supervisor skill (how to combine skills in the right sequence) because it had already been taught to use the individual skills successfully.



Read more about how AI learns in the white papers [Designing modular AI that explicitly learns skills and strategies](#) and [The secret to useful, production-ready AI](#).