

IIT Madras develops algorithms that learn like humans

Deep reinforcement learning is a way for AI to learn from its mistakes

SHUBASHREE DESIKAN

It is known that DeepMind, the company which was acquired by Google, produced an algorithm called AlphaGo that beat the world's number one at the Go game. One of the methods behind the success of AlphaGo, called deep reinforcement learning, is being further developed by IIT Madras researchers to construct their own algorithm to play not just the Go game, but for more complex tasks.

What they build into the algorithm is not just learning, but learning from mistakes as well.

"There are two parts to engineering this - one involves incorporating features into the neural network that will get the program to recognize parts of the screen [when playing a game]. The other part involves making associations between utilities and action - for instance deciding whether to move left or right based on a specific pattern on the screen," explains Prof. B Ravindran who heads the Robert Bosch Centre for Data Science and Artificial Intelligence, at IIT



Careful AI: "We are planning to build in concepts of risk-awareness through deep reinforcement learning," says Ravindran. ■ SPECIAL ARRANGEMENT

Madras.

The team trained the algorithm using "experts" that were basically programs that had mastered a method of playing the game. Apart

from this, the algorithm was also made to learn "from scratch" - that is, without the intervention of experts.

Not just this, the manner of learning mimics humans.

For instance, humans don't change their strategy too fast, usually. So if the player [a bot or an algorithm] takes a left turn, it continues to do that for a predetermined time. This incorporated smoothness into the decision making. "When we came up with algorithms that incorporated this, we observed improvement by several thousand per cent in the learning performance," says Prof. Ravindran.

Squash to tennis

If a player knew how to play squash, can she use that knowledge to play tennis? This is known as transfer learning. Within this there are various things to contend with - selective transfer, which is, in the example of tennis, akin to learning the forehand of one player and the backhand of another player. This sort of hybrid-making can come of use when the machine learns from different "experts" with different skills.

Another ability built into the program was a tendency to avoid negative transfer. That is, if the "expert" that the program was learning from is actually bad at the

game, the algorithm stops following this expert and chooses a different option - which may be following another expert or learning from scratch by itself. Prof. Ravindran explains by showing a graph in which relative performances of various programs that have been tutored with and without these features have been mapped out. The results clearly demonstrate the usefulness of incorporating the selective transfer and avoidance of negative transfers.

Having worked on the relatively simple arcade games, the team now plans to move on to more complex tasks involving higher-level skills. Humans operate at different levels of granularity in decision making, also we incorporate memory easily into learning. Can this be taught to machines?

They could be working on self-driving cars very soon: "We are planning to build in concepts of risk-awareness through deep reinforcement learning. To apply these ideas to robotics and, say, self-driving cars, there needs to be safety and risk-awareness built in. So we are working on this," he says.