Giving robots the ability to deceive

Sept. 9, 2010
Courtesy of Georgia Tech
and World Science staff

A robot tricks an enemy soldier by creating a false trail and then hiding. While this sounds like a scene from one of the Terminator movies, it's actually the scenario of an experiment conducted by researchers at the Georgia Institute of Technology as part of what is believed to be the first detailed study of robot deception.

A black and a red robot play a game of hide-and-seek. (Courtesy Georgia Tech)

Computer programs newly developed at Georgia Tech "allow a robot to determine whether it should deceive a human or other intelligent machine," said Ronald Arkin, a computer scientist at the university. They also "help the robot select the best deceptive strategy to reduce its chance of being discovered."

Techniques designed by Arkin and colleagues are designed to let a robot deceive another robot, but the principles involved would also apply to robot-human interactions, the researchers said. Results were published online on Sept. 3 in the *International Journal of Social Robotics*. The research was funded by the U.S. Office of Naval Research.

Robots capable of deception may be useful in various areas, including military and search and rescue operations, researchers say. A search and rescue robot may need to deceive in order to calm or receive cooperation from a panicking victim. Robots on the battlefield with the power of deception would be able to successfully hide and mislead the enemy to keep themselves and valuable information safe.

"Most social robots will probably rarely use deception, but it's still an important tool in the robot's interactive arsenal because robots that recognize the need for deception have advantages in terms of outcome compared to robots that do not recognize the need for deception," said the study's co-author, Alan Wagner, a research engineer at the Georgia Tech Research Institute.

For the study, the researchers focused on the actions, "beliefs" and communications of a robot trying to hide from another robot. Their first step was to teach the deceiving machine how to recognize a situation warranting deception. Wagner and Arkin used approaches known as interdependence theory and game theory to develop formulas that tested the value of deception in a specific situation. A situation had to satisfy two key conditions to warrant deception: there must be conflict between the deceiving robot and the seeker, and the deceiver must benefit from the deception.

Once a situation was deemed to warrant trickery, the robot carried it out by providing false information to benefit itself.

The researchers ran 20 hide-and-seek experiments with two autonomous robots. Colored markers were lined up along three potential pathways to places where the robot could hide. A hider robot randomly chose a hiding place and moved there, knocking down colored markers along the way. Once it reached a point past the markers, the robot changed course and hid in one of the other two locations. The presence or absence of standing markers indicated the hider's location to the seeker

robot.

“The hider’s set of false communications was defined by selecting a pattern of knocked over markers that indicated a false hiding position in an attempt to say, for example, that it was going to the right and then actually go to the left,” said Wagner. The hider robots managed to deceive the seekers in three-fourths of trials, with failures resulting from the hiding robot’s inability to knock over the markers that would produce the desired effect.

The results “weren’t perfect, but they demonstrated the learning and use of deception signals by real robots,” said Wagner. “The results were also a preliminary indication that the techniques and algorithms described in the paper could be used to successfully produce deceptive behavior in a robot.”

There are also ethical implications that need to be considered to ensure that these creations don’t harm society, the researchers said. “We have been concerned from the very beginning with the ethical implications,” explained Arkin. “We strongly encourage discussion about the appropriateness of deceptive robots to determine what, if any, regulations or guidelines should constrain the development of these systems.”