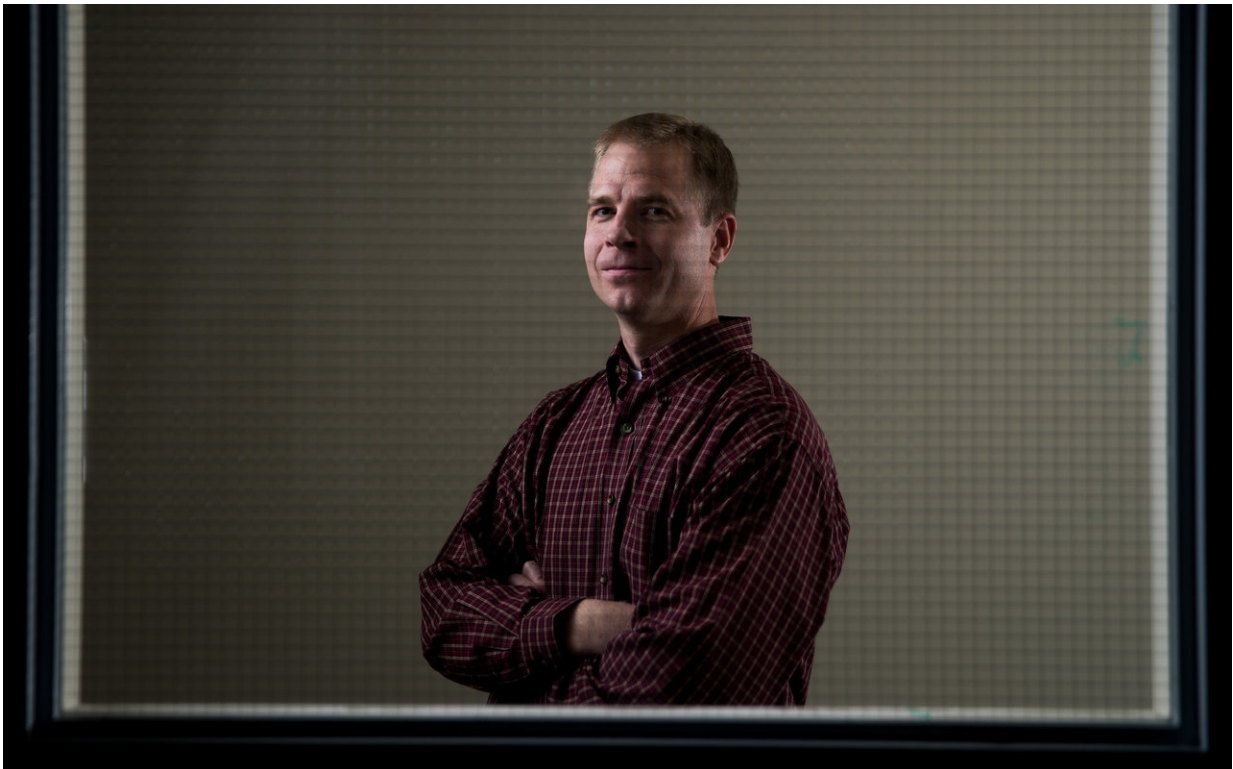


Let's make a deal: Could AI compromise better than humans?

January 19 2018, by K.c. Miller



BYU professor and lead author Jacob Crandall. Credit: Jaren Wilkey/BYU

Computers can play a pretty mean round of chess and keep up with the best of their human counterparts in other zero-sum games. But teaching them to cooperate and compromise instead of compete?

With help from a [new algorithm](#) created by BYU computer science professors Jacob Crandall and Michael Goodrich, along with colleagues at MIT and other international universities, machine compromise and [cooperation](#) appears not just possible, but at times even more effective than among humans.

"The end goal is that we understand the mathematics behind cooperation with people and what attributes [artificial intelligence](#) needs to develop [social skills](#)," said Crandall, whose study was recently published in *Nature Communications*. "AI needs to be able to respond to us and articulate what it's doing. It has to be able to interact with other people."

For the study, researchers programmed [machines](#) with an algorithm called S# and ran them through a variety of two-player games to see how well they would cooperate in certain relationships. The team tested machine-machine, [human](#)-machine and human-human interactions. In most instances, machines programmed with S# outperformed humans in finding compromises that benefit both parties.

"Two humans, if they were honest with each other and loyal, would have done as well as two machines," Crandall said. "As it is, about half of the humans lied at some point. So essentially, this particular algorithm is learning that moral characteristics are good. It's programmed to not lie, and it also learns to maintain cooperation once it emerges."

Researchers further fortified the machines' ability to cooperate by programming them with a range of "cheap talk" phrases. In tests, if human participants cooperated with the machine, the machine might respond with a "Sweet. We are getting rich!" or "I accept your last proposal." If the participants tried to betray the machine or back out of a deal with them, they might be met with a trash-talking "Curse you!", "You will pay for that!" or even an "In your face!"

Regardless of the game or pairing, cheap talk doubled the amount of cooperation. And when machines used cheap talk, their human counterparts were often unable to tell whether they were playing a human or machine.

The research findings, Crandall hopes, could have long-term implications for human relationships.

"In society, relationships break down all the time," he said. "People that were friends for years all of a sudden become enemies. Because the machine is often actually better at reaching these compromises than we are, it can potentially teach us how to do this better."

More information: Jacob W. Crandall et al, Cooperating with machines, *Nature Communications* (2018). [DOI: 10.1038/s41467-017-02597-8](https://doi.org/10.1038/s41467-017-02597-8)

Provided by Brigham Young University

Citation: Let's make a deal: Could AI compromise better than humans? (2018, January 19) retrieved 19 April 2024 from <https://techxplore.com/news/2018-01-ai-compromise-humans.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.