

How cooperative behaviour could make artificial intelligence more human

25 August 2016, by Roger Whitaker



Some innate human features can only be learned through evolution. Credit: Shutterstock/Christian Lagerek

Cooperation is one of the hallmarks of being human. We are extremely social compared to other species. On a regular basis, we all enter into helping others in small but important ways, whether it be letting someone out in traffic or giving a tip for good service.

We do this without any guarantee of payback. Donations are made at a small personal cost but with a bigger benefit to the recipient. This form of cooperation, or donation to others, is called indirect reciprocity and helps human society to thrive.

Group-based behaviour in humans originally evolved to overcome the threat of larger predators. This has led to us having a sophisticated brain with social abilities, which is [disproportionately larger in size](#) than those of other species. The [social brain hypothesis](#) captures this idea: it proposes that the large human brain is a consequence of humans evolving in complex social groups where cooperation is a distinctive component.

Indirect reciprocity is important because we see

donations happening in society despite the threat of "free riders". These are participants who readily receive but don't donate. This idea presents a complex interdisciplinary puzzle: what are the conditions in nature that promote donation over free-riding?

Economists, biologists, mathematicians, sociologists, psychologists and others have all contributed to examining donation behaviour. Investigation is challenging, however, because it involves observing evolution, but computer science can make an important contribution.

Using software, we can simulate simplified groups of humans in which individuals choose to help each other with different donation strategies. This allows us to study the evolution of donation behaviour by creating subsequent generations of the simplified group. Evolution can be observed by allowing the more successful donation strategies to have a greater chance of existing in the next generation of the group.

In modern times, cooperation is becoming increasingly important for engineering and technology. Many intelligent and autonomous devices, like driverless cars, drones and smartphones, are emerging and as these "robots" become more sophisticated we will need to address cooperative decision making for when they come into contact with other devices or humans.



Teaching social cues to robots could better integrate them into human society. Credit: Shutterstock/Michael Bednarek

How should these devices choose to help each other? How can exploitation by free-riders be prevented? By crossing the boundaries of traditional academic disciplines, our findings can provide helpful new insights for emerging technologies. This can allow the development of intelligence which can help autonomous technology decide how generous to be in any given situation.

Modelling evolution

To understand how [cooperation may evolve in social groups](#), we ran hundreds of thousands of computer-simulated "donation games" between randomly paired virtual players. The first player in each pair made a decision on whether or not to donate to the other player. This was based on how they judged their reputation. If the player chose to donate, they incurred a cost and the receiver gained a benefit. Each player's reputation was then updated in light of their action, and another game was initiated. This allowed us to observe which social comparison decisions yield a better payoff.

Social comparison is another key feature of human behaviour that we sought to include. From evolving in groups, we have become adept at comparing ourselves with others and this is highly relevant for making informed donation decisions. This is a considerable cognitive challenge when social

groups are large, so sizing up others in this way could have helped to promote the evolution of larger human brains.

The particular donation behaviour we used in our research was based on players making self comparisons of reputation. This leads to a small number of possible outcomes, for example, relative to myself, your reputation could be considered either broadly similar, higher, or lower. The major element of thinking comes from estimating someone's reputation in a meaningful way.

Our results showed that evolution favours the strategy to donate to those who are at least as reputable as oneself. We call this "aspirational homophily". This involves two main elements: first, being generous maintains a high reputation; second, not donating to lower reputation players helps to prevent free-riders.

It is important to remember that our results are from a simplified model: the donation decisions involved no exceptions that may occur in real life, and economic resources are assumed to drive behaviour rather than emotional or cultural factors. Nevertheless, such simplification allows us to gain useful clarity.

Most importantly, the social brain hypothesis is supported by our findings: the large human brain is a consequence of humans evolving in complex [social groups](#) where cooperation is a distinctive component. Understanding this through computing opens up a new line of thought for the development of sophisticated social intelligence for autonomous systems.

This article was originally published on [The Conversation](#). Read the [original article](#).

Source: The Conversation

APA citation: How cooperative behaviour could make artificial intelligence more human (2016, August 25) retrieved 16 January 2021 from <https://phys.org/news/2016-08-cooperative-behaviour-artificial-intelligence-human.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.