Report on the evolution of cooperation in relation to game-theory and different game-theoretic mechanisms to aid it

JAMES KING

Supervisor: Kostas Stathis

October 2018

Abstract

Since the early days of Darwinian evolutionary theory the phrase "Survival of the fittest" has become synonymous with much thinking on evolutionary dynamics. This idea has come along way since then but is still seen by many to promote selfish attitudes. However, we can see throughout nature that cooperation between biological agents is prevalent. So how did cooperation evolve and why has it flourished? This question fundamentally challenges what seemed concrete ideas about evolutionary dynamics, and has been approached from many angles. Game-theory is a branch of mathematics that has spawned many mathematical models of evolutionary dynamics in an attempt to solve the problem, some have been formulated programmatically to analyse the results. Study of the evolution of cooperation also has wider impacts in the world of computer science - namely on agent-based systems. A large component of agent-based system design is how interactions between agents works and how cooperation can be garnered within societies of agents. In this report, I shall explore past game-theoretic approaches to this problem. This exploration will help me gather a deeper understanding of evolutionary dynamics and the reasoning behind these mathematical approaches to the problem of the evolution of cooperation.

I. INTRODUCTION

In the seminal paper on the evolution of cooperation [1] Axelrod and Hamilton highlighted the failure of past ideas on Darwinian evolutionary theory to account for altruistic behaviour. Their paper identifies two theories proposed to solve the problem: kinship theory and reciprocation theory, focusing on the latter - particularly the Iterated Prisoner's Dilemma. Altruistic acts are actions where an individual puts others above themselves, meaning that they act in a way that benefits others and is detrimental in some way to themselves. The proposed theories attempt to explain how altruistic behaviours have evolved in our world.

Here I will explore these mechanisms that attempt to explain altruistic behaviour and attempt to relate them to multi-agent systems or even point out how they are inapplicable to this domain if appropriate.

II. CONTENT AND KNOWLEDGE

i. Kinship Theory

Kinship theory is an umbrella term for a number of theories that attempt tp solve the problem of the evolution of cooperation. The tie between these theories is that the individuals who choose to cooperate with each other are in some way 'kin'. The definition of what 'kin'
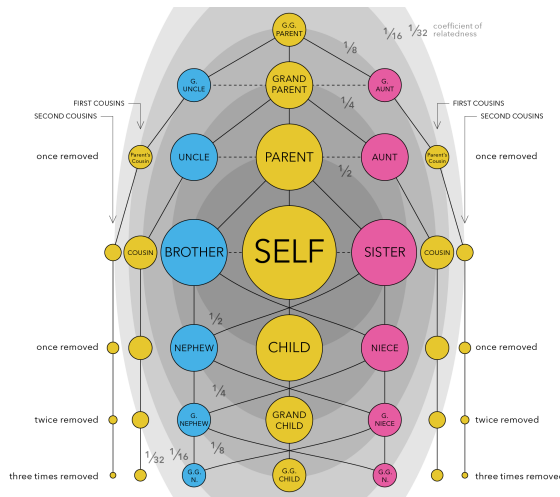
Figure 1: *By Citynoise - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=37723128*

is varies depending on the theory and the purpose for that theory.

Richard Dawkins popularized the idea of the 'selfish gene' [2]. This 'selfish gene' idea argues that as genes are the actual replicators, actors are hardwired to have a will to propagate the gene. This propagation does not only involve reproduction but also acts of cooperation to support and maintain others who share the gene. This is evident in many areas of nature especially in the family groups of different species, shown even recently in the BBC series Dynasties.

W.D. Hamilton supports a similar view [3]: that agents with shared genes will work together to preserve those genes. He argues that to do this individuals don't work to add to their own fitness, but work to improve their 'inclusive fitness', which includes the fitness of other related individuals. His model converts the extent of the relatedness of an individual to a quantitative value using Wright's Coefficient of Relationship, which is shown in figure 1 on page 2. The model then uses this coefficient of relatedness to calculate inclusive fitness.

Axelrod and Hamilton [1] highlight that although the theory works well to explain cooperation between related individual it falls short

in explaining the evolution of cooperation between unrelated individuals. This also makes it difficult to apply kinship theory as an aid to the evolution of cooperation in multi-agent systems. Kinship theory requires a reproductive drive and metrics for relatedness in order to motivate altruistic acts, both of which would likely be missing in a decentralised multi-agent system.

ii. The Iterated Prisoner's Dilemma

The Iterated Prisoner's Dilemma is one such example of a game-theoretic model that uses direct reciprocity to attempt to solve the problem that the evolution of cooperation puts forward. Direct reciprocity is the idea: "If I cooperate now, you may cooperate later" [7]. The Dilemma is in fact a game played between two individuals. The game is made up of a number of repeated rounds, in each round the two players can choose between cooperating with or defecting against each other.

A payoff matrix is provided such as the one in figure 2 on page 3. These matrices provide a temptation to defect, but a reward if both players cooperate over multiple rounds. In a single round game it is mathematically best to defect, but in repeated rounds agents are encouraged to cooperate by a mutual gain in payoff. These payoffs attempt to simulate resource gains from interactions in real life.

The dilemma has been extended into tournaments such as the one in Axelrod and Hamilton [1]. You can even run tournaments directly in Python using the Axelrod-Python library [**?**]. Tournaments can come in a variety of styles but the most popular is the round-robin tournament where every player plays a match of the iterated prisoner's dilemma against every other player. Players accumulate points throughout these games, based on the payoff matrix.

This has been even further extended to include a genetic algorithm to simulate evolution by reproducing players into a next generation. Multiple different algorithms have been used, one common one is the Moran Process (A stochastic process used to model evolution in a fi-

Figure 2: *The payoff matrix given in Axelrod and Hamilton's paper [1]*

nite, unstructured population) available in the Axelrod-Python library.

iii. Strategies

iv. Other Reciprocation Theories

[11]
[7]
[1]
[10]
[15]
[16]
[12]
[8]
[9]
[13]
[14]

III. Discussion and Conclusion

Talk about why I am choosing indirect reciprocity.
Talk about impact in real life and multi-agent systems.
[6]
[5]

References

[1] Robert Axelrod and William D. Hamilton. The evolution of cooperation. *Science*, 211:1390–1396, 1981.

[2] Richard Dawkins. *The Selfish Gene*. Oxford University Press, 2016.

[3] W. D. Hamilton. The genetical evolution of social behaviour. *Journal of Theoretical Biology*, 7:1–16, July 1964.

[4] Doug Jones. Kin selection and ethnic group selection. *Evolution and Human Behavior*, 39:9–18, August 2017.

[5] Menusch Khadjavi and Andreas Lange. Prisoner's and their dilemma. *Journal of Economic Behavior and Organisation*, 92:163–175, August 2013.

[6] M. Milinski. Tit for tat in sticklebacks and the evolution of cooperation. *Nature*, 325:433–435, January 1987.

[7] Martin A. Nowak. Five rules for the evolution of cooperation. *Science*, 314, 2006.

[8] Martin A. Nowak and Robert M. May. Evolutionary games and spatial chaos. *Nature*, 359:826–829, 1992.

[9] Martin A. Nowak and Karl Sigmund. Tit for tat in heterogenous populations. *Nature*, 355:250–252, January 1992.

[10] Martin A. Nowak and Karl Sigmund. Evolution of indirect reciprocity by image scoring. *Nature*, 393:573–577, 1998.

[11] Martin A. Nowak, Karl Sigmund, and Robert M. May. The arithmetics of mutual help. *Scientific American*, 272:76–86, 1995.

[12] Steve Phelps. An empirical game-theoretic analysis of the dynamics of cooperation in small groups. *Journal of Artificial Societies and Social Simulation*, 2016.

[13] Gilbert Roberts. Evolution of direct and indirect reciprocity. *Proceedings of The Royal Society*, 275:173–179, September 2008.

[14] Alexander J. Stewart and Joshua B. Plotkin. Extortion and cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences of the United States of America*, 109:10134–10135, June 2012.

[15] György Szabó and Gábor Fáth. Evolutionary games on graphs. *Physics Reports*, 446:96–216, 2007.

[16] Arne Traulsen and Martin A. Nowak. Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences of the United States of America*, 103:10952–10955, 2006.