

Evolution and Game Theory: Darwinian Principles as the Ultimate Player

Mehmet Ozan Baykan

ENG 102

Robin Turner

August 20, 2020

1. Introduction

Considering for a moment how behaviour, and accompanying mindware of decision-making and rationality, are rooted in biological phenomena and driven by evolutionary forces, one readily acknowledges that it is dazzling how it unfolds in nature. Moreover, behaviour translates into the macrocosm of population dynamics from the biological microcosm through interactions between organisms - a connection that incited evolutionary scientists to utilize the “theory of human strategic behavior based on an idealized picture of rational decision making” (Hammerstein and Selten 1994), namely game theory. Although game theory was originally conceived in the domain of economics, the essentiality of interactions between agents that are assumed to be rational - or in the case of evolution assumed to abide by Darwinian principles - is a common ground in both economics and evolutionary biology which renders game theory a prudential tool in analysing evolution.

In the context of game theory, games are interactions “in which at least one agent can only act to maximize his utility through anticipating (either consciously, or just implicitly in his behavior) the responses to his actions by one or more other agents” (Ross 2019). Simply put, games describe when players determine their payoff by producing strategies that govern their actions at every decision-driven move given a set of conditions put forth by other players. For example, the famous game of Prisoner’s Dilemma (PD) models a situation where two criminals are caught by the authorities, and after being isolated from one another, given the choice of either to defect by blaming the other or to cooperate by denying the blame on both parties. The payoff matrix for the game is given below, where the payoff values represent time in prison for each player.

		Player B	
		Cooperate	Defect
Player A	Cooperate	A : 1, B : 1	A : 5, B : 0
	Defect	A : 0, B : 5	A : 3, B : 3

Figure 1: Payoff matrix for Prisoner’s Dilemma

As seen above, cooperating with the other player, although leads to the best overall scenario, is risky and has a higher expected value of years in prison. On the other hand, defecting seems to be the safer scenario since it reduces the expected years in prison for each player. As Kuhn puts it, PD models “a conflict between individual and group rationality” (Kuhn 2019). In such a situation, a single choice to defect or cooperate is considered a *move*, whereas a *strategy* is a set of instructions that govern the decisions of the player given a set

of circumstances about the game. For example, starting off with cooperation and mimicking the opponent's previous move in the following iterations, namely *tit-for-tat* (Axelrod 1984), is considered a strategy.

Furthermore, observe that although cooperation is more prudential for the group, given any choice from the other player, defecting always leads to a better payoff. Such a set of moves, *action profiles*, “with the property that no single player can obtain a higher payoff by deviating unilaterally from this profile.” (Sethi 374) are called *Nash equilibria* after the pioneering game theorist John Nash who formulated the concept, which serves as a tool for understanding evolutionary stability.

In evolutionary game theory, games represent interactions between organisms that exhibit game characteristics, and payoffs represent the Darwinian notion of *fitness*, i.e the reproductive success of the organism. One essential motivation for such interdisciplinary study is analysing *frequency-dependent selection* in which the fitness of an individual with a certain trait, e.g behavioural, physiological etc., is determined by the distribution of the variations of the trait among the population, due to the role of other individuals - players - in determining the fitness - payoff - of the individual trait. (Brown 2016) Other notable benefits include the utilization of games for testing underlying assumptions about evolutionary forces, and the evaluation of rationality assumptions in game theory by comparison with strategies in nature that prove to be of greater fitness but are not rational in the conventional framework. However, there also exists arguments that render such positive conclusions doubtful, such as instances where game theoretical predictions of behaviour fail or the effect of bias in modelling evolutionary phenomena.

This paper will try to establish to what extent evolutionary phenomena align with the framework of game theory by considering cooperative and competitive behaviour, and consider philosophical restrictions on the claim that game theory explains evolutionary phenomena.

2.1 Cooperative Behaviour and Prisoner's Dilemma

Cooperative behaviour describes situations where the act provides a benefit to another individual and evolutionarily emerges at least in part due to this benefit (West et al. 2007) and altruistic behaviour is a term for actions by an individual that provide a benefit to another at the expense of its own fitness (Okasha 2020). Although cooperation is widely documented in animal behaviour, “true altruism has never been documented in nonhuman animals, presumably because such a one-way system is not evolutionarily stable.” (Wilkinson 1990), meaning that individuals with true altruistic behaviour cannot perpetuate in further generations. Instead, altruism is manifested in the form of kin selection, altruistic behaviour among relatives, and as reciprocal altruism, which Wilkinson defines on five conditions:

- (1) the behavior must reduce a donor's fitness relative to the selfish alternative,
- (2) the fitness of the recipient must be elevated relative to a nonrecipient,

- (3) performance of the behavior must not depend on receipt of an immediate benefit,
- (4) a mechanism for detecting individuals who receive benefits but never pay altruistic costs has to exist, and
- (5) a large but indefinite number of opportunities to exchange aid must exist within each individual's lifetime. (Wilkinson 1990)

Cooperation among vampire bats, *Desmodus rotundus*, is one such instance of reciprocal altruism. To test this idea, Wilkinson's research uses several setups for experiment to test for different components of reciprocal altruism as defined above. It is known that the bats risk starvation after not feeding for 60 hours, "7 to 30 percent of the bats in a cluster fail to obtain a blood meal on any given night" and that bats sometimes feed others that risk starvation by regurgitating their food (Wilkinson 1990). Observing vampire bats which feed other bats that risk starvation due to inability to obtain food by regurgitating a portion of their own food, Wilkinson's research demonstrates that the bats have the facilities to distinguish between other bats, most likely through social grooming, and usually help others which they have prior association with and which have helped them before. Moreover, their conclusion also supports additional elements of reciprocal altruistic behaviour, namely that (i) bats opt to help another only when the benefit to the receiver is considerably greater than the cost upon itself, (ii) mortality rates and rates of altruistic behaviour are positively related among bat communities. This observation implies that conditions for reciprocal altruism are fully satisfied in kin-based and buddy-based food sharing among vampire bats, as Wilkinson notes (Wilkinson 1990). The strong effect of past associations recorded by Wilkinson points out that this case reciprocal altruism is consistent with the dynamics of repeated Prisoner's Dilemma, which proves a useful tool. Substituting moves in the original table by strategies, the table below models how bats might act when the other risks starvation in a simplified way.

		Bat B	
		Shares food	Does not share
Bat A	Shares food	Reciprocal altruism	A: Loses fitness, B: Loses credibility, Cooperation does not continue
	Does not share	A: Loses credibility, B: Loses fitness, Cooperation does not continue	Starvation risk is not alleviated, fitness is reduced on both sides

Figure 2: Iterated PD representation of reciprocal altruism in vampire bats.

Modifying the outcomes to the case at hand, Wilkinson's conclusion greatly support Axelrod's argument that tit-for-tat provides an optimal equilibrium for all players, which

render the long-term strategy of vampire bats an *evolutionarily stable strategy* (ESS), a term coined by Maynard Smith and Price that describes when individuals that opt to choose a different strategy are unable to survive in the status-quo of the frequency of traits among the community (Maynard Smith and Price 1973). In this case, non-cooperative bats reduce their overall fitness in the long-term, consequently being disadvantaged in a community of cooperative bats. This is reflected in the results of Wilkinson’s research by the fact that their computer model of a non-cooperative community yields approximately a 82% mortality rate, whereas their empirical evidence shows only a 24% mortality rate in actual communities (Wilkinson 1990).

Moreover, the brown capuchin monkey, *Cebus apella*, also showcases cooperative behaviour that underlies reciprocal altruism. Experimenting in a controlled environment where the monkeys from non-interacting communities are gathered and given a task that rewards the monkeys with food only but can only be completed by reciprocal cooperation, Mendres and de Waal observe that capuchin monkeys can conceive the need to and do cooperate under certain circumstances, and suggest that this may imply a basis for understanding the cooperative behaviour of wild capuchin monkeys in hunting. (Mendres and de Waal 2000). Furthermore, de Waal and Davis’s observations that the monkeys’ cooperative behaviour is affected by the perceived benefit, which depends on the expectation that the other monkey will cooperate, suggests that the prevalent strategy among the monkeys highly corresponds to a Nash equilibrium given the other monkey will cooperate, and alternatively to an ESS of cooperation (de Waal and Davis 2003), which can also be modelled by an iterated version of Prisoner’s Dilemma.

2.2 Competitive Behaviour and Hawk-Dove Game

Competitive behaviour is “the direct or indirect interaction of organisms that leads to a change in fitness when the organisms share the same resource” (Lang and Benbow 2013). In game theory, the Hawk-Dove game (HD) is almost the canonical representation of such conflicts, where two sides are competing for a limited source V . The hawk move is characterized by aggression, where the player does not stand down until it acquires the source or is injured to the extent that it cannot continue, and the dove move is characterized by standing down in the face of a hawk and dividing the source equally when faced with another dove. The payoff matrix for the HD game is as below, where C represents the overall cost of fighting between hawks.

		Player B	
		Hawk	Dove
Player A	Hawk	A : $\frac{1}{2}(V-C)$, B : $\frac{1}{2}(V-C)$	A : V , B : 0
	Dove	A : 0, B : V	A : $V/2$, B : $V/2$

Figure 3: Payoff matrix for the Hawk-Dove game

In their seminal 1973 paper, Maynard Smith and Price utilize the HD game in modelling conflicts between animals and characterize five different strategies, from which they compute the evolutionary outcome of iterated HD among the population, thereby successfully reaching an evolutionarily stable strategy, thus laying the groundwork for such models (Maynard Smith and Price 1973).

One of the profound empirical studies on the topic is made on Gouldian finches in explaining the emergence of a stable frequency of color polymorphism, the genetic conditions of which also gives rise to different behavioural tendencies, *strategies*, due to differences in hormonal levels (Kokko et al. 2014). The Gouldian finches are competing for access to tree cavities that are the essential spaces for the finches to breed. The red finches, which exhibit higher levels of aggression due to hormonal profile, display a “hawk” characteristic in the sense that they react aggressively and do not step down when competing for a space to breed, whereas black finches, “doves”, Both characteristics can be seen in both sexes, and the strategy is almost purely genetically determined. Carefully constructing a model based on the mechanism for the HD game, and positing principles of how combinations of male and female finches act towards an ESS, Kokko et al. runs a computer program which models the outcomes of conflicts for reproducing sites and yields how the frequencies of strategies change through time. Their results reveal that hawk-dove characterizations of finches yield a highly consistent composition of frequencies with the population of finches in the wildlife, thereby validating their assumptions on the principles of competition among finches. Since such zero-sum games are not effectively explained in conventional notion of individual fitness where individual and species fitness is positively related, the game framework provides insight into how “one individual's win is another's loss, and better performance (hawk-like behaviour) does not elevate population fitness” (Kokko et al. 2014)

The existence of stabilizing strategies in the case of trade-off between benefits and costs is also predicted by Givnish in cases of “interspecific variation in leaf height among goldenrods (*Solidago*) inhabiting a virgin prairie, and the pattern of variation conform[ing] to the predictions of the [ESS] model” (Givnish 1982), where plant height distribution in the plant populations are explained through the trade-off between access to sunlight, thus greater efficiency in photosynthesis, and the cost of maintaining higher and greater physiological parts. However, he also notes that “the mechanisms by which [the ESS model] operates occur in ecological time” (Givnish 1982), thus expressing an essential difficulty that undermines the certainty of the results that are obtained from such research.

3. Criticism of Evolutionary Game Theory

Until now, we have analysed instances where cooperative and competitive behaviour exhibits characteristics that lends itself to modelling by games, and whereby starting from assumptions about different strategies among the population and the frequency thereof, we could extrapolate the actual behaviour to a considerable extent - inasmuch as the actual data supported the results. However, as mentioned by Kokko et al., “there is a remarkable lack of any empirical illustrations of the game being played out in nature by discrete morphs, with genetic inheritance of the relevant strategies”, which reflects the general problem of pointing

out the the exact molecular counterparts of evolutionary mechanisms as faced by evolutionary scientists. A precursor of this problem in evolutionary game theory is also mentioned by Maynard Smith and Price in their seminal paper titled *The Logic of the Animal Conflict* as noting that “real animal conflicts are vastly more complex than our simulated conflicts” (Maynard Smith and Price 1973). Although the models presented are well-established, they are only a fraction of the evolutionary phenomena in nature, and do only have the power to validate established facts about evolutionary dynamics and population data. Thus it is crucial to ask if the value of the models diminish if they are only validations about preexisting facts and are not proved to hold the power of prediction.

Moreover, the question of whether game theoretic models can provide information about the proximate causes of phenomena they model as opposed to serving merely as models that perpetuate the norms among populations can be raised. Alexander raises this concern about the vacuous state of game theoretic models as claiming that “it seems unlikely that an evolutionary game theoretic model would indicate a unique historical sequence suffices to bring about the phenomenon” and such explanation yet requires empirical evidence as opposed to rigorous modelling. (Alexander 2019)

4. Conclusion

All in all, the presented cases of game theoretical models and the solutions to them as evolutionarily stable strategies proves that game theory has the potential to serve as a powerful and accurate framework for explaining the effects of frequency-dependent selection and the accompanying variations on the notion of fitness. Cooperative behaviour among vampire bats and capuchin monkeys can be modelled by adapting the PD game to a great extent, and the conflicts among Gouldian finches, which are convenient to study due to correspondence between physiological and behavioural traits, also proved to be in agreement with models based on the HD game. However, as much as a narrative of the HD game to model the trade-off in plant height could be generated to a certain extent, it contained much uncertainty due to the complex and long-term nature of evolutionary phenomena. Moreover, we conclude that the realm outside instances of evolution studied under evolutionary game theory is vast, and the approach has not yet been proved as inclusive as the theory of evolution. Possible reasons for the weakening of the approach are suggested as circular nature of models that are based on preexisting evolutionary information, and the lack of insight into the proximate causes of evolutionary outcomes.

Works Cited

Alexander, J. McKenzie. 2002. "Evolutionary Game Theory", *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/sum2019/entries/game-evolutionary/>.

- Brosnan, S.F. 2011. "What Do Capuchin Monkeys Tell Us about Cooperation?." In: Forsyth D.R., Hoyt C.L. (eds) *For the Greater Good of All*. Jepson Studies in Leadership. Palgrave Macmillan, New York. https://doi.org/10.1057/9780230116269_2
- Carter, Gerald G. and Wilkinson, Gerald S. 2013. Food sharing in vampire bats: reciprocal help predicts donations more than relatedness or harassment. *Proceedings of the Royal Society*. B.28020122573 <http://doi.org/10.1098/rspb.2012.2573>
- de Waal, Frans B M, and Jason M Davis. 2003. "Capuchin cognitive ecology: cooperation based on projected returns." *Neuropsychologia*, vol. 41,2: 221-8. doi:10.1016/s0028-3932(02)00152-5
- Kim, Yong Gwan. 1995. "Status Signaling Games in Animal Contests." *Journal of Theoretical Biology* 176 (2): 221–31. doi:10.1006/jtbi.1995.0193.
- Kokko, Hanna and Griffith, Simon C. and Pryke Sarah R. 2014. "The hawk–dove game in a sexually reproducing species explains a colourful polymorphism of an endangered bird" *Proceeding of the Royal Society*. B.28120141794 <http://doi.org/10.1098/rspb.2014.1794>
- Lang, J. M. and Benbow, M. E. 2013. Species Interactions and Competition. *Nature Education Knowledge* 4(4):8. <https://www.nature.com/scitable/knowledge/library/species-interactions-and-competition-102131429/>
- Maynard Smith, J., Price, G. 1973. "The Logic of Animal Conflict" *Nature* 246, 15–18. <https://doi.org/10.1038/246015a0>
- Orr, H. Allen. 2009. "Fitness and Its Role in Evolutionary Genetics." *Nature Reviews Genetics*. doi:10.1038/nrg2603.
- Ross, Don, "Game Theory", *The Stanford Encyclopedia of Philosophy* (Winter 2019 Edition), Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2019/entries/game-theory/>.
- Sethi, Rajiv. 2008. "Nash Equilibrium", *International Encyclopedia of the Social Sciences*, 2nd edition, Vol. 5, (ed. Darity Jr., William A.)
- West, Stuart A., Ashleigh S. Griffin, and Andy Gardner. 2007. "Evolutionary Explanations for Cooperation." *Current Biology*. doi:10.1016/j.cub.2007.06.004.
- Wilkinson, Gerald S. 1990. "Food Sharing in Vampire Bats System Ensures That Food Distribution among the Bats Is Equitable." *Scientific American*, no. February.
- Wilkinson, Gerald S. (1988). "Reciprocal altruism in bats and other mammals". *Ethology & Sociobiology*, 9(2-4), 85–100. [https://doi.org/10.1016/0162-3095\(88\)90015-5](https://doi.org/10.1016/0162-3095(88)90015-5)

