

# “Selfishly altruistic” - Utilizing Agent-Based Modeling to Investigate the Evolution of Cooperation

Luke Ring, Otto Sejrskild Santesson

Faculty of Arts, Aarhus University

Bachelor of Cognitive Science

Social and Cultural Dynamics

Instructor: Riccardo Fusaroli

10 August 2022



## Abstract

Cooperation is a pervasive social behavior not solely confined to the human species but is also present across a magnitude of nonhuman taxa. Cooperative behavior might seem incompatible with Darwin's theory of 'the survival of the fittest' and Dawkins' notion of the 'Selfish Gene', since the altruistic facet of cooperative behavior displays a seemingly selfless and therefore disadvantageous evolutionary strategy. The momentous work done by evolutionary biologists, such as Robert L. Trivers and William D. Hamilton, provides tangible explanations to why cooperation can be a superior strategy compared to exploitative behavior. Such explanations include the concepts of kin selection and reciprocity which both can give an account of why such cooperative social mechanisms could be evolutionary beneficial. To assess the validity of whether these mechanisms are constituents of the emergence of cooperative behavior under natural settings, they must have generative sufficiency. The only way of investigating this is to articulate the mechanics as formal models, since the dynamics between the mechanics can not be realized if not all specifications are defined. Such articulation has already been established with the use of game theory, which simplifies such mechanisms into their most essential parts. Subsequently, the formalized mechanisms must be structured in a way, such that they dynamically simulate the natural conditions in order to provide ecological validity. For this task, traditional equation-based modeling will quickly fall short due to the immense complexity. Fortunately, agent-based modeling is an efficient way of dealing with this obstacle, making it feasible to contrive a dynamic model that is capable of achieving the generative sufficiency needed. By embedding the above mentioned mechanisms into an ABM, this paper extends the work of previous models of cooperation by modeling the mechanism analogous to reciprocity, namely TIT FOR TAT, in addition to other conventionally modeled strategies. In addition, the ABM simulates the mechanism of kin selection in order to assess its impact on the pervasiveness of cooperation in the model. It does so by assigning arbitrary traits to the agents in the ABM and making the strategy employed against other agents. The emergent properties elicited by the model deviates to some extent our expectations, although the patterns still offer explainability.

**Keywords:** ABM, Cooperation, Reciprocity, Kin Selection, Prisoner's Dilemma

# 1. Introduction

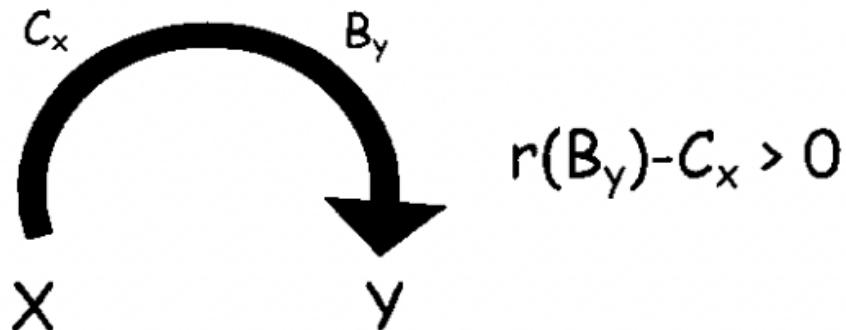
## 1.1 The Evolution of Cooperation

Inclination towards cooperation and altruistic conduct can be arguably ascribed to be one of the most pivotal attributes of human behavior in terms of evolutionary success and predominance in the modern world. It enables us to achieve feats and societal advances that otherwise would have been unachievable. Cooperation, although not idiosyncratic to the human species; is a ubiquitous phenomenon throughout a multitude of nonhuman taxa. Examples to this include ground squirrels producing alarm calls in order to alert its fellows and vampire bats regurgitating sustenance to other members of its own species. The peculiarity of these instances is that they impose a deficit upon the individual's survivability in order to increase the beneficiaries', since the ground squirrel's call attracts the predator's attention and the bat gives up on life-sustaining nourishment (Dawkins, 2006; Stevens, Cushman, et al., 2005). This exhibited that altruistic behavior appears to be discordant with the conventional tenets of evolutionary theory; how can natural selection, an intrinsically selfish process, favor behavior that provides benefits to others while simultaneously causing detrimental effects on the self?

### 1.1.1 Kin Selection

One of the most prominent mechanisms believed to sustain cooperative behavior within species is kin selection, which is the notion of individuals being biased towards cooperation with genetic relatives (Hamilton, 1964; Smith, 1964). Such a disposition can be accounted for by the gene's general incentive to maximize its propagation throughout a population – thus when a vampire bat donates food to a fellow of its species, it is simply acting within the gene's own interest, which is the maximization of its preponderance in the genetic pool. Cooperation stemming from kin selection is therefore a function of the kinship between two members of the species; the more closely related the two members are, the greater the incentive is to behave altruistically towards the other (Axelrod & Hamilton, 1981; Dawkins, 2006). More definitively put; "If the benefits to kin discounted by this coefficient of relatedness [i.e. the proportion of the genes that the individuals share] outweigh the costs of helping, altruism toward kin can evolve." (Stevens, Cushman, et al., 2005, p. 501) (see Figure 1). J.B.S. Haldane reportedly made a remark that nicely captivates

this kinship contingent behavior: “I will jump into the river to save two brothers or eight cousins” (Rand & Nowak, 2013, p. 415).



**Figure 1:**  $X$  can evolve to benefit  $Y$  if  $r(B_y) - C_x > 0$ , where  $r$  is equal to the coefficient of shared genes between  $X$  and  $Y$ ,  $C_x$  is equal to the cost of the act on  $X$  and  $B_y$  is the benefit of the act to  $Y$  (Hamilton, 1964) (Figure reprinted from Sachs et al., 2004)

An implied prerequisite to kin selection is (kin) recognition, which underlying mechanics can vary in a multitude of aspects. In general, it involves an ability to discriminate between different phenotypes which can be of different modalities, including auditory and visual, as well as odors (Sachs et al., 2004). Humans can rely on a multitude of these modalities in the kin discrimination process, although the most significant has shown to be the visual, especially in terms of face recognition (Dal Martello & Maloney, 2006).

In an investigation into the adequacy of kin selection for the emergence of cooperation, Riolo et al. (2001) constructed an agent-based model, where the agents were assigned arbitrary characteristics or ‘tags’ that were observable to other agents. In addition, a bias was introduced into the agents, which specified them to be predisposed to make costly donations (i.e. act altruistically) towards other agents with self-similar tags. After running the model, they could establish that kin contingent altruism was sufficient for the emergence of cooperation. In extension to this, Hammond and Axelrod (2006) determined that without such a preestablished bias towards kins, kin selection could not alone form the basis for the emergence of cooperation – only together with population viscosity (i.e. interactions with other agents and reproduction is local) can kin selection aid the prosperity of cooperation.

Since kin selection only relies on arbitrary characteristics, which can include cultural artifacts, such as fashion or taste in music, it only necessitates very restricted signal-detection abilities (Riolo et al., 2001). Such a superficial set of criteria for the individual's disposition towards cooperation imposes an evolutionary limitation on kin selection, since it is open for exploitation by nonkins that mimics the characteristic and thus receives unreciprocated donation by the duped individual (Alexander & Borgia, 1978; Sachs et al., 2004). Another mechanism that enables the evolutionary emergence of cooperative behavior, and is able to deal with the abovementioned shortcoming of kin selection, is reciprocity.

### **1.1.2 Reciprocity**

While kin selection can account for a multitude of cooperative behavior found in the natural environment, it still lacks explanatory power in another central aspect of human behavior. While the theory provides a clear account of why a parent would make self-sacrificing acts in order to ensure the perpetuation of their child, it does not give an explanation to why a total stranger would risk their life to save a person from drowning. To fill in the spots where Hamilton's theory of kin selection (1964) comes up short, Trivers (1971) formulated a separate model driving the emergence of cooperation in human behavior, namely the mechanism of reciprocal altruism. Reciprocal altruism, or simply reciprocity, is the notion of when "individuals pay a short-term cost of cooperation for the future benefit of a social partner's reciprocated cooperation" (Stevens, Cushman, et al., 2005, p. 501). By expanding upon the drowning person example, if the associated costs (i.e. the risk the rescuer imposes on themselves by saving the drowning person) is lower than the attained gains of the drowning person surviving, then if the drowning person would reciprocate in the future, it will have been the optimal result to each of the individuals to have risked the life for the other. Should the rescuer's efforts not be reciprocated in any way, the rescuer should then not bother saving the drowning person. A conundrum arises from the fact that the rescuer is unable to know whether their altruistic act will be paid back in the future; "[w]hy should the rescued individual bother to reciprocate?" (Trivers, 1971, p. 36). In order for such an impasse to be resolved, interactions between individuals need to occur repeatedly. If this requirement is met, then the rescuer is able to retaliate if the rescued should decide not to reciprocate the altruistic behavior. The rescued would then be worse off than if they had cooperated

and sustained the co-beneficiary relationship. Such a dynamic would then allow cooperation to be evolutionary advantageous, since the individuals would be worse off not cooperating than investing their effort into sustaining the collaboration (Dawkins, 2006; Trivers, 1971). It is important to mention that in order for reciprocity to be an evolutionary stable strategy in ecological settings, it predicated the imperative of certain cognitive capabilities (Stevens, Cushman, et al., 2005; Trivers, 1971). Such capabilities consist of the ability to distinguish between individuals and a memory system able to remember their respective past actions. Without either of these functionalities, it would not be possible to consistently punish exploitative behavior since either the individual would not be able to differentiate between cooperators and exploiters or they would not remember the past interactions with the respective interactants.

Although such cognitive prerequisites might seem tractable, Stevens et al. (2005) highlight, that even though numerous nonhuman species are able to differentiate between individuals, the fact that individuals have to balance the benefits and costs over potentially long periods of time with a multitude of partners, it demands the existence of a nontrivial memory system. Such a burdensome cognitive cost offers an explanation to why reciprocal altruism is mostly exclusive to human behavior (Milinski & Wedekind, 1998). Even though an assemblage of literature argues for the existence of reciprocal altruism in nonhuman animals (e.g. Barrett et al., 2000; Dugatkin & Alfieri, 1991; Krams & Krama, 2002; D. W. Stephens et al., 1997), Stevens et al. (2005) dispute these instances to not be representative, since they have either not been possible to replicate or can be discounted as other types of cooperative behavior, such as kin selection. Reciprocity have although been found to be the case in the behavior of some species, such as the Norway rats (Rutte & Taborsky, 2008) and the cotton-top tamarin monkey (Hauser et al., 2003), although memory of a conspecific was diminished when a third rat was introduced amidst interactions (Burman & Mendl, 2000) and the reciprocal interactions between the tamarins occurred within a short time delay - even if animalistic reciprocity is present, it is still not as temporally extensive and accommodating as in humans.

In addition to the associated load with reciprocity, temporal discounting can provide further explanation to why cooperative behavior is elusive in nonhuman animals (Stevens, Cushman, et al., 2005). Temporal discounting is analogous to reciprocity, since they both revolve around the choice between the immediate reward

or investing in the prospect of a long-term pay-off – you can either cheat your partner and relinquish on future benefits or refrain from temptation and cooperate (Green et al., 1995; Rachlin, 2000). This relationship has been validated in a study carried out by Harris and Madden (2002) which showed that participants that were more prone to immediate rewards had a significant tendency to cooperate less. Held up against research showing animals have a much higher rate of temporal discounting compared to humans (Commons et al., 2013; Rachlin et al., 1991; Richards et al., 1997; Stevens, Hallinan, et al., 2005), this tendency might provide further explanation to the lack of pervasiveness of reciprocity in nonhuman taxa.

Having established the premise of altruistic reciprocity, its requirements in order for it to be an evolutionary stable strategy and reasons to why it is not as pervasive in nonhuman as in human behavior, there is still an important aspect of reciprocity, namely the inherent structure of the interactional mechanism – without concretization, the assertion of reciprocity as a substantial mechanism for the emergence of cooperation will lack substance. Such formalization requires payoff matrices, where outcomes for both individuals for all possible actions of the interaction needs to be specified along with the assessment of which possible strategies that are the superior in terms of maximizing payoff, or in an evolutionary setting, optimizing the individual's fitness. Luckily, such an assessment has already been made possible by the formal model of the Iterated Prisoner's Dilemma.

## 1.2 Iterated Prisoner's Dilemma

The Iterated Prisoner's (IPD) is exactly analogous to “[t]he relationship between two individuals repeatedly exposed to symmetrical reciprocal situations” (Luce & Raiffa, 1989; Rapoport et al., 1965; Trivers, 1971, p. 38). It is therefore a formalization of the scenario of reciprocity, like the drowning person scenario described in the previous paragraph - both individuals stand to gain long term benefits, if they both choose to cooperate, although the temptation of the short term gains of defecting is always present. Prisoner's Dilemma (PD) is a forerunner to the IPD, where the interaction is a stand-alone event – it was devised by Melvin Dresher and Merrill Flood from the think tank RAND Corporation back in 1950 (Chiong et al., 2006). PD, unlike IPD, describes the scenario of a one time interaction between two individuals. The name was originally coined by Albert W. Tucker, who formalized the

game into a scenario, where the payoff was the length of their prison sentence (Poundstone, 1993, p. 8):

*"Two members of a criminal gang are arrested and imprisoned. Each prisoner [player A and player B] is in solitary confinement with no means of speaking to or exchanging messages with the other. The police admit they don't have enough evidence to convict the pair on the principal charge. They plan to sentence both to a year in prison on a lesser charge. Simultaneously, the police offer each prisoner a Faustian bargain."*

The descriptive specifications of the game dynamics is accompanied by a payoff matrix specifying the four different possible outcomes of the interaction (see Figure 2). These four possible outcomes are the following:

- 1) If both A and B decides to tell on each other, each of them will be sentenced to two years in prison
- 2) If A tells on B but B remains silent, A will be liberated and B will serve three years in prison
- 3) If A remains silent but B tells on A, A will be sentenced three years in prison while B will be set free
- 4) If both A and B remain silent, both of them will only be sentenced to one year in prison

		THE PRISONER'S DILEMMA	
		B stays silent (cooperates)	B betrays A (defects)
A stays silent (cooperates)	B stays silent (cooperates)	Both serve 1 year	A serves 3 years, B goes free
	A betrays B (defects)	A goes free, B serves 3 years	Both serve 2 years

**Figure 2:** Payoff matrix of the Prisoner's Dilemma, as formulated by Albert W. Tucker. The matrix specifies the outcomes for player A and B in the four different possible scenarios. The players can either choose to cooperate or defect and are ignorant of the other player's choice. (Figure reprinted from Holodny, 2018)

In order for an interaction to be regarded as a PD type game, it needs to adhere to the same generalized form (see Figure 3) as the example given above (such as that both players are unaware of each other's choices, the payoff matrix is symmetrical etc.). The specific payoff values can be arbitrarily chosen, as long as they abide the following conditions:

- 1)  $T > R > P > S$
- 2)  $2R > T + S$

The first condition specifies, that payoff for defecting when the opponent is cooperating (T) needs to be greater than the reward for mutual cooperation (R), which needs to be greater than the punishment for mutual defection (P), which in turn needs to be greater than cooperating when the other is defecting (S) (see Figure 3). The second condition requires that the payoff for mutually cooperating twice has to be greater than the sum of getting the highest payoff and the lowest payoff (although there has been some debate whether this condition should actually be constituted as a requirement; see C. Stephens, 1996) (Axelrod & Hamilton, 1981; Dawkins, 2006).

		Player B	
		C Cooperation	D Defection
		R=3 Reward for mutual cooperation	S=0 Sucker's payoff
Player A	C Cooperation	T=5 Temptation to defect	P=1 Punishment for mutual defection
	D Defection		

**Figure 3:** The generalized form of Prisoner's dilemma. The payoffs fields correspond to the outcomes for player A in each possible scenario. The variables used in the figure abides the two conditions for a PD game,  $T > R > P > S$  &  $2R > T + S$ , but are positive values in contrast to the classical payoff matrix of the PD (often referred to as a 'donation game'; see Hilbe et al., 2013)

(Figure reprinted from Axelrod & Hamilton, 1981)

The peculiarity of the Prisoner's Dilemma, is that it sheds light on an intrinsic paradox about the interaction, namely that if the game was played by purely rational agents, the agents would both defect, leaving them worse off than if they had chosen to cooperate. How can that be so? Purely rational players would have no incentive to be "nice" towards the other player, only to minimize the length of their sentence. By analyzing the optimal response for each choice the other player can take, defecting is the best in both cases; either, the opponent chooses to cooperate, which would make defecting the best response, or the opponent chooses to defect, which makes defecting the best option yet again (Axelrod & Hamilton, 1981; Dawkins, 2006; Trivers, 1971). We are therefore left in a stalemate. Since interactions between individuals in an ecological setting often adhere to a PD structure, where the payoffs correspond to changes in fitness, how should it ever be possible for cooperation to become an evolutionary stable strategy? To break free from this impasse, we need to add an additional mechanism to the interaction, namely to make it recursive; such an addition is the simple distinction between PD and IPD.

Since the IPD offers an adequate formalization of reciprocity, it facilitates an investigation into assessing whether reciprocity or cooperative behavior truly is an evolutionary viable strategy. Such assessment has already been carried out by the American political scientist Robert Axelrod, who conducted a computer simulation tournament for the Iterated Prisoner's Dilemma to see what type of strategy could thrive in a multifarious environment of strategies of different levels of complexity. The strategies submitted were contrived by game theorists in economics, sociology, and mathematics, amounting to a total of 15 strategies (14 entries and a true random strategy, which was included as a type of baseline strategy). Each strategy was paired up in turn to play IPD with every other one, including a copy of the strategy itself, totalling to 225 games. The strategy that in the end had the greatest average score was the strategy called 'TIT FOR TAT' (TFT) submitted by the game theorist and psychologist professor Anatol Rapoport. The surprisingly simple strategy operated by starting the game with cooperating, whereafter it copied whatever move the opponent did in the preceding turn (Axelrod, 1980a). An additional tournament was carried out, this time with a total of 62 strategies, including the TFT strategy. TFT yet again asserted itself as the most successful strategy (Axelrod, 1980b). Thus, the strategy that performed the best in the simulated environment of the tournament was the one that follows a reciprocally altruistic behavior; it starts out by cooperating

and continues to do so, as long as the other individual does as well. The robustness of the strategy was further established by Axelrod, who applied John Holland's genetic algorithm technique (1992) to see whether a TFT-like strategy would evolve and dominate in an ecological computer simulation, which it did (Axelrod, 1997). In addition, it was shown that TFT was capable of invading an environment and establish itself as the dominant strategy, if the probability of interaction with the same player exceeded a critical level (otherwise, an always-defecting would be able to exploit TFT's first move, which was always cooperative, see Figure 4) (Axelrod, 1984; Axelrod & Hamilton, 1981; Stevens, Cushman, et al., 2005).

For TFT to be ESS, $w \geq \frac{T-R}{T-P}$ and $w \geq \frac{T-R}{R-S}$	<b>TFT</b>	<b>All-D</b>
	<b>TFT</b> $\frac{R}{1-w}$	$S + \frac{wP}{1-w}$
	<b>All-D</b> $T + \frac{wP}{1-w}$	$\frac{P}{1-w}$

**Figure 4:** In order for TFT to be an evolutionary stable strategy (ESS), the probability of interaction again ( $w$ ) has to be greater or equal to  $\frac{T-R}{T-P}$  &  $\frac{T-R}{R-S}$  (Figure reprinted from Stevens, Cushman, et al., 2005)

One shortcoming with the otherwise superior TFT strategy in an IPD game, is that it is quite vulnerable to noise. Such noise is important to incorporate into a formal model that investigates the robustness of reciprocity as an evolutionary stable strategy, since miscommunication and mistakes can arise in the natural environment. Noise can be introduced into the IPD by making the move the players make each turn have some probability of being misimplemented (e.g. defecting instead of cooperating, thus deviating from one's strategy) – should two players with TFT strategies play against each other and one of them would erroneously end up defecting, “the echo of this mistake can go on indefinitely.” (Axelrod, 1997, p. 30). Furthermore, should another mistake occur it could lead to the two players being stuck in constant defection, ultimately making the players oscillate between all four possible outcomes, having no effective way of reestablishing a sustained cooperative relationship (Downs et al., 1985). To cope with such an inadequacy, Axelrod came up

with two iterated versions of TFT, namely the generous TFT (GTFT) and the contrite (CTFT), which work by allowing some possibility of not retaliating the other player's defection and by not reciprocating when the opponent retaliates to the defection caused by the player's erroneous defection, respectively (Axelrod, 1984). The fitness of the strategies were assessed in a simulation with a noise-level set to 1 percent (implemented as above) and an environment identical to the second IPD tournament (Axelrod, 1980b). Both strategies performed adeptly, with only a slight difference in success dependent on the level of noise introduced (Wu & Axelrod, 1995).

Reciprocity thus continues to assert its aptitude as an evolutionary stable strategy by performing proficiently in a simulation mimicking the natural environment, albeit there lies some limitations to how the simulated environment used in the assessment was constructed; for instance, in their simulation, Wu and Axelrod (1995) did not model the agents to be confined in a spatial environment nor imposed a cost of living, leading to privation of ecological validity. Therefore, to provide further substance to reciprocity as an effective behavior in a natural setting, it would be efficacious to test it in a more ecologically adhering simulation. Fortunately such undertakings have already been carried out with use of agent-based modeling, which is a fruitful tool for this endeavor.

### 1.3 Agent-Based Models

One of the main attributes to agent-based modeling is emergent behavior - by implementing the mechanisms underlying a dynamic system, it is possible to "determine the resulting microscopic and macroscopic effects of concatenations of mechanisms" (Salgado et al., 2014, p. 328). The mechanisms of the ABM denote the model micospecifications, which are simple behavioral rules that pose as a script for how the agents should behave and react to the environment they are situated in (Epstein, 2006; Epstein & Axtell, 1996). When employed while investigating social behavior, these mechanics are formalized representations of the underlying social interaction mechanics that are believed to be the constituents of the formation of a social behavior of interest (Hedstrom, 2005). One imperative which is important not to disregard, is that since the simulation is supposed to imitate evolutionary dynamics, the mechanics embedded in the ABM should also represent environmental mechanisms (e.g. cost of living) and selection mechanisms (e.g. which factors are the elicited behavior contingent on) (Ferber & Weiss, 1999; Salgado et al., 2014a).

After the specifications of these mechanisms and their associated parameters have been realized in the ABM, the emergent properties of the population will be generated after the initialization of the simulation until its termination, which then can form the grounds of analysis in order to determine whether these mechanisms are sufficient in order to actualize the macro phenomenon of interest (e.g. the pervasiveness of cooperation). Thus the maxim of ABM modeling can be put as the following "if you did not grow it, you did not explain it" (Salgado et al., 2014b, p. 330). It is therefore a necessity for an ABM's mechanisms to have the generative sufficiency in order to substantiate its explanatory power. Should the anticipated emergent properties not be realized upon the termination of the simulation, then the constituting mechanics would not be a candidate explanation behind the origin of the investigated social behavior. On the contrary, should the anticipated properties emerge after running the ABM, then the computational demonstration has provided a causal link between the given representative mechanisms and the associated investigated social behavior (Salgado et al., 2014b). Furthermore, to provide sufficient substance to the explanatory power of the ABM, it is vital that the details of the mechanics (e.g. the interaction rules and parameter values of the environment) are based on acknowledged scientific work and well-tested theories. Another asset of agent-based modeling, is that it offers a viable method to model complex systems with a lot of variables, thus facilitating such investigation of the complex dynamics of social interaction that would not have been feasible with traditional formal modeling (Helbing, 2012).

In regards to investigating the emergence of cooperation, game theory has often been applied to model the interaction between two individuals that mimicks natural conditions, forming the basic mechanism for the emergence of reciprocity. An example to this is the work done by Hammond and Axelrod (2006); they contrived an ABM that investigated the propagation of cooperation in an environment, where the agents' inclination towards cooperating with the other agent was contingent on an arbitrary trait randomly assigned to the agent population at the beginning of the simulation. They observed that by introducing arbitrary characteristics, such as color, and making the agent's behavior contingent on that characteristic could support the preponderance of cooperative behavior in a simulation mimicking natural conditions. Since these traits were transferred from the parent agent to its offspring, the mechanism mimicked the natural behavior of

kin selection, where individuals are biased towards cooperation with other kindred, whose affiliation is assessed on arbitrary characteristics. In another ABM that also simulated the emergence of cooperative behavior, Smaldino et al. (2013) showed that increased environmental harshness (up to a certain tipping point) promoted the sustainment of pure cooperators in an environment with pure defectors.

ABM therefore establishes itself as a highly applicable tool to investigate the emergence of cooperation in an ecological environment and its many different facets.

## 1.4 Present Study (LR & OS)

In continuation of the abovementioned work, building upon the theory provided in the foregoing paragraphs, we will contrive an ABM with the goal in mind to investigate whether reciprocity is an evolutionary stable strategy as well as if kin selection strengthens the pervasiveness of cooperative behavior. By employing similar mechanisms to the ABMs done by Hammond & Axelrod (2006) and Smaldino et al. (2013), we will see how these different behaviors influence the emergence of cooperative behavior under different levels of environmental harshness (i.e. cost of living and cost of movement). As a novelty, we will introduce the TIT FOR TAT (TFT) as one of the strategies used by the agents, along with ALL COOPERATION (always chooses to cooperate; ALLC), ALL DEFLECTION (always chooses to defect; ALLD) and RANDOM (randomly chooses to either cooperate or defect, RAND). The RAND strategy is introduced into the environment as a form of ‘baseline’ strategy (Axelrod, 1997). In the first model, agents will only have one strategy assigned, which it will follow when interacting with other neighboring agents. In the second model, we will model for kin selection by randomly assigning two strategies to each agent along with one of four possible traits (i.e. color). The strategy that the agent chooses to employ in an interaction with another agent is contingent on the other agent’s trait – if the trait of the opponent is identical to the agent’s own, then it plays the first strategy, otherwise it will employ the other strategy. The two models will be run under different parameter settings of the environmental harshness, as to see how this affects the pervasiveness of the different strategies. Additionally, the agents will be set in a viscous environment, meaning that they will only interact with other agents in their immediate periphery.

Based on the results in the two aforementioned papers, we anticipate to see the strategies of ALLC and TFT to perform better than ALLD and RAND when

trait-contingent strategies are introduced into the model. Furthermore, we expect to observe ALLC and TFT perform better than ALLD and RAND when the environmental harshness increases from benign to harsher, up until a certain turning point, whereafter their dominance will start to diminish (the graph depicting the population of ALLC and TFT should over environmental harshness should have the form of an inverted-U curve; see Smaldino et al., 2013, p. 456). Finally, we expect TFT to generally perform better compared to the other strategies due to its adaptive behavior, based on the literature describing its prevalent success in other simulations and that it is analogous to reciprocity, which is theoretically supported to be an effective social behavior in terms of evolutionary fitness (Axelrod, 1980a, 1980b; Dawkins, 2006; Trivers, 1971).

## 2. Methods

### 2.1 Model

Our model uses stateful agent-vs-agent prisoner's dilemma matches as the core mechanism of agent interaction. Agents compete for the energy resources and spatial positions required for reproduction. Agents must compete in game rounds with all of their neighbors, or move one space. The environment imposes a cost-of-living at the end of each simulation step and if agents move, there is an additional travel cost associated with the movement.

### 2.2 ABM Configuration

All simulations ran until they had reached 10,000 steps or until no agents were left alive. Simulation runs were run in two groups, a “pure” strategy model, in which agent strategies for co-operation or defection were consistent to all opponents, and a “contingent” strategy model, allowing for agents to have one strategy for agents sharing an arbitrary tag, and a second strategy which they employed for agents that had differing tags. Both of the simulation groups were run with increasing environmental harshness, with the cost of living value set to 0, 0.1, 0.3, 0.66, 1.33, 1.66.

## 2.3 Environment Parameters

Agents were positioned on a 512 by 512 2D grid wrapped continuous environment allowing for play, movement or reproduction across the boundaries. The environment carrying capacity was capped to half of the available spaces ( $\frac{2^{18}}{2}$ ) resulting in a hard limit of 131,072 agents. The initial number of agents was set to 16% of available locations, rounded to 41,943 agents. Across simulation runs, the harshness of the environment was varied by increasing or decreasing the cost-of-living which increased the cost of travel which was fixed to 1/2 of the cost-of-living. Agent energy resources were hard-limited to 150, and any payoffs that would go beyond that were discarded. Environmental noise and strategy mutation rate were set to 0.

### 2.3.1 Agents

Agents were spawned in random locations during the setup of the environment, each of which received a truncated normally distributed amount of starting energy with a mean of 50 and a standard deviation of 10, in the range of [5, 150]. Agent actions were limited to a Moore-neighborhood radius of 1, which means surrounded agents will play up to 8 prisoner's dilemma games per simulation step.

#### 2.3.1.1 Tags

All agents were initialized with a single, uniformly distributed arbitrary tag, there were four possible tags, and these stand in for some kind of shared property, in these simulations the tags were immutable, although they only had an effect in the contingent model, with the pure model acting as a reference model.

#### 2.3.1.2 Strategies

During simulation initialization, agents had an equal probability of 0.25 to be randomly assigned one of four game strategies and for simulations in the pure group, agents would use this strategy for any other agents they compete against. For simulations in the contingent group, agents were randomly assigned an additional strategy, which they employed for agents who did not have a matching tag. Agent strategies could be 1) always cooperate, 2) always defect, 3) tit-for-tat and 4) random.

Agents with the tit-for-tat strategy retained a memory of the last game they played with each neighboring opponent, and on subsequent games would choose the neighbor's previous response as their new response, defaulting to cooperate if no previous games had been played with the competing agent. Random strategy would generate a random number in the range of  $(0,1]$  and cooperate if the number was greater than 0.5, otherwise they would defect. During contingent model simulations, this gave a total of 16 unique strategy combinations (e.g. cooperate for agents with the same tag, random for agents with different tags).

### **2.3.1.3 Movement**

When required to travel, agents randomly selected one of the available spaces around them and attempted to claim that space with a random number in the range of  $(0, 1]$ . If another agent tried to claim the space, the agent with the highest random number would claim the land, with ties broken by the highest agent ID. Agents only move when they have no neighboring agents, meaning that movement is always possible, and as such, will attempt to claim the next space in sequence until they have successfully moved. The cost of travel is deducted before making any claim attempts, and if the agent's energy would be less than or equal to 0, the agent dies.

### **2.3.1.4 Prisoner's Dilemma Games**

Each simulation step, agents must play games of Prisoner's Dilemma with all of their neighbors. First, agents broadcast their locations along with a random number in the range of  $(0, 1]$ , indicating they are looking to play. Subsequently, agents will retrieve broadcasted messages from all of their neighboring spaces and compare, the agent with the highest roll (or highest ID in case of a tie) becomes the challenger, and the partner agent becomes the responder. Challengers would send their current energy level, their strategy (or strategies, in the contingent model) and responders would process the game, using the appropriate strategy for themselves and their opponent, with the results and new energy level sent back to the challenger. If either the challenger or responder's respective payoffs would result in their energy reaching or going below zero, they would die. Figure 5 describes the possible outcomes in energy payoff for agents based on their choice and their opponents choice.

	Cooperate	Defect
Cooperate	3, 3	-1, 5
Defect	5, -1	0, 0

**Figure 5:** Payoff matrix of the Prisoner's Dilemma interaction between agents (Figure reprinted from Aktipis, 2015).

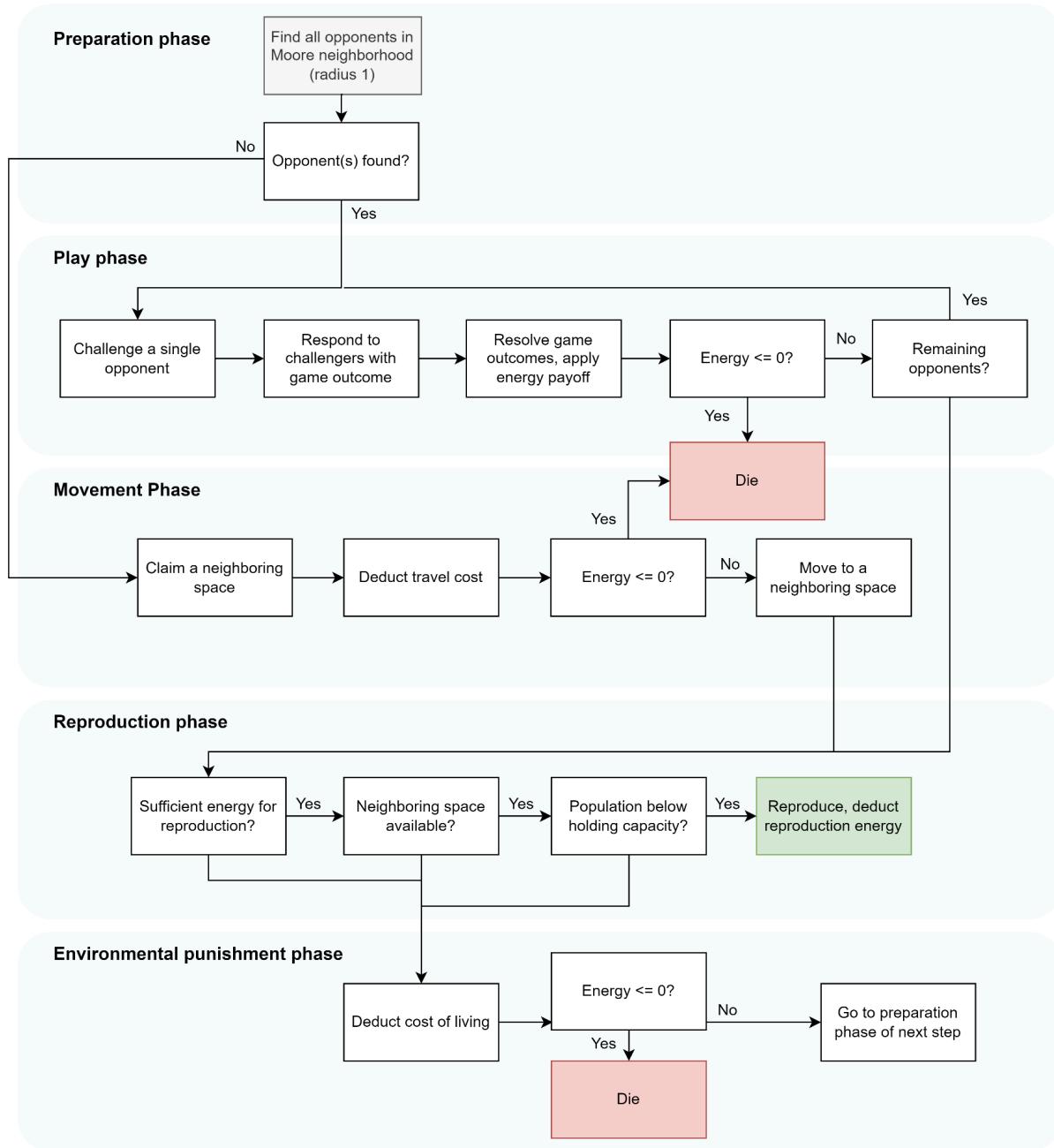
### 2.3.1.5 Reproduction

After resolving their movement or play actions, agents would attempt to reproduce if their energy met the required threshold of 100 and they have an available neighboring space. In the same manner as movements, an agent would claim the space, and if there are any other agents attempting to claim it, the one with the highest random number is successful and ties are resolved by the highest agent ID. Once the agent has succeeded in their claim, they will spawn a new agent. This new agent will inherit the parent agent's trait and strategy, and will start with random energy from the same normal distribution as during initial agent setup. After reproducing, the cost of reproduction, 50, was deducted from their energy.

### 2.3.2 Environmental Punishment

At the end of each step, the cost of living configured for the simulation would be deducted, and if an agent's energy reached 0, the agent would die. During this phase, the simulation would also remove any new agents that exceed the environment's carrying capacity. The model's sequence is described in Figure 6.

## Model Flow



**Figure 6:** Flow diagram describing the typical sequence of events for an agent during model simulation.

## 2.4 Parameters

Simulations were run for 10,000 steps, and multiple runs were performed (minimum 3, maximum 6) for each of the combinations of model and environment. There were two main models, which differed in how agents were assigned strategies. The pure strategy model assigned agents one strategy which they employed unilaterally, the

contingent strategy model used an arbitrary random tag assigned to all agents, represented by an integer in the range of [0,3], and although there were four possible tags, the contingent model assigned two random strategies per agent, regardless of the agent's tag. The first strategy was used for games when their opponent's tag matched their own, and the second strategy was used when their opponents tag differed from their own. Hammond and Axelrod (2006) showed in their work that the existence of these arbitrary tags along with viscous populations (defined as populations whose interactions are localized) can evolve contingent altruism, as it favors regional populations, and may benefit those populations when they encounter agents with differing tags and strategies. This model adds in an additional two strategies (tit-for-tat and random) in order to compare different combinations of strategies.

Environmental harshness was defined by the value of the cost-of-living energy deduction imposed upon agents at the end of each simulation step, as well as the travel cost, which was set to half of the cost of living. As such, during pilot investigation, multiple values were investigated, and harshness values of 3 and above resulted in rapid population collapse. Five final values were selected to investigate the effect on emergent strategies under increasingly harsh environments, these were 0.00, 0.15, 1.00, 1.50 and 2.50 respectively. Thus each run of the suite of simulations comprised 10 individual simulations of 10,000 steps.

## 2.5 Software and Hardware

The model created for the simulation utilized the FLAMEGPU2 ABM framework (Richmond et al., 2021), and the source code for the model is available<sup>1</sup>. This framework allowed leveraging GPU capacity for executing simulations and agent messaging, having a performance advantage over other CPU bound frameworks making it feasible to design simulations with agent populations in the millions. The model was written in python using a custom compiled version of the FLAMEGPU2 library *pyflamegpu*<sup>2</sup>. The simulation was executed on a PC with an Intel Core i5-4690K CPU and a NVIDIA GTX1070 GPU using NVIDIA CUDA 11.7.

---

<sup>1</sup> <https://github.com/zeyus/FLAMEGPU2-Prisoners-Dilemma-ABM>

<sup>2</sup> <https://github.com/FLAMEGPU/FLAMEGPU2/releases>

### 3. Analysis

The model stored population counts for each combination of agent strategies for each simulation step, these data were associated with their environmental configuration and the model that was run. Data were loaded into R for inspection and analysis, and subsequent mean and standard deviation values were calculated for the simulation runs. Table 1 (pure model) and Table 2 (contingent model) list the final population counts from their respective models based on the strategy and environmental harshness.

	Environment Harshness				
	0.00	0.15	1.00	1.50	2.50
Strategy					
<i>Co-op</i>	$1.35 \pm 0.14$	$1.66 \pm 0.21$	$3.64 \pm 0.50$	$4.88 \pm 1.28$	$4.49 \pm 2.72$
<i>Defect</i>	$3.78 \pm 0.06$	$3.17 \pm 0.36$	$1.02 \pm 0.16$	$0.17 \pm 0.07$	$0.02 \pm 0.02$
<i>Tit-for-tat</i>	$1.33 \pm 0.26$	$1.69 \pm 0.44$	$3.18 \pm 0.45$	$4.80 \pm 2.52$	$7.01 \pm 2.74$
<i>Random</i>	$6.66 \pm 0.38$	$6.59 \pm 0.32$	$5.27 \pm 0.40$	$3.26 \pm 1.42$	$1.59 \pm 1.22$

**Table 1:** Simulation results for the pure strategy model. *Mean final populations (mean agent count  $\times 10^{-4} \pm SD \times 10^{-4}$ ) across a minimum of 3 and a maximum of 6 simulations of 10,000 steps. Environment harshness is calculated from the sum of the cost-of-living + travel cost.*

	Environment Harshness				
	0.00	0.15	1.00	1.50	2.50
Strategy					
	Co-op				
<i>Pure</i>	$0.16 \pm 0.03$	$0.14 \pm 0.03$	$0.23 \pm 0.10$	$0.25 \pm 0.14$	$0.44 \pm 0.48$
<i>Contingent (defect)</i>	$3.40 \pm 0.11$	$3.52 \pm 0.46$	$4.30 \pm 0.75$	$4.41 \pm 0.47$	$6.60 \pm 4.62$
<i>Contingent (tit-for-tat)</i>	$0.14 \pm 0.04$	$0.15 \pm 0.03$	$0.24 \pm 0.10$	$0.12 \pm 0.10$	$0.09 \pm 0.15$
<i>Contingent (random)</i>	$0.65 \pm 0.10$	$0.61 \pm 0.07$	$1.03 \pm 0.34$	$1.42 \pm 0.49$	$1.32 \pm 0.50$
	Defect				
<i>Pure</i>	$0.80 \pm 0.13$	$0.72 \pm 0.08$	$0.14 \pm 0.06$	$0.05 \pm 0.04$	$0.01 \pm 0.02$
<i>Contingent (co-op)</i>	$0.16 \pm 0.02$	$0.16 \pm 0.04$	$0.08 \pm 0.04$	$0.04 \pm 0.02$	$0.01 \pm 0.01$
<i>Contingent (tit-for-tat)</i>	$0.14 \pm 0.01$	$0.14 \pm 0.03$	$0.06 \pm 0.03$	$0.04 \pm 0.02$	$0.00 \pm 0.00$

	<i>Contingent (random)</i>	$0.32 \pm 0.04$	$0.3 \pm 0.04$	$0.12 \pm 0.09$	$0.04 \pm 0.02$	$0.00 \pm 0.01$
Random						
<i>Pure</i>		$0.51 \pm 0.08$	$0.52 \pm 0.05$	$0.25 \pm 0.09$	$0.25 \pm 0.13$	$0.12 \pm 0.20$
<i>Contingent (co-op)</i>		$0.13 \pm 0.01$	$0.17 \pm 0.03$	$0.13 \pm 0.06$	$0.16 \pm 0.06$	$0.01 \pm 0.01$
<i>Contingent (defect)</i>		$2.06 \pm 0.58$	$2.05 \pm 0.31$	$0.80 \pm 0.17$	$0.37 \pm 0.21$	$0.01 \pm 0.01$
<i>Contingent (tit-for-tat)</i>		$0.21 \pm 0.05$	$0.18 \pm 0.03$	$0.14 \pm 0.04$	$0.13 \pm 0.06$	$0.02 \pm 0.03$
Tit-for-tat						
<i>Pure</i>		$0.14 \pm 0.04$	$0.16 \pm 0.03$	$0.21 \pm 0.09$	$0.23 \pm 0.10$	$0.52 \pm 0.56$
<i>Contingent (co-op)</i>		$0.14 \pm 0.03$	$0.16 \pm 0.07$	$0.22 \pm 0.06$	$0.28 \pm 0.15$	$0.16 \pm 0.14$
<i>Contingent (defect)</i>		$3.50 \pm 0.54$	$3.58 \pm 0.45$	$4.32 \pm 0.58$	$4.25 \pm 1.00$	$1.22 \pm 1.13$
<i>Contingent (random)</i>		$0.65 \pm 0.16$	$0.55 \pm 0.07$	$0.84 \pm 0.13$	$1.05 \pm 0.44$	$2.58 \pm 2.36$

**Table 2:** Simulation results for the contingent strategy model. *Mean final populations (mean agent count  $\times 10^{-4} \pm SD \times 10^{-4}$ ) across a minimum of 3 and a maximum of 6 simulations of 10,000 steps. Environment harshness is calculated from the sum of the cost-of-living + travel cost.*

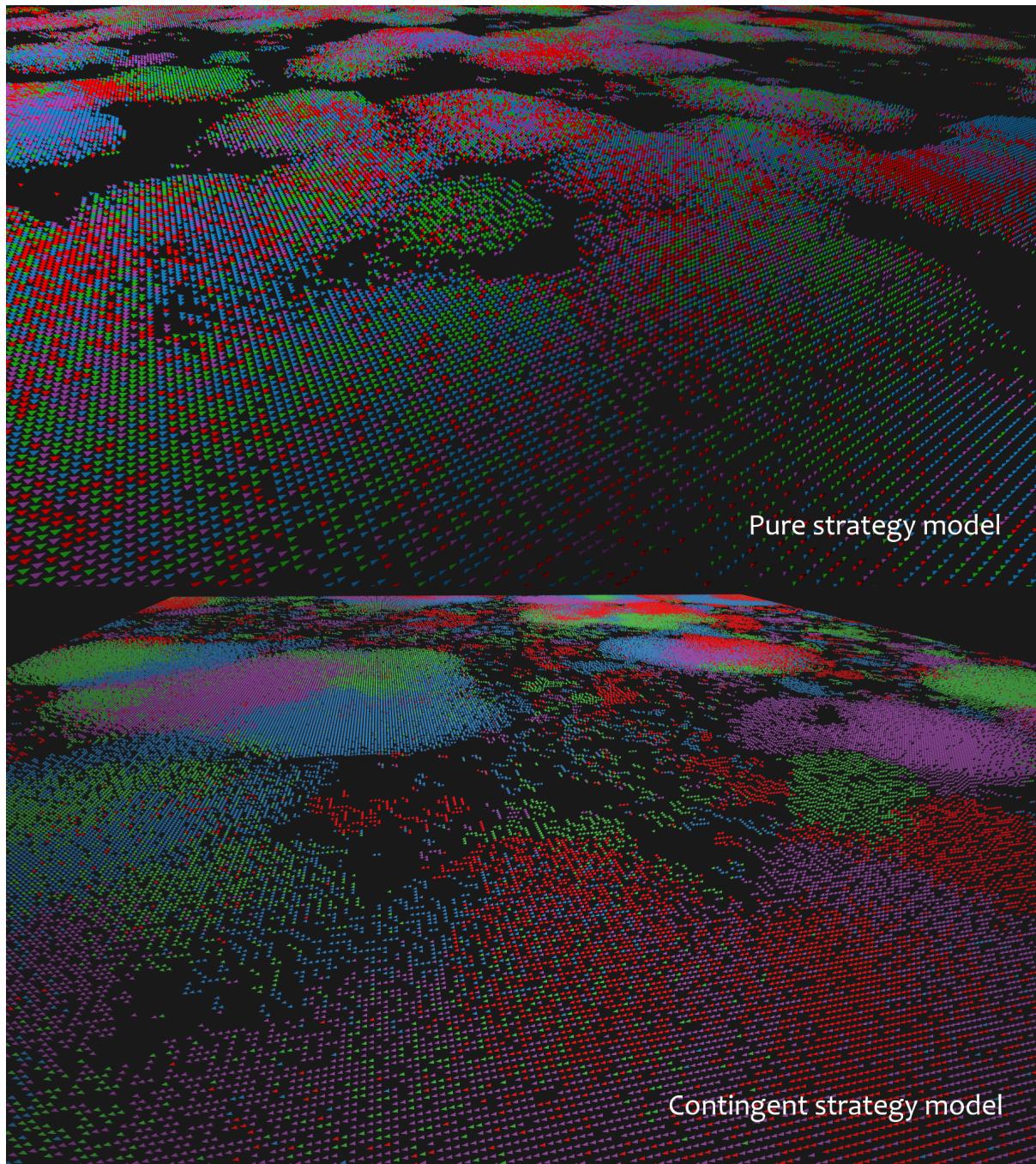
In the pure model, the most forgiving environments favor the random strategy, and as the harshness of the environment increases, cooperative strategies emerge as the most predominant and defectors die out. In the contingent model, the most effective strategies in a forgiving environment - contingent co-op (defect) and contingent tit-for-tat (defect) are both likely to cooperate with other agents that share their tag as the tit-for-tat strategy will start with cooperating, and both populations will defect against any agents who have a different tag to their own. In the harshest environment, the contingent co-operative (defect) strategy had the largest population count (mean:  $6.60 \times 10^{-4}$ , SD:  $4.62 \times 10^{-4}$ ). The contingent tit-for-tat (random) had the second largest (mean:  $2.58 \times 10^{-4}$ , SD:  $2.36 \times 10^{-4}$ ).

## 4. Discussion

The results of the simulation are indicative that contingent strategies that favor kin will emerge and thrive when there is the presence of a mechanism for arbitrary discrimination. For the contingent cooperators that defect when they encounter agents with differing tags, as well as the contingent tit-for-tat, the high level of variance reported in the final population might suggest that it is not always an effective strategy, and much will depend on the initial local conditions of the

environment, such as the tags and strategies of any nearby agents. As the simulation goes on, if an agent has survived, their likelihood of reproducing increases. During the next round, they will have to compete with that new agent, and due to the fact that no mutations were possible in this model, the existing and new agents will share a tag, making it more harmful to both agents and their future population if they were to employ defecting strategies to their own tag. Inversely, there are possible isolated defectors that can find a stable configuration within combinations of neighboring agents. Figure 7 shows a screenshot from the visualization of a simulation run from both models, highlighting the difference in tag clustering, represented by the agent's color. As expected, in both cases there is some clustering by tag, due to the tag still being inherited in the pure model, but overall appears more homogeneous than the contingent strategy model which displays some clear demarcation between clusters of colors.

Contrary to our initial expectations, the tit-for-tat agent strategy did not perform as well as defecting did for competing with non-alike agents, and this is likely due to the effect of interacting with agents which employed defect (or defected randomly) as their strategy to these agents, this gives the tit-for-tat agent an initial loss which would never be recovered in subsequent rounds with agents that defect. This kind of behavior, biased against non-kin mirrors some of the minimal group paradigm experiments on social identity where people will display in- and out-group bias based on arbitrary categorizations (Tajfel et al., 1971). Although there is no case to be made that the agents in this simulation experience any social pressure, it may speak to a genetic evolutionary advantage where naïve individuals or groups may be taken advantage of to their detriment. In our simulation, it is clear that there is some advantage to choosing a strategy of defecting in some situations, but if this is partnered with cooperation with localized populations, it can be a powerful combination.



**Figure 7:** Screenshots from simulation runs for both the pure strategy model (above) and the contingent strategy model (below). Agent colors represent their tag.

## 4.1 Limitations and Future Development

A significant limitation, that is elucidated by the results of the ABM, is the fact that for some of the simulation configurations, the standard deviance of the population mean for multiple strategy types are conspicuously high. This could indicate that the population of the given strategy type were still highly volatile, thus signifying that the simulation still had not reached stability or collapsed. In order to resolve this problem and get more accurate estimates of the population means of the different strategy types, it could be advantageous to increase the run in future simulations (this was not feasible for this current study due to restrictions on time and computational power available).

Several limitations to our model can be pointed out; firstly, as opposed to other ABMs, we do not validate our model with empirical evidence on the subject. Such procedure provides further validation of the model, since data gathered about the phenomenon in natural settings can provide values to calibrate the initial conditions of the model with, define the parameter values of the environment and the behavioral rules governing the agents. Besides the structure of the interactions that the agents undertake (see the paragraphs on reciprocity and IPD), the extent of the validation of our model is very limited. However, other ABMs modeling the emergence of cooperative behavior also lack this aspect (Axelrod, 1984; Axelrod & Hamilton, 1981; Epstein, 2006; Riolo et al., 2001). This can be explained by the nature of the subject of our investigation, namely how cooperative behavior can be an evolutionary advantageous strategy: since it is not feasible to retrieve substantial empirical data on cooperative behavior in prehistoric conditions, we must accept this fact and search out other ways to compensate for this limitation.

Another potential limitation to our model is that the reproduction mechanism specified in the ABM is asexual, meaning that reproduction only necessitated the sufficient energy and a proximate unoccupied space. Since the model is supposed to mimic the natural conditions wherein cooperation between human individuals can emerge, it thus diverges from ecological validity, since human reproduction requires two individuals, whose traits are passed onto the offspring in a more intricate process than the one modeled in this study. However, other work modeling the pervasiveness of cooperation also employs such asexual reproduction, following a conventional paradigm of agent-based modeling (Axelrod, 1984; Axelrod & Hamilton, 1981;

Epstein, 2006; Riolo et al., 2001). Taking into account the magnitude of added complexity a sexual reproduction mechanism would generate, asexual reproduction therefore seemed an adequate simplification. On the note of potential ecological inaccuracies in our model, there are several mechanisms that have been left out, which potentially could have had significant influence on the outcome of the simulation. For example, whereas other models included an environmental set mutation rate of agents' strategies (Hammond & Axelrod, 2006; Smaldino et al., 2013), ours did not incorporate such a mechanism. Since mutation rates represent naturally occurring mutations in inherited genes that can ultimately result in population dominance, omission of such a mechanism could have prevented potentially more successful strategies to invade an otherwise stable population and overtake the dominance in the population distribution, thus further diminishing the ecological validity of the model.

While it is apparent that some mechanisms are more vital than others, it is important to note that one should not attempt to unskeptically include all mechanisms that represent some aspect of the natural environment; doing so would defeat the purpose of modeling, which is to simplify a complex system by formalizing its most essential components while still retaining its general dynamics. ABM modeling is therefore a balancing act, which essence can be described by the following proverb put forth by the French writer and aviator Antoine de Saint-Exupéry: "Perfection is achieved, not when there is nothing more to add, but when there is nothing left to take away". In the pursuit of such perfection it is therefore necessary, through rigorous analysis and testing along with theoretical justification, to identify these vital substrata that are the foundation of the emergence of cooperative behavior.

## 5. References

- Aktipis, C. (2015). Principles of cooperation across systems: From human sharing to multicellularity and cancer. *Evolutionary Applications*, 9.
- <https://doi.org/10.1111/eva.12303>
- Alexander, R. D., & Borgia, G. (1978). Group Selection, Altruism, and the Levels of Organization of Life. *Annual Review of Ecology and Systematics*, 9(1), 449–474. <https://doi.org/10.1146/annurev.es.09.110178.002313>
- Axelrod, R. (1980a). Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution*, 24(1), 3–25. <https://doi.org/10.1177/002200278002400101>
- Axelrod, R. (1980b). More Effective Choice in the Prisoner's Dilemma. *Journal of Conflict Resolution*, 24(3), 379–403.
- <https://doi.org/10.1177/002200278002400301>
- Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books.
- Axelrod, R. (1997). The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration. In *The Complexity of Cooperation*. Princeton University Press. <https://doi.org/10.1515/9781400822300>
- Axelrod, R., & Hamilton, W. D. (1981). The Evolution of Cooperation. *Science*, 211(4489), 1390–1396. <https://doi.org/10.1126/science.7466396>
- Barrett, L., Henzi, S. P., Weingrill, T., Lycett, J. E., & Hill, R. A. (2000). Female baboons do not raise the stakes but they give as good as they get. *Animal Behaviour*, 59(4), 763–770. <https://doi.org/10.1006/anbe.1999.1361>
- Burman, O. H. P., & Mendl, M. (2000). Short-term social memory in the laboratory rat: Its susceptibility to disturbance. *Applied Animal Behaviour Science*, 67(3), 241–254. [https://doi.org/10.1016/S0168-1591\(99\)00120-3](https://doi.org/10.1016/S0168-1591(99)00120-3)
- Chiong, R., Wong, D. M. L., & Jankovic, L. (2006). Agent-based economic modelling

with iterated prisoner's dilemma. *2006 International Conference on Computing & Informatics*, 1–6.

<https://doi.org/10.1109/ICOICI.2006.5276488>

Commons, M. L., Mazur, J. E., Nevin, J. A., & Rachlin, H. (2013). *The Effect of Delay and of Intervening Events on Reinforcement Value: Quantitative Analyses of Behavior, Volume V*. Psychology Press.

Dal Martello, M. F., & Maloney, L. T. (2006). Where are kin recognition signals in the human face? *Journal of Vision*, 6(12), 2. <https://doi.org/10.1167/6.12.2>

Dawkins, R. (2006). *The Selfish Gene* (30th anniversary ed). Oxford University Press.

Downs, G. W., Rocke, D. M., & Siverson, R. M. (1985). Arms Races and Cooperation. *World Politics*, 38(1), 118–146. <https://doi.org/10.2307/2010353>

Dugatkin, L. A., & Alfieri, M. (1991). Guppies and the TIT FOR TAT strategy: Preference based on past interaction. *Behavioral Ecology and Sociobiology*, 28(4), 243–246. <https://doi.org/10.1007/BF00175096>

Epstein, J. M. (2006). *Generative Social Science: Studies in Agent-Based Computational Modeling* (STU-Student edition). Princeton University Press.  
<https://www.jstor.org/stable/j.ctt7rxj1>

Epstein, J. M., & Axtell, R. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings Institution Press.

Ferber, J., & Weiss, G. (1999). *Multi-agent systems: An introduction to distributed artificial intelligence* (Vol. 1). Reading: Addison-wesley.

Green, L., Price, P. C., & Hamburger, M. E. (1995). Prisoner's Dilemma and the Pigeon: Control by Immediate Consequences. *Journal of the Experimental Analysis of Behavior*, 64(1), 1–17. <https://doi.org/10.1901/jeab.1995.64-1>

Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of*

*Theoretical Biology*, 7(1), 17–52.

[https://doi.org/10.1016/0022-5193\(64\)90039-6](https://doi.org/10.1016/0022-5193(64)90039-6)

Hammond, R. A., & Axelrod, R. (2006). Evolution of contingent altruism when cooperation is expensive. *Theoretical Population Biology*, 69(3), 333–338.  
<https://doi.org/10.1016/j.tpb.2005.12.002>

Harris, A. C., & Madden, G. J. (2002). Delay Discounting and Performance on the Prisoner's Dilemma Game. *The Psychological Record*, 52(4), 429–440.  
<https://doi.org/10.1007/BF03395196>

Hauser, M. D., Chen, M. K., Chen, F., & Chuang, E. (2003). Give unto others: Genetically unrelated cotton-top tamarin monkeys preferentially give food to those who altruistically give food back. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 270(1531), 2363–2370.  
<https://doi.org/10.1098/rspb.2003.2509>

Hedstrom, P. (2005). *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge University Press.

Helbing, D. (2012). Agent-Based Modeling. In D. Helbing (Ed.), *Social Self-Organization: Agent-Based Simulations and Experiments to Study Emergent Social Behavior* (pp. 25–70). Springer.  
[https://doi.org/10.1007/978-3-642-24004-1\\_2](https://doi.org/10.1007/978-3-642-24004-1_2)

Hilbe, C., Nowak, M. A., & Sigmund, K. (2013). Evolution of extortion in Iterated Prisoner's Dilemma games. *Proceedings of the National Academy of Sciences*, 110(17), 6913–6918. <https://doi.org/10.1073/pnas.1214834110>

Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. <https://doi.org/10.7551/mitpress/1090.001.0001>

Holodny, E. (2018). *A Professor Explains How 1980s Game Theory Could Explain*

*How Morality Evolved.* ScienceAlert.

<https://www.sciencealert.com/a-professor-explains-how-1980s-game-theory-could-explain-how-morality-evolved>

Krams, I., & Krama, T. (2002). Interspecific reciprocity explains mobbing behaviour of the breeding chaffinches, *Fringilla coelebs*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 269(1507), 2345–2350.

<https://doi.org/10.1098/rspb.2002.2155>

Luce, R. D., & Raiffa, H. (1989). *Games and Decisions: Introduction and Critical Survey*. Courier Corporation.

Milinski, M., & Wedekind, C. (1998). Working memory constrains human cooperation in the Prisoner's Dilemma. *Proceedings of the National Academy of Sciences*, 95(23), 13755–13758. <https://doi.org/10.1073/pnas.95.23.13755>

Poundstone, W. (1993). *Prisoner's dilemma* (1st Anchor Books ed). Anchor Books.

Rachlin, H. (2000). *The Science of Self-Control*. Harvard University Press.

Rachlin, H., Rainieri, A., & Cross, D. (1991). Subjective Probability and Delay. *Journal of the Experimental Analysis of Behavior*, 55(2), 233–244.

<https://doi.org/10.1901/jeab.1991.55-233>

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, 17(8), 413–425. <https://doi.org/10.1016/j.tics.2013.06.003>

Rapoport, A., Chammah, A. M., & Orwant, C. J. (1965). *Prisoner's Dilemma: A Study in Conflict and Cooperation*. University of Michigan Press.

Richards, J. B., Mitchell, S. H., de Wit, H., & Seiden, L. S. (1997). Determination of Discount Functions in Rats with an Adjusting-Amount Procedure. *Journal of the Experimental Analysis of Behavior*, 67(3), 353–366.

<https://doi.org/10.1901/jeab.1997.67-353>

Richmond, P., Chisholm, R., Heywood, P., Leach, M., & Kabiri Chimeh, M. (2021).

*FLAME GPU*. Zenodo. <https://doi.org/10.5281/zenodo.5769677>

Riolo, R. L., Cohen, M. D., & Axelrod, R. (2001). Evolution of cooperation without reciprocity. *Nature*, 414(6862), 441–443. <https://doi.org/10.1038/35106555>

Rutte, C., & Taborsky, M. (2008). The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): Direct vs generalised reciprocity. *Behavioral Ecology and Sociobiology*, 62(4), 499–505.

<https://doi.org/10.1007/s00265-007-0474-3>

Sachs, J. L., Mueller, U. G., Wilcox, T. P., & Bull, J. J. (2004). The Evolution of Cooperation. *The Quarterly Review of Biology*, 79(2), 135–160.

<https://doi.org/10.1086/383541>

Salgado, M., Noguera, J. A., & Miguel, F. J. (2014a). Modelling Cooperation Mechanisms: Some Conceptual Issues. *Journal of Archaeological Method and Theory*, 21(2), 325–342. <https://doi.org/10.1007/s10816-013-9186-3>

Salgado, M., Noguera, J. A., & Miguel, F. J. (2014b). Modelling Cooperation Mechanisms: Some Conceptual Issues. *Journal of Archaeological Method and Theory*, 21(2), 325–342. <https://doi.org/10.1007/s10816-013-9186-3>

Smaldino, P., Schank, J., & McElreath, R. (2013). Increased Costs of Cooperation Help Cooperators in the Long Run. *The American Naturalist*.

<https://doi.org/10.1086/669615>

Smith, J. M. (1964). Group Selection and Kin Selection. *Nature*, 201(4924), 1145–1147. <https://doi.org/10.1038/2011145ao>

Stephens, C. (1996). Modelling Reciprocal Altruism. *The British Journal for the Philosophy of Science*, 47(4), 533–551.

Stephens, D. W., Anderson, J. P., & Benson, K. E. (1997). On the spurious occurrence of Tit for Tat in pairs of predator-approaching fish. *Animal Behaviour*, 53(1), 113–131. <https://doi.org/10.1006/anbe.1996.0283>

Stevens, J. R., Cushman, F. A., & Hauser, M. D. (2005). Evolving the Psychological Mechanisms for Cooperation. *Annual Review of Ecology, Evolution, and Systematics*, 36(1), 499–518.

<https://doi.org/10.1146/annurev.ecolsys.36.113004.083814>

Stevens, J. R., Hallinan, E. V., & Hauser, M. D. (2005). The ecology and evolution of patience in two New World monkeys. *Biology Letters*, 1(2), 223–226.

<https://doi.org/10.1098/rsbl.2004.0285>

Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149–178.

<https://doi.org/10.1002/ejsp.2420010202>

Trivers, R. L. (1971). The Evolution of Reciprocal Altruism. *The Quarterly Review of Biology*, 46(1), 35–57. <https://doi.org/10.1086/406755>

Wu, J., & Axelrod, R. (1995). How to Cope with Noise in the Iterated Prisoner's Dilemma. *Journal of Conflict Resolution*, 39(1), 183–189.

<https://doi.org/10.1177/0022002795039001008>