A Framework for Indirect Reciprocity

John Maloney
malo0052@umn.edu

With *direct reciprocity*, an agent responds in-kind to its direct experiences with another agent. The game theoretic framework of the prisoner's dilemma is often used to investigate the evolution of cooperation through direct reciprocity. In a seminal work on the evolution of cooperation through direct reciprocity, the authors of [1] showed that, under certain conditions, cooperation based on direct reciprocity could become established in a population of unrelated individuals. This work prompted many follow on studies where the results have been debated and refined.

With *indirect reciprocity* an agent responds based on observations of interactions involving agents other than itself. The driving force behind indirect reciprocity is reputation [2]. An agent's reputation contributes to its perception as trustworthy by other agents. Agents learn the reputations of other agents and thereby learn which agents can be trusted to cooperate and which are likely to defect.

In a seminal work on the subject, the authors of [3] introduced a theoretical framework for analyzing the evolution of cooperation through indirect reciprocity. Through contributions made in publications following its original introduction, a refined framework has emerged and become a common foundation for research on indirect reciprocity. In what follows, the major components of that framework are presented.

1 Indirect Reciprocity Game

This section describes the game that is used to evaluate the evolutionary stability of indirect reciprocity.

1.1 Basic Game

The indirect reciprocity game forms the foundation for the framework. In this game, two agents are paired together with one agent taking the role of the donor and the other agent taking the role of the recipient. The donor can choose to donate at a cost of c to itself or abstain from donating. If the donor donates, the recipient receives a benefit $b > c$. Otherwise, both agents receive zero payout. The payouts for the game are described in the following table:

Donor Action	Donor Payout	Recipient Payout
Donate	$-c$	b
Do Not Donate	0	0

Alternatively, the game can be formalized as a prisoner's dilemma game in which each player has the option to donate to the other player. In this case, the game has the following payout matrix for the row player:

	Donate	Do Not Donate
Donate	b-c	-c
Do Not Donate	b	0

This alternative formalization is used in [9], [10] and [11] and does not impact the results of the analysis or simulations.

In order to ensure that agent payouts do not become negative, the cost c is added to the payout for both players involved in an interaction.

1.2 Repeated Game Procedure

The indirect reciprocity game is used in the context of evolutionary game theory to evaluate the fitness of strategies followed by agents in a large population. The fitness of the agents in a generation is determined by playing multiple rounds of the game. The payouts earned by each agent determine the agent's fitness. After a sufficient number of rounds have been played, a new generation is created based on the fitness demonstrated by each agent in the previous generation.

The process used to pair agents during a generation is designed such that the probability that two agents meet more than once is zero or very small and thus the game can be used to investigate indirect reciprocity. Several methods are used to identify pairs of agents and to determine when the generation is complete.

- Random pairing with fixed number of rounds – In this case, for each round, one agent is selected at random to play the donor role and a second agent is randomly selected to play the recipient role. After a fixed number of pairs have interacted, the generation ends.
- Random pairing with random number of rounds – This is similar to the previous case except that the generation doesn't end after a fixed number of pairs have interacted. Instead, after the interaction of each pair is complete, there is a fixed probability w that another pair will be selected for another round. Therefore, after the interaction of each pair, the generation ends with probability $(1 - w)$.
- Round-robin pairing – In this case, each agent is paired once with every other agent. For each pair, one agent is randomly selected to play the role of the donor and the other plays the role of the recipient. The generation ends when all possible pairs of agents have interacted once.
- All agents paired exactly once – In this case, all agents are paired with exactly one other agent. For each pair, one agent is randomly selected to play the role of the donor and the other plays the role of the recipient. The generation ends when all agents have interacted exactly once.

1.3 Population Structure

The agent population can be unstructured (well-mixed) or structured. In the unstructured case, an agent can be paired with any other agent in the population while in the structured case an agent is limited to being paired only with its neighbors. Common organizations for structured populations are lattices and graphs.

The agent population can be optionally divided into groups[1]. When the population is divided into groups, an agent is limited to being paired with agents in its own group. For some purposes, it is necessary to identify which groups are neighbors. In this case, the groups can be structured or unstructured. In an unstructured group organization each group can interact with any other group while a structured group organization limits interactions to neighboring groups.

1.4 Agent Evolution Process

At the end of a generation, a new generation of agents is created. The strategy followed by each agent in the following generation is determined based on the fitness demonstrated by the agents in the previous generation. The evolution process may optionally include the possibility of mutation. In the case of a mutation, the strategy followed by an agent in the following generation is selected uniformly from among all possible strategies.

Let f_i be the fitness of agent i and μ represent the probability that a mutation has occurred. If mutations are not included in the evolution process then $\mu = 0$. Given a population of N agents, where each agent is following one of M strategies, the following two methods are used to generate the population for the next generation:

- Individual-based: An agent in the next generation inherits a strategy directly from an individual in the previous generation. The probability that a child agent inherits its strategy from a particular parent agent is equal to the normalized fitness of that parent agent's fitness across all agents in the population. Mathematically, the probability p_i that a child agent inherits its strategy from agent i is given by the following equation:

$$p_i = (1 - \mu)\frac{f_i}{\sum_{j=1}^{N} f_j} + \frac{\mu}{N}$$

- Strategy-based: An agent in the next generation inherits a strategy based on the average fitness achieved by all agents that followed that strategy in the previous generation. The probability that an agent inherits a particular strategy is equal to the normalized average fitness achieved by all agents within the group that followed that strategy during the previous generation. Let $S^{(k)}$ be the set of agents that followed strategy k in the previous generation. Mathematically, the probability p_k that a child agent inherits a particular strategy k is given by the following equation:

$$p_k = (1 - \mu)\frac{\sum_{i \in S^{(k)}} f_i}{\sum_{j=1}^{N} f_j} + \frac{\mu}{M}$$

When the population is divided into groups, an agent in the next generation can inherit its strategy locally from its own group or globally from the entire population of agents. In this case, the agent inherits its strategy from its local group with probability p and inherits its strategy from the entire population of agents with probability $(1 - p)$.

[1] Sometimes these groups are referred to as islands or tribes.

Let A_g be the set of agents from the previous generation that belong to group g, N_g be the size of group g and $I_g(\cdot)$ be a function, defined as follows, that indicates whether an agent is a member of group g:

$$I_g(i) = \begin{cases} 1, & i \in A_g \\ 0, & i \notin A_g \end{cases}$$

If the individual-based fitness method is used then the probability p_i that a child agent destined to be a member of group g in the next generation inherits its strategy from agent i is given by the following equation:

$$p_i = pI_g(i)\left[(1-\mu)\frac{f_i}{\sum_{j\in A_g} f_j} + \frac{\mu}{N_g}\right] + (1-p)\left[(1-\mu)\frac{f_i}{\sum_{j=1}^{N} f_j} + \frac{\mu}{N}\right]$$

Let $S_g^{(k)}$ be the set of agents from the previous generation that belong to group g and followed strategy k. If the strategy-based fitness method is used then the probability p_k that a child agent destined to be a member of group g in the next generation inherits strategy k is given by the following equation:

$$p_k = p\left[(1-\mu)\frac{\sum_{i\in S_g^{(k)}} f_i}{\sum_{j=1}^{N_g} f_j} + \frac{\mu}{M}\right] + (1-p)\left[(1-\mu)\frac{\sum_{i\in S^{(k)}} f_i}{\sum_{j=1}^{N} f_j} + \frac{\mu}{M}\right]$$

1.5 Group Evolution Process

When the agent population is divided into groups, the groups may have one or more characteristics that can be evolved. For example, in [13] each group possesses an assessment rule that is shared by all members of the group. In this case, the group characteristics can optionally be evolved before starting the next generation. The average payout earned by all agents in a group is used as the group's fitness measure.

For each possible pair of neighboring groups, the pair is selected to participate in the group evolution process with probability p_G. If a pair is selected for evolution, then based on the fitness of the two groups, one group is selected to be evolved. The characteristics of the selected group are modified to become more similar to the characteristics of the other group.

1.6 Examples

The following are examples of game configurations used in the reviewed literature:

# Groups	Group Structure	Pairing Procedure	# Rounds per Generation	Fitness Calculation	Mutations Occur	References
Single	Well-mixed	TBD	Fixed	Individual-based	No	[3]
Single	Well-mixed	TBD	Fixed	Individual-based	Yes	[3]
Single	Well-mixed	TBD	Fixed	?	No	[4]

# Groups	Group Structure	Pairing Procedure	# Rounds per Generation	Fitness Calculation	Mutations Occur	References
Single	Well-mixed	TBD	Random	Individual-based	No	[3][7][8]
Single	Well-mixed	TBD	Random	?	No	[4][10][11]
Multiple	Well-mixed	Round Robin	Fixed	Individual-based	No	[12]
Multiple	Well-mixed	Round Robin	Fixed	Individual-based	Yes	[13]
Multiple	Well-mixed	Random	Fixed	Individual-based	Yes	[8]
Multiple	Well-mixed	Random	Fixed	Strategy-based	No	[6]
Multiple	Well-mixed	Random	Fixed	Strategy-based	Yes	[6]
Multiple	Well-mixed	TBD	Random	Strategy-based	No	[6]

2 Reputation Model

Reputation plays a key role in indirect reciprocity. Agents use reputation to decide how to act when paired with another agent as well as how to assess the moral value of agent actions. The reputation score is an integer value used to measure the reputation of an agent. In most cases, lower and upper bounds are defined for the score. Common choices for bounds are [-5, 5] and [0, 1].

An agent's action is deemed either good or bad by an assessment rule (see section 4). An agent's reputation score is increased by one if its action is deemed to be good and decreased by one otherwise. If changing the agent's score would cause it to move outside the established bounds then its score remains unchanged.

An agent's reputation score is compared to a threshold to determine whether an agent's moral status is good or bad. If the agent's reputation score is greater than or equal to the threshold then the agent's moral status is good otherwise it is bad.

The most common combination is a reputation score whose bounds are [0, 1] and a threshold equal to one. This defines a binary reputation score where the agent's moral status only depends on that last action it has taken. In this case, the reputation score is called the agent's *image* and the two scores are labeled "bad" and "good".

Common reputation scores and thresholds used in the reviewed literature are the following:

Name	Lower	Upper	Threshold	References
Image	0	1	1	[3][4][6][7][8][10][11][12][13]
Image Score	-5	+5	Various	[3][6]

3 Observation Model

An agent must acquire knowledge of other agents' reputations in order to use those reputations to guide its decisions. The following observation models are used to model how agents acquire knowledge of agent reputations.

- Direct observation model: In this model, an agent obtains knowledge of agent reputations by directly observing agent actions. The model is parameterized by two probabilities. The first parameter defines the probability that an agent observes an interaction. The second parameter defines the *perception error* rate – the probability that an agent misperceives the action taken by a donor agent. Since not every agent witnesses every interaction and the agents' observations are subject to a perception error, each agent has a private opinion of the reputation of the other agents.
- Indirect observation model: In this model, for each interaction, one agent is randomly selected to observe that interaction. The observer agent applies its assessment rule to the observed situation and shares its assessment with all other agents in the population. Although the observer is susceptible to *perception errors*, the observer's assessment is faithfully propagated throughout the agent population. Therefore, all agents agree on the reputation of each agent in the population.
- Intermittent observation model: In this model, the agents have a shared opinion of the reputations of all agents in the population. However, this reputation knowledge is intermittent due to factors such as forgetfulness. Due to the intermittent nature of each agent's reputation knowledge, when an agent is selected to play the donor role, there is a probability that the agent will not remember the reputation of the recipient.

In each model, the agents share a prior belief in the goodness of the other agents in the population. This prior belief is used to assign a reputation to an agent in the absence of any knowledge of that agent's reputation.

The following table describes the observation models that appear in the reviewed literature.

Type	Perception Error	P(Observes)	P(Good)	References
Direct	No	1	N/A[2]	[6][8]
Direct	No	<1	1	[3]
Direct	Yes	1	1	[6]
Direct	Yes	<1	1	[8]
Indirect	No	N/A	<1	[4]
Indirect	Yes	N/A	?	[10][11]
Indirect	Yes	N/A	1	[12][13]
Intermittent	N/A	N/A	1	[4][7]
Intermittent	N/A	N/A	<1	[3][7]

4 Agent Strategy

Each agent in the population follows a particular strategy. A strategy is composed of two parts: an assessment module and an action model. Given the framework described in the next two sections, there are 256 possible assessment modules and 16 possible action models leading to a total of 4096 possible agent strategies.

4.1 Assessment Module

When observing an interaction between a donor and a recipient, an agent uses an assessment module to assign a moral value to the action taken by the donor. An assessment module assigns a moral value to all possible situations that an agent may observe. The order of an assessment rule depends on the level of granularity that is used to distinguish situations to be assessed. A first-order assessment module only considers the donor action and therefore only distinguishes two situations while a third-order assessment module considers the donor action, donor reputation and recipient reputation and therefore distinguishes eight situations. In general, an n-order assessment module can distinguish 2^n situations. An n-order assessment module assigns one of the two moral values (good or bad) to each of the 2^n situations it can distinguish. Therefore, there are 2^{2^n} different n-order assessment modules.

Some assessment modules incorporate the observation that it can be difficult to recover a good reputation after acquiring a bad reputation. In this case, when the donor agent has a bad reputation, the assessment module incorporates a probability that the agent's action is considered bad even though the assessment module would normally consider it good.

In some cases, all agents in the population share the same assessment module while in others each agent has its own private assessment module. In the case of a population divided into groups, each group may have a shared assessment module.

[2] The authors use a reputation score that ranges from -5 to +5. Each agent starts out with a reputation score of zero.

The following table describes the common assessment modules that appear in the reviewed literature.

Observed Situation			**Assigned Reputation**		
Donor Reputation	**Recipient Reputation**	**Donor Action**	**Scoring**	**Standing**	**Judging**
Good	Good	Donate	Good	Good	Good
Good	Good	Refuse	Bad	Bad	Bad
Good	Bad	Donate	Good	Good	Bad
Good	Bad	Refuse	Bad	Good	Good
Bad	Good	Donate	Good	Good	Good
Bad	Good	Refuse	Bad	Bad	Bad
Bad	Bad	Donate	Good	Good	Bad
Bad	Bad	Refuse	Bad	Bad	Bad
References			[3][4][6][7][8]	[6][7][8]	[8]

In [9], [10], [11], [12] and [13], the authors perform exhaustive searches that consider all possible third-order assessment modules.

4.2 Action Module

When selected to play the donor role, an agent uses an action module to decide which action to take. An action module determines the action an agent will take in all possible situations that the agent may encounter. As with the assessment module, the order of an action module depends on the level of granularity that is used to distinguish situations. A first-order action module only considers the reputation of one of the agents while a second-order module considers the reputations of both agents. There are two different kinds of first-order action modules, those that consider the donor's reputation and those that consider the recipient's reputation. Similar to assessment modules, there are 2^{2^n} possible n-order action modules.

Some action modules incorporate the observation that agents are not always able to correctly execute the action specified by the action module. For example, the donor may lack the resources necessary to make a donation to the recipient. In this case, when the donor interacts with a recipient, the agent is susceptible to an *execution error*. There are two variations of the execution error that appear in the literature:

- Two way execution error: When interacting with a recipient, there is a probability that the donor takes the opposite action specified by its action module. [6][8]
- One way execution error: When interacting with a recipient, there is a probability that the donor fails to donate when its action module specifies that it should donate. In this case, the donor never fails to not donate when its action module specifies that it should not donate. [10][11][12][13]

The following table describes the common action models that appear in the reviewed literature.

Situation		Action to Take					
Donor Rep.	**Recipient Rep.**	**SELF**	**CO[3]**	**AND**	**OR[4]**	**ALLC**	**ALLD**
Good	Good	Refuse	Donate	Refuse	Donate	Donate	Refuse
Good	Bad	Refuse	Refuse	Refuse	Refuse	Donate	Refuse
Bad	Good	Donate	Donate	Donate	Donate	Donate	Refuse
Bad	Bad	Donate	Refuse	Refuse	Donate	Donate	Refuse
References		[3][6][8]	[3][4][6][7][8]	[3][6][8]	[3][6][7][8]	[3][4][6][7][8]	[3][4][6][7][8]

In [9], [10], [11], [12] and [13], the authors perform exhaustive searches that consider all possible second-order action modules.

[3] The CO strategy is also referred to as the "discriminator" strategy.
[4] The OR strategy is also referred to as the "contrite tit-for-tat" strategy.

5 References

[1] Axelrod, R., and W. D. Hamilton, "The evolution of cooperation," Science, vol. 211, pp. 1390-1396, 1981.

[2] Nowak, M. A., "Five Rules for the Evolution of Cooperation," Science, vol. 314, pp. 1560-1563, 2006.

[3] Nowak, M. A., and K. Sigmund, "Evolution of indirect reciprocity by image scoring," *Nature*, vol. 393, pp. 573-577, 1998.

[4] Nowak, M. A., and K. Sigmund, "The Dynamics of Indirect Reciprocity," *Journal of Theoretical Biology*, vol. 194, pp. 561-574, 1998.

[5] Sugden, R., *The economies of rights, co-operation and welfare*, Oxford, UK: Basil Blackwell, 1986.

[6] Leimar, O., and P. Hammerstein, "Evolution of cooperation through indirect reciprocity," *Proceedings of the Royal Society London B*, vol. 268, pp. 745-753, 2000.

[7] Panchanathan, K., and R. Boyd, "A tale of two defectors: the importance of standing for evolution of indirect reciprocity," *Journal of Theoretical Biology*, vol. 224, pp. 115-126, 2003.

[8] Brandt, H., and K. Sigmund, "The logic of reprobation: assessment and action rules for indirect reciprocation," *Journal of Theoretical Biology*, vol. 231, pp. 475-486, 2004.

[9] Ohtsuki, H., and Y. Iwasa, "How should we define goodness? – reputation dynamics in indirect reciprocity, " *Journal of Theoretical Biology*, vol. 231, pp. 107-120, 2004.

[10] Ohtsuki, H., and Y. Iwasa, "The leading eight: Social norms that can maintain cooperation by indirect reciprocity," *Journal of Theoretical Biology*, vol. 239, pp. 435-444, 2006.

[11] Ohtsuki, H., and Y. Iwasa, "Global analyses of evolutionary dynamics and exhaustive search for social norms that maintain cooperation by reputation, " *Journal of Theoretical Biology*, vol. 244, pp. 518-531, 2007.

[12] Chalub, F. A. C. C., F. C. Santos, and J.M. Pacheco, "The evolution of norms," *Journal of Theoretical Biology*, vol. 241, pp. 233-240, January 2006.

[13] Pacheco, J. M., F. C. Santos, and F. A. C. C. Chalub, "Stern-Judging: A Simple, Successful Norm Which Promotes Cooperation under Indirect Reciprocity," *PLoS Computational Biology*, vol. 2, issue 12, December 2006, pp. 1634-1638.