IMPLICIT COOPERATION IN CONFLICT RESOLUTION FOR SIMPLE AGENTS

P. SCHERMERHORN, University of Notre Dame, IN[1]
M. SCHEUTZ, University of Notre Dame, IN

ABSTRACT

Conflicts over resources can be resolved in many ways, from fighting to sharing. We introduce here a very simple mechanism for implicitly taking turns, the *2-turn-taking rule*. Agents adjust their tendencies to fight over a resource based on previous encounter outcomes. Agents possessing this mechanism are shown to be effective in competition with agents lacking the mechanism, indicating that there is some benefit to fairness, particularly when it comes at such a low computational cost.

Keywords: Conflict resolution, turn-taking, cooperation

INTRODUCTION

In real life, agents (humans and animals) have different needs and desires of different urgency, which they attempt to satisfy. These needs and desires can range from very basic ones in all animals such as the need to eat, survive or procreate, to more complex ones such as the desire to be respected or to need to have social relationships. Needs and desires typically involve resources (the objects of the need or desire). When fewer resources are available than there are agents needing/desiring them, agents will be likely in conflict over these resources. In its most general form, the conflict will end in one of three ways: (1) some agents win, the others retreat, (2) nobody gets the resource (everybody loses), or (3) the resource can be shared (everybody gets a part, but not the full resource). In this paper, we examine encounters of the first kind and study a mechanism that allows agents that implement (1) to reach (3) (over the course of multiple encounters), while avoiding (2), which will turn out to be beneficial for the whole population.

Previous work with agents that display their action tendencies–whether to continue an encounter or whether to abort it–has shown that taking other agents' displayed action tendencies into account leads to better group outcomes (Scheutz and Schermerhorn, 2003). For example, if it is obvious that an opponent is very likely to continue to fight over the resource (i.e., that it has a high action tendency to fight) and ultimately win the encounter, then it is not in an agent's best interest to enter the fight when it is less likely to continue to fight (i.e., it has a lower action tendency to fight) and win the encounter, thereby wasting resources fighting while gaining no benefit. Retreating immediately may also be costly, but compared to the cost of prolonged fighting, it is in the agent's best interest to retreat. Furthermore, it is in the more aggressive agent's best interest for its opponent to leave early, because prolonged fights reduce the net benefit of the resource being contested.

So-called "rational" agents that can figure out immediately who is likely to win an encounter (by comparing their own action tendencies with their opponents' displayed tendencies) tend to have higher survival rates than agents whose ability to predict is worse, given a random

[1]Corresponding author address: P. Schermerhorn, Artificial Intelligence and Robotics Laboratory, Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN 46556, USA; pscherm1@cse.nd.edu.

distribution of action tendencies (Scheutz and Schermerhorn, 2003). This advantage is especially pronounced in competition with "asocial" agents, i.e., agents that do not attempt to use their opponents' action tendencies into account. In an evolutionary environment, however, perfect prediction can lead to problems. Perfect predictors with high action tendencies will win more encounters than agents with lower action tendencies, and will, therefore, have access to more resources than others. They will then be more likely to reproduce and pass on the higher action tendencies to their offspring, causing the average tendency of the population to skew upward. As time goes on, successive generations will continue this arms race until every agent has an action tendency of 1, at least in the absence of external pressures. However, once every agent has a 100% probability of fighting, fights will go on until one or more agents is unable to continue. The extended conflicts are very costly, and the population will dwindle or become extinct as a consequence.

One strategy that could overcome this problem is turn-taking (Neill, 2003). In a two-agent environment, for example, if each player would gain the resource benefit every other turn while paying only the retreat cost every other turn, both would benefit while avoiding costly fights. Thus players get the benefit of perfect prediction (i.e., fewer resources spent on fighting) *and* the added benefit of fairness, so that resources are not hoarded by the most aggressive agents, thereby heading off the arms race scenario described above. Generalizing the turn-taking strategy to environments with more than two agents is made difficult by the fact that agents will not be entering conflicts with the same opponents every time. It is possible that both agents in a conflict took their turn at losing in the previous encounter, so both will expect to win in the current encounter. An effective turn-taking strategy must resolve this issue. This paper describes a strategy for implicit turn-taking based on modifying action tendencies in response to wins and losses for multi-agent environments.

THE TURN-TAKING RULE

There are a number of ways to implement a fair turn-taking rule (e.g., Iizuka and Ikegami, 2002). An agent could dedicate memory and processing resources to remembering with whom it has interacted, and whose turn it is to win next. Over the lifetime of the agent, however, it may may interact with dozens or hundreds of agents. Requiring an agent to set aside resources for all of these agents in such an explicit turn-taking mechanism could be quite burdensome, especially for very simple agents. A mechanism that ensures fairness without consuming a substantial proportion of the available resources could flourish in "selfish" populations, given its potential benefits and low cost.

We introduce the *2-turn-taking rule* (2TTR), which allows agents to keep track of their wins and losses using a simple computational procedure:

Definition [2-turn-taking-rule] Let r be the rest value of agent A and let m be the action tendency. Then $2TTR(m)^+$ is defined (for losses) as follows: if $m \geq r$, then $2TTR^+(m) = m + (1-m)/2$; if $m \leq r/2$, then $2TTR^+(m) = 2{*}m$; else $2TTR^+(m) = r + (2m-r)(1-r)/2r$ (this maps values in the interval $(r/2, r)$ into $(r, (1-r)/2)$). Similarly, $2TTR(m)^-$ is defined (for wins) as follows: if $m \geq r + (1-r)/2$, then $2TTR^-(m) = m - (1-m)$; if $m \leq r$, then $2TTR^-(m) = m/2$; else $2TTR^-(m) = r/2 + r(m-r)/(1-r)$ (this maps $(r, (1-r)/2)$ into $(r/2, r)$).

Corollary 1 *Let $2TTR^{+,n}(m)$ ($2TTR^{-,n}(m)$) denote the n-fold (recursive) application of $2TTR^+$ ($2TTR^-$) to m. Then $2TTR^{-,n}(2TTR^{+,n}(m)) = m$ and $2TTR^{+,n}(2TTR^{-,n}(m)) = m$ (where $2TTR^{+,1} := 2TTR^+$ and $2TTR^{-,1} := 2TTR^-$).*

Proof We show only the first part by induction on n, the second being analogous. For $n = 1$, we consider all three cases: if $m \geq r$, then $2TTR^{+,1}(m) = m + (1 - m)/2 \geq r + (1 - r)/2$, hence $2TTR^{-,1}(m + (1 - m)/2) = m$, if $m \leq r/2$, then $2TTR^{+,1}(m) = 2 * m < r$, hence $2TTR^{-,1}(2 * m) = m$, and finally, $2TTR^{+,1}(m) = r + (2m - r)(1 - r)/2r$ and, hence, $2TTR^{-,1}(2TTR^{+,1}(m)) = r/2 + r((r + (2m - r)(1 - r)/2r) - r)/(1 - r) = m$.

Now suppose the statement is true for $k = n - 1$. Then $2TTR^{-,n}(2TTR^{+,n}(m)) = 2TTR^{-,1}(2TTR^{-,k}(2TTR^{+,k}(2TTR^{+,1}(m)))) = 2TTR^{-,1}(2TTR^{+,1}(m)) = m$ by definition, induction hypothesis, and base case, respectively. $\dashv$

For the following, let P denote a population of rational agents with the 2-turn-taking rule and let $|P| = n$ be its size. Furthermore, let $@A$ denote the action of an agent A and let $@(t)$ denote the set of action tendencies of P at time t, called "configuration". Finally, we assume that the action tendencies of the initial population of agents are at their rest value, all of which are between $2TTR^{+}(min)$ and $2TTR^{-}(max)$, where max is the largest and min the lowest action tendency/rest value in P, and define *region 0* to be the interval given by $(2TTR^{-}(max), 2TTR^{+}(min))$. *Positive regions* $k > 0$ are then defined inductively by $(2TTR^{+,k}(min), 2TTR^{+,k+1}(min))$, and similarly, negative regions are defined by $(2TTR^{-,k+1}(max), 2TTR^{-,k}(max))$. The regions are defined such that if an agent competes against another agent in the same region k looses, then the 2-turn-taking rule will update the agent's action tendency such that the losing agent's action tendency will be in region $k+1$, and the winner's action tendency will be in region $k - 1$, as shown by the following Corollary.

Corollary 2 *Let a be an action tendency in region k. Then $2TTR^{+}(a)$ is in region $k + 1$ and $2TTR^{-}(a)$ is in region $k - 1$.*

Proof Let a be an action tendency in k. We distinguish three cases. Suppose $k > 0$, then $min^k = 2TTR^{+,k}(min) \leq a < min^{k+1} = 2TTR^{+,k+1}(min)$. Then applying $2TTR^{+}$ to all part of the inequality given that $2TTR^{+}(a)$ is strictly monotone, $min^k = 2TTR^{+,k+1}(min) \leq 2TTR^{+}(a) < min^{k+2} = 2TTR^{+,k+2}(min)$, i.e., $2TTR^{+}(a)$ is in region $k + 1$. The other cases are shown analogously. $\dashv$

We can now show that the 2-turn-taking rule in combination with the rational agents is fair in a clearly specified sense: the difference between wins and losses is bound by $int(log(n) + 1)$, where $|P| = n$ is the size population of competing agents. First observe, that all action tendencies are less than $2TTR^{+,int(log(n)+1)}(min)$ and greater than $2TTR^{-,int(log(n)+1)}(max)$ for rational agents (this lemma essentially uses the rational agent's decision rule and is not true of other agents, e.g., probabilistic agents):

Lemma 3 *For every agent A in population P size $|P| = n$ of rational agents, all action tendencies $@A$ are less than $2TTR^{+,int(n/2)+1}(min)$ and greater than $2TTR^{-,int(n/2)+1}(max)$.*

Proof [Sketch] Lemma: For each n there is exactly one configuration, in which each positive and negative region $n/2$ inhabits exactly one agent (with the 0-region also inhabiting one agent for odd n and empty for even n) and the configuration can be reached from the initial configuration (with all agents inhabiting region 0).

The lemma shows that the spread is at least $n/2$ in each direction. To see that it is at most that suppose that there is an agent A's whose action tendency $@A$=a is greater than $2TTR^{+,int(n/2)+1)}(min)$, i.e., $a \in region\ int(n/2) + 1$. Then A must have lost a fight against another agent with a higher action tendency than $2TTR^{+,int(n/2)}(a)$ in region $int(n/2)$ given the rational decision rule. However, that means that two agents were in region $int(n/2)$, which is not possible: suppose it were possible, then by backwards induction using the above argument, at least two agents must have inhabited region $int(n/2) - 1$ at some point, as so forth. Eventually, after $int(n/2)$ steps we reach region 1, which also must have two agents in it, but that is impossible, since it means that there is a total of $2 \cdot (int(n/2)+1) > n$ agents (given that positive and negative regions are symmetric for the maximal spread). $\dashv$

Now we can prove that the difference between wins and losses is bounded by a fixed parameter d for all agents for any number of interactions.

Lemma 4 *For an agent population P of size n, there exists a d such that $|wins - losses| < d$ for all agents in P for any configuration.*

Proof First observe that $losses - wins$ denotes the region the agent is in at any time, given that the agent started out in region 0 (as we assume for all agents): whenever an agent wins/loses, the agent is put in a lower/higher region and its number of wins/losses is increased/decreased. At the best/worst, the agent can be in the highest/lowest region $int(n/2)$/$-int(n/2)$ corresponding to $int(n/2)$ losses/wins. Hence, the largest possible difference between wins and losses, $|wins - losses|$, is bounded by $int(n/2)$ for every agent. $\dashv$

This fixed bound d implies that in the long run the difference between the agents accumulated utility is a most the constant given by $int(n/2) * (w - l)$, where w is the utility of winning, and l the utility of losing. This leads to a definition of what it means for an agent's action tendency update rule to be *balanced* for an agent group P:

Definition Given an agent group P with a uniform update rule $R = @A$ of A's action tendency for every agent $A \in P$. A is balanced if there exists a fixed bound d such that for any number of random interactions of agents in the agent group such $|wins - losses| < d$, for all agents in P.

The above definition captures the best balance we could hope for in an agent populations that is continuously engaged in competitive interactions: although the best distribution would be achieved if all agents inhabited region 0, which would imply that $|wins - losses| = 0$ for all agents, this is not always going to be the case, given that there are competitions. The worst possible distribution would be one, where $|wins - losses| = d$ for all agents, but this is fortunately only possible in the 2 agent population (where there are only three classes), in any other population, only two agents can have d at any given time (the other ones, must lower differences). There is nothing in principle that prevents an agent population from assuming any of the possible configurations determined by the initial distribution of agents and their population size if interactions are chosen at random.

We now can state the main theorem of this section, which follows from the above lemmas and corollaries:

Theorem 5 *The 2-turn-taking rule is balanced for rational agents.*

Theorem 5 is valid for *rational agents* as described above that make a determined decision based on the action tendencies of both agents in a conflict. These agents basically treat the action tendency as a counter that keeps track of how many wins or losses an agent has. The *asocial* agents described above, on the other hand, use the action tendency as a probability that they will decide to fight. For this reason, it is still possible for both agents to flee, for both agents to fight, or even for the agent with the lower action tendency to stay while the other retreats. The 2-turn-taking rule serves as a nondeterministic place-keeper whose behavior may in the short run appear unfair, but in the long run (as the number of encounters approaches infinity) should come out fair. In fact, experimental results described below indicate that even over the relatively short run, in which agents average only a handful of conflicts over their lifetimes, the probabilistic version of the turn-taking mechanism provides its holders with an advantage over "selfish" agents. The details of the formal argument are currently under investigation, while at present the results outlined below support this hypothesis.

EXPERIMENTAL RESULTS

A series of experiments were run to test the effectiveness of the turn-taking mechanism described above. The experiments took place in the SimWorld artificial life environment (Scheutz, 2001). SimWorld provides a world in which agents forage for food, procreate asexually when they are sufficiently old and have enough energy stores, and interact with one another during the pursuit of these goals. When agents come into close proximity, they are considered to be in conflict over whatever resources are present (if any). In a given cycle during a conflict, an agent may decide to retreat or fight. Each action carries with it a substantial cost. The benefit of retreating is that the cost is paid only once, whereas the cost of fighting is paid as often as both agents decide to fight. Agents who fight until their opponents retreat (or die!) will obtain the benefit of whatever resources are in the immediate vicinity.

Agents decide how to react in an encounter based on their action tendencies for conflicts. Each action tendency is mapped onto the range from 0 to 1, and can be thought of as the likelihood that an agent will decide to stay and fight in an encounter. Some agent types (so-called "asocial" agents) consider only their own action tendencies when making decisions about fighting, so their action tendencies map directly onto their probability of fighting. Others, however, mediate their probability of fighting based on a comparison of their own tendencies with those of their opponents. For these agent types, the probability of fighting is increased when the agent's action tendency is higher than that of its opponent and decreased otherwise. The "rational" agent is an extreme case of this type that will fight if and only if its action tendency is higher than its opponent's.

In addition to these "rational" and "asocial" agent types, turn-taking versions of each type based on the 2-turn-taking rule described above are tested. The first set of experiments tests the agent types in homogeneous environments. Each agent type is placed in an environment containing only agents of its own kind. The environment is unbounded, however, food is randomly generated in only a 1440 by 1440 area; agents may wander out of this area, but must return replenish their energy stores. New food sources are generated at random locations within the food region with a probability of 0.5 per cycle. Agents reproduce in roughly 350-cycle generations, and the simulations were run for 10,000 cycles. 20 agents are generated and placed at random locations within the food region at the beginning of each experimental run; performance is measured by counting the number of survivors at the end of an experimental run. The numbers given here are averaged over 80 experimental runs with differing random initial conditions.

Each agent type was tested using two methods to determine initial action tendencies or action tendency rest values (for non-turn-taking agents and turn-taking agents, respectively). The first method assigns a value randomly chosen in a Gaussian distribution centered at 0.5. This ensures that there will be a diversity in the action tendencies and rest values throughout the experimental run. The results from experiments using this method are given in Figure 1. The first four columns

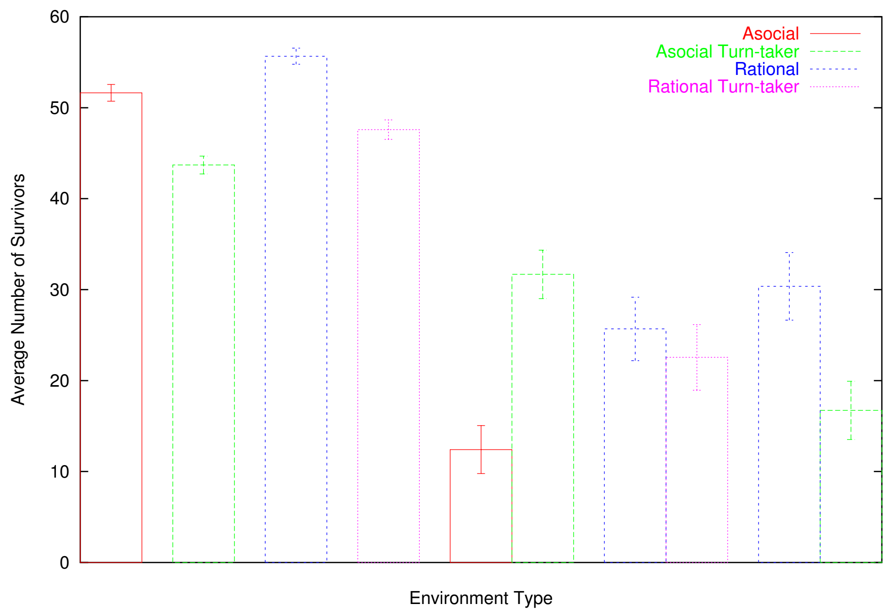

Figure 1: Survival Rates for Gaussian-distributed Action Tendencies for Asocial, Rational, Asocial Turn-taker and Rational Turn-taker (the four bars to the left indicate results from homogeneous environments, while the six on the left are results from heterogeneous environments)

depict the average number of survivors of each of the four types (Asocial, Asocial Turn-taker, Rational, and Rational Turn-taker) in homogeneous environments where they compete against only agents of their own kind. For both Asocial and Rational agent types, the normal agents outperformed their turn-taker counterparts. In the absence of turn-takers, it would appear that the unfair approach is more efficient.

Placing normal and turn-taking agents in the same environment yields mixed results. The next six columns are paired to indicate that these results are from heterogeneous environments in which two agent types competed (Asocial vs. Asocial Turn-taker, Rational vs. Rational Turn-taker, and Rational vs. Asocial Turn-taker). The first pair compares Asocial agents with Asocial Turn-takers; the turn-taking agents enjoy a pronounced advantage over the normal Asocial agents. The next pair shows that there is no significant difference between Rational agents and Rational Turn-takers. The third pair is of interest in light of previous results (Scheutz and Schermerhorn, 2003) in which normal Asocial agents failed to average even one survivor against normal Rational agents. The improvement to averaging over fifteen survivors with the addition of the turn-taking mechanism is testimony to the benefit of the mechanism.

Initial experiments used only the Gaussian method for determining action tendencies and rest values. However, our prediction was that in an evolutionary context, agents would engage in a self-destructive arms race, driving their action tendencies up to the point where the population could not sustain itself. To explore this possibility, further experiments were conducted in which agents of both kinds inherited their tendencies and rest values. This gives the normal agents a chance to raise their own action tendencies, albeit over a longer time scale than the turn-takers, and in all likelihood in only one direction. The agents in the initial group of a simulation run were given values using the same Gaussian distribution, but thereafter the values were inherited by their

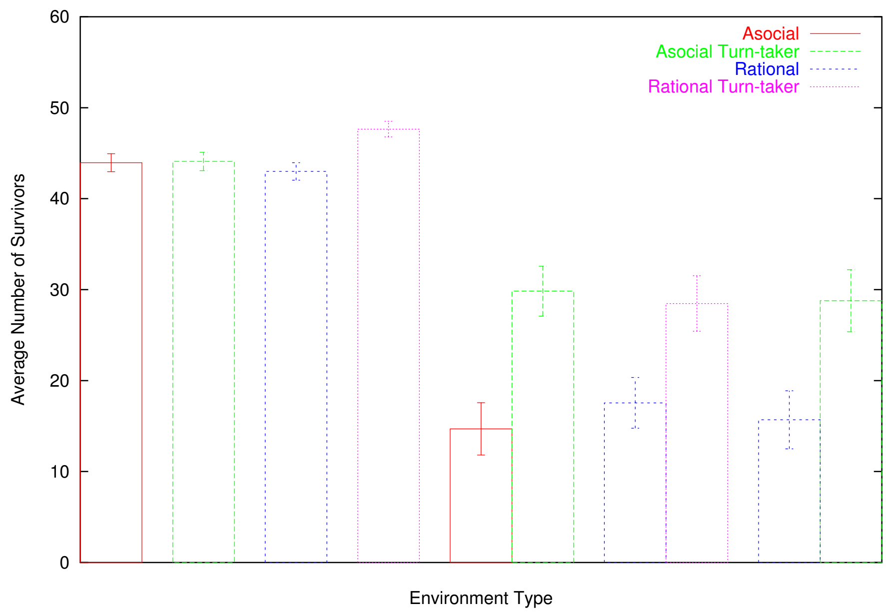

Figure 2: Survival Rates for Inherited Action Tendencies for Asocial, Ration, Asocial Turn-taker and Rational Turn-taker (the four bars to the left indicate results from homogeneous environments, while the six on the left are results from heterogeneous environments)

offspring (without modification). The idea here is to test whether this turn-taking mechanism is really an effective method of conflict resolution that could avoid the trap of a destructive arms race.

Figure 2 presents the results of this set of experiments. What we find is encouraging. The advantage the normal agents has vanished in the homogeneous environments. In fact, Rational Turn-takers outperform normal Rational agents. This is in line with our prediction that unrestrained increase in the average action tendency would eventually prove costly. In the mixed environments, the Asocial Turn-takers retain their advantage over normal Asocial agents (although the gap closes somewhat). The other two pairings are more dramatic. Rational Turn-takers now perform significantly better than their non-turn-taking counterparts. Asocial Turn-takers also now perform significantly better than normal Rational agents (recall that normal Asocial agents perform very poorly against Rational agents). These results support the hypothesis that fair turn-taking can be an effective strategy for avoiding inflation of action tendencies to destructive levels.

CONCLUSIONS

We have described here a simple method for implicit cooperation via turn-taking. The 2-turn-taking rule specifies that an agent should modify its behavioral tendencies in response to winning or losing an encounter with another agent. The 2 turn-taking rule is guaranteed to be fair for encounters between two agents that employ a deterministic decision procedure (i.e., that use their action tendency values as counters to keep track of wins and losses). When implemented in simple artificial agents, the mechanism provides an advantage to its possessors, in particular, when agents are allowed to take part in an "arms race" of inheriting the action tendencies of their parents. Such an arms race leads to progressively higher action tendencies, on average, which is destructive to non-turn-takers who have no mechanism to back down from time to time, engaging them in costly extended conflicts. The results of our simulation indicate that fairness can be a winning strategy against selfish agents, and that, moreover, the proposed mechanisms–the 2-turn-taking-rule–can even be implemented by agents that do not or cannot take other agents' action tendencies into account, while retaining its benefits for the whole agent population.

It is important to note that the turn-taking mechanism introduced here is a very simple one whose cost is small enough that for all practical purposes, it can be ignored. In the evolutionary environment, this is key, because it indicates that the substantial benefits outlined above will not be offset by equally substantial costs, as would be the case for a more sophisticated turn-taking mechanism that, for example, relied on memory to recall the history of an agent's interactions in order make good decisions about whose turn it is to win. This extremely low-cost method of ensuring fairness would almost surely invade any unfair population in which it arises, based on the figures above.

We are currently exploring further the theoretical properties of the probabilistic version of the 2-turn-taking rule, and expect that long-term evolutionary investigations (including the ability to mutate) will reveal its advantage in a great variety of environmental settings and conflict types.

REFERENCES

Iizuka, H. and Ikegami, T. (2002), "Simulating turn-taking behaviors with coupled dynamical recognizers," In Mark A. Bedau Russell K. Standish and Hussein A. Abbass, editors, *Artificial Life VIII*.

Neill, D. (2003), "Cooperation and coordination in the turn-taking dilemma," In *14th International Conference on Game Theory at Stony Brook*.

Scheutz, M. and Schermerhorn, P. (2003), "The role of signaling action tendencies in conflict resolution," *Journal of Artificial Societies and Social Simulation (under review)*.

Scheutz, M. (2001), "The evolution of simple affective states in multi-agent environments," In Dolores Cañamero, editor, *Proceedings of AAAI Fall Symposium*, pp. 123–128, Falmouth, MA. AAAI Press.