SELF-INTERESTED AGENTS CAN BOOTSTRAP SYMBOLIC COMMUNICATION IF THEY PUNISH CHEATERS

EMILY WANG

AI Laboratory, Vrije Universiteit Brussel, Pleinlaan 2,
Brussels, 1050, Belgium
emily@arti.vub.ac.be

LUC STEELS

Sony Computer Science Laboratory, 6 Rue Amyot,
Paris, 75005, France
AI Laboratory, Vrije Universiteit Brussel, Pleinlaan 2,
Brussels, 1050, Belgium
steels@arti.vub.ac.be

We examine the social prerequisites for symbolic communication by studying a language game embedded within a signaling game, in which cooperation is possible but unenforced. Despite incentives to cheat, and even with persistent cheating, the lateral inhibition dynamics commonly used in language game models remain resilient, as long as sufficient mechanisms are in place to detect deceit. However, unfairly antagonistic strategies can undermine lexical convergence. Symbolic communication, and hence human language, requires a delicate balance between restrained deception and revocable trust, but unconditional cooperation is unnecessary.

1. The Reciprocal Naming Game

Sociality is generally regarded as a prerequisite for symbolic communication. In contrast to animal communication, which is automatic and costly, symbolic communication is controlled consciously and conventional. This makes cheating easy and hence raises the question how the necessary trust relations could have emerged (Dessalles, 2000; Knight, 1991; Steels, 2008). Kin selection could have played a role, but one of the greatest strengths of language is precisely the ability to communicate with unrelated strangers. A well known mechanism for establishing sociality is reciprocal altruism, which has been shown to emerge if individuals can recognise each other, keep a record of cooperative behavior, and direct their own altruistic behavior towards those who have shown cooperation in the past (Trivers, 1971). Axelrod and Hamilton (1981) have shown that a tit-for-tat policy can achieve this. With tit-for-tat, a player initially cooperates, and subsequently plays the same move against another player as this other player did in their prior interaction. In this paper we study the interaction between reciprocal altruism, established and maintained by a tit-for-tat strategy, and the emergence of symbolic communication. We combine two well studied models, the Naming Game and the Signaling Game, to make the *Reciprocal Naming Game*, and then study how well agents must cooperate in order to agree.

The Naming Game has been introduced (Steels, 1996) as a minimal model for studying the conventionalization of names in a population of agents having only local peer-to-peer interactions and no global coordination whatsoever. The game has been studied both from a computational and theoretical point of view (Vylder & Tuyls, 2006). A widely used strategy relies on a *lateral inhibition mechanism* to reward successful words and inhibit their competitors.

The Crawford-Sobel model of strategic information transmission (1982) defines a two-player strategy game called a *signaling game*. For convenience, we denote the *signaler* as S, and the *receiver* as R. S is better informed than R, with private information t about the environment. S transmits a message m to convey either t, or something misleading. Based on m, R takes an action a that determines the payoff for both players. If S adopts a strategy of lying about t, then R adapts by ignoring information in m.

In the Naming Game, the *speaker* utters a word to best convey the intended referent to the *hearer*. But in a signaling game, the signaler need not transmit $m \cong t$. We combine the two games into a single game by presenting two players, randomly chosen in each iteration, with a context of two items, one of which is the *target*, and the other a *distracter*. S has access to this information, but may choose either item as the referent. This situation can be conceived as a shell game (similar to Three-card Monte), where a set of shells forms the context, and a dealer (the simulation) has hidden a *pea* under one of the shells. The better (the receiver) wins by guessing correctly the shell that contains the pea. S is a third party who may act either as an informant and truthfully indicate the target to R, or as a shill by indicating the distracter. So S may use m to deceive and R must decide whether to believe m. Thus the interaction scheme is similar to that of the regular Naming Game, but without feedback from explicit pointing. With the Reciprocal Naming Game, the signaler's intended meaning is never revealed to the receiver. Adding this layer of uncertainty preserves the privacy of each player's choice whether to cooperate or defect.

The remainder of this paper studies the Reciprocal Naming Game. We first introduce a minimal agent architecture needed to play the game, and then some different strategies. Next we report on the result of computational simulations that examine key questions about the social prerequisites of symbolic communication.

2. Agent Architecture

To remember names, each agent is equipped with a *lexical memory* associating words with meanings and scores. Multiple lexicon entries may share the same word or meaning, and these competing conventions can be ordered by preference according to their score. Scores are governed by lateral inhibition, that is, incremented following successful interactions and decremented following failed interactions, or the successful use of a competing association. *Coherence*, a measure of the similarity between individual agent lexicons, represents agreement in a population. The *group lexicon* summarizes the most popular words, but this lexicon is only known to an external observer.

To identify other agents in the population for following a tit-for-tat strategy, each agent also has a *social memory*, associating each other individual with a rating. One agent can *regard* another either positively, with the intent to cooperate, or negatively, with the intent to defect. Two agents who regard each other in the same way share *mutual regard*, otherwise their relationship is *one-sided*.

There are eight possible outcomes of a game. Each outcome can be coded as a binary string of four bits $(a_S\, c\, a_R\text{-}p)$. The actions of the signaler and receiver are a_S and a_R, where cooperation and trust are coded as 1, and defection and disbelief as 0. The predicate c indicates whether R comprehended the message correctly, and p indicates whether R successfully located the pea. So p is set like an even parity bit, with $p = 1$ only when an odd number of the bits in $\{a_S, c, a_S\}$ are 1.

Three levels of information govern the players' knowledge. Actions a_S and a_R are kept *private* by each player. The result p is *public* information, displayed to both players, but the result c is not revealed to any player; it is known only by virtue of experimenter introspection. Players cannot inspect each others' internal processes, so they cannot know for certain whether partners cooperate or defect. Nevertheless, S and R can each estimate the action of the other, given knowledge of their own actions, and their observation of p.

Fig. 1 shows the eight distinct combinations of events, leading to four distinct outcomes, that can occur in one iteration of the Reciprocal Naming Game. Arrows between the states denote an indistinguishability relation (e.g., given its personal knowledge, S cannot distinguish between states 'a' and 'b'). With four versus two indistinguishability relations, S has less knowledge of the game outcome than R.

3. Player Strategies

Under the general condition of *complete reciprocity*, the signaler chooses $a_S = \text{regard}(S, R)$ and the receiver chooses $a_R = \text{regard}(R, S)$.

An *empty strategy* was implemented to refute the null hypothesis, which is that the system can converge when the agents possess deception, but not cheater detection. In this condition, S behaves as above. R assumes that the target $\cong m$, but if R cannot interpret m, then it looks for the pea under a random context item.

In another condition with only *partial reciprocity*, we relax the requirement that $a_S = \text{regard}(S, R)$, and instead allow $a_S = 0$ with some probability, even when S has positive regard for R. Each agent has a constant *fairness* parameter f, such that $0 \leq f \leq 1.0$. Whenever $\text{regard}(S, R) = 1$, a random value v is drawn

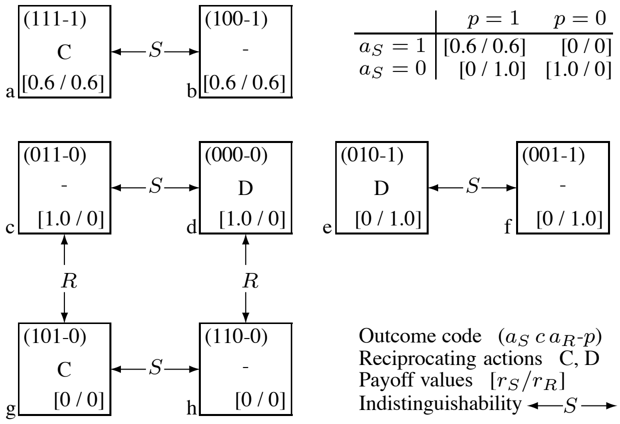

Figure 1. Payoff matrix and player knowledge of possible outcomes

such that $0 \leq v < 1.0$. Then, $a_S = 0$ if $v > f$, or if regard$(S, R) = 0$. Otherwise, $a_S = 1$. With lower f, an agent is more likely to succumb to the temptation to defect, which in the prisoner's dilemma is the high reward for defecting against a partner who cooperates and receives the sucker's payoff. A *fair agent* has $f = 1.0$, and behaves with complete reciprocity. When $f = 0$, the agent acts as a *free rider*, and can never cooperate when playing as S. Thus the signaler reciprocates previous cooperation only with probability equal to f. Adding this randomness allows S to defect when it has not been provoked by its counterpart, and this makes the behavior of S less predictable.

The agents can also employ different strategies for updating their memories. For the lexicon, both players promote the association that was applied in the interaction when they have received a nonzero reward, and they demote associations resulting in zero payoff. With a *short-term memory* strategy, associations reaching the minimum score threshold are deleted from the lexicon, but such entries are kept when using *long-term memory*.

Updates for social regard are less symmetric. The signaler's sole criteria for updating its regard for R is whether or not the receiver chose the object that was intended, thus S assumes $c = 1$. When $a_S = 1$, the intended object is the target, and when $a_S = 0$, it is the distracter. So the receiver's choice matches the signaler's intention when $p = a_S$. The receiver employs a more complex method to estimate whether the signaler cooperated in the interaction. As shown in Fig. 1, R can sometimes deduce c and a_S, given a_R and p. When $r_R = 0.6$ it is certain that $a_S = 1$, even if R did not cooperate. R responds by cooperating with S next. When $r_R = 1.0$, both players defected, and R continues to defect against S. When $r_R = 0$, there is some uncertainty, and R responds by modifying its regard for S by a bit-flip.

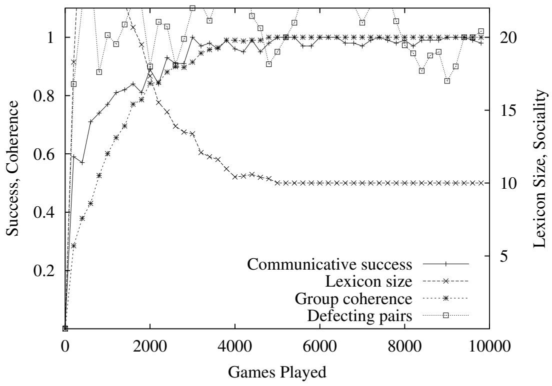

Figure 2. Simulation where cheaters are punished effectively

4. Experimental Results

Figs. 2 and 3 show a Reciprocal Naming Game with ten objects and ten agents using short-term memory, which converges around 5,000 games. (Color versions of figures can be viewed at http://arti.vub.ac.be/~emily/evolang7-figures.) In the first graph we see the running average of communicative success, as well as lexicon size, which initially shows the typical explosion when words get invented and propagated, followed by an approach towards the optimal lexicon. Even under the harsher conditions of the Reciprocal Naming Game, the agent population is capable of reaching agreement on a set of lexical associations. However, communicative success remains less than perfect, even when coherence is full. The graph also shows that the number of defecting pairs fluctuates due to continuing communicative failures. The second graph displays some general indicators of success, such as whether R found the pea p, and whether any reward was given, which

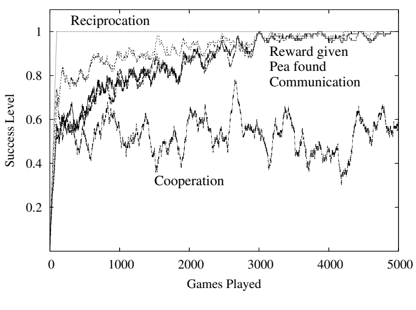

Figure 3. Reciprocation in Fig. 2

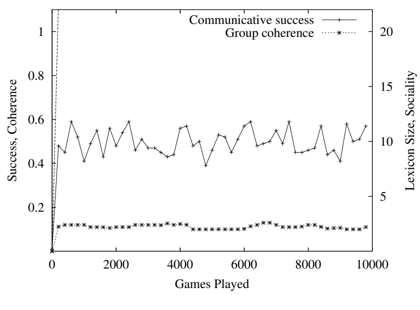

Figure 4. Simulation where R has no strategies

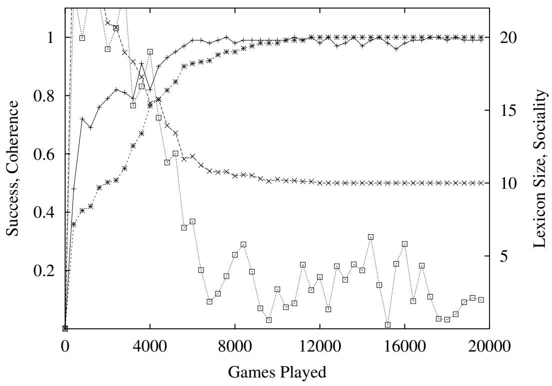

Figure 5. Simulation with long-term memory

is 0 if $r_S = r_R = 0$, and 1 otherwise. "Reciprocation" is 1 if $r_S = r_R$, and 0 otherwise. "Cooperation" is stricter, and it is 1 if $r_S = r_R = 1$, or 0 otherwise.

We now examine the importance of sociality by discussing four major issues:

4.1. *Retaliation allows deception to be tolerated*

In Fig. 4, the receivers employ the empty strategy, and simply assume that signalers are truthful. Coherence is not realized because misinterpreted messages cause too many homonyms (multiple meanings represented by the same word) that pollute the lexicon. Even though the initial population is fully cooperative, R guesses randomly when it does not know m, and this introduces negative regard into the system. Fig. 2 exhibits lingering symptoms of the same lexical inefficiency. By inspection, we can see that occasional communicative failures are caused by the imprecision of one persistent homonym. Due to the lack of pointing, agents cannot distinguish between a a zero payoff due to failure of communication from the same result due to a defecting partner.

But convergence can still carry on when the agents are equipped retaliate, as they are in Fig. 2, but not in Fig. 4. Therefore lexical convergence depends not upon a complete lack of deception, but rather upon striking a balance between the ability to deceive, and the ability to detect deception. Thus, cheater detection is essential, even if it is fallible. Since R cannot deduce the true value of a_S in all cases, it seems an approximation of the speaker's honesty suffices.

4.2. *More memory prevents the death spiral*

One weakness of tit-for-tat, cited for the iterated prisoner's dilemma, is the problem of the death spiral in noisy environments, because a single mistake can destroy a mutually cooperative relationship (Axelrod & Hamilton, 1981). The Reciprocal Naming Game seems far less prone to this potential pitfall, and this is especially

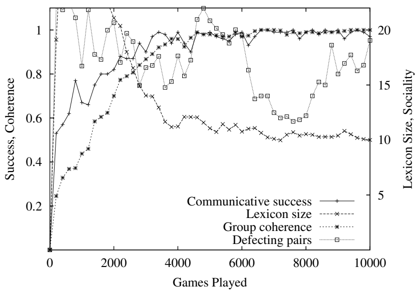

Figure 6. Simulation with one free rider

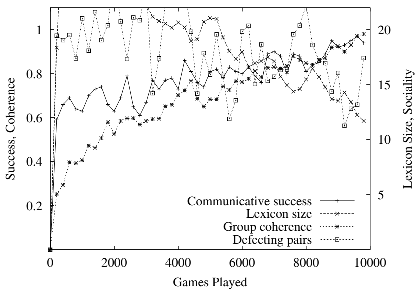

Figure 7. Simulation with three free riders

true when the agents use long-term lexical memory. When obsolete associations are kept, the agents can interpret more messages, and the increased level of comprehension seems to suppress mutually defecting pairs, nearly to zero, as shown in Fig. 5. However, convergence time doubles.

Mutually cooperative relations are more constructive and stable, since shared reward results in synchronous score promotions, while defection virtually guarantees that the players will make mismatched lexical updates. Complete equilibrium in the system would only be achieved by having an optimal lexicon and zero negative regard simultaneously, but this is highly unlikely given the current strategies.

4.3. *Limited numbers of free riders are bearable*

Fig. 6 shows that a population of mostly fair agents can accurately retaliate against a single free rider. But retaliation becomes less effective as the number of free riders grows, as shown in Fig. 7, where coherence is significantly reduced from Fig. 2. The advantage of the free rider strategy depends on how many other agents in the population are following the same strategy. Too many free riders leads to a tragedy of the commons, since free riders depend on having enough fair agents on which to prey. Reciprocating fairly becomes a losing strategy once the free riders equal or outnumber the fair agents. Free riders detract from the common good (total payoffs), since mutually cooperative interactions benefit from a 0.2 bonus. Individual utility is best served by taking part in the majority, that is, to cease reciprocating when there are more free agents than fair agents in the population.

4.4. *Reciprocation produces coherence in spite of deception*

While the agents never form explicit agreements, each agent's personal utility depends on its ability to establish reciprocal relationships. Acting without reciprocity is costly. Cooperating with a partner who defects results in the sucker's payoff. Defecting against a partner who cooperates prevents future cooperation.

But we must distinguish between failing to reciprocate and choosing not to cooperate. If two agents have established a pattern of repeated, mutual defection, then they receive roughly equal cumulative payoff. In a sense, one player sacrifices itself in each interaction, to provide the other with a large reward, and they take turns since roles are randomly assigned. This way, sharing takes place not within each interaction, but over the course of multiple interactions, resulting from adherence to tit-for-tat. But free riders never reciprocate, except by accident.

In this model, the group lexicon is itself a sort of social contract, since participation in the shared lexicon makes an agent vulnerable to those who deceive using shared words. Ostensibly, it would be every agent's goal to avoid coherence with unfair partners. However, R can still derive a nonzero reward when dealing with a deceitful partner, by interpreting m correctly, and then choosing to disbelieve it. Therefore coherence contributes to personal utility when cheaters can be detected, and this supports convergence in the face of deception.

5. Conclusion

These simulations demonstrate that peer-to-peer negotiation of conventions in language games remains viable in a social environment where deception is possible, as long as additional mechanisms exist to encourage cooperation and deter unprovoked deceit. Results presented here demonstrate that trust need not be permanent or unconditional for language development, although reciprocity remains an important aspect of these social exchanges.

References

Axelrod, R., & Hamilton, W. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390–1396.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*.

Dessalles, J.-L. (2000). Language and hominid politics. In C. Knight, M. Studdert-Kennedy, & J. Hurford (Eds.), *The evolutionary emergence of language: Social function and the origins of linguistic form*. Cambridge: Cambridge University Press.

Knight, C. (1991). *Blood relations: Menstruation and the origins of culture*. New Haven and London: Yale University Press.

Steels, L. (1996). Self-organizing spatial vocabularies. *Artificial Life Journal*, *2*(3).

Steels, L. (2008). Sociality is a crucial prerequisite for the emergence of language. In R. Botha & C. Knight (Eds.), *The cradle of language*. Oxford: Oxford University Press.

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Journal of Biology*(46), 35–57.

Vylder, B. D., & Tuyls, K. (2006). How to reach linguistic consensus: A proof of convergence for the naming game. *Journal of Theoretical Biology*, 818–831.