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

EMILY WANG

Artificial Intelligence 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
Artificial Intelligence 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, and agents have incentive to deceive. Despite this incentive, and even with persistent cheating, naming conventions can still arise from strictly local interactions, as long as agents employ sufficient mechanisms to detect deceit. However, unfairly antagonistic strategies can undermine lexical convergence. Simulated agents are shown to evolve trust relations simultaneously with symbolic communication, suggesting that human language need not be predicated upon existing social relationships, although the cognitive capacity for social interaction seems essential. Thus, language can develop given a balance between restrained deception and revocable trust. Unconditional cooperation and outright altruism are not necessary.

1. The Reciprocal Naming Game

Sociality is generally regarded as a prerequisite for symbolic communication (Steels, 2008), but given the pressure of natural selection, there remains the question of how honest communication can be evolutionarily stable when individuals might gain an advantage by deceiving others (Dessalles, 2000). In hunter-gatherer societies, imparting personal knowledge to others about the location of food can be of negligible cost and may bring extra benefits if collaboration is required to harvest the food, or if the other individuals are likely to return the favor at a later time (Knight, 1991). Reciprocity has been put forward as a mechanism that sufficiently elicits altruism directed at unrelated individuals given Darwinian constraints, as long as individuals encounter each other repeatedly over the course of many interactions, and are exposed symmetrically to opportunities for altruism, as in the prisoner's dilemma strategy game (Trivers, 1971). With a tit-for-tat policy, a player remembers each opponent's previous action so that cooperation is only directed towards those who did not defect in the previous interaction, and this has been shown to foster reciprocity because it is punishing yet forgiving (Axelrod & Hamilton, 1981). Thus, we present a computational model where individuals can recognize each other, keep a record of cooperative behavior, and direct their own altruistic behavior towards those who previously offered cooperation.

We combine two well-studied models, the Naming Game and the Signaling Game, to make the *Reciprocal Naming Game*, which we use to study the interaction between optional altruism and the emergence of symbolic communication. The Naming Game (Steels, 1995) was introduced as a minimal model for studying the conventionalization of names in a population of agents, using only peer-to-peer interactions. The goal is to develop globally accepted naming conventions from only the sum experience of many local interactions. The Crawford-Sobel model of strategic information transmission (1982) defines a Signaling Game, which is a two-player strategy game in which the players communicate using signals. 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 create a single game out of these two by presenting two players, randomly chosen out of the population 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, where a set of shells forms the context, and a dealer has hidden a *pea* under one of the shells. R is like a player who places a bet, and wins by correctly guessing which shell contains the pea. S is a third party that may act as an informant and truthfully indicate the target to R, in which case S takes a share of R's winnings. Or, S may act as a shill by indicating the distracter, and receive a payment from the dealer if R guesses incorrectly. So S may use m to deceive and R must decide whether to believe m. This 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 the players' choices 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 object 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 usage and decremented following failed interactions, or the successful use of a competing association. *Group coherence* represents agreement in the population, and this is summarized by a group lexicon of the most widely accepted words, but this measure is only known to an external observer. The agents themselves receive only local information.

To identify other agents in the population and to record previous experiences, each agent also has a *social memory*, associating each other individual with a rating. One agent can *regard* another with the intent to cooperate, regard$(a_j, a_k) = 1$, or with the intent to defect, regard$(a_j, a_k) = 0$. Two agents that regard each other in the same way share *mutual regard*, regard$(a_j, a_k) =$ regard(a_k, a_j), but otherwise their relationship is *one-sided*.

The outcome of one iteration of the Reciprocal Naming Game depends upon three binary parameters, a_S, c, and a_R. 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. A fourth value p depends on the other three, and indicates whether R successfully located the pea, which can occur on purpose or by accident, depending on c. So p is set like an even parity bit, with $p = 1$ only when an odd number of the bits in $\{a_S, c, a_R\}$ are 1, and this collapses the eight possible combinations into four distinct outcomes. These outcomes are summarized by the payoff matrix,

	$p = 1$	$p = 0$
S cooperates, $a_S = 1$	0.6, 0.6	0, 0
S defects, $a_S = 0$	0, 1.0	1.0, 0

where u denotes utility, and each entry gives u_S, u_R. Note that p is used to decide the payments instead of a_R, since the dealer or R only pay S based on the final outcome of the shell game.

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 inspection. Players cannot inspect each others' internal processes, so they cannot know for certain whether their opponents 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.

For an agent-knowledge formulation of the Reciprocal Naming Game, as well as further results not presented here, see `http://arti.vub.ac.be/~emily/msc/`.

3. Player Strategies

Under the general condition of *complete reciprocity*, the signaler chooses $a_S = \text{regard}(S, R)$ and the receiver chooses $a_R = \text{regard}(R, S)$, in accordance with tit-for-tat. An *empty strategy* was implemented to refute the null hypothesis, which would be that cheater detection has no effect on the ability of the population to agree upon lexical conventions. 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 $a_S = \text{regard}(S, R)$. Instead we allow $a_S = 0$ even when $\text{regard}(S, R) = 1$, by introducing a constant *fairness* parameter f for each agent. A *fair agent* has $f = 1.0$, and behaves with complete reciprocity. When $f = 0$, the agent acts as a *free rider*, and always defects when playing as S, although it can still choose to believe the signaler when playing as R.

The agents also employ specified 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 considers the size of u_R to estimate whether the signaler cooperated in the interaction. As illustrated by the payoff matrix, R can sometimes deduce c and a_S, given a_R and p. When $u_R = 0.6$, it is certain that $a_S = 1$, even if R did not cooperate. R responds by cooperating with S next. When $u_R = 1.0$, both players defected, and R continues to defect against S. When $u_R = 0$, R cannot be certain about a_S, and responds by modifying its regard for S by a bit-flip, since the payoff was not favorable.

4. Experimental Results

Figure 1 shows a Reciprocal Naming Game with ten objects and ten agents using short-term memory. Measures are shown as running averages. Figures 2–5 are meant to be read in direct comparison to Fig. 1 (and so they have been simplified, and afforded less space; complete color versions can be viewed at `http://arti.vub.ac.be/~emily/evolang7/`). In successful systems, an initial lexical explosion due to the rapid invention of new words is followed by an approach towards high group coherence and communicative success as the lexicon becomes more efficient. Even under the more challenging conditions of the Reciprocal Naming Game, the agent population is capable of reaching complete agree-

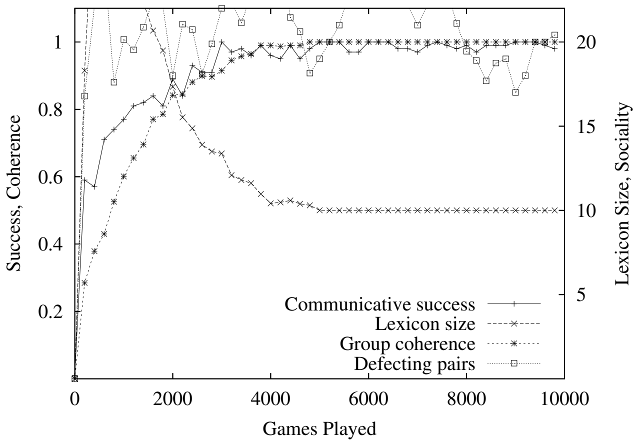

Figure 1. Lexical agreement is not hindered by cheating in a simulation where the agents employ tit-for-tat and have short-term memory. The lexicon becomes optimal and stable after 5,000 games, with complete group coherence fixed at 1.0, and lexicon size at 10. Communicative success is near perfect, but fluctuates just below 1.0. Reciprocating relationships are split about equally, and fluctuating.

ment on a set of lexical associations, despite the persistence of mutually defecting pairs. However, communicative success remains less than perfect, even when coherence is full, due to homonyms that are propagated following games where m was misunderstood. Because of the lack of pointing, agents cannot distinguish between a a zero payoff due to failed communication, and the same result due to a defecting partner. Thus communicative success and social relationships fluctuate continuously as a result of lexical inefficiency.

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

4.1. *Retaliation allows deception to be tolerated*

In Fig. 2, R employs the empty strategy and simply assumes that S is truthful, while S follows tit-for-tat. Coherence is not realized because misinterpreted messages pollute the lexicon with many homonyms. Even though the initial population is fully cooperative, R guesses randomly when it does not know m, and this introduces uncooperative regard into the system. So agreement can form when the agents are equipped retaliate, as they are in Fig. 1, but not in Fig. 2. This clearly rejects the null hypothesis since the population only develops group coherence when the receivers, as well as the speakers, follow a policy of reciprocation. Therefore lexical convergence depends not upon a complete lack of deception, but rather upon balance between deception and the ability to detect it. Given this, individuals can direct their altruism accordingly. But since R cannot always deduce the true value of a_S, it seems even an approximation of the speaker's honesty suffices. Thus, cheater detection is essential, even if it is fallible.

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, where a single mistake can destroy a mutually cooperative relationship (Axelrod & Hamilton, 1981). The Reciprocal Naming Game tends to resist this pitfall since the true actions, a_S and a_R, remain private, and players must deal with doubt when estimating these values. Cooperative relations become even more robust with long-term lexical memory, when obsolete associations remain accessible to R for interpreting m. This increases the chance of comprehension, and suppresses defecting pairs to much lower numbers, as shown in Fig. 3. The time to reach convergence doubles, but mutually cooperative relations are more constructive and stable since a shared reward results in synchronous score promotions, while defection virtually guarantees that the players will make mismatched lexical updates.

4.3. *Limited numbers of free riders are bearable*

Figure 4 shows that a population mostly composed of 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. 5 where coherence is significantly more difficult to achieve, and unstable. Free riders detract from the common good in total utility, since mutually cooperative interactions benefit from a 0.2 bonus. The advantage of the free rider strategy depends on how many other agents in the population are following the same strategy. 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 reci-

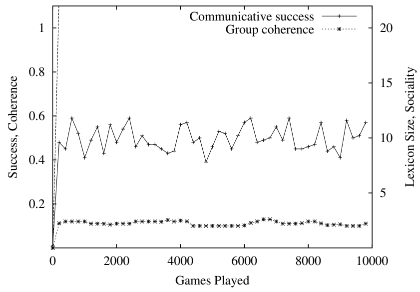

Figure 2. Agents perform at random when R has no strategies for detecting deceit. Lexical agreement under these conditions is not possible.

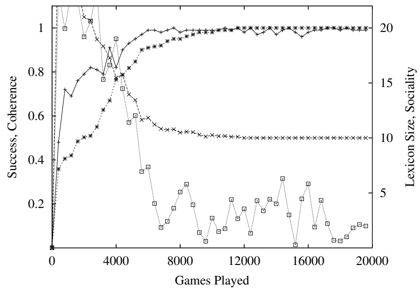

Figure 3. Defection is suppressed when agents have the added capacity of long-term memory. The learning curve compares with that of Fig. 1.

procity is costly. Cooperating with a partner who defects results in the sucker's payoff. Defecting against a partner who cooperates precludes 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 doing this since roles are randomly assigned. This way, cooperation takes place not within each interaction, but over the course of multiple interactions, emerging from tit-for-tat.

The level of information sharing found in human language use suggests that speakers must be motivated to share personal knowledge by some direct payoff (Scott-Phillips, 2006). In the context of the Reciprocal Naming Game, a speaker can be seen to derive utility from the propagation of its own words, because later in the receiver role, this agent will deal better with the social situation when it is able to interpret the linguistic situation. Ostensibly, it would be every agent's goal to avoid coherence with unfair partners if coherence renders an agent vulnerable to deception perpetrated by shared words. But coherence contributes to personal utility when cheaters can be detected, and this supports convergence in the face of deception. Although an opponent might use a word to deceive once, the word cannot be used against the same agent to cheat repeatedly if the meaning of the word is shared, since an agent who has been deceived will choose to disbelieve the message in the next round, if playing by tit-for-tat. Thus in the long run, comprehension of messages elevates receiver performance above chance, and it is in an agent's interest to share the words it knows, and to learn the words spoken by other players. This way, the group lexicon serves as a neutral tool and as a sort of social contract, especially because it would be difficult for a single agent to deviate unilaterally from the agreed naming conventions. In this system, the language remains a constant fixture because the opportunity to brandish it for deceit is no greater than the opportunity to engage it for cooperation.

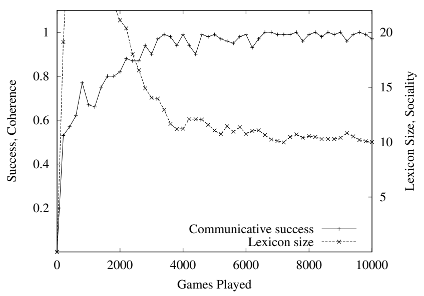

Figure 4. With only one free rider, lexical agreement and stability nearly matches Fig. 1.

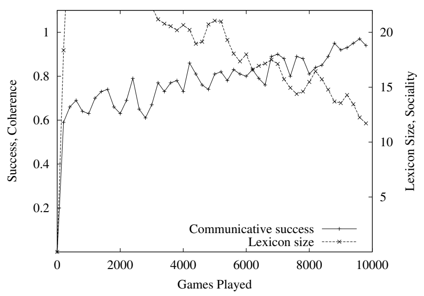

Figure 5. With three free riders, the ability to build agreement becomes greatly diminished.

5. Conclusion

In simulations guided by a model of selfish communication, we experimented by endowing agents with a tit-for-tat policy, as well as some other policies for guiding altruistic behavior. With tit-for-tat, the agents' selfishness did not impede lexical agreement. But without sufficient reciprocation, deception prevented consensus. These simulations show that peer-to-peer negotiation of conventions in language games remains viable in a social environment where deception is prevalent, as long as a socially-informed mechanism governs the agents' choices between co-operation and deception. Bootstrapping a symbolic system of communication can even occur in parallel with the formation of trust relations. This demonstrates that trust need not be permanent or unconditional for communication to develop and remain stable. Rather, reciprocity may serve as a proxy for honesty.

Acknowledgments

This research has been conducted at the AI Laboratory of the Vrije Universiteit Brussel, with funding from FWO project AL328. Emily Wang visited the AI Lab during the 2006–07 academic year on a Fulbright fellowship sponsored by the U.S. Department of State. We would like to thank both Pieter Wellens and Joris Bleys for their insights on Naming Game dynamics.

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*, *50*(6), 1431–1451.

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* (pp. 62–79). Cambridge, UK: Cambridge University Press.

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

Scott-Phillips, T. C. (2006). Why talk? Speaking as selfish behaviour. In *Proceedings of the 6th international conference on the evolution of language* (pp. 299–306).

Steels, L. (1995). A self-organizing spatial vocabulary. *Artificial Life*, *2*(3), 319–332.

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

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