Convergence Analysis for Collective Vocabulary Development *

Jun Wang
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
junwang4@uiuc.edu

Les Gasser
Graduate School of Library and Information Science
University of Illinois at Urbana-Champaign
gasser@uiuc.edu

Jim Houk
Department of Physiology
Northwestern University Medical School
Chicago, IL
j-houk@northwestern.edu

ABSTRACT

We study how decentralized agents can develop a shared vocabulary without global coordination. Answering this question can help us understand the emergence of many communication systems, from bacterial communication to human languages, as well as helping to design algorithms for supporting self-organizing information systems such as social tagging or ad-word systems for the web. We introduce a formal communication model in which senders and receivers can adapt their communicative behaviors through a simple reinforcement learning mechanism that adjusts each agent's *vocabulary:* the ways it associates words with meanings. We analyze the model's dynamics in terms of collective convergence conditions and convergence speed. Our main result on the convergence conditions is that for a given number of meanings, there exists a threshold of the number of words below which the agents can't converge to a shared vocabulary. We also give the time needed for the agents to converge to a fully communicable, shared vocabulary system; specifically, we show that we can lower convergence time by allowing the agents to use more words than necessary in order to jump out of the threshold point. Finally, the effect of reinforcement learning rates of the agents on the convergence are analyzed, showing there exists a range of learning rates for the agents to achieve the best convergence time.

1. INTRODUCTION

We study the question how decentralized agents can develop a shared vocabulary without global coordination. This is a fundamental question that underpins how we understand the emergence of many communication systems arising in nature and in human society, e.g., how do different organisms develop the communication needed to coordinate their activities[4], how do encoding/decoding networks arise in the brain[2], or how did human languages emerge[3]? The study of this question can also help us design algorithms that support self-organizing information systems such as social tagging, web ad-words, or peer-to-peer retrieval systems[1, 14, 16, 19].

*Work supported by NSF Grant IIS-0340996 and UIUC Beckman Summer Fellowship.

Two frameworks, *evolutionary* and *self-organizing*, have been proposed to study the emergence of communication systems. The evolutionary framework assumes that vocabularies of agents are inherited genetically and/or learned culturally from their parents or teachers. The driving force for the emergence of shared vocabularies comes from mechanisms in which agents who can communicate well produce more offspring with similar vocabularies [5, 8, 11, 12, 13]. In contrast, in the self-organizing framework decentralized agents actively develop a shared vocabulary without global knowledge in a short span of time by changing their own representations via distributed learning and positive feedback loops [1, 6, 7, 9, 15, 18]. The work reported here can be seen as an effort in the self-organizing camp.

Most existing work on the development of shared vocabularies has been done using computer simulations and only intuitive interpretations. While this is often suggestive, to achieve a thorough understanding of vocabulary self-organization we need to augment simulation studies with mathematical characterizations of specific conditions and limitations under which the emergent behaviors are stable. Recent work by Baronchelli et al. [1] provides one example; they explain how a sharp transition to a shared vocabulary observed in simulations can occur as the density of certain inter-agent vocabulary hypotheses grows beyond a "tipping point."

In this paper, we focus on a formal communication model in which senders and receivers can adapt their communicative behaviors through a simple reinforcement learning mechanism. Our purpose is to develop a dynamical model for vocabulary development. To simplify the characterization of the dynamics, we study a system that consists of two agents. We give the convergence conditions under which the agents can develop a stable shared vocabulary, and the convergence time that the agents need to develop a fully communicable system. The effects of the learning rates of the agents are also discussed.

2. BASIC MODEL

First, by *shared vocabulary*, we mean a set of stable meaning-word agreements among the agents. We assume that the only way in which the agents convey meanings to other agents is through words. Given a meaning-word pair, there is an agreement among a group of agents, if every agent uses the word to represent the meaning and uses the meaning as interpretation when perceiving the word. At the beginning, there may be no agreement among the agents about how to express meanings, or about the meanings of words. But the agents can update their representation and interpretation functions by learning from the results of the interaction between them, and this collective learning dynamics constitutes the vocabulary development process.

We focus on a communication model that consists of two agents, one of which is called *sender* and the other one *receiver*. In our model, there are m meanings $X = \{x_1, \cdots, x_i, \cdots, x_m\}$, and n words $Y = \{y_1, \cdots, y_j, \cdots, y_n\}$. The sender's communicative behavior is defined by its *send function* s that maps from a meaning to a word, $s : X \mapsto Y$, and the receiver's behavior is defined by its *receive function* r that maps from a word to a meaning, $r : Y \mapsto X$. With this notation, we can give an explicit definition of a stable meaning-word agreement. A meaning-word pair (x, y) is an agreement between the sender and receiver if $s(x) = y$ and $r(y) = x$. An agreement is stable if there exists a time point t_0, and for any time $t > t_0$, we have $s^{(t)}(x) = y$ and $r^{(t)}(y) = x$, where $s^{(t)}$ denotes the send function of the sender at time t, and $r^{(t)}$ the receive function of the receiver at time t.

Now we specify how the sender's send and receiver's receive functions are implemented. The send function is determined by the sender's *association matrix*, which specifies the strength of the association between a meaning and a word, and similarly the receive function is determined by the receiver's association matrix.

Let the association matrix of the sender be $S = (s_{ij})$ which specifies the weight between meaning x_i and word y_j, and the matrix of the receiver be $R = (r_{ji})$ which specifies the weight between word y_j and meaning x_i. Then the sender's send function s, which maps from meaning x_i to word y_j, is given by the following *winner-take-all* selection mechanism

$$s(x_i) = y_j \qquad \text{where} \qquad j = \arg\max_k s_{ik}, \tag{1}$$

and the sender's receive function r, which maps from word y_j to meaning $x_{i'}$, is given by

$$r(y_j) = x_{i'} \qquad \text{where} \qquad i' = \arg\max_k r_{jk}. \tag{2}$$

Then we specify how the send/receive functions change as a result of the update of their association matrices. When the communication on some meaning x_i is successful, i.e., the meaning $x_{i'}$ interpreted by the receiver is the same as the meaning x_i referred by the sender (via some word, say y_j), the two agents update their corresponding association weights according to the following *reward rules* [1]

$$\begin{cases} s_{ij} & \leftarrow \quad s_{ij} \quad +1 \qquad //sender \\ r_{ji'} & \leftarrow \quad r_{ji'} \quad +1 \qquad //receiver. \end{cases} \tag{3}$$

When the communication fails, the two agents update their weights according to the following *punishment rules*

$$\begin{cases} s_{ij} & \leftarrow \quad s_{ij} \quad -1 \qquad //sender \\ r_{ji'} & \leftarrow \quad r_{ji'} \quad -1 \qquad //receiver. \end{cases} \tag{4}$$

These four update rules are very simple reinforcement learning rules: they assume all learning rates are 1. We discuss the effects of the learning rates of the agents in Section 5.

We call one communication event between the sender and receiver a *communication game*, and thus we model the vocabulary development process as repeated play of communication games (Fig. 1). Each game is initiated by the sender, which has a meaning to convey to the receiver. We suppose at each round of game a meaning x_i is randomly chosen for play according to probability $p(x_i)$. For abbreviation, we use p_i to denote $p(x_i)$.

---

1. initialize the sender and receiver's association matrices with random values in $[0, 1]$
2. for each round from 1 to stopping time T, do
3. a random meaning x_i is chosen with probability $p(x_i)$
4. sender produces a word y_j for meaning x_i: $y_j = s(x_i)$
5. receiver interprets the word y_j as meaning $x_{i'}$: $x_{i'} = r(y_j)$
6. if the interpreted meaning $x_{i'}$ is correct[2], i.e., $x_{i'} = x_i$
7. $s_{ij} = s_{ij} + 1$ // *reward rule of the sender*
8. $r_{ji'} = r_{ji'} + 1$ // *reward rule of the receiver*
9. otherwise
10. $s_{ij} = s_{ij} - 1$ // *punishment rule of the sender*
11. $r_{ji'} = r_{ji'} - 1$ // *punishment rule of the receiver*

Figure 1: The basic framework of the repeated play of communication games.

3. CONVERGENCE CONDITIONS

In this section, we first show some striking results obtained by simulations. To understand the results, we introduce an analysis framework to approximate the dynamics specified in the above section, and then based on the framework we investigate two cases of meaning distribution: uniform and Zipf distribution.

3.1 A simulation experiment

We begin by illustrating a striking result obtained in a series of simulations. Fig. 2 shows the result of the communicative performance (or success ratio) of the system in 51 simulations, where the number of meanings m is fixed to 50 and the number of words n varies from 10 to 60. The purpose of

[1]Note that we don't require the association weights to be within the range of $[0, 1]$. A weight could be infinitely large; it could also be a negative number.

[2]To avoid unrealistic and circular assumptions about communicative feedback through direct supervision, we assume language develops in the service of a joint activity. Thus *successful task performance* provides feedback on communicative success (e.g., Steels' mutual pointing task[17], Werner and Dyer's mating task[20]).

the experiment is to see if there is any pattern on how the communicative performance changes as the number of word n varies. Strikingly, there exists a dramatic *phase transition* when n is at a *threshold* n^*. In the case of $m = 50$ as shown in the figure, the threshold is $n^* = 37$. When n is above the threshold, the two agents can develop a communication system with converged communicative performance given by $\min\{\frac{n}{m}, 1\}$; that is, the sender and receiver can converge on $\min\{n, m\}$ pairs of meaning and word. However, when the number of words is below the threshold, it rapidly approaches random guess—the sender and receiver can't develop any effective communication system, or a shared vocabulary.

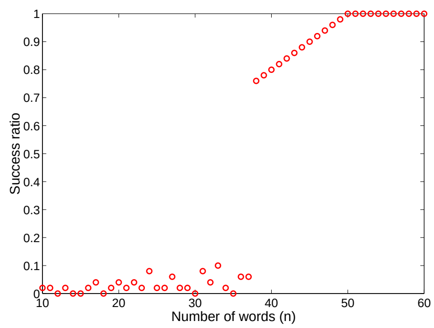

Figure 2: Communicative performance (success ratio) as a function of the number of words when the number of meanings is held constant at 50. A dramatic phase transition occurs when the number of words n is around a threshold $n^* = 37$, above which the agents can develop a communication system with communicative performance given by $\min\{\frac{n}{m}, 1\}$.

3.2 General analysis

In this subsection we give a general analysis framework to understand why the above phase transition phenomenon happens. The basic idea is that if there are enough available words a temporary (or lucky) meaning-word agreement between the sender and receiver will have a greater chance to be reinforced by the reward rule than to be weakened by the punishment rule. Once a temporary agreement has a better chance to be reinforced than to be weakened, random walk theory[10] can tell us that the temporary agreement will have a positive probability of becoming a stable agreement. Therefore, our task is to find, given a fixed number of meanings, how many words are sufficient for a temporary agreement to be reinforced and eventually to become a stable agreement.

To facilitate our arguments and analysis, we give the following informal notations, which will be defined formally later.

Reward probability. The probability that a temporary agreement is reinforced by some reward rule.

Punishment probability. The probability that a temporary agreement is weakened by some punishment rule.

Without loss of generality, let the first temporary agreement be on the meaning-word pair (x_1, y_1). This means for the sender $y_1 = s(x_1)$, and for the receiver $x_1 = r(y_1)$. In terms of association matrix, for the sender $s_{11} = \max_k s_{1k}$, and for the receiver $r_{11} = \max_k r_{1k}$ (see Eqs. (2) and (1) for reference). That the agreement on (x_1, y_1) is reinforced means both s_{11} and r_{11} will increase, and that the agreement is weakened means at least one of s_{11} and r_{11} will decrease.

If at the next round what is played is again the meaning x_1, then the communication on meaning x_1 will succeed, so the reward rule will be applied to both s_{11} and r_{11}. Since the probability of playing meaning x_1 is p_1, the *reward probability* of the agreement will be p_1.

If at the next round what is played is not the meaning x_1, let it be x_2 for example, then there is a chance that the word produced by the sender to represent the meaning x_2 happens to be the word y_1, and this chance can be written as $p(y_1 = s(x_2))$. If this doesn't happen, then there is neither reward nor punishment, so it is of no interest. If this happens, the communication will fail, since the interpreted meaning by the receiver for y_1 must be x_1, according to the winner-take-all selection mechanism. And then, the association weight r_{11} between word y_1 and meaning x_1 in the receiver's matrix will decrease by 1, according to the punishment rule of the receiver. Similarly, at the same time, the weight s_{21} between meaning x_2 and word y_1 in the sender's matrix will decrease by 1. This means the punishment rule of the sender has no effect on the weight s_{11}. Therefore, the *punishment probability* can be represented as $(1 - p_1)p(y_1 = s(x_2))$, where $(1 - p_1)$ is the probability of playing a meaning other than x_1.

Now we want to compute $p(y_1 = s(x_2))$, the probability that the word y_1 is the one used by the sender for the meaning x_2. Considering that in our model words are produced by winner-take-all mechanism, a nonlinear function, it is hard to give a precise estimation. So, we make the following assumption.

Assumption 1. (*Random word production assumption.*) The word produced by the sender for a meaning that is not in any (temporary or stable) agreement is assumed to be randomly generated.

With the assumption, we can suppose that the probability of the word y_1 being the one produced by the sender for meaning x_2, $p(y_1 = s(x_2))$, can be approximated by $\frac{1}{n}$, provided that there are totally n words. (We will show by simulations that this is a good approximation, or the assumption is a good one.)

Then, our task is to solve the following inequality, where p_1 is the reward probability, and $(1 - p_1)\frac{1}{n}$ is the punishment probability:

$$p_1 > (1 - p_1)\frac{1}{n}. \qquad (5)$$

Eq.(5) is just the condition for the agents to develop the first stable agreement. (Actually it is very likely that more than one agreements can be established parallel). Now we need to find the conditions for the other agreements to be established. Before we go further, we make the following assumption about the order of developing agreements.

Assumption 2. (*Sequential agreement establishment assumption.*) Agreements are established in the sequential order of the frequency that their meanings are used. If the probabilities of the meanings have the relationship: $p_1 \geq p_2 \geq \cdots$, then we assume that the agreements will be established in an order $(x_1, y_{l_1}), (x_2, y_{l_2}), \cdots$.

Now suppose we already have l established stable agreements that satisfy the above sequential assumption. Then we need to find the condition for a new agreement, the $l+1^{th}$ agreement, to be established. Note that $l = 0$ is the case we have just discussed above. Let the probabilities of those meanings associated with the already established agreements be $p_1 \geq \cdots \geq p_l$. By applying the similar arguments as above, for the $l+1^{th}$ temporary agreement to become stable, we require the following inequality to hold

$$p_{l+1} > (1 - \sum_{k=1}^{l+1} p_k)\frac{1}{n}, \qquad (6)$$

where the term $\sum_{k=1}^{l+1} p_k$ in the inequality comes from the fact that these $l+1$ meanings can guarantee their sender won't generate a word that would cause the temporary agreement to be weakened.

In general, for the agents to establish L agreements, we must require the following L inequalities to hold

$$p_l > (1 - \sum_{k=1}^{l} p_k)\frac{1}{n}, \qquad (l = 1, 2, \cdots, L),$$

which are equivalent to

$$n > (1 - \sum_{k=1}^{l-1} p_k)/p_l - 1, \qquad (l = 1, 2, \cdots, L). \qquad (7)$$

When the agents can develop L stable agreements on meanings $x_1, \cdots, x_L$, the communicative performance (or success ratio) of the system will be

$$\phi(L) = \sum_{k=1}^{L} p_k.$$

3.3 When meanings are uniformly distributed

Above we have shown that, for the agents to establish stable communication on L meanings $x_1, \cdots, x_L$, the inequalities in Eq.(7) should be satisfied for $l = 1, 2, \cdots, L$. Now we are ready to solve the inequalities to obtain the convergence conditions for some specific distributions over meanings. We consider two common distributions: uniform and Zipf distribution. We discuss the case of Zipf distribution in the next subsection.

When the meanings are uniformly distributed, we have $p_i = \frac{1}{m}$ for $i = 1, ..., m$. For the agents to establish L agreements, by substituting $p_l = \frac{1}{m}$ into Eq.(7), we have

$$n > m - l, \qquad (l = 1, 2, \cdots, L)$$

So, if $n > m - 1$, then for all $l \geq 1$, the above L inequalities will hold. In other words, if we let $n^* = m - 1$, then n^* will be the number of words threshold, above which the agents can develop a communication system.

Clearly, the two agents can develop as more as $\min\{m, n\}$ stable agreements, since there are at most $\min\{m, n\}$ distinct pairs of meaning-word. And thus, the maximum converged communicative performance is $\min\{m, n\}\frac{1}{m}$ or equivalently $\min\{\frac{n}{m}, 1\}$.

Fig. 3 illustrates the simulation results on how the number of words threshold changes as the number of meanings varies from 10 to 100, showing that there exists a linear relationship between the number of words threshold n^* and the number of meanings m. The linear relationship is given by

$$n^* = 0.82m - 3.5.$$

Comparing to the analytical number of words threshold function $n^* = m - 1$, we can say that the two assumptions, the *random word production assumption* and the *sequential agreement establishment assumption*, are good enough to characterize our adaptive communication model.

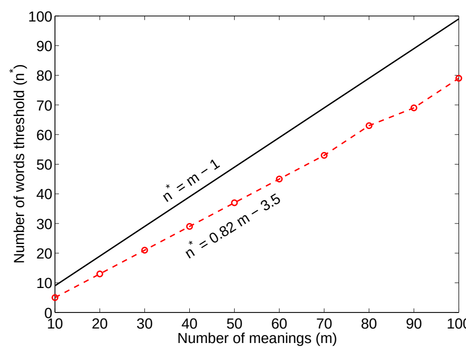

Figure 3: There exists a linear relationship between the number of words threshold and the number of meanings. The dashed line is obtained by fitting simulation results (circle marks), and the smooth line is obtained by analysis.

3.4 When meanings are Zipf distributed

Now we move to another distribution called *Zipf* distribution, a common distribution observed in nature and in human activities[21]. Without loss of generality, we can assume that the probabilities of the m meanings have the relationship: $p_1 \geq p_2 \geq \cdots \geq p_m$. Then under Zipf distribution, the probability of the meaning x_k is

$$p_k = Ck^{-\alpha}, \qquad (k = 1, 2, \cdots, m),$$

where $\alpha \geq 0$ and C is given by $1/\sum_{k=1}^{m} k^{-\alpha}$, to satisfy the property of probability distribution. Note that when $\alpha = 0$, the Zipf distribution degenerates to uniform distribution, i.e., $p_k = 1/m$.

The substitution of $p_k = Ck^{-\alpha}$ into Eq.(7) yields

$$n > \frac{1 - \sum_{k=1}^{l-1} k^{-\alpha} / \sum_{k=1}^{m} k^{-\alpha}}{l^{-\alpha} / \sum_{k=1}^{m} k^{-\alpha}} - 1, \qquad (l = 1, 2, \cdots, L).$$

This time we are not that lucky as in the case of uniform distribution, where the condition, $n > m - 1$, can cover all the L conditions, $n > m - l$ $(l = 1, \cdots, L)$, and thus we only have one threshold: $n^* = m - 1$. Here, for each different value of l between 1 and L, we may need a different threshold. Since the threshold (the right side of the above inequality) depends on l, we denote it by $\eta(l)$, and call it the threshold function of l.

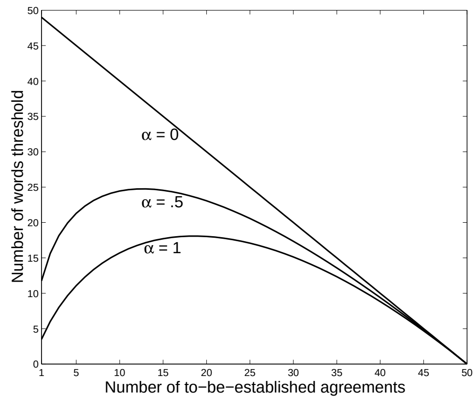

Figure 4: Illustration of the number of words threshold function for three settings of the Zipf distribution parameter: $\alpha = 0$, 0.5 and 1. (The number of meanings m is fixed to 50).

Before we move ahead, let us take a look at the behavior of the "complicated" threshold function $\eta(l)$. Fig. 4 gives an illustration of how the threshold function $\eta(l)$ behaves as the number of to-be-established agreements, l, changes from 1 to 50, while the number of meanings m is held constant at 50. We can see when $\alpha = 0$, it degenerates to the case of uniform distribution. In this case, if the agents can establish the first agreement (i.e., $l = 1$) whose threshold is 49, then there is no problem for them to establish the other agreements whose thresholds are less than 49. However, when $\alpha > 0$, the maximum of the threshold function value is not located at $l = 1$; instead, it falls on somewhere between 1 and m. For example, when $\alpha = 1$, the maximum value 18 of the threshold function is obtained when l is around 18. [3]

What can we tell from the hump shape obtained when $\alpha > 0$? Again, take $\alpha = 1$ as an example. When the number of available words n is less than $\eta(l)_{l=1} \approx 4$, no agreement can be established. When n is larger than $\eta(l)_{l=18} \approx 18$, then n stable agreements[4] can be established. When n is between 4 and 18, the agents can only develop $\eta^{-1}(n)$ stable agreements, which are less than n agreements. This is because, to establish l agreements, they need at least $\eta(l)$ terms; so if there are only $\eta(l) = n$ terms, they can at most develop $l = \eta^{-1}(n)$ agreements. For instances, when $n = 10$, $\eta^{-1}(10) \approx 4$.

The above analysis can be supported by simulation results shown in Fig. 5. We can see there are three phases. In phase 1 ($n < 4$), the agents can't develop any agreements. In phase 2 ($4 \leq n \leq 18$), the agents can develop $\eta^{-1}(n)$ agreements, which are less than n agreements. In phase 3 ($n > 18$), the agents can develop n agreements and achieve the best performance. If there were no threshold, the agents would develop n agreements and achieve the best performance in all phases. In the figure, the green smooth (best performance) line is obtained by computing $\phi(n) = \sum_{k=1}^{n} p_k$ under the premise that there is no threshold; the black, partially staired (analytical performance) line is obtained by computing $\phi(n) = \sum_{k=1}^{\eta^{-1}(n)} p_k$; the red circle sequence is obtained by simulations.

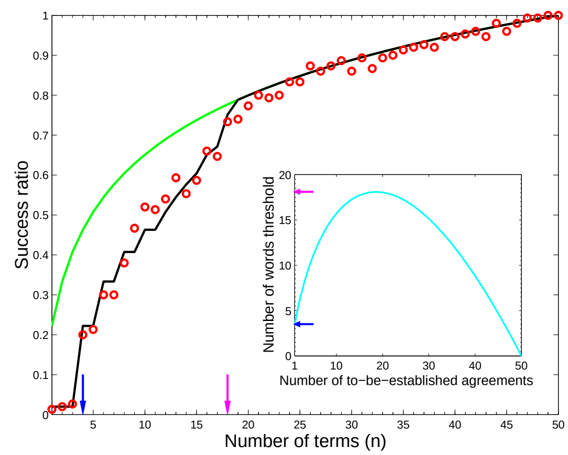

Figure 5: The performance (success ratio) of the system shows different behaviors in three phases. The separation of the three phases is marked by two vertical arrows that fall on $n = 4$ and $n = 18$, which are also indicated in the inset figure by two horizontal arrows. The parameter settings are: $\alpha = 1$ and $m = 50$.

4. CONVERGENCE TIME

We have given the convergence conditions for the agents to develop a communication system. Now we ask the question: how much time will the agents need to converge given that the convergence conditions are satisfied. We focus only on the case of uniform meaning distribution. We first present an analytic solution for computing the convergence time, then we do some simulations to verify the analytic result.

Recall that in the section of convergence conditions analysis, we have shown that for a temporary agreement on meaning x_l and word y_l to become a stable one, its reward probability, p_l, should be larger than its punishment probability, $(1 - \sum_{k=1}^{l} p_l)/n$. We denote by p the reward probability, and by q the punishment probability. Then at each round of play, the association weight r_{ll}, the weight between meaning x_l and word y_l in the receiver's association matrix[5], will increase by 1 with probability p and decrease by 1 with probability q.

The problem of computing convergence time can be rendered to the hitting time problem in Markov chains or random walk theory: How many rounds do the agents need to play in order for the weight r_{ll} to reach a certain number that is necessary for a temporary agreement to become stable. We denote by N such a number.

[3]An interesting observation is that the maximum value $\hat{n}$ of the threshold function and its corresponding $\hat{l}$ has the relation: $\hat{l} = \alpha\hat{n}$.

[4]Suppose n is equal to or less than m.

[5]We have shown in Section 3.2 that what matters for the convergence is the weight r_{ll} of the receiver.

Random walk theory tells us that the average hitting time for r_{ll} to reach N is

$$T(l) = \frac{N}{p-q}. \quad (8)$$

We know for uniform distribution $p = \frac{1}{m}$ and $q = (1 - \frac{l}{m})\frac{1}{n}$. Substituting p and q into Eq.(8), we have

$$T(l) = \frac{Nm}{1-(m-l)/n}.$$

We are concerned about how much time needed by the agents to develop a fully communicable system. A communication system is *fully communicable* if every meaning sent by the sender can be correctly interpreted by the receiver. In other words, for m meanings, the agents need to develop m stable agreements. (Clearly, for a system to be fully communicable, we must have $n \geq m$.) The total time for m stable agreements to be established will be

$$T = \sum_{l=1}^{m} T(l) = Nm \sum_{l=1}^{m} \frac{1}{1-(m-l)/n}.$$

Now we need to determine the value of N, the hitting number that guarantees a temporary agreement to become stable. Our current approach is to estimate N by simulations. We find when $m = n$, by taking $N = 3$, it is good enough to approximate the convergence time for success ratio to achieve 90%. [6] Therefore, our estimation of the time that the agents need to develop a communication system of 90% success ratio, is given by (which is well supported by simulation results as shown in Fig. 6)

$$T = 3m \sum_{l=1}^{m} \frac{1}{1-(m-l)/n}. \quad (9)$$

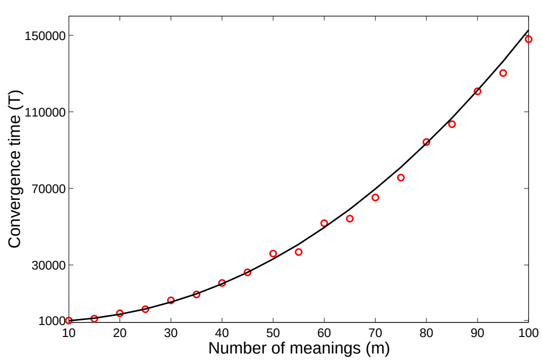

Figure 6: Convergence time as a function of the number of meanings m (given that the number of meanings is equal to that of words). It shows that the convergence time increases at the rate of m^2. Simulation results are shown as circle marks.

[6]In simulations, on average it will take much longer time to reach a 100% success ratio, and correspondingly N will take a much larger number. We leave for future research the question of why N is around 3 for the success ratio to reach 90%, or generally what value N should take in order for the success ratio to reach a given number.

We are also interested in asking the question: Can we reduce the convergence time by increasing the number of words while the number of meanings is fixed? To answer this question, we define a parameter $\rho = n/m$, called *word-meaning ratio*, to measure the ratio of the number of words to the number of meanings. Substituting $n = \rho m$ into Eq.(9), we have

$$T = 3m \sum_{l=1}^{m} \frac{1}{1-(1-l/m)/\rho}$$

When $\rho \to \infty$, we see $T = 3m^2$, which suggests that the convergence time increases at the rate of m^2. But when ρ approaches the threshold point $1 - 1/m$ (i.e., $n = m - 1$), T becomes infinitely large. So, we can gain a lot on convergence time by increasing the number of words around the threshold point. Fig. 7 shows when ρ is small, we do gain a lot, and when ρ is large, the convergence time will converge to $3m^2 = 7500$ for $m = 50$.

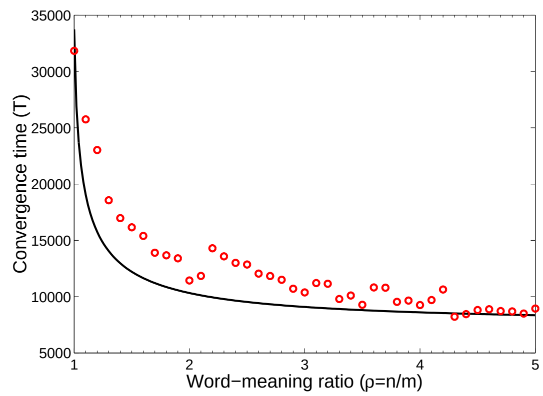

Figure 7: Convergence time as a function of word-meaning ratio (given that the number of meanings is fixed). The convergence time obtained by simulations (shown as circle marks) is generally larger than the time obtained by analysis (shown as the smooth line). This is because that in our analysis we didn't take into consideration the time needed for the agents to form a (lucky) temporary agreement, and the more words the more this time.

5. EFFECTS OF LEARNING RATE

In our model, after the play of a communication game, the agents update their association weights by reinforcement learning. We have two agents, the sender and the receiver, and each agent has two update rules: reward rule and punishment rule. Together there are four update rules that involve four learning rates: $(\gamma_a^+, \gamma_b^+, \gamma_a^-, \gamma_b^-)$, which are shown in the following. In the previous analysis that is based on the model given in Section 2, all the four learning rates are set to 1. Here, we want to study how other settings of these learning rates affect the convergence conditions and time.

$$\begin{cases} s_{ij} & \leftarrow & s_{ij} + \gamma_a^+ & \textit{// sender's reward rule} \\ r_{ji'} & \leftarrow & r_{ji'} + \gamma_b^+ & \textit{// receiver's reward rule} \\ s_{ij} & \leftarrow & s_{ij} - \gamma_a^- & \textit{// sender's punishment rule} \\ r_{ji'} & \leftarrow & r_{ji'} - \gamma_b^- & \textit{// receiver's punishment rule .} \end{cases}$$

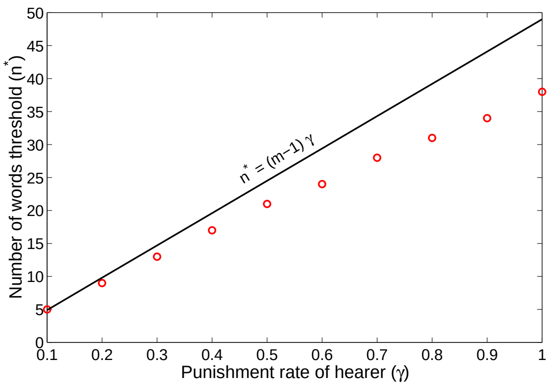

Figure 8: The effect of the punishment learning rate of the receiver on the convergence condition. The small the learning rate, the small the threshold, and so it is easier for the agents to converge to a shared vocabulary. Simulation results are shown as a sequence of circle marks. (The number of meanings is fixed as $m = 50$.)

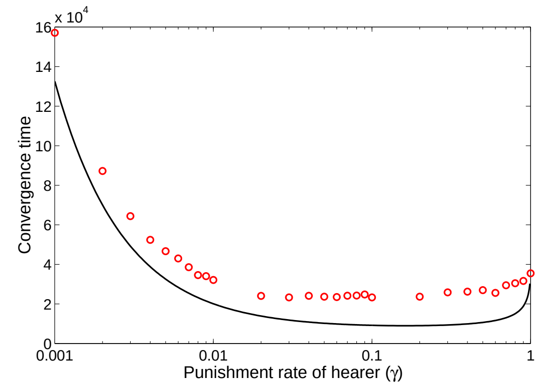

Figure 9: The effect of the punishment rate of the receiver on the convergence time. It shows there exists a range of learning rates for the agents to achieve the best convergence time.

During the process of convergence condition analysis, we have found two learning rates play important roles for the convergence. They are the receiver's reward and punishment rates, γ_b^+ and γ_b^-. Specifically, what is important is the relation between the two rates; we will see this later in Eq.(10). So we only need to consider one learning rate by fixing the other one. We will fix the receiver's reward rate $\gamma_b^+ = 1$, and study how the convergence changes as the receiver's punishment rate γ_b^- varies. For convenience, we denote γ_b^- by γ, since it is now the only learning rate that concerns us.

Again, we will focus on the case of uniform meaning distribution. First we study the convergence conditions. It is easy to see that the new convergence condition will be (we only need to consider the inequality condition in terms of p_1, the probability of the first to-be-established meaning, for the case of uniform distribution)

$$p_1\gamma_b^+ > \frac{1-p_1}{n}\gamma_b^- \qquad (10)$$

which can be rewritten as, since $\gamma_b^+ = 1$ and $\gamma = \gamma_b^-$

$$p_1 > \frac{1-p_1}{n}\gamma.$$

The substitution of $p_1 = 1/m$ yields the following new convergence condition that depends greatly on the receiver's learning rates (see Fig. 8 for support from simulations)

$$n > n^* = (m-1)\gamma.$$

Now we move to investigating how the receiver's punishment rate γ affects convergence time. First let us consider two extreme cases of the punishment rate. At the one end, the punishment rate could be very big such that the convergence condition can not be satisfied, resulting in infinite convergence time. At the other end, if the punishment rate is zero, the receiver's receive function will never change (even when its reward rate is positive). If two words are mapped by the receiver to a same meaning, they will always be mapped to the same meaning, therefore the agents can never develop a fully communicable system.

The above brief analysis tells us there must exist an optimal setting of the punishment rate. Fig. 9 shows that, when there are 50 meanings and 50 words, the convergence time is the best if punishment rate is around 0.1. In the figure, the circle marks present the simulation results, and the smooth line shows an analytical estimation of the convergence time that takes into consideration the dual role of the receiver's punishment rate, which is given in the following equation

$$T = m(N + 0.05/\gamma)\sum_{l=1}^{m}\frac{1}{1-(1-l/m)\gamma}. \qquad (11)$$

In the equation, the term $0.05/\gamma$, where the number 0.05 is obtained empirically, indicates that as the punishment rate gets smaller, the convergence time will get longer, and the term $(1-(1-1/m)\gamma)$ indicates the small punishment rate can improve the convergence condition and so can help to speed up the convergence time.

6. CONCLUSION

We have presented an adaptive communication model in which a sender and a receiver can adapt their communicative behaviors through a simple reinforcement learning mechanism. Through mathematical analysis and computer simulations, we find that for a given number of meanings, there exists a threshold for the number of words below which the agents can't converge to a shared vocabulary. We also have shown that we can gain a lot on the convergence time by allowing the agents more words to use than necessary in order to jump out of the threshold point. Finally, we find that the learning rates of the receiver agent can greatly affect the convergence condition; in addition, there exists a range of learning rates for the agents to achieve the best convergence time.

Though our analysis in the current work is focused on the communication model of two agents, we believe it can be extended to the case of multiple agents—one sender and many receivers, many senders and one receiver, or many senders and many receivers. Of other interesting questions for future study, we would like to explore what will happen if the word production or the meaning interpretation function is implemented by probability-based selection, rather than the currently used winner-take-all selection mechanism.

7. REFERENCES

[1] A. Baronchelli, M. Felici, E. Caglioti, V. Loreto, and L. Steels. Sharp transition towards shared vocabularies in multi-agent systems. http://arxiv.org/abs/physics/0509075, 2005.

[2] D. G. Beiser and J. C. Houk. Model of cortical-basal ganglionic processing: encoding the serial order of sensory events. *J. Neurophysiol*, 79:3168–3188, 1998.

[3] M. H. Christiansen and S. Kirby, editors. *Language Evolution: The States of the Art*. Oxford University Press, 2003.

[4] E. P. Greenberg. Bacterial communication: Tiny teamwork. *Nature*, 424:134, 2003.

[5] J. Hurford. Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, 77(2):187–222, 1989.

[6] E. Hutchins and B. Hazlehurst. How to invent a lexicon: the development of shared symbols in interaction. In G. N. Gilbert and R. Conte, editors, *Artificial Societies: The computer simulation of social life*. UCL Press, London, 1995.

[7] F. Kaplan. Semiotic schemata: Selection units for linguistic cultural evolution. In M. Bedau, J. McCaskill, N. Packard, and S. Rasmussen, editors, *Artificial Life VII*. The MIT Press, 2000.

[8] N. L. Komarova and M. A. Nowak. The evolutionary dynamics of the lexical matrix. *Bull. Math. Biol.*, 63(3):451–485, 2001.

[9] T. Lenaerts, B. Jansen, K. Tuyls, and B. de Vylder. The evolutionary language game: An orthogonal approach. *Journal of Theoretical Biology*, 235(4):566–582, August 2005.

[10] J. R. Norris. *Markov Chains*. Cambridge University Press, 1997.

[11] M. A. Nowak, J. B. Plotkin, and D. C. Krakauer. The evolutionary language game. *Journal of Theoretical Biology*, 200:147–162, 1999.

[12] M. Oliphant and J. Batali. Learning and the emergence of coordinated communication. *The newsletter of the Center for Research in Language*, 11(1), 1997.

[13] K. Smith. The evolution of vocabulary. *Journal of Theoretical Biology*, 228(1):127–142, 2004.

[14] S. Staab. Emergent semantics. *IEEE Intelligent Systems*, 17(1):78–86, 2002.

[15] L. Steels. Self-organizing vocabularies. In C. Langton and T. Shimohara, editors, *Artificial Life V*, Nara, Japan, 1996.

[16] L. Steels. The origins of ontologies and communication conventions in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, 1(2):169–194, October 1998.

[17] L. Steels. The evolution of communication systems by adaptive agents. In D. K. Alonso, E. and D. Kazakov, editors, *Adaptive Agents and Multi-Agent Systems: Adaptation and Multi-Agent Learning. LNAI 2636*, pages 125–140. Springer Verlag, Berlin, 2003.

[18] P. Vogt and H. Coumans. Investigating social interaction strategies for bootstrapping lexicon development. *Journal of Artificial Societies and Social Simulation*, 6(1), 1 2003.

[19] J. Wang and L. Gasser. A mutual adaptation model for information retrieval. submitted, 2005.

[20] G. Werner and M. Dyer. Evolution of communication in artificial organisms. In C. Langton, C. Taylor, D. Farmer, and S. Rasmussen, editors, *Artificial Life II*, pages 659–687, Redwood City, CA, 1992. Addison-Wesley Pub.

[21] G. K. Zipf. *Human Behaviour and the Principle of Least Effort: An Introduction to Human Ecology*. Addison-Wesley, Cambridge, MA, 1949.