

# A Distributed Learning Algorithm for Communication Development

Edwin D. de Jong\*

Luc Steels†

*Vrije Universiteit Brussel,  
Artificial Intelligence Laboratory,  
Pleinlaan 2, B-1050 Brussels, Belgium*

---

We study the question of how a local learning algorithm, executed by multiple distributed agents, can lead to a global system of communication. First, the notion of a perfect communication system is defined. Next, two measures of communication system quality are specified. It is shown that maximization of these measures leads to perfect communication production. Based on this principle, local adaptation rules for communication development are constructed. The resulting stochastic algorithm is validated in computational experiments. Empirical analysis indicates that a mild degree of stochasticity is instrumental in reaching states that correspond to accurate communication.

---

## 1. Introduction

---

Recently, the problem of how individual agents may adapt their behavior in order to maximize a global utility function has received substantial attention [1–10]. We study this problem in the particular context of communication development. The question that will be considered is: How can a distributed group of agents arrive at a shared system of communication, mapping internal meanings of agents to public words or signals in a consistent manner [11]?

Within different communities, a number of approaches to the problem of communication development have been explored. One such approach is to use evolutionary methods [12–16]. Such methods make random, and possibly large, changes to the behavior of agents and use selection to drive the population towards a shared convention. Here, in contrast, we are interested in learning algorithms where each individual agent makes small directed changes to its behavior. A substantial amount of work concerns supervised learning. There, learners are instructed by teachers that specify the desired communicative be-

---

\*Current address: Utrecht University, Decision Support Systems Group, Utrecht, The Netherlands. Electronic mail address: dejong@cs.uu.nl.

†Electronic mail address: steels@arti.vub.ac.be.

havior [17–19]. Thus, the learners themselves have no influence on the resulting communication system. The question of how a communication system can be formed by means of distributed learning has received relatively little attention. One example is work by Oliphant [20], building on earlier work by Hurford [21]. There, the behavior of an agent is adapted in a single step to correspond to its observations of the production and, for some methods, interpretation behavior of the population. Here, the algorithm that will be developed employs the observed production behavior of other agents, and operates by making small incremental adaptations. Methods for communication development that rely on the interaction between distributed entities, rather than receiving control information from outside the system, are based on self-organization [22–24]. Game-theoretic approaches [25–28] can provide valuable information about what global communication systems are stable, but do not typically provide algorithms for individual agents to arrive at such systems.

An interesting approach to optimizing a global utility function by means of distributed learning is to let agents optimize a local reward function that has the desired global effect [10]. Methods from reinforcement learning can then be used to determine how individual agents should adapt their own local behavior. While this approach can in principle be applied to communication problems, we take a different approach here by employing specific knowledge of the problem of communication development. Based on criteria for communication system quality, we will directly derive local adaptation rules for agents, rather than specifying a reward function to be optimized.

First, the notion of a perfect communication system will be defined. Based on this definition, two measures for communication system quality can be specified. It is then shown that maximization of these criteria is sufficient to guarantee a perfect communication system. Next, these measures are used to develop an algorithm. The algorithm is stochastic, and consistently leads to accurate communication systems in computational experiments. To study to what extent the use of stochasticity in the algorithm contributes to its performance, the relation between the degree of stochasticity and the resulting performance is investigated.

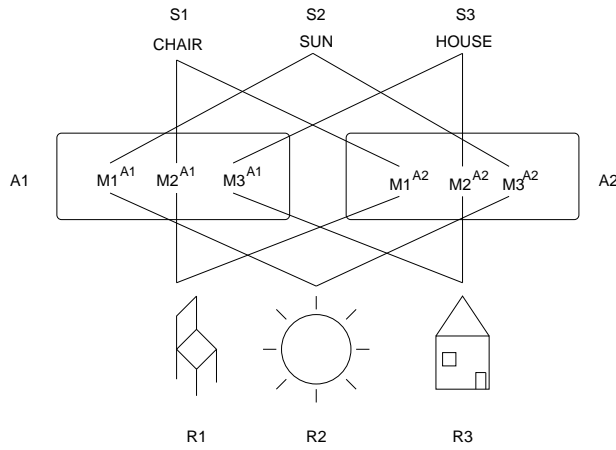
The model of communication that will be used consists of *words*, public symbols or signals that are exchanged during communication; *referents*, entities in the environment to which the words used in communication refer; and *meanings*, entities internal to agents that represent referents. It is assumed that each agent has formed a conceptual system, that is, a system relating referents to meanings. Thus, whenever a referent is present in the environment, this leads to the activation of a unique corresponding meaning in each agent. The agents need to form a mapping between their internal meanings and words

they can use in communication. The formation of the conceptual system is outside the scope of this article. However, the algorithm for communication development that will be presented here has also been studied in combination with concept formation, with successful results [29]. While that work grouped collections of similar feature vectors into concepts that were then associated with words, words can also be directly related to feature vectors [30, 31]. Since messages consist of single words, they are unstructured, and the use of compositional messages, ultimately leading to grammar, will not be studied here; see [32–36] for recent approaches to this topic. Finally, since the aim is to maximize global utility, agents have a shared goal and are therefore cooperative. When this is not the case, the issue of honesty in signaling must be considered [13, 16, 37].

The structure of the article is as follows. Section 2 describes the model of communication that is employed. In section 3, measures for the quality of conceptual systems and communication systems are introduced. In section 4, the algorithm is developed based on these measures, and experimental results are presented. In section 5 it is shown that a stochastic version has a regime of moderate exploration in which accurate communication is consistently achieved. Finally, section 6 presents conclusions. The algorithm is available in Matlab<sup>®</sup> format (see appendix).

## 2. Model

We assume the presence of an environment  $E$  that contains a set of agents  $\mathcal{A} = \{A_1, A_2, \dots, A_{n_a}\}$ . The environment contains a set of referents  $\mathcal{R} = \{\rho_1, \rho_2, \dots, \rho_{n_r}\}$  that can be the subject of communication. For example, the environment can be in different states or situations at different points in time, or it may contain different objects (see Figure 1) that agents are to refer to by means of communication. Each referent creates a corresponding sensorial impression, that is, it generates or activates some internal representation of the referent. It will be assumed that referents are recognized on multiple occasions of observing them, so that each referent is reliably associated with its corresponding internal representation. The latter will be called the agent's *meanings* or *concepts*  $\mathcal{M}_{A_i} = \{\mu_1^{A_i}, \mu_2^{A_i}, \dots, \mu_{n_m}^{A_i}\}$ . These meanings are assumed to be the result of some generalization or concept formation process. Since meanings are internal to each agent, they cannot be exchanged directly. Rather, to communicate about a referent, the corresponding meaning must be associated with an observable signal or word that can be uttered. Words are chosen from some fixed, possibly large set  $\Sigma = \{\sigma_1, \sigma_2, \dots, \sigma_{n_w}\}$ . Thus, the purpose of communication development is for agents to form mappings between the available words and their private meanings, such that each word becomes connected to



**Figure 1.** Illustration of the model for communication development. The figure shows three physical objects that function as referents (R1..R3). Both agents A1 and A2 have a conceptual system that maps these external referents to internal meanings (M1..M3 for each agent). The aim is to form a mapping between meanings and words (S1..S3) so that information about objects or situations in the environment can be shared.

meanings that correspond to the same referent. A communication system for which this mapping between referents and words is one-to-one will be called a *perfect* communication system.

Regardless of how the mapping between referents and meanings and the mapping between meanings and words are represented internally, the communicative behavior of an agent can be studied by considering the resulting production and interpretation behavior. For each combination of word  $\sigma_i$  and referent  $\rho_j$ , the production matrix  $\mathbf{P}_{\text{prod}}^{A_i}(\sigma|\rho)$  specifies the probability that to communicate about referent  $\rho_j$ , the agent would use word  $\sigma_i$ . Conversely, the probability that referent  $\rho_i$  was observed given that an agent produces word  $\sigma_j$  is given by  $\mathbf{P}_{\text{prod}}^{A_i}(\rho|\sigma)$ . Concerning interpretation, the probability that an agent  $A_i$  will interpret the signal  $\sigma_j$  as representing referent  $\rho_k$  equals the *interpretation* probability  $\mathbf{P}_{\text{int}}^{A_i}(\rho|\sigma)$ . Any agent-specific probability can be averaged over all agents to yield the average population behavior, for example,

$$\mathbf{P}_{\text{prod}}(\sigma|\rho) = \frac{1}{n_a} \sum_{A_i \in \mathcal{A}} \mathbf{P}_{\text{prod}}^{A_i}(\sigma|\rho).$$

Throughout this article, variables denoting matrices are set in bold. The columns of the conditional probability matrix  $\mathbf{P}(x|y)$  represent values of the dependent variable  $x$ . Variables denoting specific probabilities rather than matrices will be set in plain. For example, the

probability that an agent will produce a particular word  $\sigma$  to represent a particular referent  $\rho$  is written as  $P_{\text{prod}}(\sigma|\rho)$ , while the matrix containing all such probabilities is written  $\mathbf{P}_{\text{prod}}(\sigma|\rho)$ .

Using the production and interpretation probability matrices, we can now define a perfect system of communication as a combination of perfect production and interpretation. Production is perfect when there is a one-to-one mapping between referents and words in production.

**Definition 1 (Perfect communication production)** *Communication production, characterized by the average probabilistic mappings  $\mathbf{P}_{\text{prod}}(\sigma|\rho)$  and  $\mathbf{P}_{\text{prod}}(\rho|\sigma)$  between  $n_r$  referents and the  $n_s$  signals, is perfect if and only if:*

$$\forall \rho : \exists \sigma : P_{\text{prod}}(\sigma|\rho) = 1$$

and

$$\forall \sigma : \exists \rho : P_{\text{prod}}(\rho|\sigma) = 1.$$

The communication matrix  $\mathbf{P}_{\text{comm}}(\rho|\rho)$  represents the effect of encoding a referent into a word (production) and subsequently decoding this word back into a referent (interpretation). Communication interpretation is perfect if this process always yields the original referent, in which case the communication matrix equals the identity matrix.

**Definition 2 (Perfect communication interpretation)** *Given communication production behavior  $\mathbf{P}_{\text{prod}}(\sigma|\rho)$ , communication interpretation, characterized by the average probabilistic mapping  $\mathbf{P}_{\text{int}}(\rho|\sigma)$  between  $n_r$  referents and  $n_s$  signals is perfect if and only if:*

$$\mathbf{P}_{\text{comm}}(\rho|\rho) = \mathbf{P}_{\text{prod}}(\sigma|\rho) \cdot \mathbf{P}_{\text{int}}(\rho|\sigma) = \mathbf{I}.$$

Using these definitions, we can define a perfect system of communication as follows.

**Definition 3 (Perfect system of communication)** *A system of communication is perfect if and only if both communication production and communication interpretation are perfect.*

### 3. Measures for the evaluation of conceptual and communication systems

In this section, three measures for the quality of communication systems are introduced. Specificity and consistency evaluate the production behavior of a population of agents. *Specificity* measures the degree to which a word specifically identifies a single referent. *Consistency* measures to what degree an agent consistently uses the same word

for a referent. We are furthermore interested in whether communication is also interpreted correctly. This is expressed by the measure of *fidelity*, which gives an overall measure for the quality of a communication system. The fidelity of communication is the average probability that a referent, when encoded into a word by one agent and decoded back into a referent by another, yields the same referent. All measures take values between zero and one, where zero indicates low quality and one high quality. It will be shown that the measures are correct and complete, in that maximal values for the measures imply a perfect communication system and *vice versa*.

### ■ 3.1 Specificity

The specificity of a word is the degree to which it identifies a single referent. This can be measured by computing the relative decrease of uncertainty in determining the referent given a word that was produced, where uncertainty is defined by the information-theoretic concept of entropy. The definitions are based on the average population production behavior. They can be calculated for individual agents by using the individual production probabilities:

$$H(\rho|\sigma_i) = \sum_{j=1}^{n_r} -P_{\text{prod}}(\rho_j|\sigma_i)\log P_{\text{prod}}(\rho_j|\sigma_i) \quad (1)$$

$$H(\rho) = \sum_{j=1}^{n_r} -P(\rho_j)\log P(\rho_j) \quad (2)$$

where  $P(\rho_j)$  is the relative frequency with which referent  $\rho_j$  occurs.

The specificity of a lexicon is thus defined as the specificities of the words weighted by their occurrence probabilities:

$$\text{spec}(\sigma_i) = \frac{H(\rho) - H(\rho|\sigma_i)}{H(\rho)} = 1 - \frac{H(\rho|\sigma_i)}{H(\rho)} \quad (3)$$

$$\text{spec} = \sum_{i=1}^{n_s} P_{\text{prod}}(\sigma_i)\text{spec}(\sigma_i). \quad (4)$$

### ■ 3.2 Consistency

*Consistency* means that agents consistently uses the same word for a particular referent. It can be calculated as follows:

$$H(\sigma|\rho_i) = \sum_{j=1}^{n_s} -P_{\text{prod}}(\sigma_j|\rho_i)\log P_{\text{prod}}(\sigma_j|\rho_i) \quad (5)$$

$$H(\sigma) = \sum_{j=1}^{n_s} -P_{\text{prod}}(\sigma_j)\log P_{\text{prod}}(\sigma_j) \quad (6)$$

$$\text{cons}(\rho_i) = 1 - \frac{H(\sigma|\rho_i)}{H(\sigma)} \quad (7)$$

$$\text{cons} = \sum_{i=1}^{n_r} P(\rho_i) \text{cons}(\rho_i). \quad (8)$$

where  $P_{\text{prod}}(\sigma_j)$  is the relative frequency with which word  $\sigma_j$  occurs.

MacLennan [13] used entropy to measure the uniformity of the joint probability distribution represented by the co-occurrence probabilities of words and referents, and notes that perfect communication requires a nonuniform distribution. Since the measure is calculated over all co-occurrence probabilities however, it does not distinguish such communication systems from nonperfect communication systems that have nonuniform distributions. The current measures address this issue by considering the entropies of the conditional probability matrices per row. Thus, both completeness and correctness criteria are achieved, as expressed by the following theorems.

**Theorem 1 (Production measures)** *Communication production is perfect if and only if both specificity and consistency are equal to one.*

*Proof.* The first condition of the definition for perfect communication production requires that for each referent  $\rho_i$ , a word must exist  $\sigma_j^A$  for which  $P_{\text{prod}}(\sigma_j|\rho_i) = 1$ . Thus,  $P_{\text{prod}}(\sigma_k|\rho_i) = 0$  for  $k \neq j$ . Therefore,  $H(\sigma|\rho_i) = 0$  for each referent  $\rho_i$ , which implies consistency equals one. Likewise, using the second condition of the definition, for each word there exists a referent for which  $P_{\text{prod}}(\rho_j|\sigma_i) = 1$ , hence  $H(\rho|\sigma_i) = 0$ , and so specificity equals one. This proves the implication.

If consistency is equal to one, then for each referent  $\rho_i$ ,  $H(\sigma|\rho_i) = 0$ . Since each referent is represented by some word,  $P_{\text{prod}}(\sigma_j|\rho_i)$  must be nonzero for some  $\sigma_j$ , and the only nonzero value consistent with  $H(\sigma|\rho_i) = 0$  is one. Thus, for each referent  $\rho_i$  there exists a word  $\sigma_j$  for which  $P_{\text{prod}}(\sigma_j|\rho_i) = 1$ , which satisfies the first condition of Definition 3. The case for the second condition is analogous: if specificity is equal to one, then for each word  $\sigma_i$ ,  $H(\rho|\sigma_i) = 0$ , and there must exist a referent for which  $P_{\text{prod}}(\rho_j|\sigma_i) = 1$ . Thus, both conditions of the definition are satisfied. This demonstrates the reverse implication, which completes the proof. ■

### ■ 3.3 Fidelity

Fidelity is based on the matrix  $P_{\text{comm}}(\rho|\rho)$ , see section 2. The diagonal of this matrix specifies the probability that a referent, when encoded by a word and decoded back into a referent, yields the same referent. Ideally, the diagonal should contain ones. *Fidelity* is defined as the

average of this diagonal:

$$\text{fid} = \frac{\sum_{i=1}^{n_r} P_{\text{comm}}(\rho_i, \rho_i)}{n_r}. \quad (9)$$

**Theorem 2 (Interpretation measures)** *Communication interpretation is perfect if and only if fidelity is equal to one.*

*Proof.* If communication interpretation is perfect, then  $\mathbf{P}_{\text{comm}}(\rho|\rho) = \mathbf{I}$ . Thus, the fidelity equals one. This proves the implication. If the fidelity equals one, the diagonal elements of the matrix  $\mathbf{P}_{\text{comm}}(\rho|\rho)$  must all be one. Since each row must sum up to one, all remaining entries must equal zero. Thus,  $\mathbf{P}_{\text{comm}}(\rho|\rho)$  is the identity matrix  $\mathbf{I}$ , which proves the reverse implication and completes the proof. ■

**Theorem 3 (Communication measures)** *A communication system is perfect if and only if consistency, specificity, and fidelity are equal to one.*

*Proof.* If a communication system is perfect, then by definition both communication interpretation and communication production are perfect, which as shown implies that consistency, specificity, and fidelity are equal to one. If consistency, specificity, and fidelity are all equal to one, then both communication interpretation and communication production are perfect, which implies the system of communication is perfect. ■

This section has introduced two measures of communication production (specificity and consistency), and one measure of overall communication quality (fidelity). These measures can be used to analyze the performance of algorithms that aim to establish a communication system. Moreover, they can be used to devise such algorithms, as shown in the following section.

#### 4. An algorithm for communication development

The previous section described a method for analyzing communication systems and their development. In this section, we show how the goals that were identified for communication systems can be translated into an algorithm.

Two criteria for production behavior were specified: specificity and consistency. To address these goals, each agent will sample the communicative behavior of the population, and bias its current behavior towards choices that improve the specificity and consistency of the observed population behavior. The following sections describe the development of an algorithm, and experimental results with the algorithm are reported at the end of the section.



#### ■ 4.1 Achieving consistency

A focused distribution on the rows of  $\mathbf{P}(\sigma|\rho)$  corresponds to consistency, that is, using only a single word for each referent and, consequently, not using any other word for that referent. A complication is that the links between words and referents cannot be controlled directly by the agents; rather, agents must adapt associations between their meanings and the words. If concept formation has stabilized however, considering the relation between words and meanings will be sufficient, since there will be a good, but not necessarily one-to-one correspondence between meanings and referents.

The *use* of an association is an estimate of  $\mathbf{P}(\sigma|\mu)$  and thereby, assuming a reasonable conceptual system, of  $\mathbf{P}(\sigma|\rho)$ . Favoring associations with a high use value in word production leads to a positive-feedback process. Words that are used somewhat more frequently for a meaning than other words will tend to be produced even more often in its context, at the cost of less strongly associated words whose associations will become even weaker, and which eventually will not be produced anymore. This process favors consistency. Thus, word selection should depend on the *use* of associations.

#### ■ 4.2 Achieving specificity

A focused distribution on the rows of  $\mathbf{P}(\rho|\sigma)$  corresponds to specificity, ultimately restricting the usage of a word to a single referent only. The approach here is similar to the one taken previously for consistency. For each association, an estimate of  $\mathbf{P}(\mu|\sigma)$  is maintained. Since associations with high values for this estimate contribute to specificity, this value is called the *specificity* of the association. If for each meaning, words that are specific to it are favored in word production, this can further enhance the specificity of those words. Thus, word selection should depend on the *specificity* of associations.

#### ■ 4.3 A selection mechanism for word production

The above reasoning suggested that the selection of the word  $\sigma$  to be produced in the context of a given meaning  $\mu$  should depend on both the use and the specificity of that association. A simple way of achieving this is to use a linear combination of the estimated use and specificity of an association to represent its strength  $a(\mu, \sigma)$ :

$$a(\mu, \sigma) = \beta \text{spec}(\mu, \sigma) + (1 - \beta) \text{use}(\mu, \sigma). \quad (10)$$

Word production in the context of a given meaning  $\mu_i$  can then be based on the strengths of the associations of  $\mu_i$  with all possible words. The simplest selection method is greedy selection, which would always select the strongest association. However, greedy selection can

easily lead to deadlock situations, corresponding to local optima in the space of global communication systems; for example, if one word, due to random initial conditions, happens to be the preferred word for two different referents, pure greedy selection would continue to select this word for both referents, thus never achieving specificity. Such a deadlock can be avoided by using *exploration*, that is, occasionally making apparently suboptimal choices.

One way of achieving exploration is to occasionally make a completely random choice. This has the disadvantage that choices that are known to be bad will continue to be made. An elegant solution to this is to select each element with a probability that is proportional to the value of the element. This removes the distinction between greedy and exploratory choices. The degree of stochasticity in such a choice mechanism can be parameterized, so that selection may range from greedy (the highest element has probability one) to random (all  $n$  elements have equal probability  $1/n$ ). This functionality is provided by the Boltzmann distribution, which is used in various selection schemes in, for example, genetic algorithms [38] and reinforcement learning [39].

The greediness of selection based on the Boltzmann distribution can be adjusted by choosing the *temperature* parameter  $T$ . The probability of selecting the word  $\sigma_i$  using a Boltzmann distribution equals:

$$P_{\text{prod}}(\sigma_i) = \frac{e^{v(\sigma_i)/T}}{\sum_{j=1}^{n_s} e^{v(\sigma_j)/T}}. \quad (11)$$

By choosing a high value, exploration is encouraged, eventually leading to a uniform distribution in the limit of infinite temperature. By lowering the temperature parameter, the differences between the values are magnified, and a more greedy choice mechanism is obtained.<sup>1</sup>

#### Cycle(sensor\_data)

1.  $m := \text{determine\_meaning\_from\_sensors}(\text{sensor\_data});$
2.  $\mathbf{mw} := \beta * \mathbf{spec} + (1 - \beta) * \mathbf{use};$
3.  $w := \text{select}(\text{Boltzmann}(\mathbf{mw}(m, 1:nw)', t));$
4.  $\text{produce\_word}(w);$
5.  $\text{words} := \text{receive\_words}();$
6.  $\text{wordfreq} := \text{compute\_frequencies}(\text{words});$
7.  $\mu_{\max} := \arg \max_{\mu} \sum_{\sigma \in \text{words}} \mathbf{P}_{\text{int}}(\sigma, \mu) \cdot \text{wordfreq}(\sigma);$
8.  $\sigma_{\max} := \arg \max_{\sigma \in \text{words}} \{\mathbf{P}_{\text{int}}(\sigma, \mu_{\max}) \cdot \text{wordfreq}(\sigma)\};$
9. with probability  $P = \mathbf{P}_{\text{int}}(\sigma_{\max}, \mu_{\max})$  {

<sup>1</sup>The temperature parameter is called such because the Boltzmann distribution is the distribution of the possible states of a gas at high temperature.

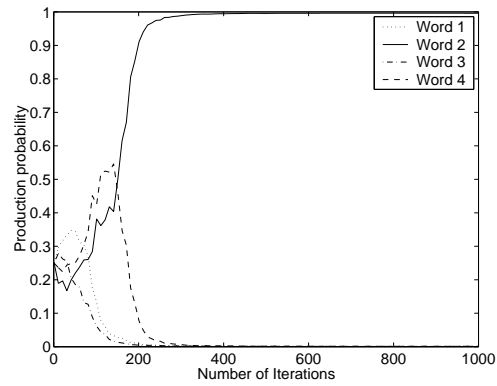
10.  $\mathbf{spec}(\mu_{\max}, \sigma_{\max}) := \alpha * 1 + (1-\alpha) * \mathbf{spec}(\mu_{\max}, \sigma_{\max});$
11.  $\forall \sigma \in \{\sigma_1, \sigma_2, \dots, \sigma_{n_w}\} \setminus \sigma_{\max}$
12.  $\mathbf{spec}(\mu_{\max}, \sigma) := \alpha * 0 + (1-\alpha) * \mathbf{spec}(\mu_{\max}, \sigma);$
13.  $\forall \mu \in \{\mu_1, \mu_2 \dots \mu_{n_m}\} \setminus \mu_{\max}$
14.  $\mathbf{spec}(\mu, \sigma_{\max}) := (1-\alpha) * \mathbf{spec}(\mu, \sigma_{\max});$
15. }
16.  $\forall \sigma \in \text{words}$
17.  $\mathbf{use}(m, \sigma) := \alpha * 1 + (1 - \alpha) * \mathbf{use}(m, \sigma);$
18.  $\forall \sigma \in \{\sigma_1, \sigma_2, \dots, \sigma_{n_w}\} \setminus \text{words}$
19.  $\mathbf{use}(m, \sigma) := (1 - \alpha) * \mathbf{use}(m, \sigma);$
20.  $\forall \sigma \in \text{words}\{$
21.  $\mathbf{P}_{\text{int}}(\sigma, m) := \alpha * 1 + (1 - \alpha) * \mathbf{P}_{\text{int}}(\sigma, m);$
22.  $\forall \mu \in \{\mu_1, \mu_2 \dots \mu_{n_m}\} \setminus m$
23.  $\mathbf{P}_{\text{int}}(\sigma, \mu) := (1 - \alpha) * \mathbf{P}_{\text{int}}(\sigma, \mu);$
24. }

#### ■ 4.4 Description of the resulting algorithm

The above ideas can be translated into the following algorithm. Each agent estimates the contribution of meaning–word associations to consistency and specificity, which together determine association strength. The strength of an association determines the probability that it will be selected in word production. Thus, associations that promote the aims of specificity and consistency are reinforced, while a potential for exploration is maintained.

The algorithm describes the behavior of a single agent upon receiving sensory information about the current referent (situation or object) in the environment. Based on this information, the agent first determines the meaning its sensors indicate. It then uses its current associations to select a word associated with this meaning, and produces it. Since other agents perform the same steps, each agent then receives the words produced by other agents and itself. This word is used to determine a likely meaning based on the received words, called *signal-based meaning determination*. If the likelihood of this meaning is sufficient, the specificity of the associations with this meaning is updated. Next, the use of the associations with the meaning that was indicated by the sensors (*sensor-based meaning determination*), and the interpretation probabilities  $\mathbf{P}_{\text{int}}(\mu|\sigma)$ , are adapted. To let a group of agents develop communication, this cycle is to be repeated so that the gradual adaptations made by individual agents can direct the behavior of the agents towards a shared system of communication.

*Spec* and *use* represent the matrices containing the specificity and use of each association between meaning and word.  $\mathbf{P}_{\text{int}}$  is the ma-



**Figure 2.** Development of the associations of all words with the first referent of agent number two. After an initial period in which the agents settle on a shared mapping between referents and words, the chosen associations become consolidated while the use of other words diminishes.

trix estimating  $P_{\text{int}}(\mu|\sigma)$  and  $t$  is the temperature parameter of the Boltzmann distribution.

The algorithm only requires samples of the production behavior of other agents. In the following section, it will be seen that this is sufficient as feedback for the development of communication.

#### ■ 4.5 Experimental investigation

Having described the algorithm, we proceed by investigating its behavior in a simulation experiment. In this simulation, an environment is simulated by randomly selecting one of the referents at each time step to be the subject of communication. Each agent owns a set of  $n_m = n_r$  meanings that bear a one-to-one correspondence to the referents. The set of words is a fixed set of size  $n_w = n_r$ . The use and specificity values of the associations of the agents are all initialized at 0.5, so that no word is preferred over another. The parameters are as follows: the number of agents  $n_a = 5$ , the temperature variable regulating stochasticity  $t = 0.15$ , the learning rate  $\alpha = 0.05$ , and the specificity-use ratio  $\beta = 0.5$ .

##### 4.5.1 Example run

To describe the behavior of the algorithm, we study a single run of it.<sup>2</sup> Figure 2 shows the production probability of each word for the first referent for agent number two over time. We consider the calculation

<sup>2</sup>The example is the first run in the series of 100 that was performed, and can be reproduced by choosing random seed 0 in the algorithm, which is available as an online appendix to the article.

of the consistency measure of this particular agent, which is based on this word production probability matrix. The matrix at the beginning of the experiment, at time step zero, is shown in Table 1 (top). The uniformly initialized values are the least consistent configuration, and should result in a minimal consistency value. To calculate consistency, we first calculate the entropy over the first row of the matrix, using the base two log:

$$H(\sigma|\rho_1) = \sum_{j=1}^{n_s} -P_{\text{prod}}(\sigma_j|\rho_1)\log P_{\text{prod}}(\sigma_j|\rho_1) \quad (12)$$

where, for example,  $P_{\text{prod}}(\sigma_1|\rho_1) = 0.25$ . Thus,  $H(\sigma|\rho_1) = 2$ , and

$$\text{cons}(\rho_1) = 1 - \frac{H(\sigma|\rho_1)}{H(\sigma)} = 0.$$

Performing the same calculation for each referent and taking the sum of the outcomes, weighted by the occurrence probability of each referent (which in this experiment, since each referent is equally likely, amounts to averaging) yields the consistency for this agent. Since all values are equal, this average

$$\text{cons}(A) = \sum_{i=1}^{n_r} P(\rho_i)\text{cons}(\rho_i)$$

also equals zero, and thus consistency is indeed minimal.

In the course of the experiment, as Figure 2 showed for referent 1, the association strengths first fluctuate while the global system of communication settles on a single mapping that links referents to words. Once each referent becomes associated with a single unique word, there are no more conflicts between the word choices of the agents, and the associations are consolidated. This leads to polarized values in the production matrix: association strengths are either very low or very high. In the current experiment, the production matrix for agent two has changed into the highly focused matrix of Table 1 (bottom) by the end of the experiment.

As the matrix shows, each referent is associated with a single and unique word, and thus consistency should be high. Indeed, the same calculation now yields a value of 0.96 for the consistency measure. In the following section, we move from the study of an individual agent to the behavior of the population.

#### 4.5.2 Experimental results

The simulation experiment described above was run 100 times for a duration of 1000 iterations of the algorithm. Figure 3 shows the graph of the average fidelity over time. As the figure shows, communication systems of very high quality are obtained, with fidelity converging to a

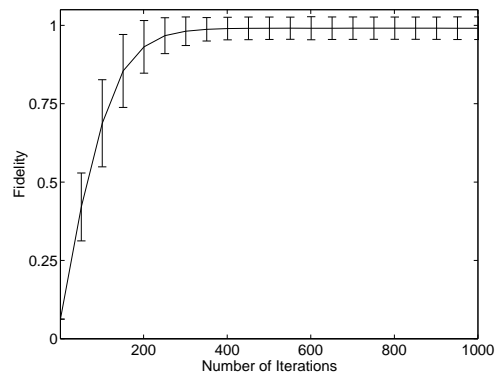
$P(\sigma \rho)$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_4$
$\rho_1$	0.25	0.25	0.25	0.25
$\rho_2$	0.25	0.25	0.25	0.25
$\rho_3$	0.25	0.25	0.25	0.25
$\rho_4$	0.25	0.25	0.25	0.25

$P(\sigma \rho)$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_4$
$\rho_1$	0.0013	<b>0.9961</b>	0.0014	0.0013
$\rho_2$	0.0013	0.0013	0.0013	<b>0.9962</b>
$\rho_3$	0.0014	0.0013	<b>0.9960</b>	0.0013
$\rho_4$	<b>0.9962</b>	0.0013	0.0013	0.0013

**Table 1.** Production matrix  $\mathbf{P}^{A_2}(\sigma|\rho)$  for agent number two at the first (top) and last (bottom) time step of the experiment. Whereas initially the matrix is highly dispersed, the result of the adaptations of each agent to its production behavior in response to the observed co-occurrences between meanings and words leads to a shared communication system, characterized by a focused and stable mapping between words and referent at the end of the experiment.

value of just below one. Given furthermore that the minimal number of words is used in this experiment, this indicates the communicative behavior is very close to that in a perfect system of communication. Since the temperature parameter is positive, some stochasticity necessarily remains in word production, and poses a theoretical limit on fidelity that depends on the temperature. This poses no practical problem, since fidelity is near its maximum for the current choice of  $t = 0.15$ . When an annealing schedule is used to decrease the temperature parameter over time, fidelity can increase further to values arbitrarily close to one, as has been confirmed in additional experiments. However, in order to allow for the learning of new meaning–word associations, or changes to the current language used by agents, it is preferable to avoid such an annealing mechanism, thus permitting further adaptation of the behavior of agents after an initial system of communication has formed.

In related work, a variant of the algorithm described is applied in a setting consisting of a group of reinforcement learning agents that form concepts about predators in their environment [29]. There it was demonstrated that using communication, agents are able to reduce the uncertainty about their environment resulting from incomplete perception.



**Figure 3.** Development of a communication system in the simulation experiment, averaged over 100 runs. Fidelity consistently rises to high values; some stochasticity in word production remains to permit further adaptation, thus posing a theoretical limit to fidelity just below one.

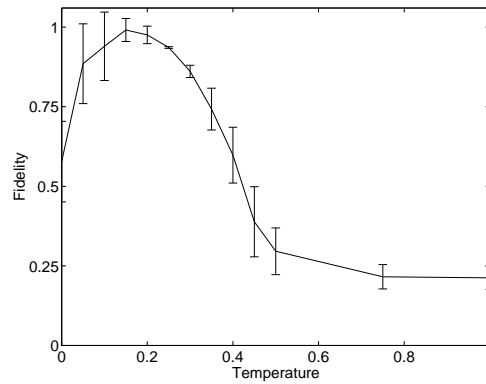
### 5. The role of stochasticity in communication development

In this section, we study the effect of stochasticity on communication development. It is expected that stochasticity plays an important role by overcoming deadlocks. To study how stochasticity affects the development of communication, its influence will be considered by varying the temperature parameter of the Boltzmann distribution used in word production (see section 4.1).

For each of a number of different temperatures, a hundred experimental runs are carried out. Using these experiments, the average fidelity at the end of a run can be determined as a function of temperature. The results are shown in Figure 4.

A low temperature corresponds to little exploration, and in the limit of zero becomes equal to a deterministic system. As the leftmost data point of the graph shows, this deterministic system does not always converge for different initial conditions, which is in line with the above explanation. Increasing the temperature to around 0.15 results in the stable development of communication that has been observed in the previous sections. By still further increasing the amount of exploration, the fidelity drops again, as a result of the increase of the random factor in word production. In the limit, increasing temperature results in word production behavior in which each word is equally likely to be produced, precluding any possibility of information transfer.

Analysis of the consistency and specificity measures provides an informative addition to these observations. For low temperatures, specificity is low, while consistency is high. This means that some words are used in combination with multiple meanings. For high



**Figure 4.** The effect of stochasticity on the quality of resulting communication systems. The curve shows that mild exploration is instrumental in reaching accurate systems of communication.

temperatures, specificity is also low, but consistency is low as well. This is precisely what would be expected for high temperatures; since word production involves more exploration, meanings will be expressed by different words at different times. For intermediate temperatures ( $0.15 \leq T \leq 0.25$ ), the fidelity of communication is high. Thus, by using a moderate amount of exploration, accurate systems of communication can be reliably developed.

The experiment provides an explanation of why a deterministic system can be less effective in developing communication, namely because it lacks exploration. When some word is the preferred (most strongly associated) word for multiple referents, greedy selection results in a deadlock. The specificity components of this word's associations with each meaning are increased in some interactions, but, due to its association with the other meanings, decreased in others. In such cases, exploration allows new words to be used sporadically, and since these are not restrained by their associations with other meanings, their association values can increase and surpass those of the ambiguous word.

Although the state space of the complete system is far too large to be sampled completely, the experimental findings on the test problem indicate that the stochasticity parameter controls the development of communication and that there exists a regime for which this development consistently occurs.

## 6. Conclusions

A distributed learning algorithm for communication development has been described. The algorithm adapts the associations of individual agents such that a global system of communication results. The only



form of feedback used by the algorithm are the observed co-occurrence frequencies between meanings and words.

Furthermore, the agents make small and directed changes to their communicative behavior. Thus, their behavior is relatively stable even during learning. This is a useful property when agents are to perform useful tasks while learning to communicate.

Since the agents have equivalent roles, all agents may equally influence the developing communication system. Thus, the resulting system adapts to the requirements of the learning agents, rather than being fixed in advance.

It has been shown that there exists a parameter regime for which communication systems with high fidelity consistently develop. This is realized at moderate amounts of exploration, allowing the system to try alternatives for conflicting associations and thus settle into a shared mapping between referents and words.

The method was demonstrated in an experiment that focuses on the development of a shared mapping between referents and words. A variant of the communication development algorithm has furthermore been applied in a setup involving concept formation [29]; there, it was found that communication can be used to reduce uncertainty about the state of the environment.

The implications of the work are as follows. First, a principle for communication development has been identified that makes directed and incremental adaptations and that uses a readily available form of feedback, consisting of the observed co-occurrence frequencies of meanings and words. Furthermore, principled measures of communication system quality have been provided. These measures have been shown to be correct and complete. Finally, this work exemplifies an approach to the construction of distributed learning algorithms which consists of defining global utility measures and translating these into local adaptation rules for individual agents.

---

## Appendix

The algorithm that has been described is available in Matlab<sup>®</sup> format from the homepage of the first author.<sup>3</sup>

---

## Acknowledgments

The first author gratefully acknowledges a Fulbright grant. The authors would like to thank Bram Bakker, Bart de Boer, Michiel de Jong, Dean Driebe, Paul Vogt, Richard Watson, and Jelle Zuidema for useful comments on this work that have improved the article.

---

<sup>3</sup><http://www.cs.uu.nl/~dejong>

## References

- [1] Craig Boutilier, "Learning Conventions in Multiagent Stochastic Domains Using Likelihood Estimates," in *Proceedings of the Twelfth Conference on Uncertainty in Artificial Intelligence (UAI-96)*, edited by Eric Horvitz and Finn Jensen (Morgan Kaufmann, San Francisco, 1996).
- [2] Michael Bowling and Manuela Veloso, "Rational and Convergent Learning in Stochastic Games," in *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, Seattle, WA, 2001.
- [3] Caroline Claus and Craig Boutilier, "The Dynamics of Reinforcement Learning in Cooperative Multiagent Systems," in *Proceedings of the Fifteenth National Conference on Artificial Intelligence (AAAI-98) and of the Tenth Conference on Innovative Applications of Artificial Intelligence (IAAI-98)*, Menlo Park, July 26–30, 1998 (AAAI Press).
- [4] Junling Hu and Michael P. Wellman, "Multiagent Reinforcement Learning: Theoretical Framework and an Algorithm," in *Proceedings of the Fifteenth International Conference on Machine Learning* (Morgan Kaufmann, San Francisco, 1998).
- [5] Michael L. Littman, "Markov Games as a Framework for Multi-agent Reinforcement Learning," in *Proceedings of the Eleventh International Conference on Machine Learning* (Morgan Kaufmann, San Francisco, CA, 1994).
- [6] John W. Sheppard, "Colearning in Differential Games," *Machine Learning*, **33** (1998) 201.
- [7] Yoav Shoham and Moshe Tennenholtz, "On the Emergence of Social Conventions: Modeling, Analysis and Simulations," *Artificial Intelligence*, **94**(1–2) (1997) 139–166.
- [8] R. Sun and T. Peterson, "Multi-agent Reinforcement Learning: Weighting and Partitioning," *Neural Networks*, **12**(4–5) (1999) 727–753.
- [9] Ming Tan, "Multi-agent Reinforcement Learning: Independent vs. Cooperative Agents," in *Proceedings of the Tenth International Conference on Machine Learning* (Morgan Kaufmann, San Francisco, 1993).
- [10] Kagan Tumer and David Wolpert, "Collective Intelligence and Braess' Paradox," in *Proceedings of the Seventh Conference on Artificial Intelligence (AAAI-00) and of the Twelfth Conference on Innovative Applications of Artificial Intelligence (IAAI-00)* (AAAI Press, Menlo Park, CA, 2000).
- [11] Luc Steels, "The Origins of Ontologies and Communication Conventions in Multi-agent Systems," *Autonomous Agents and Multi-Agent Systems*, **1**(2) (1998) 169–194.

- [12] Ezequiel A. Di Paolo, "Social Coordination and Spatial Organization: Steps Towards the Evolution of Communication," in *Proceedings of the Fourth European Conference on Artificial Life ECAL'97*, edited by Phil Husbands and Inman Harvey (The MIT Press, Cambridge, MA, 1997).
- [13] Bruce MacLennan, "Synthetic Ethology: An Approach to the Study of Communication," in *Artificial Life II*, volume X of *SFI Studies in the Sciences of Complexity*, edited by C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (Addison-Wesley, Redwood City, CA, 1991).
- [14] Jason Noble, *The Evolution of Animal Communication Systems: Questions of Functions Examined through Simulation*, Ph.D. Thesis, University of Sussex, Sussex, UK, November, 1998.
- [15] Martin A. Nowak and David C. Krakauer, "The Evolution of Language," in *Proceedings of the National Academy of Sciences of the United States of America*, **96** (1999) 8028–8033.
- [16] Gregory M. Werner and Michael G. Dyer, "Evolution of Communication in Artificial Organisms," in *Artificial Life II*, volume X, edited by C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen (Addison-Wesley, Redwood City, CA, 1991).
- [17] Aude Billard and Gillian Hayes, "Learning to Communicate through Imitation in Autonomous Robots," in *Proceedings of the Seventh International Conference on Artificial Neural Networks ICANN 97* (Springer, Berlin, 1997).
- [18] Aude Billard and Gillian Hayes, "Drama, a Connectionist Architecture for Control and Learning in Autonomous Robots," *Adaptive Behavior*, **7**(1) (1999).
- [19] H. Yanco and L. Stein, "An Adaptive Communication Protocol for Cooperating Mobile Robots," in *From Animals to Animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, edited by H. L. Roitblat, J-A. Meyer, and S. Wilson (The MIT Press, Cambridge, MA, 1993).
- [20] Michael Oliphant, *Formal Approaches to Innate and Learned Communication: Laying the Foundation for Language*, Ph.D. Thesis, University of California, San Diego, CA, 1997.
- [21] J. Hurford, "Biological Evolution of the Saussurean Sign as a Component of the Language Acquisition Device," *Lingua*, **77** (1989) 187–222.
- [22] Bart De Boer, *The Origins of Vowel Systems: Studies in the Evolution of Language*. oup, 2001.
- [23] L. Steels, "A Self-organizing Spatial Vocabulary," *Artificial Life Journal*, **2**(3) (1996).

- [24] Luc Steels, "The Synthetic Modeling of Language Origins," *Evolution of Communication*, **1**(1) (1997) 1–34.
- [25] David K. Lewis, *Convention: A Philosophical Study* (Harvard University Press, Cambridge, MA, 1969).
- [26] Martin J. Osborne and Ariel Rubinstein, *A Course in Game Theory* (The MIT Press, Cambridge, MA, 1994).
- [27] Ariel Rubinstein, "Why Are Certain Properties of Binary Relations Relatively More Common in Natural Language?" *Econometrica*, **64**(2) (1996) 343–355.
- [28] Thomas C. Schelling, *The Strategy of Conflict* (oup, New York, 1963).
- [29] Edwin D. De Jong, *Autonomous Formation of Concepts and Communication*, Ph.D. Thesis, Vrije Universiteit Brussel, Brussels, Belgium, June 2000. Available from: <http://www.cs.uu.nl/~dejong/>.
- [30] Karl F. MacDorman, "Partition Nets: An Efficient On-line Learning Algorithm," in *ICAR 99: Ninth International Conference on Advanced Robotics*, 1999.
- [31] Karl F. MacDorman, "Proto-symbol Emergence," in *Proceedings of IROS-2000: International Conference on Intelligent Robots and Systems*, 2000.
- [32] J. Batali, "Computational Simulations of the Emergence of Grammar," in *Approaches to the Evolution of Language: Social and Cognitive Bases*, edited by J. R. Hurford, M. Studdert-Kennedy, and C. Knight (Cambridge University Press, 1998).
- [33] Takashi Hashimoto and Takashi Ikegami, "Emergence of Net-grammar in Communicating Agents," *BioSystems*, **38**(1) (1996) 1–14.
- [34] Simon Kirby, *Function, Selection, and Innateness* (Oxford, New York, 1999).
- [35] Simon Kirby, "Spontaneous Evolution of Linguistic Structure—An Iterated Learning Model of the Emergence of Regularity and Irregularity," *IEEE Transactions on Evolutionary Computation*, **5**(2) (2001) 102–110.
- [36] Luc Steels, "The Origins of Syntax in Visually Grounded Robotic Agents," *Artificial Intelligence*, **103**(1–2) (1998) 133–156.
- [37] Seth Bullock, "A Continuous Evolutionary Simulation Model of the Attainability of Honest Signalling Equilibria," in *Proceedings of the Sixth International Conference on Artificial Life*, edited by Christoph Adami, Richard K. Belew, Hiroaki Kitano, and Charles Taylor (The MIT Press, Cambridge, MA, 1998).
- [38] Melanie Mitchell, *An Introduction to Genetic Algorithms: Complex Adaptive Systems* (The MIT Press, Cambridge, 1996).

- [39] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction* (The MIT Press, Cambridge, MA, 1998).

