The Development of a Lexicon Based on Behavior

Edwin D. de Jong
Artificial Intelligence Laboratory
Vrije Universiteit Brussel
Pleinlaan 2, B-1050 Brussels
edwin@arti.vub.ac.be

Abstract

This paper investigates whether a group of agents may develop a common lexicon relating words to situations by a process of self-organization. Each agent independently decides which situations are useful to distinguish, based on its experience with the environment. It then starts to associate signals with each situation. The agents adapt their own associations based on the signals they received from other agents. The system is monitored using measures which reflect the development of the lexicons over time. The result of the distributed activities of the agents is that a coherent shared lexicon emerges linking signals to situations.

Introduction

A central aspect of human intelligence is our use of language. The aim of the currently expanding research field of language evolution is to gain insight into principles that may explain the development of language and communication. In this paper we pose the following question: Can a common lexicon be established by a process of self-organization? That is, can a group of agents converge to using the same set of associations between concepts and words? This question has already been addressed for the case where the signals refer to objects and the interactions are organized by appointing a speaker and a hearer in [Steels, 1996a]. In this work however, we study the formation of concepts denoting situations encountered during interaction with the environment. An agent determines which situations to distinguish according to the effects of its actions. Successful communication provides potential benefit to the receiver since it complements the basic sensory data. Both the conceptualization of sensory data and the adaptation of a lexicon which associates signals with situations are investigated in this paper.

Agents build up their own lexicon by associating their individually determined situations with the signals they hear. On each turn, an agent produces a signal, depending on its situation. This signal is audible to all agents. An agent may use these signals to decide that its actual situation is probably different from the one indicated by its sensors. This will affect the action it selects. If however the effect of this action is different from what it expected, then the relation between the signals received and the hypothesized situation is deceptive and needs to be adapted. We investigate the global effects of these locally adapted lexicons.

In the general model of communication used here, agents develop a set of meanings (which in these experiments denote situations) and for each of these meanings a set of signals that are associated with it. The bare result that agents develop a lexicon would not be anything new. The value of such results can only be appreciated in view of the *commitments* that are adhered to. In this work, these are:

- We do not assume that agents use the same meanings. Each agents develops its own set of concepts based on its interaction with the environment.
- Meanings are not passed directly from one agent to another. This would be implausible as a general model of communication, and is implied by the first point.
- The agents are not appointed different roles; there is no separation between teaching and learning agents, and thus all agents may influence the development of the communication.
- Agents adapt their communication during their lifetime. In this sense the work differs from most approaches based on genetic algorithms.
- There is no direct feedback to agents about the success of their communication; only feedback about their behavior *excluding* communication is available, as is advocated in [Werner and Dyer, 1991]. Thus, a coherent system of communication can only come about if agents use the signals they perceived to choose their actions and at the same time use the feedback on their behavior to adapt their lexicon.

Several authors describe the development of communication where agents are either teacher or learner, possibly alternating. Some examples are [Yanco and Stein, 1993], [Billard and Hayes, 1997] and [Steels and Vogt, 1997], all of which are interesting in that real robots are used to investigate the development of communication. An example of work where agents are not appointed different roles is [MacLennan, 1991].An investigation of the issue of altruism, i.e. how can language evolve if only *receiving*, and not *producing* truthful signals has a clear benefit, is presented in [Ackley and Littman, 1994].

The structure of the paper is as follows. In the following section, the environment of the agents is described. Section 2 describes the categorization method. The mechanisms used by the agents to adapt their lexicon are described in section 3. In section 4, we describe methods which allow one to measure the development of a lexicon. The results of the experiments are reported in section 5, which is followed by a discussion.

1 The problem: how to hide for predators

In order to investigate the categorization of sensory information, it is necessary that the effect of the actions available to an agent bears some relation to the sensor values. Conceptualization then is the process of determining which sensor readings should be treated differently, and which can be treated equally. In the environment used in the experiments, the effects of the different actions depend on the presence of a predator and its type. This information, combined with the horizontal and vertical position of the agent itself, determines the sensory data. The available actions amount to horizontal (staying or moving one step left or right) and vertical movement (choosing a new position). There are three vertical positions. These can be viewed as abstractions for hiding places, since each row is a safe hideaway for one type of predator. Predators appear and disappear at random intervals. Each type of predator has the same probability of appearing, and roughly half of the time no predator is present.

If, during the presence of a predator, an agent is not in the safe row, it will receive a zero success value. Moving to the safe row yields a success of 1.0. When no predator is present, staying in the same place is rewarded as 1.0, and the other choices for vertical movement yield only 0.5. Apart from this sensor information and these actions, the agents may choose to produce a signal. All signals produced at time t are perceived by the agents at time $t+1$. Thus, the information contained in these signals reflects the situations as viewed by the agents on the previous turn.

The problem was inspired by the fascinating system of alarm calls used by vervet monkeys [Seyfarth et al., 1980], which allows the animals to respond appropriately to calls indicating the class of a predator: bird of prey, large mammal or snake. However, we want to state clearly that the simulations are not intended to model or describe vervet monkey alarm calls; in fact, the relation between signal and class or predator is thought to be innate, whereas in the experiments reported here, both the categories and the lexicon are learned.

2 Categorization of situations

The concepts formed in these experiments represent situations which an agent has learned to distinguish. Initially, all experiences are treated equally. As the agent interacts with the environment and thus gains knowledge about the effects of its actions, it starts to distinguish between different regions of the sensor-action space. The rationale behind this is that a distinction between two parts of a region in the sensor-action space should only be introduced if this enables a more accurate prediction of the results of the different actions in the area. This decision is based on a comparison of the probability distributions of the results in these two parts.

The idea of introducing increasingly precise distinctions based on utility has previously resulted in a method for *discrimination games* [Steels, 1996b] where utility was determined by the ability to tell apart the different objects sensed by a robot [De Jong and Vogt, 1998]. This method yields a tree which stores information with adaptive resolution. Here, the upper nodes determine intervals for sensors, allowing the agent to discriminate between different situations, and lower nodes specify intervals determining sets of actions. Adaptive resolution methods have been used in [Moore and Atkeson, 1995], and the idea of splitting based on utility is also used in [McCallum, 1996].

Figure 2 shows a typical tree learned by one of the agents during an experiment. The first choice at the root is whether S3, the value of sensor 3 is smaller than 2 or not. If so, then it can be either smaller or larger than 1. The decisions in the left part of the tree concern sensors. After the dashed line, only action distinctions occur. This organization of the tree presents the agent with a set of relevant situations that should be distinguished in this environment. These are represented by rightmost nodes on the left side of the line. Note that all sensor distinctions concern sensor 3; the agent has correctly learned that only this sensor, indicating the type of predator that is present, determines the success of each action. The different options for choosing an action are limited in the lower nodes of the tree. Here, only the second action dimension, which specifies vertical movement,

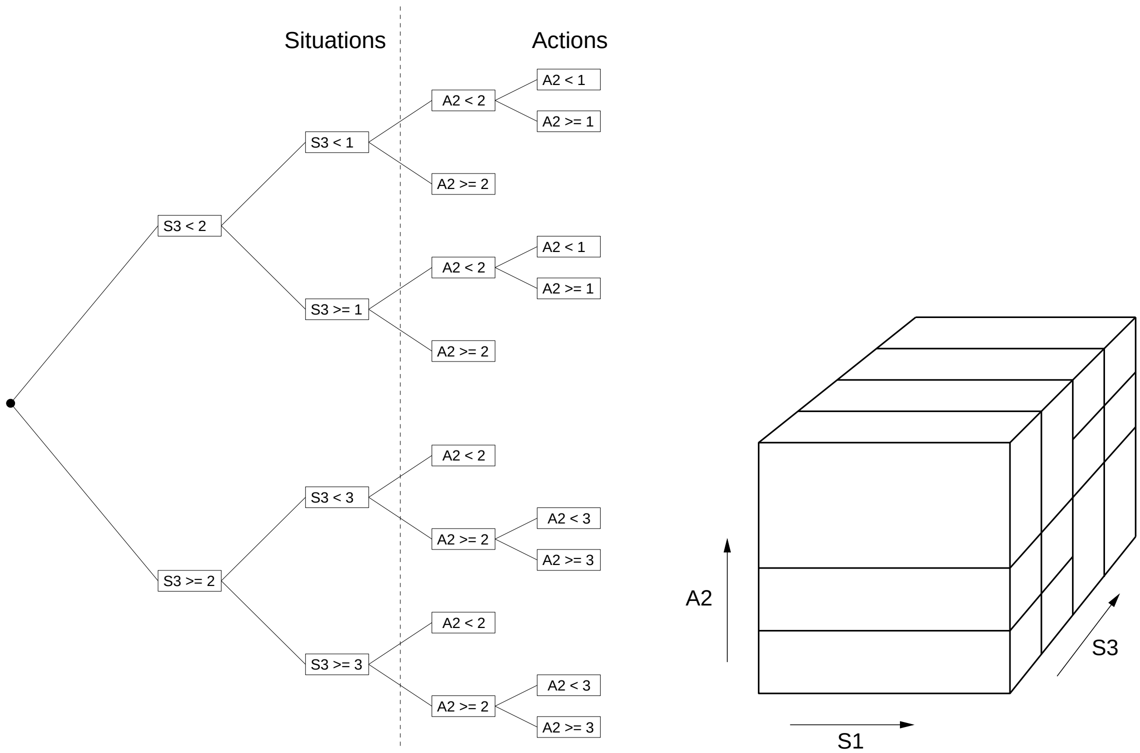


Figure 1: Left: a typical tree with situations and actions as learned by an agent in one of the experiments. Right: an equivalent representation of the tree in figure 2. All distinctions concern sensor S3 and action A2. In other dimensions (only S1 is shown), extra distinctions are unnecessary, as the agent has successfully learned using the statistical tests.

appears in the comparisons. Hence, the interval for horizontal movement is unrestricted.

The agent chooses its actions randomly from the intervals determined by the nodes of the tree. When the sensors are used to determine the agent's situation and action, the selection of the node is straightforward; the first decisions are determined by the sensor values observed, and subsequent choices maximize the predicted success, which is kept in the leaves. Alternatively, the agent may use the signals it perceives to hypothesize that its actual state is different from the one indicated by its sensors. This is described in section 3. In this case, action selection is analogous to the sensor-based case.

3 Lexicon formation

The tree built by an agent using the mechanisms that have been described provides a collection of situations that should be distinguished, since the predicted success values of the actions differ. These situations are represented by the rightmost internal nodes that still concern sensor dimensions. The subtrees under the situation nodes contain predictions of the success after selection of the corresponding actions. Each situation is associated with a set of signals, which are represented as two letter 'words' consisting of a consonant and a vowel (e.g. `PA`, `LU`, `MO`, etc.).

Since all agents inhabit the same environment, the situations, or concepts, developed by the agents should be very similar, and inspection of the data shows that these are mostly, but not always, equivalent. During each turn, every agent produces a signal which it selects from the list of signals associated with that situation with probabilities according to the association strengths. Thus, the number of signals an agent perceives is equal to the number of agents. When a signal is perceived in a situation, the strength of its association with that situation is increased. Likewise, it is decreased when the situation is encountered without hearing the signal. The probabilities of hearing a signal given the current situation are thus estimated by successive approximation. Furthermore, successive estimations of the overall signal probabilities are made.

Apart from the probability of each signal in general and given a situation, agents estimate the probability of each situation, which is computed as its relative frequency multiplied by the fraction of the sensors that correspond to that situation's intervals. Therefore, they are able to compute the probability of being in a certain situation given the signals perceived using Bayes' rule:

$$P(\mu \mid \sigma) = \frac{P(\mu \wedge \sigma)}{P(\sigma)} \tag{1}$$

$P(\mu \wedge \sigma)$ can be computed using

$$P(\mu \wedge \sigma) = P(\mu) \cdot P(\sigma \mid \mu)$$

σ is a signal that was perceived, and μ is a possible situation or, more general, a meaning. This probability is computed for every situation, given the current signals. Depending upon the maximum of these probabilities, the agent either assumes to be in the situation indicated by its sensor values, or in the situation indicated by the signals. In the latter case, a way is needed to determine whether the signal-based determination of the situation was correct. Although agents are endowed with complete perception in these experiments, we do not want to use this as an assumption. Therefore, this test cannot be based on sensor values, even though these do provide this information. Instead, the success of taking an action is compared to the predicted value. If there is a discrepancy between these, then probably the information of the signals was misleading. This criterion has been operationalized by testing whether the prediction error is smaller or larger than the average prediction error. In the early phases of an experiment, this happens frequently, since all agents initially produce random signals. Whenever this is the case, the association between the signal that determined the choice of the situation is decreased. If on the other hand the action success was near the prediction, then the association between signal and situation is reinforced. In both cases, the association of the signal with other situations is adapted in the opposite direction.

4 Measuring the Development of the Lexicon

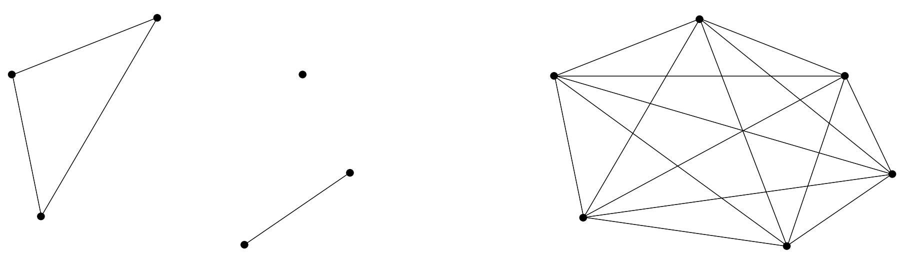
To investigate whether the signals used by the agents provide useful information, we monitor the experiment using two measures: *coherence* and *specificity*. This methodology for investigating communication is motivated in [De Jong, 1998]. The coherence indicates to what extent agents use the same signal for a certain concept. Since an agent may hear different words on different instances of what it considers to be the same situation, it will generally associate each situation with several words. For a particular situation, the word with the strongest association is called the *preferred word*. Both measures that are described here are based on preferred words. For each concept, the coherence is measured by computing the fraction of agents that has a word as preferred word for that concept,

	m1	m2	m3	m4
a1	PI	PA	PO	PU
a2	PI	PA	PA	PA
a3	LU	PA	PO	PU
a4	PI	PA	PO	PA
Max. freq.	3	4	3	2
Coherence	0.75	1.0	0.75	0.5

Table 1: An example calculation of coherence for four agents ($a1 \ldots a4$) and four meanings ($m1 \ldots m4$). For each meaning, the highest frequency of a word is determined, shown in the bottom row. The average of these figures is the coherence.

and then taking the maximum of these fractions. If no word is associated with a concept, the frequency of that concept's word is zero. An example of the calculation is shown in table 1.

In the extreme situation that all agents would use a single signal for every situation, coherence would be perfect. Clearly, another factor plays a role in determining the utility of the communication system. In order to be of use, a signal should limit the number of possible situations. Ideally, a signal brings down the number of possible situations to 1. This factor is expressed by another measure called *specificity*, which we will now introduce.



A concept graph: each node represents a concept
A vertex indicates two concepts have the same word

The corresponding fully connected graph:
All concepts are associated with the same word

Specificity can be computed by constructing a graph which has as its nodes the meanings, or situations in this case, of a certain agent. In this graph, two meanings are connected if a single word is primarily associated with these meanings. The optimal graph then is a graph without any edges at all. The worst case, in which a single word is used for every situation, results in the fully connected graph. An exception occurs when no word is associated with a certain meaning. Then, that meaning is connected with all other meanings, since the agent has no word to distinguish it from these. This exception appropriately decreases the specificity when one or more meanings lack words.

The specificity σ is inversely proportional to the connectivity of the graph, i.e. the fraction of edges that are present in the graph:

$$\sigma \stackrel{def}{=} 1 - \frac{v}{\frac{n^2-n}{2}} \qquad (2)$$

where v is the number of edges and n is the number of nodes in the graph. Interestingly, the same figure can be obtained in a more easily implementable fashion by adding the

	m1	m2	m3	m4	$\sum f$	*Specificity*
a1	1	1	1	1	4	1.0
a2	1	3	3	3	10	0.5
a3	1	1	1	1	4	1.0
a4	1	2	1	2	6	0.83

Table 2: An example calculation of specificity. The words are those shown in table 1. For each agent, frequencies of the words are summed. The specificity of an agent's signals can then be calculated using equation 3.

agent-specific frequencies of the symbols and computing the difference with the maximum sum relative to the maximum value of this difference:

$$\sigma = \frac{n^2 - \sum_{k=1}^{n} f_k}{n^2 - n} \tag{3}$$

The equivalence of the two calculation methods can be seen as follows. The frequency f_k of a concept's word decreased by 1 equals the number of other concepts associated with this word, and hence the number of outgoing edges from this concept's node. Accumulating these numbers and dividing by 2 yields the total number of edges v in the graph:

$$v = \frac{\sum_{k=1}^{n}(f_k - 1)}{2} = \frac{-n + \sum_{k=1}^{n} f_k}{2} \tag{4}$$

Substituting (4) for v in (2) shows the equivalence of (2) and (3):

$$1 - \frac{2}{n^2 - n} \cdot \frac{-n + \sum_{k=1}^{n} f_k}{2} = \frac{n^2 - \sum_{k=1}^{n} f_k}{n^2 - n} \tag{5}$$

$$n^2 - n + n - \sum_{k=1}^{n} f_k = n^2 - \sum_{k=1}^{n} f_k \tag{6}$$

These measures assume the simplification that meanings are shared between agents. In reality, it is my conviction that our concepts are built from elements that are, somewhat like in these experiments, based on experiences. It is the simple structure of the environment and agents that allows us to map meanings developed by agents onto eachother; this would not be possible with human subjects.

5 Results

The experiments start with a conceptualization phase. During this period, the agents investigate the effects of their actions and use this experience to develop a set of useful distinctions between sensor values and action values, using the mechanisms described in section 2.

After game no. 2,000, the agents start to produce signals, and base the determination of the situation, and hence their action selection, on the signals instead of the sensors depending on the Bayesian estimation given by (1). Initially, this decreases success considerably. However, in the course of time, the relation between signals and situations

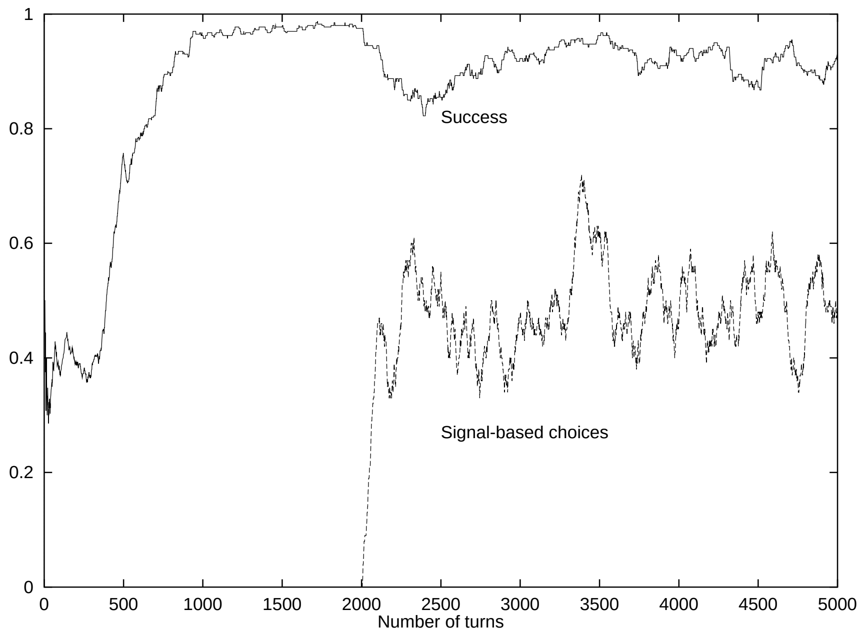


Figure 2: Success averaged over 200 turns as a function time

becomes more solid, and situation prediction according to signals becomes more accurate. The fraction of turns on which the agent determined its situation based on the signals it received is shown as well. This occurs roughly half of the time. The fact that the success decreases only slightly combined with the information that signals were used in determining the state on half of the occasions makes clear that the signal-based choice is quite accurate. If the actual situation is different from the one determined by the agent, the success of its action is only 0.5 in the cases where no predator is present, and 0.0 in the cases where some predator is present; this would have affected the success considerably.

Figure 5 shows the value of the coherence and the specificity over time during the same run. The coherence increases considerably before the specificity does so. Analysis of the data explains this. When the agents first learn to associate a word with a situation, that word is produced far more often than any other word, also by the other agents. When the situation changes, this signal is almost certainly perceived, since the signals always reflects the situation of the agents on the previous turn. Thus, this signal will immediately be associated with the new situation. Now, a single word is associated with two situations. This explains why the coherence can increase rapidly. It also shows why specificity initially remains low; the word is now associated with two situations, resulting in a low specificity score. This undesirable effect is repaired by the tendency of each agent to reduce the strength of the association between a word and a situation when the success differs considerably from the prediction. This induces variation, temporarily decreasing the coherence, but convergence to a new word for the new situation has the desired effect of increasing both coherence and specificity.

After a while, a shared lexicon emerges. Both specificity and coherence are 1.0 for all agents during a substantial period of time. This combination of measures can only occur when all agents have converged to using the same word for each situation, and simultaneously using a distinct word for every situation that needs to be distinguished. Inspection of the data showed that this is indeed the case; every agent used `ki` for situations without a predator, and `ta`, `bu` and `ga` for the three different types of predator.

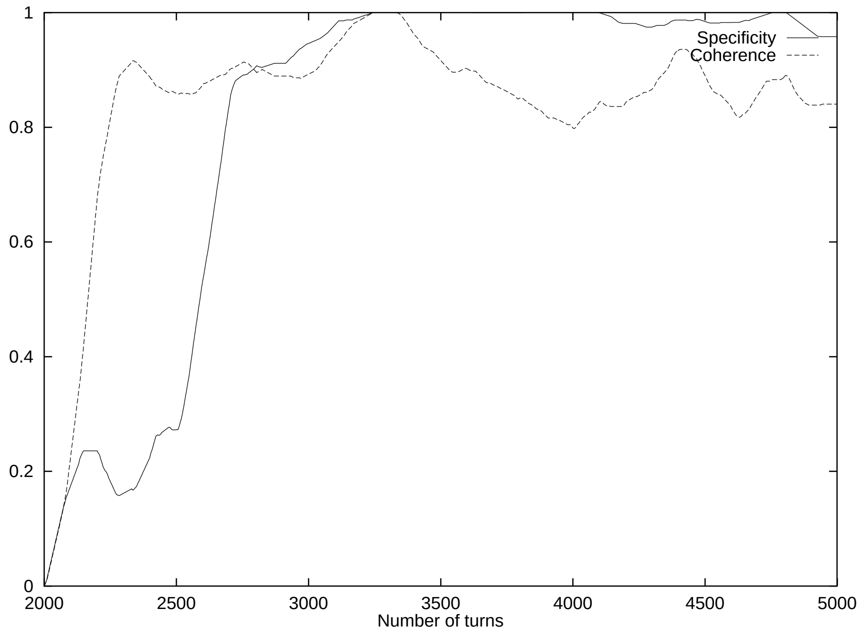


Figure 3: Coherence and specificity averaged over 200 turns as a function of time

One might expect that this ideal configuration of associations would be stable, since both occurence of signals and the accurate success predictions result in positive feedback. This is however not the case. The instability is the result of two factors. First, as noted above, signals are always delayed one turn. The other cause of change is the stochasticity in the signal selection process. An agent chooses to produce a signal using the strength of the association between the signals and the current situation as the relative probability of selecting that signal. Since several signals are associated to each situation, this effect causes continuous variation. Nonetheless, after the decrease corresponding to this effect, the coherence rises again and continues to vary between 0.8 and 1.0.

6 Discussion

In this paper, we investigated the question of whether a process of self organization may result in the development of a shared lexicon. The experiments have shown that this question can be answered positively. Local decisions of agents to adapt the relation between their perceived situation and the signals associated with it yielded a lexicon containing a different word for each situation that is relevant given the properties of the environment. Moreover, this lexicon was identical for all agents. The stochasticity involved in the choice of signals results in gradual but continuous changes of the lexicon over time.

The use of signals complements the information provided by the sensors. In future work, it would be interesting to see whether this principle can be observed in experiments. A possible setup for testing this involves partial perception, since this brings about a potential benefit for using signals to compensate for the missing information. In the example presented here, this could be investigated by limiting the visual field of the agents. In that case, not all agents will see the predator, and successfully guessing the presence of a predator based on the signals would allow these agents to take the appropriate action.

References

[Ackley and Littman, 1994] Ackley, D. H. and Littman, M. L. (1994). Altruism in the evolution of communication. In Brooks, R. A. and Maes, P., editors, *Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*. MIT Press/Bradford Book.

[Billard and Hayes, 1997] Billard, A. and Hayes, G. (1997). Learning to communicate through imitation in autonomous robots. In *Proceedings of the 7th International Conference on Artificial Neural Networks ICANN 97*. Springer Verlag.

[De Jong, 1998] De Jong, E. D. (1998). Coordination developed by learning from evaluations. AI-MEMO 98-13, Vrije Universiteit Brussel, AI Lab.

[De Jong and Vogt, 1998] De Jong, E. D. and Vogt, P. (1998). How Should a Robot Discriminate Between Objects? A comparison between two methods. In *Proceedings of the Fifth International Conference of The Society for Adaptive Behavior 98 (SAB '98)*, volume 5, Cambridge, MA. MIT Press.

[MacLennan, 1991] MacLennan, B. (1991). Synthetic ethology: An approach to the study of communication. In Langton, C. G., Taylor, C., Farmer, J., and Rasmussen, S., editors, *Artificial Life II, SFI Studies in the Sciences of Complexity*, volume X. Addison-Wesley.

[McCallum, 1996] McCallum, A. K. (1996). Reinforcement learning with selective perception and hidden state. *Ph.D. Thesis.*

[Moore and Atkeson, 1995] Moore, A. W. and Atkeson, C. G. (1995). The parti-game algorithm for variable resolution reinforcement learning in multidimensional state-spaces. *Machine Learning*, 21(3):199–233.

[Seyfarth et al., 1980] Seyfarth, R., Cheney, D., and Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210:801–803.

[Steels, 1996a] Steels, L. (1996a). Emergent adaptive lexicons. In Maes, P., editor, *From Animals to Animats 4: Proceedings of the Fourth International Conference On Simulating Adaptive Behavior*, Cambridge MA. The MIT Press.

[Steels, 1996b] Steels, L. (1996b). Perceptually grounded meaning creation. In Tokoro, M., editor, *Proceedings of the International Conference on Multi-Agent Systems*, Menlo Park CA. AAAI Press.

[Steels and Vogt, 1997] Steels, L. and Vogt, P. (1997). Grounding adaptive language games in robotic agents. In Husbands, C. and Harvey, I., editors, *Proceedings of the Fourth European Conference on Artificial Life*, Cambridge MA and London. MIT Press.

[Werner and Dyer, 1991] Werner, G. M. and Dyer, M. G. (1991). Evolution of communication in artificial organisms. In Langton, C. G., Taylor, C., Farmer, J., and Rasmussen, S., editors, *Artificial Life II, SFI Studies in the Sciences of Complexity*, volume X. Addison-Wesley.

[Yanco and Stein, 1993] Yanco, H. and Stein, L. (1993). An adaptive communication protocol for cooperating mobile robots. In Meyer, J-A., H. R. and Wilson, S., editors, *From Animals to Animats 2. Proceedings of the Second International Conference on Simulation of Adaptive Behavior*, pages 478–485, Cambridge MA. The MIT Press.