The learning barrier: Moving from innate to learned systems of communication

Michael Oliphant
Language Evolution and Computation Research Unit
Department of Linguistics, University of Edinburgh
George Square, Edinburgh EH8 9LL, UK
+44 131 650 3958 – FAX: +44 131 650 3962
oliphant@ling.ed.ac.uk

Submitted to Adaptive Behavior

Abstract

Human language is a unique ability. It sits apart from other systems of communication in two striking ways: it is syntactic, and it is learned. While most approaches to the evolution of language have focused on the evolution of syntax, this paper explores the computational issues that arise in shifting from a simple innate communication system to an equally simple one that is learned. Associative network learning within an observational learning paradigm is used to explore the computational difficulties involved in establishing and maintaining a simple learned communication system. Because Hebbian learning is found to be sufficient for this task, it is proposed that the basic computational demands of learning are unlikely to account for the rarity of even simple learned communication systems. Instead, it is the problem of *observing* that is likely to be central – in particular the problem of determining what meaning a signal is intended to convey.

1 The learning barrier

There is a long-standing tradition of treating the evolution of human language as being roughly synonymous with the evolution of syntax (Bickerton, 1981; Lieberman, 1984; Pinker and Bloom, 1990; Newmeyer, 1991). This position presumably reflects the assumption that, since other animals possess vocabulary-like systems of communication, all that is left to be explained is how humans evolved the ability to use syntactic structures. Other species, such as the vervet monkey, have quite adequate systems for communicating whether an intruding predator is a snake or an eagle, but they have no means of combining simple signals with each other to form more complex meanings. Put simply, they lack syntax.

Human language differs from the communication systems of other animals in another important way, however. Human language is learned, while the vast majority of existing communication systems involve an innate, genetically specified mapping between signal and meaning (Oliphant, 1998a). Although some communication systems involve a learned component, there are surprisingly few cases where the actual *mapping* from meaning to signal is truly plastic. Song in male oscine birds, for example, will not develop normally without exposure to the song of others (Marler, 1970). The learned intricacies of bird song do not, however, seem to be used make semantic distinctions. Aspects of the signal are being learned, rather than a relationship between signal and meaning. Another example of communicative behavior in which learning is involved is the alarm call system used by vervet monkeys. In this case, the animals learn to fine-tune the use of alarm calls through experience (Seyfarth and Cheney, 1986), but the basic mapping between a vervet alarm call and the predator it indicates appears to be specified innately (Hauser, 1996).

That communicative behavior tends to be encoded genetically is, perhaps, largely explained by the fact that natural selection provides a very effective means of tuning innate systems of communication. Game-theoretic approaches (Warneryd, 1993; Blume, Kim, and Sobel, 1993; Kim and Sobel, 1995; Skyrms, 1996) and an increasingly large literature of computational modeling work (Werner and Dyer, 1991; Oliphant, 1993; Oliphant, 1996; MacLennan and Burghardt, 1994; Ackley and Littman, 1994; Levin, 1995; Cangelosi and Parisi, 1996; Bullock, 1997; Werner and Todd, 1997; de Bourcier and Wheeler, 1997; Di Paolo, 1997; Noble, 1998), have given us a good understanding of how such innate mappings can be tuned by selection.

If the system is to be learned, rather than being directly specified genetically, the learning mechanism will now be responsible for establishing and maintaining coordination. The rarity of learned communication seems to indicate that this transition from an innate system to a learned one is non-trivial. This paper is an attempt to begin to understand why this might be the case. Using various network models, I will explore the computational demands of learned communication, looking at what abilities are required if a population of individuals is to be able to construct and use a simple system of communication.

2 Signaling systems and communicative accuracy

Because this paper is concerned with the basic requirements of learning *any* system of communication, a very simple model will be used. The model, introduced by Lewis (1969), is of a *signaling system*, which involves a simple mapping between *signals* and *meanings*. In this model (variants of which form the basis for most computational work on simple communication systems), meanings are unstructured, and signals are discrete tokens that cannot be combined. Successful communication depends on a conventional, unambiguous form/meaning mapping.

In such a model, the basic unit of analysis is a *communicative interaction*, which consists

Transmission

	a	*b*	*c*
1	1.0	0.0	0.0
2	0.0	0.6	0.4
3	0.0	0.4	0.6

Reception

a	*b*	*c*	
1.0	0.0	0.0	*1*
0.0	0.4	0.6	*2*
0.0	0.6	0.4	*3*

Figure 1: An example of a population's communication behavior for a system that uses three meanings (1, 2, 3) and three signals (a, b, c). The probability that an individual in the population will transmit a given signal for a given meaning is shown in the upper table. The lower table gives the probability a given signal will be interpreted as a given meaning. The communicative accuracy for this population is 0.65.

of an exchange between two individuals: a *sender* and a *receiver*. The sender, given a meaning, produces a signal. The receiver is then given this signal and interprets it as a meaning. The communicative interaction is said to be successful if the meaning the receiver interprets the sender's signal as the same meaning the sender was given.

Populations of interacting individuals will be evaluated by determining their *communicative accuracy* – the probability that communicative interaction will be successful between a randomly chosen sender and a randomly chosen receiver for a randomly chosen meaning. Figure 1 shows the behavior of an example population communicating about three meanings using three signals. The two tables give the average transmission and reception behavior of individuals in the population. Members of this population will, for example, always (with probability 1.0) send signal a in response to meaning 1, and will interpret signal b as meaning 2 with probability 0.4. Overall, the communicative accuracy of this population is 0.65, meaning that a communicative interaction can be expected to succeed 65% of the time.

A population will be said to communicate *optimally* if its communicative accuracy is 1.0 – the case when every individual communicates accurately with every other individual for every meaning. An example of an optimally communicating population is shown in Figure 2.

3 Observational learning

The majority of simulation approaches to the study of learned communication use some form of reinforcement learning paradigm (Yanco and Stein, 1993; Hutchins and Hazelhurst, 1995; Steels, 1996; Murciano and Millan, 1996). While it seems likely that reinforcement learning is involved to some degree, questions such as where the reinforcement signal comes

Transmission

	a	*b*	*c*
1	0.0	1.0	0.0
2	1.0	0.0	0.0
3	0.0	0.0	1.0

Reception

a	*b*	*c*	
0.0	1.0	0.0	*1*
1.0	0.0	0.0	*2*
0.0	0.0	1.0	*3*

Figure 2: A population that communicates optimally (a communicative accuracy of 1.0). Each meaning is expressed unambiguously with a single signal and the reception behavior of the population is such that all signals are correctly interpreted.

from and how costly it might be, as well as speed issues make it problematic as a primary learning mechanism (see Oliphant (1997) for a more in-depth discussion).

Because of the potential problems with reinforcement learning, the simulations I present in this paper will explore the computational issues that arise if a purely observational learning paradigm is used. In this framework, no reinforcement signal is used. The communicative behavior of a learning individual is based solely on observations of the behavior of others.

The model of observational learning I use is similar to that used by Hurford (1989) and Oliphant and Batali (1997). I assume that the life of an individual proceeds in two stages: a learning stage and a behaving stage. During the learning stage, an individual observes the behavior of the other individuals in the population, and uses these observations to construct its own communication system. After learning, this communication system remains fixed. At no point during the learning stage does an individual use its forming communication system and modify it based on feedback regarding its success (to do so would be to add an element of reinforcement learning). During the behaving stage, individuals interact with one another, providing the basis for the learning of new individuals.

Beginning with an initial population (generally seeded with randomized communication systems), a new individual is introduced at each time-step, or *round*, learning to communicate and becoming a behaving member of the population. In addition, an randomly chosen individual is removed, keeping the population at a constant size. This occurs in a continuous cycle, as is shown in Figure 3.

4 Evaluating a learning mechanism

There are many possible ways in which an individual can learn, using observations of the behavior of others to determine its own behavior. In evaluating a learning mechanism, I will require more of it than is generally demanded. A suitable learning mechanism must

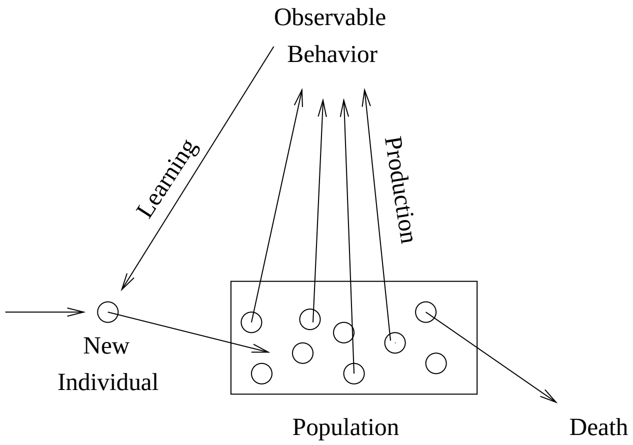

Figure 3: The learning cycle. Old individuals are continually replaced by new ones in the population. New individuals learn from an observed sample of behavior produced by the existing population.

satisfy three requirements:

- **Acquisition:** A learning mechanism must be able to acquire a preexisting optimal system of communication when it is introduced into a population that uses it.
- **Maintenance:** A learning mechanism must be able to maintain a preexisting optimal system of communication against reasonable levels of noise.
- **Construction:** A learning mechanism must improve a non-optimal system of communication in such a way that, as new individuals who use the learning mechanism are added to the population, the communicative accuracy of the population increases and eventually reaches an optimal state.

Thus we are not only interested in how a new individual might acquire an existing system, but also how such systems are created in the first place and maintained over time. The problems of acquisition, construction and maintenance are seen as being intrinsically linked to one another.

5 Network learning mechanisms

In the simulations I will present, learning takes place using simple networks operating within an associative learning framework. In this case, the association to be learned is between signals and meanings. The general network architecture can be seen in Figure 4. A signal is represented on one layer of the network by activating a single unit. Meanings are represented in similar fashion on the other layer. Associations between signal an meaning are represented by the bidirectional weights that connect the units of the signal layer to the units of the meaning layer. More formally, the networks consist of a set of signal units,

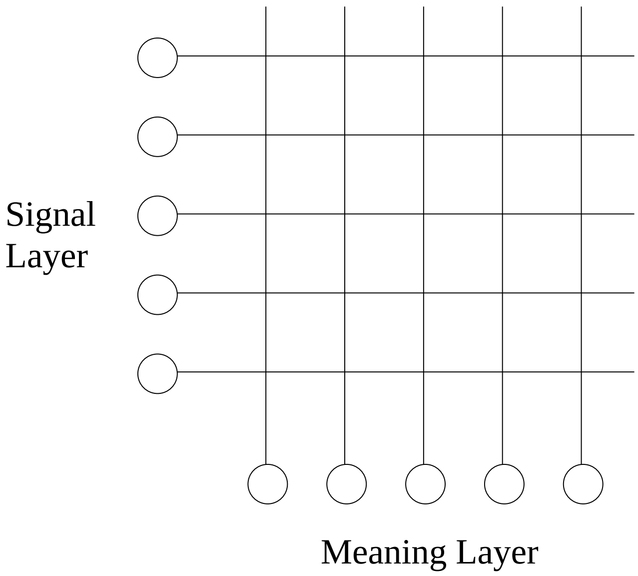

Figure 4: The associative network model. Signals and meanings are each represented by a layer of units, with the interconnecting weights storing the association between them.

S, and a set of meaning units, M. Individual units will be referred to as S_i and M_j, with w_{ij} designating the weight connecting signal unit S_i and meaning unit M_j.

New networks begin with all weights set to zero, producing random initial behavior through absence of any particular bias. All learning is done based on observed samples of transmission behavior (an individual producing a signal for a given meaning). Because the bidirectional weights in the network imposes an inherent link between transmission and reception behavior, observations of either behavior are sufficient. New individuals entering the population are exposed to three samples of signals produced in response to each meaning.[1] When an individual is used as a model for a learner, it is also trained on its own response. This is necessary to ensure consistency in cases where the network has no bias in the current situation (something that happens very often in the early rounds, where the population consists almost entirely of the initial, unbiased networks used to begin the simulation).

While the learning rule that modifies the weights in response to the presentation of signal/meaning pairs will vary depending on the type of network, all of the simulations presented use a winner-take-all output strategy. Given a particular input pattern, the most highly active output unit is set to be active, while the other units are turned off. Thus, to use the network to interpret a particular signal, the unit corresponding to signal is activated (set to 1.0), and the output unit, with the largest net input:

$$a_j = \sum_i S_i w_{ij} \tag{1}$$

[1] Three exposures is the minimum number required to allow discrimination between competing patterns based on their frequency.

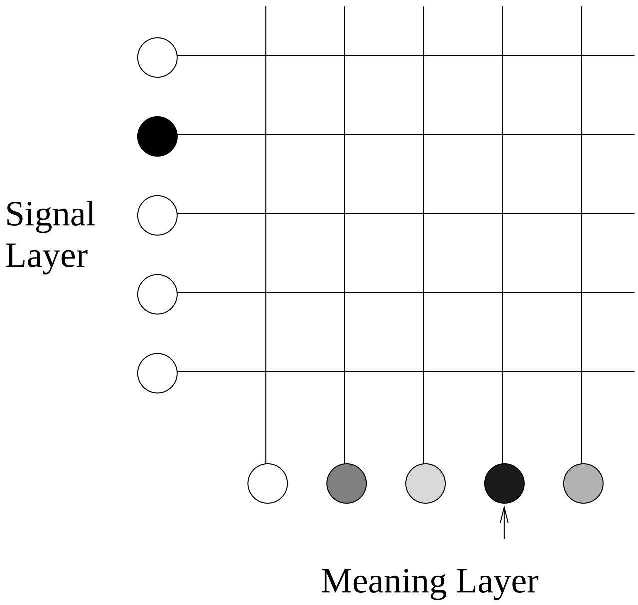

Figure 5: Winner-take-all recall. A particular signal is being interpreted by the network. The resulting meaning is the mostly highly activated unit, as indicated by the arrow.

is the winner. This procedure is diagrammed in Figure 5. Recall operates in a corresponding way for transmission behavior.

6 Willshaw networks

Perhaps the most basic kind of associative networks are Willshaw networks, designed to associate pairs of sparse binary patterns (Willshaw, Buneman, and Longuet-Higgins, 1969; Willshaw, 1971). The learning rule used by these networks simply sets a weight, w_{ij}, to 1 if both S_i and M_j are activated for a given pair of input and output patterns. This learning rule is diagrammed in Figure 6.

Used as the learning/production mechanism for a population of individuals in an observational learning simulation as outlined in Section 3, Willshaw networks are capable of learning an existing communication system. If the initial population is seeded with individuals that are communicating successfully, the communication system will be accurately learned by new individuals as long as they are given sufficient exposure (at least one observation of a signal being sent for each meaning is required). This is unsurprising, as the data set that the networks need to learn is of the simplest possible form. The sets of vectors representing both the signals and the meanings are orthogonal, resulting in no inter-correlations and a very easy learning task.

If, however, a modest amount of noise is introduced into the system, performance degrades over time. Figure 7 shows how communicative accuracy quickly drops to chance levels if a learner is prone to misobserve a signal/meaning pair 5% of the time. This occurs

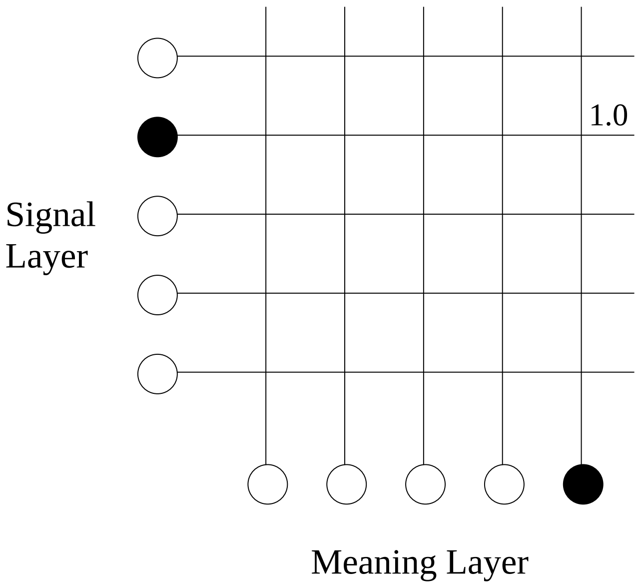

Figure 6: Training in a Willshaw network. A weight is set to 1 if the input and output units it connects are both active.

because there is no way for a Willshaw network to represent one of two existing associations as being stronger than another. The network is unable to distinguish a low level of noise from the stronger, true behavior of the population. The effects of noise accumulate over time, resulting in an eventual destruction of the previously optimal communication system given even the smallest amount of noise.

For similar reasons, Willshaw networks are unable to improve the communicative accuracy of a population that is initially communicating randomly, even given a noise-free environment. Performance remains at chance levels, as can be seen in Figure 8.

7 Cumulative-Association networks

To correct the problem that Willshaw networks have in discriminating different levels of association, one can simply increment a weight when an association is perceived, rather than just setting the weight to 1. The weight update rule for such a network, which I will call a *Cumulative-Association* network, is as follows:

$$\Delta w_{ij} = \begin{cases} 1 & \text{if } S_i = 1 \text{ and } M_j = 1; \\ 0 & \text{otherwise;} \end{cases} \tag{2}$$

This new learning rule is able to maintain a communication system against reasonable rates of noise (where the noise does not drown out the signal). It also improves the communicative accuracy of an initially random population of networks, as can be seen in solid line of the graph shown in Figure 9.

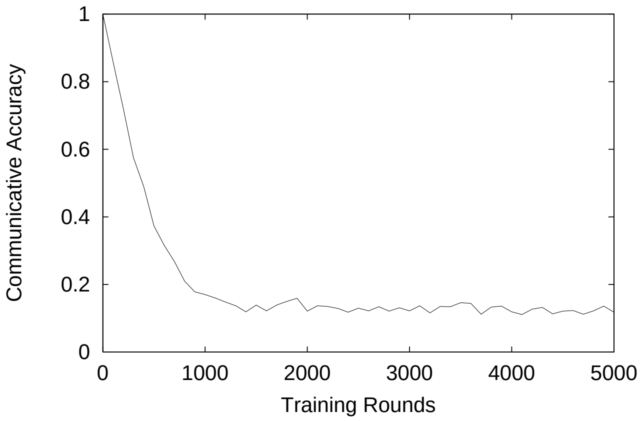

Figure 7: Failure of the Willshaw network to maintain a communication system under noisy conditions. The simulation is initially seeded with an optimally communicating population. Given a 5% chance of misobservation, a perfect communication system degrades to chance performance. Ten signals and ten meanings are used. Results are averaged over ten simulation runs. All simulations involve populations of 100 individuals, unless otherwise specified.

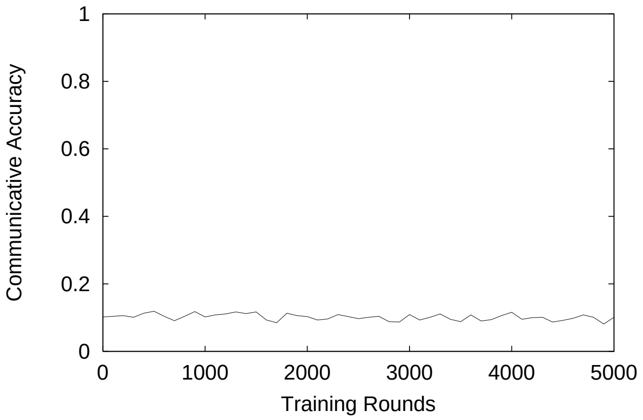

Figure 8: Failure of the Willshaw network to construct a communication system. Beginning with a population of blank networks, the Willshaw learning rule is unable to improve communicative success. Ten signals and ten meanings are used. Results are averaged over ten simulation runs.

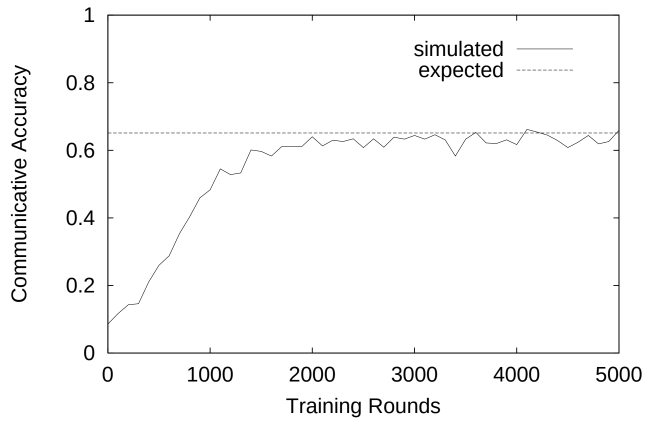

Figure 9: Performance of the Cumulative-Association network, and the expected value based on Equation 3. Performance is better than chance, but the Cumulative-Association learning rule fails to reliably produce an optimal system from random initial conditions. Ten signals and ten meanings are used. Results are averaged over ten simulation runs.

The Cumulative-Association learning rule does not, however, result in a population that communicates optimally. It fails because it has no pressure to avoid ambiguously using the same signal for multiple meanings. Given any random initially tendency to use a particular signal for a particular meaning, this tendency will be exaggerated in the population until that signal is used conventionally. There is nothing to prevent the signal of choice being the same for more than one meaning, however. An example of the kind of sub-optimal system that results is shown in Figure 10.

The expected performance of the Cumulative-Association learning rule can be mathematically predicted. Because biases in the random communication behavior of the initial population will simply be exaggerated over time, this effectively results in the population using a randomly selected signal for each meaning. The expected number of unique signals resulting from randomly choosing from s signals to represent m meanings can be calculated as follows:

$$n_u = s(1 - (1 - (1/s))^m) \tag{3}$$

This allows for accurate communication about n_u of the m meanings, resulting in an expected communicative accuracy of n_u/m. This expected value is plotted with the simulated results in Figure 9. The simulation results conform very closely to the level of accuracy that is predicted.

8 Hebbian networks

It turns out that a very simple way to provide a tendency to use a unique signal for every meaning in a network is the addition of lateral inhibition to the weight update rule. Adding lateral inhibition to the Cumulative-Association rule described in the previous

Transmission

	a	b	c
1	1.0	0.0	0.0
2	0.0	1.0	0.0
3	0.0	1.0	0.0

Reception

a	b	c	
1.0	0.0	0.0	*1*
0.0	0.5	0.0	*2*
0.0	0.5	0.0	*3*

Figure 10: An example of the communicative behavior converged upon by a population of Cumulative Association networks for a system of three signals and three meanings. While meaning 1 is uniquely conveyed by signal a, signal b is used for both meanings 2 and 3, while signal c is not used at all. Individuals in this population will communicate successfully with each other two times out of three.

section results in a form of Hebbian learning (Hebb, 1949). Inhibition is implemented by decreasing the strength of a connection between a signal unit and a meaning unit if one, but not both of them are active. This results in the following weight update rule:

$$\Delta w_{ij} = \begin{cases} 1 & \text{if } S_i = 1 \text{ and } M_j = 1; \\ 0 & \text{if } S_i = 0 \text{ and } M_j = 0; \\ -1 & \text{otherwise;} \end{cases} \tag{4}$$

This new update rule is diagrammed in Figure 11.

It is important to note that this update rule does not increase the weights if both units are not firing, as is done in the most common formulation of the Hebbian learning rule.[2] This results in better performance, and is in fact more compatible with the original hypothesis of Hebb (1949). It is also important that the networks use binary units rather than the signed (+1,−1) units that are used for mathematical convenience in the most standard formulation of Hebbian networks. Aside from the resulting problems in making analogies with real neural activity, using signed units in a task where patterns are sparse creates a great deal of spurious correlation. Two different patterns, each with one unit active, will be correlated in all but two of their units. This makes the task of the network unnecessarily difficult, and use of signed units results in very poor performance.

The performance of populations of networks using this formulation of the Hebbian learning rule as shown in (4) is presented in Figure 12. The addition of lateral inhibition to the network learning rule results in the ability to produce an optimal communication system in a population that begins with random behavior. This ability to construct a communication system necessarily entails the ability to acquire and maintain an existing

[2] A learning rule similar to the one used here was used by Billard and Dautenhahn (1997) to coordinate the activity of two robotic agents – one the teacher and one the learner – engaged in a simple following task.

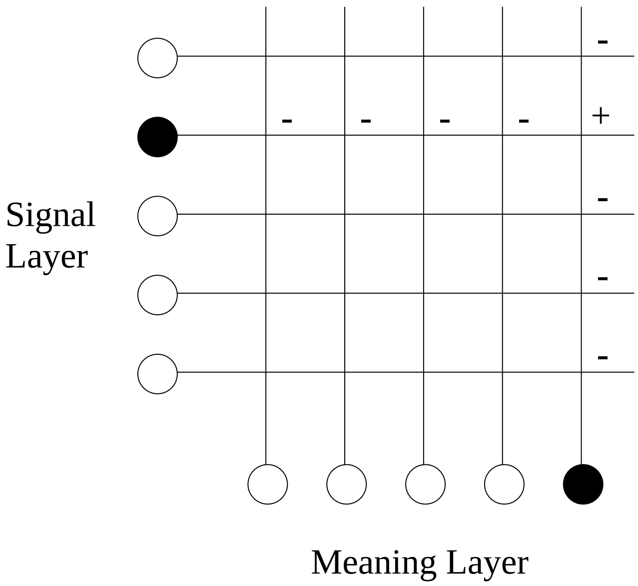

Figure 11: Training in a Hebbian network. A weight is increased if the input and output units it connects are both active. The weight is decreased if one, but not both of the units are active.

system. After a communicative accuracy of 1.0 is reached soon after round 1000, new individuals are accurately acquiring this optimal system. Maintenance against noise can essentially be thought of as continual reconstruction of the system each time it is degraded by noise. The Hebbian learning rule, then, satisfies all three of the requirements set out in Section 4, providing a possible mechanism for the use of a simple system of learned communication.

8.1 Scalability

The performance of the Hebbian network scales well for increased numbers of meanings. Figure 13 shows that, while the time it takes a population to converge on an optimal system increases with the number of meanings, optimal communication is nevertheless always achieved. The increase in convergence time is likely due to a lowered starting point – the chance level of communicative accuracy decreases as the number of meanings goes up.

It is important to be clear that the increase in time shown here is in terms of the number of rounds (replacements of an existing individual with a blank, new individual) it takes the population to reach an optimal state, not the amount of time that it takes each new individual to learn. Learning involves observing three communicative interactions based on the transmission of each meaning. Thus, as the number of meanings increase, so does the number of observed samples of behavior. This is a result of the number of observations required to get a representative sample of a larger system, and is not reflected by the

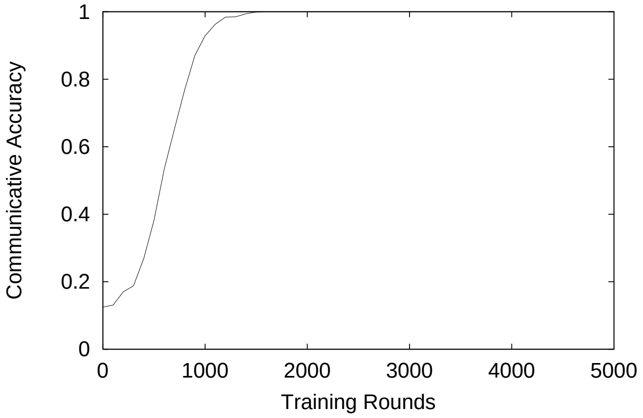

Figure 12: Performance of the Hebbian network. Populations of Hebbian networks are able to construct optimal communication systems. Ten signals and meanings are used. Results are averaged over 10 simulation runs.

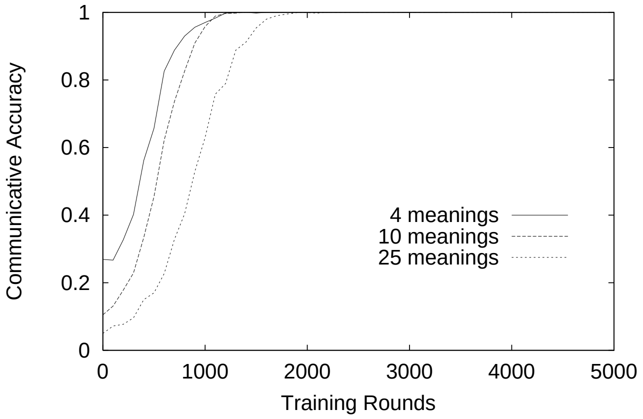

Figure 13: Performance of the Hebbian network using varying numbers of signals and meanings. The number of signals and meanings is equal in each case. Results for each plot are averaged over ten simulation runs.

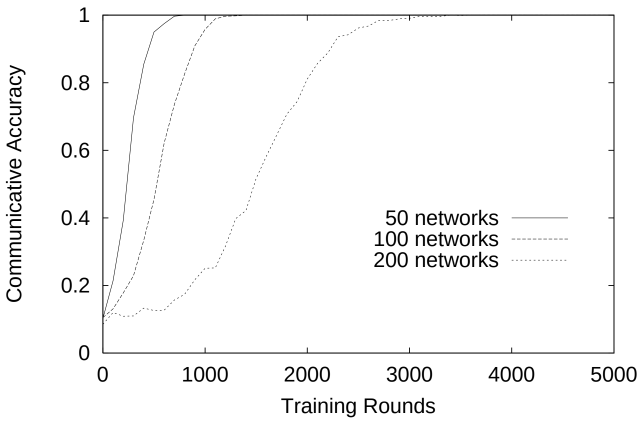

Figure 14: Network populations of varying sizes establishing a conventional encoding for ten input patterns. All plots are averaged over ten simulation runs.

number of rounds.

Time to convergence is also affected by the size of the population, as can be seen in Figure 14. As the number of individuals in the population increases, so does the number of rounds required to reach an optimal state of communication. This increase seems to be linear in the size of the population, and likely reflects the additional time to reach consensus in a larger population.

8.2 Multiple-unit meanings

Using Hebbian networks, it is also possible to deal with more complex meaning patterns than the ones used in the previous simulations. Instead of turning a single unit on in the vector to represent a particular meaning, a pattern across multiple units can be used. This allows the meanings to be structured according to some task-based metric. It also provides a way in which meanings can be more or less similar to each other as patterns can overlap, sharing some of the same units.

The addition of structure transforms the problem into a form of vector quantization (Kohonen, 1989). The difficulties that arise because of correlations between the new structured meanings can be avoided by adding a conscience mechanism to the winner-take-all output threshold, as is done in other vector quantization tasks (Grossberg, 1976; Bienenstock, Cooper, and Munro, 1982). Additional details about this type of network, and results demonstrating its performance can be found in Oliphant (in preparation).

9 Discussion

The simulations that I have presented demonstrate that associative network learning is capable of supporting a learned system of communication in a population of individuals. This result is particularly significant in that the learning framework used was purely observational. Although I have shown plots of communicative success of the simulated populations, at no point are such measures used to provide learners with a reinforcement signal. The success of an observational strategy indicates that, while an error signal might be used if one was available, reinforcement is not strictly required to learn to communicate.

Not just any observational learning mechanism is sufficient, however. Learning strategies that simply exaggerate initial random tendencies in the populations behavior will fail to reliably result in populations that communicate successfully. Instead, a learning mechanism must be such that it pushes the population's communicative behavior toward a state where each meaning is expressed unambiguously. This is exactly what the Hebbian learning rule does.

Hebbian learning is perhaps the most simple, biologically plausible learning mechanism one could ask for. This would seem to place the ability to utilize learned communication within the means of virtually any animal species. Given that such a simple form of learning seems sufficient both to acquire an existing system of communication and to construct such a system in the first place, why are learned communication systems so rare?

One possible answer to this question lies in one of the assumptions that is inherent in the model of observational learning that I have presented. I assume that a learner has access to the behavior of others in the form of signal/meaning pairs. This assumption, while making it possible to actually do the simulation, is likely to be unjustified in general. In particular, it seems unrealistic for a learner to have easy access to the meanings that others are using their signals to convey. If a child is out for a walk with its mother, and hears the word "dog", how is she to know that her mother is referring to the furry, brown creature wagging its tail, rather than the wagging tail itself, or even the fact that it is a sunny day?

The difficult aspects of observational learning might then have less to do with *learning* than they do with *observing*. This paper has given an analysis of the computational requirements of the learning mechanism, but it remains an open question how learners extract the signal/meaning pairs that they are to associate. To begin to answer this question, we should turn to the large body of literature that exists on child language acquisition. When children learn words, they seem to simplify the task of deciding what a word denotes through knowledge of the existence of taxonomic categories (Markman, 1989), awareness of the pragmatic context (Tomasello, 1995), and reading the intent of the speaker (Bloom, 1997).

The ability that children have to accurately determine what meaning an utterance might refer to is of critical importance in the learning task. Perhaps it is an inability to constrain the possible space of meanings that prevents animals from using learned systems of communication, even systems that are no more complicated than existing innate signaling systems. If so, the development of an effective observational learning strategy may represent an important evolutionary milestone in moving from an innate system of communication to one that is learned.

It is possible that the ability to communicate by combining sequences of signals requires cognitive skill that other animals do not have, significantly above and beyond the ability to learn observationally. It is also possible, however, that observational learning is the primary factor limiting the evolution of language ability. Perhaps, once animals can learn to communicate by observing others, the achievement of syntactic communication is an extension that is comparatively less difficult. Recent work involving computational simulations of the origins of syntax supports this view (Batali, 1997; Hutchins and Hazelhurst, 1997; Worden, 1997; Kirby, 1998).

In either case, it is a mistake to take the existence of simple innate communication systems in other species to imply that the problem of the lexicon in language evolution is solved. To do so is to ignore the important distinction between innate and learned communication. The rarity of simple, learned systems of communication suggests that shifting from an innate system to an equally simple learned one is a difficult task, and one worthy of careful study.

References

Ackley, D. and M. Littman (1994). Altruism in the evolution of communication. In R. Brooks and P. Maes (Eds.), *Proceedings of the fourth artificial life workshop*, Cambridge, MA, pp. 40–48. MIT Press.

Batali, J. (1997). Computational simulations of the emergence of grammar. In J. Hurford, C. Knight, and M. Studdert-Kennedy (Eds.), *Evolution of Language: Social and Cognitive Bases for the Emergence of Phonology and Syntax*. Cambridge University Press.

Bickerton, D. (1981). *Roots of language*. Ann Arbor, MI: Karoma.

Bienenstock, E., L. Cooper, and P. Munro (1982). Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *Journal of Neuroscience 2*, 32–48.

Billard, A. and K. Dautenhahn (1997). The social aspect of communication: a case study in the use and usefulness of communication for embodied agents. In *Fourth European Conference on Artificial Life*, Brighton, England.

Bloom, P. (1997). Intentionality and word learning. *Trends in cognitive sciences 1*(1), 9–12.

Blume, A., Y. Kim, and J. Sobel (1993). Evolutionary stability in games of communication. *Games and Economic Behavior 5*, 547–575.

Bullock, S. (1997). An exploration of signaling behaviour by both analytic and simulation means for both discrete and continuous models. In *Fourth European Conference on Artificial Life*, Brighton, England, pp. 454–463.

Cangelosi, A. and D. Parisi (1996, January). The emergence of a 'language' in an evolving population of neural networks. Technical Report NSAL-96-004, Institute of Psychology, National Research Council, Rome.

de Bourcier, P. and M. Wheeler (1997). The truth is out there: The evolution of reliability in aggressive communication systems. In *Fourth European Conference on Artificial Life*, Brighton, England, pp. 444–453.

Di Paolo, E. (1997). An investigation into the evolution of communication. *Adaptive Behavior 6*(2), 285–324.

Grossberg, S. (1976). Adaptive pattern classification and universal recoding: Ii. feedback, expectation, olfaction, illusions. *Biological Cybernetics 23*, 187–202.

Hauser, M. (1996). *The evolution of communication*. Cambridge, Massachusetts: MIT Press.

Hebb, D. (1949). *The organization of behavior*. New York: John Wiley & Sons.

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

Hutchins, E. and B. Hazelhurst (1995). How to invent a lexicon: the development of shared symbols in interaction. In N. Gilbert and R. Conte (Eds.), *Artificial Societies: The computer simulation of social life*. London: UCL Press.

Hutchins, E. and B. Hazelhurst (1997). The emergence of propositions from the coordination of talk and action in a shared world. forthcoming.

Kim, Y. and J. Sobel (1995). An evolutionary approach to pre-play communication. *Econometrica 65*(5), 1181–1193.

Kirby, S. (1998). Language evolution without natural selection: From vocabulary to syntax in a population of learners. Technical Report EOPL-98-1, Department of Linguistics, University of Edinburgh.

Kohonen, T. (1989). *Self-organization and associative memory* (3rd ed.). Berlin: Springer-Verlag.

Levin, M. (1995). The evolution of understanding: a genetic algorithm model of the evolution of communication. *BioSystems 36*, 167–178.

Lewis, D. (1969). *Convention: A philosophical study*. Harvard University Press.

Lieberman, P. (1984). *The biology and evolution of langauge*. Havard University Press.

MacLennan, B. and G. Burghardt (1994). Synthetic ethology and the evolution of cooperative communication. *Adaptive Behavior 2*(2), 161–187.

Markman, E. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Marler, P. (1970). A comparative approach to vocal learning: song development in white-crowned sparrows. *Journal of Comparative and Physiological Psychology 71*, 1–25.

Murciano, A. and J. Millan (1996). Learning signaling behaviors and specialization in cooperative agents. *Adaptive Behavior 5*(1), 5–28.

Newmeyer, F. J. (1991). Functional explanation in linguistics and the origins of language. *Language and Communication 11*, 3–28.

Noble, J. (1998). Evolved signals: Expensive hype vs. conspiratorial whispers. Paper to be presented at the ALive VI conference.

Oliphant, M. (1993). Conditions for the evolution of saussurean communication. Master's thesis, University of California, San Diego.

Oliphant, M. (1996). The dilemma of saussurean communication. *BioSystems 37*(1-2), 31–38.

Oliphant, M. (1997). *Formal approaches to innate and learned communication: Laying the foundation for language.* Ph. D. thesis, University of California, San Diego.

Oliphant, M. (1998a). Rethinking the language bottleneck: Why don't animals learn to communicate? Paper presented at the 2nd International Conference on the Evolution of Language.

Oliphant, M. (1998b). Self-organized coordination in populations of interacting associative networks. Manuscript in preparation.

Oliphant, M. and J. Batali (1997). Learning and the emergence of coordinated communication. *Center for Research on Langauge Newsletter 11*(1).

Pinker, S. and P. Bloom (1990). Natural language and natural selection. *Behavioral and Brain Sciences 13*, 707–784.

Seyfarth, R. and D. Cheney (1986). Vocal development in vervet monkeys. *Animal Behavior 34*, 1640–1658.

Skyrms, B. (1996). *Evolution of the social contract.* Cambridge University Press.

Steels, L. (1996). Self-organizing vocabularies. In *Proceedings of the V Alife Conference*, Nara, Japan.

Tomasello, M. (1995). Pragmatic contexts for early verb learning. In M. Tomasello and W. Merriman (Eds.), *Beyond Names For Things: Young Children's Acquisition of Verbs*. Lawrence Erlbaum.

Warneryd, K. (1993). Cheap talk, coordination, and evolutionary stability. *Games and Economic Behavior 5*, 532–546.

Werner, G. and M. Dyer (1991). Evolution of communication in artificial organisms. In C. Langton, C. Taylor, J. Farmer, and S. Rasmussen (Eds.), *Artificial life II*, pp. 659–687. Redwood City, CA: Addison-Wesley.

Werner, G. and P. Todd (1997). Too many love songs: Sexual selection and the evolution of communication. In *Fourth European Conference on Artificial Life*, Brighton, England, pp. 434–443.

Willshaw, D. (1971). *Models of distributed associative memory.* Ph. D. thesis, University of Edinburgh.

Willshaw, D., O. Buneman, and H. Longuet-Higgins (1969). Non-holographic associative memory. *Nature 222*, 960–962.

Worden, R. (1997). The evolution of language from social intelligence. In J. Hurford, C. Knight, and M. Studdert-Kennedy (Eds.), *Evolution of Language: Social and Cognitive Bases for the Emergence of Phonology and Syntax*. Cambridge University Press.

Yanco, H. and L. Stein (1993). An adaptive communication protocol for cooperating mobile robots. In J. Meyer, H. Roitblat, and S. Wilson (Eds.), *From Animal to Animats 2: Proceedings of the Second International Conference on Simulation of Adaptive Behavior*. MIT Press.