National University of Ireland, Galway
Ollscoil na hÉireann, Gaillimh

Department of Information Technology

technical report NUIG-IT-131002

Language Evolution in Artificial Systems

D. Curran (NUI, Galway)
C. O'Riordan (NUI, Galway)

itreports@nuigalway.ie http://www.it.nuigalway.ie/

Language Evolution in Artificial Systems

Dara Curran and Colm O' Riordan
Dept. of Information Technology,
National University of Ireland, Galway.

Abstract

Questions regarding language and its evolution have been addressed in many disciplines. In this paper we discuss various approaches and review work in the domain of language evolution in artificial systems. Using artificial systems to study concepts regarding language evolution is useful in that we can control the system and hence study language evolution in a closed controllable environment.

1 Introduction

Language provides human beings with a capacity of adaptation and creativity unrivaled in the natural world. The relatively sudden burst in human activity, giving rise to civilisations, is largely attributed to the discovery or refinement of language. There are many approaches to the study of language both in the human and animal world, each with its own inherent difficulties. The rise of artificial intelligent systems powerful enough to simulate artificial, yet complete, worlds is providing new insights into how language may have evolved and its effect on the performance of both individuals and populations.

This paper will examine some of the traditional approaches to the study of language, introduce synthetic ethology - a new artificial intelligence approach to linguistics, and examine in turn a number of experiments showing interesting results and directions for future research.

2 Synthetic Ethology

An approach known as synthetic ethology[1, 2] argues that the study of language is too difficult to perform in real world situations and that more meaningful results can be produced by modelling organisms and their environments in an artificial manner. The following sections briefly outline the more traditional approaches to the study of language from which synthetic ethology stems.

2.1 Philosophical

The philosophical approach to the study of language is sub-divided into two theories. The first, denotational theory, argues that the meaning of a word is the same as the object or concept it denotes. Such a system is highly restrictive and does not allow for ambiguities which naturally occur even in the most simple language systems[3].

The second argument states that meaning emerges from a shared pool of beliefs, (also termed a cultural background). While this approach may be closer to the truth, the difficulty of extracting language specific insights from culture is, particularly for humans, extremely difficult.

2.2 Behaviourist

In order to better understand the complexity and origins of human cultural interactions, much research has been undertaken in the domain of animal communication. This is usually accomplished by observing animal behaviour in captivity. The main drawback is that this tends to give rise to skewed results as animals never behave normally in such conditions[3]. Typically, such approaches only examine very restricted portions of animal behaviour, particularly as a result of the difficulty in providing stable environments, where the experimenter is in control of the vast number of variables involved.

2.3 Ethological

Ethology recognises the need to assess an individual's behaviour with relation to its environment. It has been argued that these two elements are tightly coupled and that any study of animal or human communication should take into account the type of environment in which such communication occurs[4]. Hence, ethological studies must take place in an animal's natural environment, leading to difficulties with regards to logistics and reliability.

2.4 Neuro-Psychological

The neuro-psychological approach examines the actual physical mechanism by which communication or language can occur. Due to its inherent biological nature, this type of study is highly complex, particularly in view of the fact that only portions of the brain's inner workings are understood. Also, focusing solely on biological factors does not take account of environmental factors which may have a substantial role in both the development and emergence of the very biological systems the approach attempts to study.

2.5 Artificial Intelligence

As a result of the shortcomings of such approaches and thanks to the increase in efficiency and accessibility of computing power, artificial intelligence methods have begun to significantly explore the question of language evolution. Artificial intelligence systems can create tightly controlled environments where the behaviour of artificial organisms can be readily observed and modified. Using evolutionary computation techniques, inspired by Darwinian evolution, and the computing capacity of neural networks, artificial intelligence researchers have been able to achieve very interesting results. However, before illustrating these results, it is worth examining the theories which have been proposed from earlier works relating to the origin of language.

3 Theoretical Origins of Language

3.1 Genetic Evolution

The genetic make-up of all organisms provides for the transmission of information necessary to build up a complex system. Through crossover and mutation, there are enough novelty generating factors to explain, to some degree, where our language capability possibly arose. The Chomskian school of linguistics argues that there is a language organ within our brain structure which includes a Language Acquisition Device (LAD) and the mechanisms for producing and parsing sentences[5].

The LAD is not refined to a particular language, or even a particular language structure, but is rather a basic system which gives humans their innate ability to learn any language as infants. The argument for the LAD stems from the poverty of stimulus theory. This theory proposes that human children are not exposed to enough samples of language during their early learning period to be able to produce novel sentences. The fact that they are able to do so suggests that there is an evolved mechanism which has already performed much of the required groundwork. Further evidence towards its existence is the alleged discovery of language genes which, when damaged or missing, severely impair language learning ability, suggesting that they are perhaps contributing to the growth of the LAD.

The genetic evolution of language is known as the phylogenetic approach, and while popular, has been challenged[2].

3.2 Adaptation and Self-Organisation

Whilst the phylogenetic approaches concentrate on the physical genetic explanation for language evolution, the following approaches focus on the language community as a whole. The argument is that the language community behaves as a complex adaptive system and that no language organ exists. All information regarding a language is stored in memory and is passed on through cultural transmission, not through genetic evolution.

Novelty, which is provided by crossover and mutation in the genetic approach, is achieved through errors in the cultural transmission process which comprises imitation learning. The structure of an evolved language is shaped in a selective fashion by a number of criteria including the maximisation of communicative success, the minimisation of memory required to learn the language and the capability of the sensory organs at the organism's disposal.

Language achieves its coherence through self-organisation, much like other parts of an organism's body. As the language improves, the communicative success of the language community increases, generating a positive feedback loop. The evolution and growth of the language is a side effect of this cultural transmission and adaptation. These proposals are classed generally as glossogenetic.

3.3 Genetic Assimilation

The last main approach takes portions of the preceding two and combines them with a phenomenon first explored by Baldwin named the Baldwin effect[6]. The Baldwin effect takes place when a phenotypically fit population of organisms transmits a phenotypically acquired trait onto its genetic code by a combination of mutations or crossover. The transmission produces a number of organisms which have a innate, genetic ability to perform the action that had to be previously acquired through lifetime learning. Such organisms are able to outperform others which have not taken this step and thus the gene propagates through natural selection, increasing the population fitness.

It has been argued that the LAD could have initially developed through cultural transmission and eventually became genetically encoded, creating the innate ability towards language exhibited by human beings.

4 Experiments

The following experiments involve a setup comprising a population of artificial organisms usually consisting of neural networks or finite state machines. Each organism typically has a genetic code which is used for reproduction through mutation and recombination. In addition, the experiments usually follow the guidelines first set out by Hutchins and Hazlehurst[7, 8]:

- The limited rationality constraint: No mind may influence another except via a mediating structure (no telepathy).
- The distributed system constraint: no agent has a complete view of all other agents.
- The open system constraint: there should be an influx/out flux of agents throughout simulations to examine the stability of an emerging communication system.
- No social mind can become organised except via interaction with other or its environment
- The nature of a mental representation cannot be assumed, but must be explained.

4.1 Evolution of Communication

MacLennan et al[1] performed experiments to examine how communication could evolve and, in particular, stabilise in a population of artificial organisms. Agents in the population are born with an innate ability to produce and receive communication signals. The fitness of each is determined by its ability to understand and respond correctly to signals produced by others. Each agent has a signal-situation table which maps a signal to its meaning, or situation. These tables are genetically encoded, but once an agent is born, modifications to the table are possible.

The linguistic interaction between agents takes place as follows: an agent perceives a situation in its environment, maps it to a signal using its table and produces the appropriate signal in response. Another agent then perceives this signal and performs the same look-up task. Communication is deemed successful if the agent's mapping corresponds to the actual situation.

The experiment makes a number of assumptions which have been criticised. Firstly, a successful communication always has a direct benefit. This is not emergent but is hard coded into the experiment from the beginning. Others have made this feedback more indirect by associating it with a specific task (such as foraging food or finding a mate)[9, 8, 10]. Another assumption is that the same signal is used for the reception and emission of a particular message. This restricts the possible variability of the communication system.

4.2 Phonetics and Phonology

These experiments are concerned with the evolution of sound producing and perceiving mechanisms required for communication. It has been argued that humans possess an innate ability, which may form part of the LAD, to produce and understand certain sound features[2]. Previous experiments focusing on this domain have employed a genetic approach and have largely been unsuccessful. This can be explained by viewing the situation where a particular individual creates a new sound. The new sound does not bring the individual selective advantage because the remaining population is not yet familiar with it. Only when a number of individuals create the same sound does it become adopted by the community. This implies that a number of individuals would have to undergo similar mutations, which makes such a situation unlikely to occur.

Newer approaches have chosen an adaptive systems architecture. These argue that the mapping between the signal and meaning evolves through cultural transmission and adaptation rather than genetics. Boers[11] conducted an experiment where a sound system was evolved through self organisation through repeated imitation interactions.

The organisms in the experiment have no innate phonetic structure but do have the ability to produce and receive sounds. In other words, they are capable of hearing sounds, but have no innate structure which understands them. The interactions take place as follows: one agent produces a sound from its repertoire. Initially this will be empty so the agent produces some random sound. Another agent, the imitator, attempts to mimic the sound it has heard. The initiator interprets the received sound and gives negative or positive feedback. If the interaction fails, the imitator attempts to create a new sound if what it has heard is too distant from its existing repertoire, or attempts to modify a sound in its repertoire to fit the perceived sound.

The experiment showed that a sound system can emerge without genetic interaction or innate perceptual ability, through a process of self organisation.

4.3 Lexicon Development

The study of communication in artificial populations has led some researchers to include fixed lexicons as a part of their experiment[12, 13]. While this has provided a useful starting point, others argue that the use of a fixed lexicon is not representative of real world language development. Much research has been done focusing on a dynamic lexicon in a population of communicating organisms[8, 14, 15, 16, 17, 18].

The population in an experiment conducted by Hutchins and Hazlehurst[8] consists of neural networks with an additional structural feature. In addition to a hidden layer, the networks contain a hidden verbal input/output layer. This extra layer performs the task of feature extraction used to distinguish different physical inputs. It is responsible for both the perception and production of signals for the agent. The experiment has a number of situations to which each agent can respond. Agents have no innate semantic or lexical knowledge.

Interactions occur in the form of the increasingly popular teacher/pupil scenario[19]. An agent is selected from the population to be the teacher and brings with it one or more pupil agents which follow the teacher and observe its interaction with the environment. In this experiment error backpropagation is used for both the agent's interaction with the environment and its interaction with a teacher agent.

The results of the experiment showed that populations achieved a consensus regarding the differentiation of objects and their associated signals. In other words, populations shared the same lexicon for describing situations in their environment.

4.4 Syntax Development

Syntax is arguably the most complex part of a language system. It has also been suggested that human communication evolved in such a way as to maximise the efficiency of a language by restricting sentence structure, making it more easily parsable. Briscoe[20] has conducted some experiments on the evolution of syntax, focusing on word order. The experiment comprises a population of agents each of which are born with a large universal grammar. Certain parameter values describing this grammar need to be set for the agents to correctly order words. The goal of the experiment was to ascertain whether agents were capable of learning the correct values for these parameters through progressive genetic assimilation.

The linguistic interactions take place as follows: an agent generates a number of grammar categories which are perceived by a listener agent. The listener attempts to parse the sentence. The fitness of the agents depends on their ability to correctly generate the appropriate parse tree. The results of the experiment suggest that communicative success depends on the complexity of the parsing.

4.5 Language Dynamics

A very interesting area of language development research is that of language dynamics. How do population or environmental shifts affect the course of language development? Among the areas of interest to be explored is the situation arising from two populations with different languages coming in contact with each other. A number of results can occur from such interaction: language death, creation of a new language (like Creole), mutual influences or even lexicon or syntax replacement. While some of these phenomena have been observed in real world languages, the artificial intelligence approach is still in its infancy. However some interesting results have been reported.

Hutchins and Hazlehurst[8] have observed a number of language dynamics phenomena. Firstly, they found that adding individuals with completely random weights in their neural networks to populations with evolved lexicons had effects dependent on the population size. For small communities, the effect of such disruption was disastrous - even annihilating the existing language. Larger communities were far more robust, forcing the individual to adopt the established language.

During their experiments it was found that certain communities were unable to find a common language as a result of initial conditions. Randomly generated neural networks were sometime physically incompatible and thus were never able to reach a consensus. As a result, experiments were carried out where the teacher agent responsible for imparting much of the lexicon, was chosen only if the pupil's utterances where reasonably similar to the teacher's. The experiment resulted in clusters of the population sharing lexicons which were slightly different from the rest, or in other words, dialects.

4.6 Linguistic Constraints

There exist certain constraints which are present in all human languages. These are known in linguistics as universals. There are two theories to explain how these arose. The first proposal is the phylogenetic argument, which states that they are a result of the LAD's evolution over time. The glossogenetic view states that universals emerged through acquisition and use over generations[2].

Kirby et al[21, 22] investigated the phenomenon of universal emergence by conducting experiments where agents had access to a grammar encoded as a binary string. The genome comprised 0s, 1s and wild-card flags. Agents with defined portions of the genome (0s and 1s) could only acquire a grammar with similar properties, thus the system allowed for various degrees of constraint to be imposed on the grammar learned by an agent.

Interactions between agents would cause an agent that had not understood an incoming sentence to change parameters in its LAD and attempt to re-process the sentence. A reinforcement tool successively tested various parameters in the LAD to ascertain which parameter change best suited the understanding of the sentence. The experiment's results suggest that a phylogenetic approach alone does not sufficiently explain how language constraints could have emerged over time and that it is essential to consider glossogenetic factors.

4.7 Cultural Convergence

Oliphant[23] suggested three criteria for examining the worth of a communication system in an artificial environment.

- Acquisition - Given a certain set of rules, can an agent acquire an optimal communication system?
- Maintenance - Can acquisition occur in the presence of noise?
- Construction - Can the set of rules build on a non-optimal system in order to improve its performance?

While Oliphant's system had its drawbacks, it has been built upon by others, most notably Smith[24, 25] who introduced breeding into a similar experimental setup. The results showed that it was possible to speed up the construction process due to increased fitness payoff. In the original experiment, it was difficult for a population to achieve a good construction rate due in part to the hiding effect[26]. Smith argues that a communication system can be maintained and constructed even if meaning observance is relatively low.

The experiment also observed that the construction rate could be speeded up again by splitting the population into sub-populations and by introducing more learning cycles. The key is to accelerate cultural convergence: in other words, speed up the process whereby a population moves towards an 'agreed' communication system.

4.8 Critical Periods

The argument for critical periods[27], where agents are allowed to learn a language for only a portion of their lifetime, stems from the question of whether the increased time spent training agents is viable from an evolutionary point of view. In the natural world there are relative advantages and disadvantages to delayed maturity. The advantages are that the creatures that survive until the maturity age will likely be quite fit, thus will have a higher fertility. Also, there will be a lower infant mortality rate due to the length of time the parent cares for its offspring. However, this increased time is an additional cost to the parent and limits the number of offspring it can produce. Another disadvantage is that both population and individual fitness will decrease, because of increased generation time and decreased probability of reaching maturation age.

To investigate the effect of critical periods, experiments have been carried out on agents using the Latent Energy Environment (LEE) model where agents accumulate energy throughout their lifetime by performing certain energy giving tasks. Once the energy levels have reached a threshold, the agents are deemed mature enough to begin reproducing.

In one experiment[28], an addition element was introduced. Each agent's genome contains a maturity gene which determines the age at which the agent is allowed reproduce. This gene is allowed to evolve using mutations over generations. The results show that this gene tends to converge towards no maturation, unrelated to the type or amount of learning used in the lifetime of an agent. The learning required by agents in their environment was eventually encoded into their genome, thus no longer requiring a maturity phase to reach the high energy levels required, displaying an example of the Baldwin effect.

4.9 Indexed Memory and Cultural Artifacts

It has been proposed that instead of agents communicating directly with one another, sometimes in a seemingly random fashion, it may be useful to have them share information through a specified medium[29]. This medium is more easily observed by the experimenter and direct effects can be produced by modifying its properties.

A population may share its information through means of a centralised memory repository where individuals can write and read information about their perceived environment. Because each agent may write to the shared memory, there is a risk that agents not well suited to their environment could be disrupting others by sharing erroneous information. However, researchers conducting such experiments found that the relative advantage of good agents writing useful information outweighed that of the incorrect portion of the population.

A similar approach is that of cultural artifacts[7, 9]. In a simulated world there can exist one or more cultural artifacts which are in essence portions of information gleaned from this world by its population and left for further generations. This approach differs from shared memory in that artifacts placed in the environment must not be direct mappings of signal/meaning pairs, but should be symbolic.

4.10 Noise as a Source of Diversity

The success of genetic algorithm approaches to function optimisation problems is due in part to the algorithm's capability for novelty arising from mutations. To investigate whether a similar scheme could be provided for cultural transmission experiments was conducted[10] where a population of agents underwent a process of cultural imitation using the teacher/pupil scenario. The teacher's output to a given situation became the pupil's input to allow the pupil to associate a situation in its environment with a given signal. The experiment used a purely cultural evolution scheme, so no genetic information was passed on to further generations.

It was found that if a population taught the successor generation in the fashion described above, the cultural information passed on would dissipate over generations. This could be reduced by applying a selective process to the choice of teacher, but this only seemed to delay dissipation which was in the end inevitable.

It was suggested that this may have been because of the lack of novelty in the cultural transmission and that an equivalent to genetic algorithm's mutation operator could be the addition of noise in the signal from teacher to pupil. The results showed that the populations were able to sustain communication systems over successive generations with the inclusion of random noise.

5 Conclusions

Synthetic ethology provides a controlled and flexible method for the study of evolutionary linguistics. The approach has brought to light possible explanations for the development of the Language Acquisition Device, the impact of the Baldwin effect on language evolution, the relative advantages and disadvantages of critical periods during learning and the role of noise in the sustenance of a communication system.

Perhaps the most interesting area of research highlighted by current trends is that of language dynamics. For the first time in linguistics, it is possible to observe the result of clashing languages in a controlled environment. Disruptive elements can be introduced, dialects can be seen to form and the role of population size observed.

Many of the experiments discussed concentrated on the cultural development of a population, while effectively ignoring its genetic evolution. Rarely do such experiments attempt to incorporate language and genetic development. Much work has recently been done on the coupling of neural networks and genetic algorithms (including some of our previous work[31]), the tools most prevalent in the experiments, but none of the experiments take advantage of these new methods. While their inclusion would increase the complexity of the experiment, it would also yield more fruitful results and generate populations of greater flexibility.

6 Acknowledgments

This research is funded by the Irish Research Council for Science, Engineering and Technology.

References

[1] B. MacLennan and G. Burghardt. Synthetic ethology and the evolution of cooperative communication. In *Adaptive Behavior 2(2)*, pages 161–188, 1993.

[2] L. Steels. The synthetic modeling of language origins. In *Evolution of Communication*, pages 1–34, 1997.

[3] B. MacLennan. Synthetic ethology: An approach to the study of communication. In *Artificial Life II: The Second Workshop on the Synthesis and Simulation of Living Systems, Santa Fe Institute Studies in the Sciences of Complexity*, pages 631–635, 1992.

[4] P.J.B. Slater. Message meaning analysis. In *Animal Behavior Volume 2: Communication*, pages 9–42. W.H. Freeman, 1983.

[5] N. Chomsky. On the nature of language. In *Origins and evolution of language and speech*, pages 46–57. Annals of the New York Academy of Science, New York. Vol 280, 1976.

[6] J.M. Baldwin. A new factor in evolution. In *American Naturalist 30*, pages 441–451, 1896.

[7] E. Hutchins and B. Hazlehurst. Learning in the cultural process. In *Artificial Life II, ed. C. Langton et al.* MIT Press, 1991.

[8] E. Hutchins and B. Hazlehurst. How to invent a lexicon: The development of shared symbols in interaction. In N. Gilbert and R. Conte, editors, *Artificial Societies: The Computer Simulation of Social Life*, pages 157–189. UCL Press: London, 1995.

[9] A. Cangelosi. Evolution of communication using combination of grounded symbols in populations of neural networks. In *Proceedings of IJCNN99 International Joint Conference on Neural Networks (vol. 6)*, pages 4365–4368, Washington, DC, 1999. IEEE Press.

[10] D. Denaro and D. Parisi. Cultural evolution in a population of neural networks. In *M.Marinaro and R.Tagliaferri (eds), Neural Nets Wirn-96.New York: Springer*, pages 100–111, 1996.

[11] B. De Boer. Emergent vowel systems in a population of agents. In *ECAL-97*, 1997.

[12] H. Yanco and L. Stein. An adaptive communication protocol for cooperating mobile robots, 1993.

[13] C. Angelo and D. Parisi. The emergence of a language in an evolving population of neural networks. *Technical Report NSAL–96004, National Research Council, Rome*, 1996.

[14] L. Steels. Self-organising vocabularies. In *Proceedings of Artificial Life V*, 1996.

[15] L. Steels. Emergent adaptive lexicons. In *P. Maes, editor, Proceedings of the Simulation of Adaptive Behavior Conference*. MIT Press, 1996.

[16] Luc Steels. Perceptually grounded meaning creation. In Victor Lesser, editor, *Proceedings of the First International Conference on Multi–Agent Systems*. MIT Press, 1995.

[17] Angelo Cangelosi. Modeling the evolution of communication: From stimulus associations to grounded symbolic associations. In *European Conference on Artificial Life*, pages 654–663, 1999.

[18] E.D. De Jong. Analyzing the evolution of communication from a dynamical systems perspective. In *Proceedings of the European Conference on Artificial Life ECAL'99*, pages 689–693, 1999.

[19] Billard A. and Hayes G. Learning to communicate through imitation in autonomous robots. In *7th International Conference on Artificial Neural Networks*, pages 763–738, 1997.

[20] E. Briscoe. Grammatical acquisition: Co-evolution of language and the language acquisition device. In *Proceedings of the Diachronic Generative Syntax*. Oxford University Press, Oxford, 1999.

[21] S. Kirby and J. Hurford. Learning, culture and evolution in the origin of linguistic constraints. In *4th European Conference on Artificial Life*, pages 493–502. MIT Press, Cambridge, MA., 1997.

[22] H. Brighton and S. Kirby. The survival of the smallest: Stability conditions for the cultural evolution of compositional language. In *J. Kelemen and P. Sosk, editors, ECAL01. Springer-Verlag*, pages 592–601, 2001.

[23] M. Oliphant and J. Batali. Learning and the emergence of coordinated communication. Technical report, Department of Cognitive Science, University of California at San Diego, 1996.

[24] K. Smith. The cultural evolution of communication in a population of neural networks. *Connection Science*, 14(1), 2002.

[25] Kenny Smith. The importance of rapid cultural convergence in the evolution of learned symbolic communication. In *European Conference on Artificial Life*, pages 637–640, 2001.

[26] Giles Mayley. Guiding or hiding: Explorations into the effects of learning on the rate of evolution. In Phil Husbands and Inman Harvey, editors, *Fourth European Conference on Artificial Life*, Cambridge, MA, 1997. MIT Press.

[27] F. Cecconi, F. Menczer, and R. Belew. Maturation and the evolution of imitative learning in artificial organisms. In *Adaptive Behavior, 4(1)*, pages 29–50, 1995.

[28] Batali J. Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In *Proceedings of Artificial Life IV*, pages 160–171. The MIT Press, Cambridge MA., 1994.

[29] Lee Spector and Sean Luke. Culture enhances the evolvability of cognition. In *Cognitive Science (CogSci) 1996 Conference Proceedings*, 1996.

[30] Gregory M. Saunders and Jordan B. Pollack. The evolution of communication in adaptive agents.

[31] D. Curran and C. O'Riordan. Learning in artificial life societies, report number nuig-it-220202. Technical report, Dept. Of IT, NUI, Galway, 2002.