Robotic Experiments on the Emergence of a Lexicon

Joris Van Looveren

Artificial Intelligence Laboratory
Vrije Universiteit Brussel
Pleinlaan 2, 1050 Brussels, Belgium
e-mail: joris@arti.vub.ac.be

Abstract

The Talking Heads experiment is a robotic version of earlier experiments on the emergence of a lexicon. A speaker and a hearer agent chosen randomly from a larger population try to communicate with each other about objects they see in their environment; the hearer's task is to guess the topic that the speaker talks about. The agents use bodies equipped with pan-tilt cameras to perceive their environment, which consists of a whiteboard containing geometrical shapes that form the topics of the interactions between the agents. The paper examines data gathered during a four-month run involving several interconnected Talking Heads sites.

1 Introduction

The lexicon is a very important part of language; without it, complex communication would not be possible. In its simplest form, it can be considered to be a list of associations between symbols and representations of concepts. The symbols are essentially arbitrary; the majority of words we know have no direct relationship with the concepts they represent.

A consequence of the arbitrariness is that, when a child grows up and learns its native language, it has to learn all associations from scratch, without any prior knowledge. An important question is then how these associations can be learned. Many experiments have already been done in the area of lexicon formation, but most of these were done in simulation, using the game paradigm from economics and game theory.

This paper will describe an experiment in which a mechanism for lexicon creation and learning (which has been tested in simulations before) has been implemented on robotic bodies, that live in a real-world environment, using the sensors present in the body.

2 Language Models

Many linguistic theories implicitly assume a single, ideal speaker; they do not take into account the fact that language is a social phenomenon. Nevertheless, this is an important aspect: no single language user has perfect knowledge of the language, and language is not centrally controlled (i.e. there is no single entity that decides how the language should be), which suggests that the dynamics of the interactions between individuals has an important bearing on the language that exists in a community of language users. For example, any individual in a community can coin a new word, which may or may not spread in the community depending on whether the other individuals decide to adopt the innovation.

A new way of creating and testing theories about lexicon formation was inspired by game theory [1]. A community of language users (agents) is implemented in a computer program, which simulates the interactions between the agents. Every interaction includes a linguistic component, that may help the agent in accomplishing its task. The task itself depends on the experiment.

One of the first examples of simulation models focusing on lexicon learning is the model by Hurford [5], where he compares three strategies for acquiring a lexicon from examples. Oliphant studies the conditions in which a Saussurean communication system can evolve in a population of agents [8]. Steels researches the emergence of a lexicon through self-organisation [10], and Cangelosi and Parisi studied the usefulness and of a lexicon as an aid in solving a foraging task [3].

With lexicon studies well underway, attention is also being turned towards grammar, most notably by Batali [2], Kirby [6], and also by Steels [11]. De Boer has done experiments in the area of phonetics and phonology [4].

A step beyond simulations is the implementation of a linguistic theory in real robotic agents. These robots live in a real environment, and are subject to the constraints of this real world. In working with real robots, the experimenter is not any more in (full) control of many parameters of the environment (such as light conditions, external interferences, drifting calibrations...). This puts stronger demands on the performance of the language model that is built into the robots, and thus is a good test of the model in question.

Language experiments in which real robots are used are still relatively rare; an example can be found in [13]. An experiment in which a robot derives relevant semantic features from the words' syntactic environments and its sensor values is described in [7]. This paper examines the "Talking Heads" experiment, which is described in [9].

2.1 Talking Heads

The Talking Heads experiment was a public experiment which was run at two different occasions. The data examined in this paper is from the first experiment, which ran from July 1999 through October 1999. In the experiment, there were 3 permanent sites (at an exhibition in Antwerp, in Paris at Sony CSL and at the VUB in Brussels) set up as shown in fig. 1 with two pan-tilt cameras looking a whiteboard (a) on which geometrical shapes are pasted (b).

The second run of the experiment started at the end of January 2000 and lasted until August 2000. The setup of the sites was the same as in the first experiment but they were located at different places (at the UvA in Amsterdam, in the UK in galleries in London and Cambridge, at Sony CSL in Paris and at the VUB in

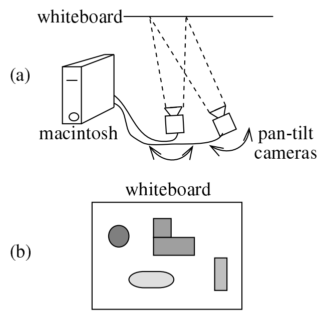

Figure 1: Schema of a Talking Heads installation.

Brussels).

Through a public web site, anyone could create new individuals (called *agents*), introduce them into the system, send them to the different sites of the Talking Heads network, and even teach their agents new words.

2.2 Language Games

In our experiments, the interactions between agents are called “language games.” The task of the agents is to identify an object in the environment solely using linguistic means. A language game proceeds as follows.

1. Two agents are chosen randomly from the population. One is the speaker, the other one the hearer.

2. The speaker chooses a topic from its environment. In the Talking Heads case, the speaker moves its camera around the whiteboard until it finds an area with two objects or more (one object will be the topic; the rest will be the background). The hearer looks at the same area of the whiteboard.

3. The speaker tries to find a meaning for this topic, which is a unique description of the topic in terms of its properties. (Unique in the sense that it should not apply to any other object in the environment.) An example of a meaning could be *green*, when the topic is the only green object that was observed. (In the early simulation experiments, topic and meaning were the same.)

4. Subsequently the speaker tries to find a word for this meaning in its lexicon.

5. The speaker utters the word it found.

6. The hearer hears the speaker’s word and perceives the same objects in the environment. Based on the word, it will try to guess the object that the speaker chose to be the topic, by looking it up in its lexicon and verifying to which object(s) the meaning it found is applicable.

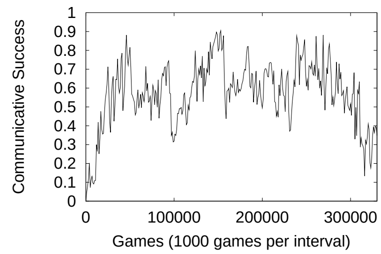

Figure 2: Global communicative success.

7. If the hearer successfully identified the topic, the speaker and the hearer increase the score of the meaning-word associations they used.

8. If the hearer did not find the topic, the speaker will point to it. The hearer then knows both the topic and the word, and can find a meaning for the topic, associate it to the speaker's word, and add it to its lexicon.

3 Results & Discussion

3.1 Communicative Success

Communicative success is the simplest possible measure for our language model. It measures the fraction of interactions that was successful, i.e. in which the hearer succeeded in finding the topic based on the speaker's linguistic hint.

As can be seen in fig. 2, communicative success is around 60%, with very large fluctuations. In the simulation experiments, success consistently rose to 100% quickly.

There are several external factors that influence the success of the agent population. On each individual site, the population changes continuously. Agents arrive from other sites, and other agents leave after having completed all their games at that site. There is also a continuous influx of new agents. Figure 3 shows the number of new agents in the system per interval of 1000 games. These new agents have to learn the language from scratch, which causes a number of unsuccessful games until they learn enough of the language.

Simulation experiments by Steels & Kaplan [12] have investigated the influence of population change. Their conclusion is that, as long as the rate of population change is not too high, both communicative success and coherence drop as the new agents are learning the language, but the language itself does not change.

There are a number of other external events that can cause communicative success to drop, such as changes in the layout of the geometrical shapes on the whiteboard, breakdowns of the equipment or errors in the calibration, etc.

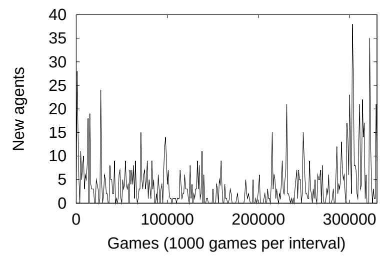

Figure 3: Number of new agents every 1000 games.

3.2 Coherence

Coherence measures the extent to which all agents in the population use the same words for the same objects. It is measured by counting, for every object, how many agents prefer to use the same word for it. Generally, coherence will not be total, because some agents will prefer different words than others. In this case, the word that most agents prefer will be considered the one that the population prefers. Averaging this over all meanings gives a measure for the quality of the language that the agents developed.

For the Talking Heads, the notion of coherence has to be extended, because there is no direct relationship any more between objects and words. Since the Talking Heads have to use their robotic bodies to perceive the environment, they have to use their own internal representation of each object instead of the object itself in their associations (as can be done in simulations). Additionally, their task is to find a unique description for the topic that is not applicable to other objects in their environment, which means that in different interactions, the same object may be represented by different meanings.

In this case, coherence can be calculated not only between words and objects, but also between words and meanings, and meanings and objects. Unfortunately, in the Talking Heads experiment calculation of meaning-object or word-object coherence is not possible, because there is not enough information in the database to reconstruct the referents of the interactions.

When meaning-word coherence is calculated in the standard way, averaging over all meanings, the Talking Heads experiment scores a mere 43.2%. This is not very much compared to the 90–100% found in the basic simulation experiments, suggesting that indirect reference to objects results in much worse performance (which would be corroborated by the low success scores). However, figure 4 shows meaning-word coherence in an alternative way. Every bar in the graph shows the average coherence for 5 meanings, with the error bar showing the standard deviation. The meanings are sorted according to their frequency of use in the interactions; the most used meanings are on the left side. It can be seen clearly that for the most frequently used meanings, coherence (almost) reaches the levels achieved in the basic language game experiments. Only the meanings that are less

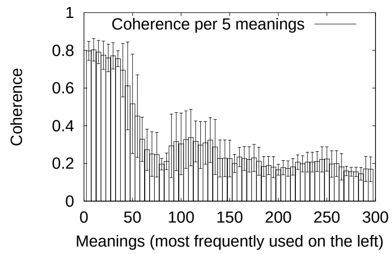

Figure 4: Meaning-word coherence.

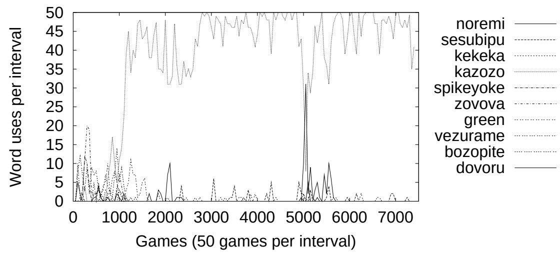

Figure 5: Evolution of the lexicon for the meaning *green*.

frequently or almost never used, have a very low coherence. Figure 5 shows the meaning-word coherence for the meaning *green* throughout the experiment. In the beginning when no word is dominant yet, many words are competing with each other. Later on, the word *kazozo* becomes dominant, and remains the preferred word for the rest of the experiment, except for a short peak when other words momentarily become more successful.

Figure 6 shows how many interactions are covered by how many of the most frequent meanings. It can be seen that in 98% of the interactions, one out of the 50 most frequent meanings is used. This confirms that there is a small number of meanings that are used very often. The extra meaning selection step that is performed by the Talking Heads agents introduces a lot of meanings that are used only once or very few times, which is not the case in the simulation experiments. Since the agents are not capable of removing unused associations from their lexicons, they remain in the lexicon for the duration of the experiment.

4 Conclusion

The Talking Heads experiment was an experiment to study the organisation of a lexicon in a population of agents. While previous experiments have almost exclusively used simulations to test mechanisms for lexicon emergence and acquisition,

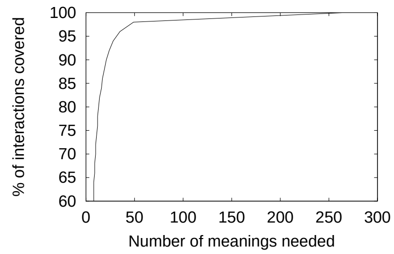

Figure 6: Percentage of interactions covered.

the Talking Heads experiment was publically accessible, and used real-world robotic bodies for perception of the environment. People from all over the world were able to create agents, launch them into the network of sites that were available at different places and teach them new words, thus influencing the experiment beyond the control of the experimenter.

The experimental data reveal the influence of the real world. Communicative success was much lower than in the idealised simulation experiments because of many external factors, such as the inflow of new agents in the system, and the movements of agents between different sites. The influence of these factors has also already been shown in simulation experiments.

Despite the comparatively low communicative success, the language that developed in the early stages of the experiment remained stable. One reason for this is that in the majority of the language games, only relatively few meanings are used: in 98% of the interactions, one of the 50 most frequent meanings is used. The robotic agents, which have to do a meaning selection process next to the word selection process, collect many associations that turn out to be not useful later on. For the core set of 50 meanings, coherence is almost 80%, while for the other meanings coherence is between 20 and 35%, which shows that the language is resilient against changes.

From these results, it is clear that a real-world robotic experiment shows more subtle behaviour than simulation experiments. While simulation experiments are extremely useful to study the influence of individual factors on the core of a model, robotic experiments are more demanding as they impose a great many different factors at the same time on the model. This makes robotic models a good testing ground for theories that have been developed using simulation models.

5 Acknowledgements

The research presented in the paper was funded by the Flemish Institute for the Promotion of Innovation by Science and Technology (IWT). It is based on data gathered from an experiment done by the VUB AI-lab and Sony CSL Paris, together with the Laboratorium exhibition in Antwerp (first experiment), and with the NO1SE exhibition (UK) and Ben Kröse (UvA) for the second experiment.

References

[1] Robert Axelrod. *The Evolution of Cooperation.* Basic Books, New York, 1984.

[2] John Batali. Computational simulations of the emergence of grammar. In *Approaches to the Evolution of Language: Social and Cognitive Bases*, pages 405–426. Cambridge University Press, Cambridge, 1998.

[3] A. Cangelosi and D. Parisi. The emergence of language in an evolving population of neural networks. In *Proceedings of the 18th Conference of the Cognitive Science Society.* San Diego, 1996.

[4] Bart de Boer. Emergent vowel systems in a population of agents. In *Proceedings of ECAL-97, Brighton UK, 1997.* The MIT Press, Cambridge, Ma., 1997.

[5] James R. Hurford. Biological evolution of the saussurean sign as a component of the language acquisition device. *Lingua*, 77:189–222, 1989.

[6] Simon Kirby. *Function, Selection and Innateness: the Emergence of Language Universals.* Oxford University Press, 1999.

[7] T. Oates, Z. Eyler-Walker, and P. Cohen. Using syntax to learn semantics: An experiment in language acquisition with a mobile robot. Technical Report 99-35, Computer Science Department, University of Massachusetts, 1999.

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

[9] L. Steels and F. Kaplan. Bootstrapping grounded word semantics. In T. Briscoe, editor, *Linguistic evolution through language acquisition: formal and computational models.* Cambridge University Press, Cambridge, UK, 1999.

[10] Luc Steels. Emergent adaptive lexicons. In P. Maes, editor, *Proceedings of the Simulation of Adaptive Behaviour Conference.* The MIT Press, Cambridge, Ma., 1996.

[11] Luc Steels. The origins of syntax in visually grounded robotic agents. In M. Pollack, editor, *Proceedings of the 15th International Joint Conference on Artificial Intelligence.* Morgan Kauffman, 1997.

[12] Luc Steels and Frédéric Kaplan. Stochasticity as a source of innovation in language games. In *Proceedings of Artificial Life VI.* The MIT Press, Cambridge, Ma., 1998.

[13] Luc Steels and Paul Vogt. Grounding adaptive language games in robotic agents. In *Proceedings of the 4th European Conference on Artificial Life.* The MIT Press, Cambridge, Ma., 1997.