Developmental exploration in the cultural evolution of lexical conventions

Pierre-Yves Oudeyer and Frédéric Delaunay
INRIA Bordeaux Sud-Ouest

The objective of this poster is to show that developmental mechanisms can impact importantly the efficiency of cultural evolution dynamics in the formation of shared lexicons. Thus, this poster tries to gather two threads of research: semiotic dynamics and developmental sciences.

On the one hand, many models of language formation and self-organization were developed in the last decade, under the general theoretical umbrella of semiotic dynamics (Steels, 2003). One of the best studied model is the naming game (Kaplan, 2005), in which a population of agents builds a shared lexicon, associating words and meanings, through cultural evolution and self-organization. In these simulations, agents interact in a peer-to-peer manner, negotiating each time the word that shall be associated to a randomly chosen meaning. Experiments have shown that simple feedback mechanisms lead to the self-organization of a globally shared set of associations (Steels, 2003; Kaplan, 2005).

On the other hand, recent years have also seen the multiplication of models showing that a developmental approach to learning can be very efficient in real-world spaces. The general idea is based on the gradual control of the complexity of the task or skill to be learnt, "starting simple" and progressively becoming more complicated.This can be realized through the control of the number of degrees of freedom of the motor apparatus, the sensitivity of sensors, or the complexity of the learning situations, which are progressively and actively increased (Berthouze and Lungarella, 2004; Elman, 1993; Oudeyer et al., 2007).

In this poster, we propose to show that the introduction of active mechanisms of complexity control at the level of individuals in naming game experiments allows to drastically improve the speed of convergence of the global population to a shared lexicon. A related coupling was presented in (Steels and Wellens, 2007), but it concerned the evolution of grammatical forms and was combined with other new complex mechanisms, making the study of the specific impact of complexity control not directly addressed.

The "standard" naming game

In the naming game (Kaplan, 2005), a population of N agents tries to negotiate a lexicon, associating for each agent each meaning $m_i \in M$ to a single word $w_j \in W$, avoiding both synonymy and homonymy. The set M of meanings and the set W of words are finite and predefined, represented by discrete symbols, with $|W| \geq |M|$. Initially, each agent starts with an empty lexicon, which grows through peer-to-peer interaction. In an interaction, a speaker and a hearer are randomly chosen. The speaker chooses randomly a meaning among those that are possible, and then looks for the word which is most strongly associated to this meaning in its internal memory. If there is no word already associated to it, then it chooses randomly a word among those that are not used yet in another association. Then, it utters the word and the hearer finds the meaning in its own memory which is most closely associated to it, and shows it to the speaker. If this is the same meaning as the one initially chosen by the speaker, then the interaction is a success, and both agents reinforce the association, and weaken all associations involving only the word or only the meaning used in the interaction. If the interaction is not a success, the speaker tells the hearer what meaning he intended and the hearer decreases the score of the association he used, and increases the score of the association proposed by the speaker. After a certain number of interactions, (Steels, 2003) and (Kaplan, 2005) showed that a globally shared lexicon was formed.

Controlling complexity in the naming game

In the standard version of the naming game, as meanings are chosen randomly across the set of all possible meanings by the speakers, all meanings and many words are quickly introduced in the population. The consequence is that when $|M|$ is large, this leads to a massive competition among associations right from the few first dozens of interactions: the complexity of the task grows nearly to its maximum in a very short time.

We propose to introduce a mechanism, local to each agent, in order to permit the regulation of this growth of complexity. A mechanism which we experimented consists in having speakers choose actively the meanings they will talk about instead of choosing them randomly. This choice of meaning depends on the subjective level of success in communication that each agent currently measures: a given agent decides to talk about a meaning it has never talked about before if and only if the mean of the best scores of all associations involving each already used meaning is over a predefined threshold (90 percent), i.e. when they have enough success in interaction with others when talking about the meanings they already talked about. Otherwise, the speaker chooses randomly a meaning he has already used. The consequence is that the number of meanings, whose association with a word is negotiated in the population, gradually and actively increases instead of being maximal right from the start.

We have conducted systematic experiments to investigate the impact of such an active policy for complexity control. Results show that the speed of convergence is one order of magnitude faster when a developmental mechanism is used as opposed to the standard version of the naming game. For example, figure 1 shows comparisons of the speeds of convergence for both strategies in various experiments where the number of agents = the number of available words = the number of available meanings = N, where N is varied from 50 to 300. These curves show that the active strategy is not only faster, but that the algorithmic complexity changes (first approximations show that it shifts from $O(N^2 \log(N))$ to $O(Nlog(N))$). Figure 2 shows the evolution of the mean number of meanings and words introduced in the population at a given time both for the standard and developmental version of the naming game: this shows that complexity increases more gradually in the developmental case.

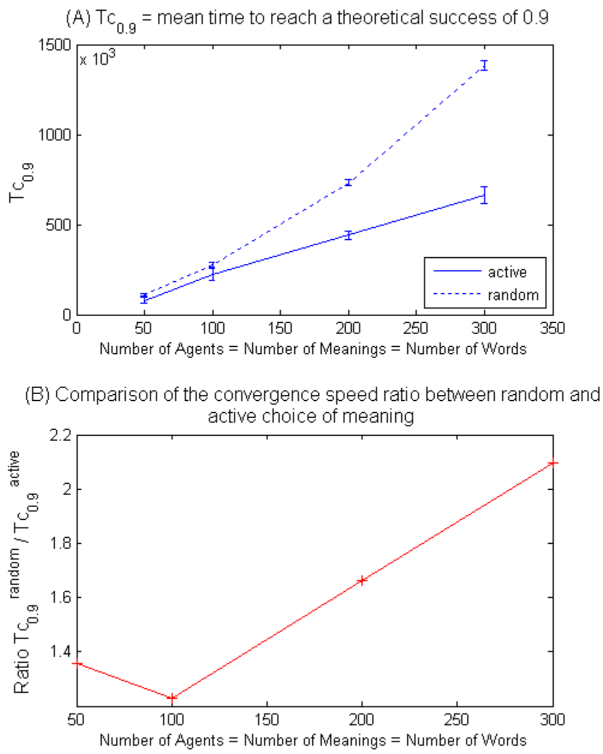

Figure 1 Comparison of the speed of convergence on a shared lexicon in population of agents with the standard naming game (random choice of meaning) and with the developmental naming game (active choice of meaning). The theoretical success is a measure described in (Kaplan, 2005) which allows us to monitor how far agents in the population are from globally sharing the same non-ambiguous lexicon. When total convergence is reached, its value is 1. We used here the value 0.9, which is quasi-convergence, because it allowed the simulation to be fast enough to obtain statistically significant results for the random strategy for N =|M|=|W| = 300 (each simulation already takes several hours on a standard PC).

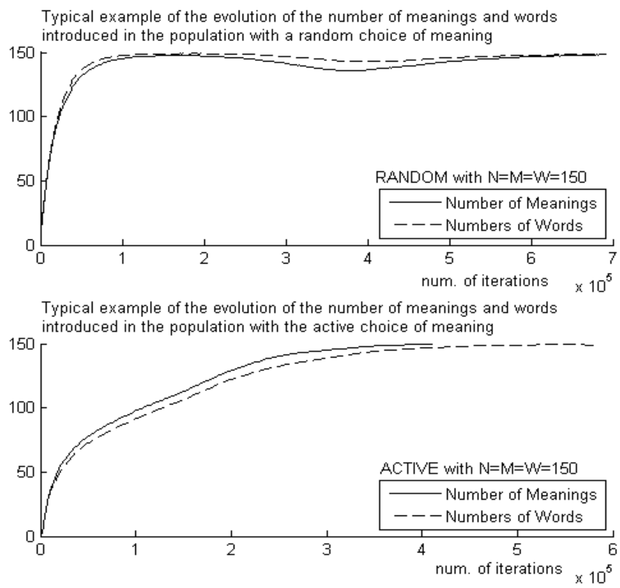

Figure 2 With the active choice of meaning, the growth of the number of meanings and words is dynamically controlled.

References

Berthouze, L., and Lungarella, M. (2004). Motor skill acquisition under environmental perturbations: On the necessity of alternate freezing and freeing of degrees of freedom. *Adaptive Behavior* 12(1):47-63.

Elman, J. (1993) Learning and development in neural networks: The importance of starting small. Cognition, v. 48. p. 71-89.

Kaplan, F. (2005) Simple models of distributed co-ordination, Connection Science, 17 (3-4): 249-270.

Oudeyer P-Y & Kaplan , F. & Hafner, V. (2007) Intrinsic Motivation Systems for Autonomous Mental Development, IEEE Transactions on Evolutionary Computation, 11(2), pp. 265--286.

Steels, L. and Wellens P. (2007) Scaffolding language emergence using the autotelic principle, Proceedings of the 2007 IEEE Symposium on Artificial Life, p. 325-332, Honolulu, HI, USA.

Steels, L. (2003) Evolving grounded communication for robots. Trends in Cognitive Science, 7(7):308-312 July 2003.