Scaling the impact of active topic choice in the Naming Game

William Schueller, Pierre-Yves Oudeyer

Scaling the impact of active topic choice in the Naming Game

Flowers Project-team, Inria Bordeaux Sud-Ouest
William Schueller, Pierre-Yves Oudeyer

How does language emerge, evolve and gets transmitted between individuals? What mechanisms underly the formation and evolution of linguistic conventions, and what are their dynamics? Computational linguistic studies have shown that local interactions within groups of individuals (e.g. humans or robots) can lead to self-organization of lexica associating semantic categories to words [13]. However, it still doesn't scale well to complex meaning spaces and a large number of possible word-meaning associations (or lexical conventions), suggesting high competition among those conventions.

In statistical machine learning and in developmental sciences, it has been argued that an active control of the complexity of learning situations can have a significant impact on the global dynamics of the learning process [2, 4, 3]. This approach has been mostly studied for single robotic agents learning sensorimotor affordances [6, 5]. However active learning might represent an evolutionary advantage for language formation at the population level as well [8, 12].

Naming Games are a computational framework, elaborated to simulate the self-organization of lexical conventions in the form of a multi-agent model [11]. Through repeated local interactions between random couples of agents (designated *speaker* and *hearer*), shared conventions emerge. Interactions consist of uttering a word – or an abstract signal – referring to a topic, and evaluating the success or failure of communication.

However, in existing works processes involved in these interactions are typically random choices, especially the choice of a communication topic.

The introduction of active learning algorithms in these models produces significant improvement of the convergence process towards a shared vocabulary, with the speaker [7, 9, 1] or the hearer [10] actively controlling vocabulary growth.

We study here how the convergence time and the maximum level of complexity scale with population size, for three different strategies (one with random topic choice and two with active topic choice) detailed in table 1.

Naive (random)	Success Threshold	Minimal counts
$m \leftarrow$random($\mathcal{M}$)	**if** mean$\left(\frac{succ(i)}{succ(i)+fail(i)}\right)_{i\in\mathcal{LM}} \geq \boldsymbol{\alpha}$: $m \leftarrow$random($\mathcal{UM}$) **else**: $m \leftarrow$argmin$_{i\in\mathcal{LM}}\left(\frac{succ(i)}{succ(i)+fail(i)}\right)$	**if** $\forall i \in \mathcal{LM}\ succ(i) > \boldsymbol{n}$: $m \leftarrow$random($\mathcal{UM}$) **else**: $m \leftarrow$argmin$_{i\in\mathcal{LM}}(succ(i))$
$\mathcal{M}$: all meanings, $\mathcal{LM}$: labeled meanings, $\mathcal{UM}$: unlabeled meanings, μ: vocabulary size (# word-meaning associations) $succ$: # successful interactions per meaning, $fail$: # failed interactions per meaning		

Figure 1: Strategies: Choice of meaning m. Both active strategies use a parameter (α and n), which is each time chosen optimal in our simulations.

As for the version of the Naming Game used in our work, the scenario of the interaction is described in [10]. Vocabulary is updated as described in the Minimal Naming Game, detailed in [14]. In our simulations, we choose to set $N = M = W$, where N is the population size, M the number of meanings, and W the number of possible words. The computed theoretical success ratio of communication is used to represent the degree of convergence toward a shared lexicon for the whole population. A value of 1 means that the population reached full convergence. Complexity level of an individual lexicon is measured as the total number of distinct associations between meanings and words in the lexicon, or in other words: memory usage.

We show here (see figures 2,3) that convergence time and maximum complexity are reduced with active topic choice, a behavior that is amplified as larger populations are considered. The minimal counts strategy yields a strictly minimum complexity (equal to the complexity of a completed lexicon), while converging as fast as the success threshold strategy. Further work will deal with other variants of the Naming Game (with different vocabulary update, population replacement, and different ratio for N, M and W). For the moment only the hearer's choice scenario is studied, because of its high robustness to changes in parameter values for the different strategies [10].

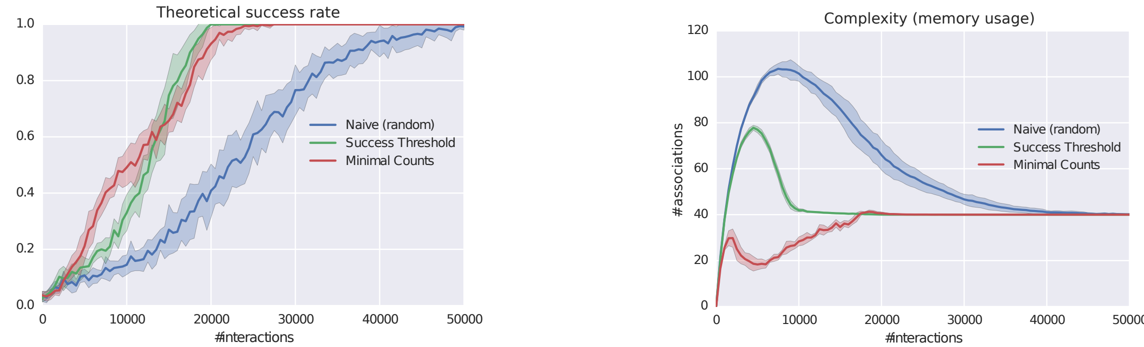

Figure 2: Strategy comparisons, in terms of convergence time (theoretical success ratio) and complexity level (memory usage). In this case, the hearer is the one choosing the topic. M=W=N=40, averaged over 8 trials

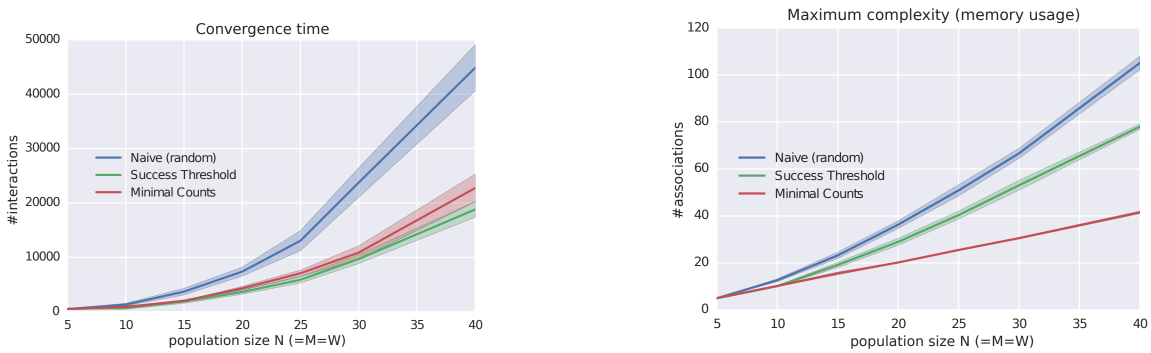

Figure 3: Scaling of maximum memory usage and convergence time for the different strategy, in fonction of population size. In this case, the hearer is the one choosing the topic. M=W=N, averaged over 8 trials.

References

[1] Miquel Cornudella, Paul Van Eecke, and Remi Van Trijp. How intrinsic motivation can speed up language emergence. In *Proceedings of the European Conference on Artificial Life*, pages 571–578, 2015.

[2] Jacqueline Gottlieb, Pierre-Yves Oudeyer, Manuel Lopes, and Adrien Baranes. Information-seeking, curiosity, and attention: computational and neural mechanisms. *Trends in cognitive sciences*, 17(11):585–593, 2013.

[3] Frédéric Kaplan, Pierre-Yves Oudeyer, and Benjamin Bergen. Computational models in the debate over language learnability. *Infant and Child Development*, 17(1):55–80, 2008.

[4] Manuel Lopes and Luis Montesano. Active Learning for Autonomous Intelligent Agents: Exploration, Curiosity, and Interaction. *CoRR*, abs/1403.1:40, 2014.

[5] Clément Moulin-Frier and P-Y Oudeyer. Exploration strategies in developmental robotics: a unified probabilistic framework. In *proceedings of the 3rd International Conference on Development and Learning and on Epigenetic Robotics*, pages 1–6. IEEE, 2013.

[6] P-Y Oudeyer, Frédéric Kaplan, and Verena Vanessa Hafner. Intrinsic motivation systems for autonomous mental development. *Evolutionary Computation, IEEE Transactions on*, 11(2):265–286, 2007.

[7] Pierre-Yves Oudeyer and Frédéric Delaunay. Developmental exploration in the cultural evolution of lexical conventions. In *proceedings of the 8th international conference on epigenetic robotics : modeling cognitive development in robotic systems*, 2008.

[8] Pierre-Yves Oudeyer and L Smith. How evolution may work through curiosity-driven developmental process. *Topics Cogn. Sci*, 2014.

[9] William Schueller and Pierre-Yves Oudeyer. Active Learning Strategies and Active Control of Complexity Growth in Naming Games. In *proceedings of the 5th International Conference on Development and Learning and on Epigenetic Robotics*, pages 220–227, 2015.

[10] William Schueller and Pierre-Yves Oudeyer. Active control of complexity growth in naming games: Hearer's choice. In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11)*, 2016.

[11] Luc Steels. Language games for autonomous robots. *Intelligent Systems, IEEE*, 16(5):16–22, 2001.

[12] Luc Steels. The Autotelic Principle. *Science*, 3139:1–16, 2004.

[13] Luc Steels, Frédéric Kaplan, Angus McIntyre, and Joris Van Looveren. Crucial factors in the origins of word-meaning. *The transition to language*, 12:252–271, 2002.

[14] Pieter Wellens. Adaptive Strategies in the Emergence of Lexical Systems, 2012.