Talking urns, or the emergence of linguistic conventions in a population of agents

keywords: computer simulations, multi-agent model, emergence of conventions, signalling game, reinforcement learning.

Recently, emergence of signalling conventions draws interdisciplinary interest ranging from game theory (Skyrms, 2010) to Artificial Life (Steels and Loetzsch, 2012) to evolutionary linguistics (Steels, 1995; Smith, 2002). Signalling behaviour is found in many species, however, only human communication system is open-ended. This unique property may partly be due to the fact that words, its basic elements, are not innate signals as in other animal communication systems but are learned socially (Spike *et al*., 2017).

A population of individuals communicating with each other can be treated as a complex adaptive system (Steels, 2000). Without any central control and purely as a result of local interactions between individuals, such a system can self-organize and shared linguistic conventions emerge globally. Hence the emergence and evolution of language can be regarded as an effect of its learning and cultural transmission (Steels, 2000).

Computer modelling is becoming an increasingly important tool for studying the problem of origin and evolution of language (de Boer, 2006; Cangelosi and Parisi, 2002). An efficient technique for studying complex adaptive systems is multi-agent modelling, in which a population of dynamically interacting individuals called agents is simulated. Agents are equipped with strictly defined properties and rules governing their behaviour or evolution. As a result of local interactions between agents or between agents and the environment, some new global features might emerge.

We examine the emergence of signalling conventions in a multi-agent reinforcement learning model. In its single-object version, we have a population of agents, which try to establish a name for a given object. Each agent's inventory contains the same words, to which they assign their own weights. Local interactions between neighbouring agents are repeated such that a randomly selected speaker communicates to a hearer a word chosen with probability proportional to its weight. After interaction, both the speaker and the hearer increase their weights of the communicated word.

Our simulations show that topology of the network of interactions between agents plays a very important role: while on a complete graph or sufficiently dense random graphs, a global consensus is typically reached (all agents communicate using the same word), on finite-dimensional lattices, the model gets trapped in a disordered configuration with only a local consensus (with clusters of agents using the same word) – differently than, e.g., in the Naming Game model (Baronchelli *et al*., 2006).

In a multi-object version of the model, agents try to establish names for a number of objects. On a square lattice, this model behaves similarly to the original one, namely it gets trapped in a disordered configuration with only some local consensus. On a complete graph, a much better consensus is reached, however, only when the number of words in agents' inventories is large enough, in which case the resulting language provides a nearly perfect one-to-one mapping between objects and words. Otherwise, there are synonyms and homonyms in the emerging language.

The underlying framework of our model is an urn model (Pemantle, 2007; Spike *et al*., 2017), in which after a large number of iterations, it is almost impossible to shift the balance of weights. However, to allow for this and avoid getting trapped in disordered configurations, we implemented a population renewal, which proved to restore coarsening of clusters and considerably enhanced formation of efficient signalling in both versions of the model.

word count: 547

References

Baronchelli, A., Felici, M., Loreto, V., Caglioti, E., and Steels, L. (2006). Sharp transition towards shared vocabularies in multi-agent systems. *Journal of Statistical Mechanics*, P06014.

Cangelosi, A. and Parisi, D. (eds.) (2002). *Simulating the Evolution of Language*. London: Springer Verlag.

de Boer, B. (2006). Computer modelling as a tool for understanding language evolution. In: N. Gonthier et al. (Eds.) *Evolutionary Epistemology, Language and Culture – A Non-adaptationist, Systems Theoretical Approach* (pp. 381–406). Dordrecht: Springer.

Pemantle, R. (2007). A survey of random processes with reinforcement. *Probability Surveys* 4,1-79.

Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford, UK: Oxford University Press.

Smith, K. (2002). The cultural evolution of communication in a population. *Connection Science*, 14(1), 65-84.

Spike, M., Stadler, K., Kirby, S. and Smith, K. (2017). Minimal requirements for the emergence of learned signalling. *Cognitive Science*, 41(3), 623-658.

Steels, L. (1995). A self-organizing spatial vocabulary. *Artificial Life*, 2(3), 319–332.

Steels, L. (2000). Language as a Complex Adaptive System. In M. Schoenauer (Ed.), Proceedings of PPSN VI (Lecture Notes in Computer Science). Berlin: Springer-Verlag.

Steels, L. and Loetzsch, M. (2012). The grounded naming game. In L. Steels (Ed.), *Experiments in Cultural Language Evolution* (pp. 41-59). Amsterdam: John Benjamins.