SOCIO-ENVIRONMENTAL CONSTRAINTS IN CATEGORY LEARNING

KATARZYNA OKTABA[*1,2] and DARIUSZ KALOCIŃSKI[1]

[*]Corresponding Author: katarzyna.oktaba@wp.pl
[1]Institute of Philosophy, University of Warsaw, Warsaw, Poland
[2]Warsaw School of Economics, Warsaw, Poland

According to Christiansen and Chater (2008), language is primarily shaped by communicative and cognitive pressures. In response, Wallentin and Frith (2008) claim that these constraints might be insufficient for explaining, for example, why we find variability and sub-optimality in certain systems of linguistic categorization. As a possible remedy, they propose to look at social motivations in language learning, which may run orthogonally to pressures for optimality.

To investigate the plausibility of the above proposal, we modify the Lewis signalling game (Lewis, 1969), which provides a model of linguistic conventionalization as shaped by environmental constraints. To include social pressures in our model, the payoffs of the game depend also on the similarity between agents' behaviour, even if it runs contrary to the environmentally optimal strategy.[1]

Our model consists of two players, $M = 2$ terms (signals) and $N > M$ objects, thus giving rise to categories (Skyrms, 2010; Barrett, 2007). There is a stable partition of the objects into two classes (natural groupings). Each player associates objects with signals probabilistically. Objects that are associated with the same term comprise one of the player vague categories. Players engage in repeated exchanges with sender-receiver roles. Each round the sender has some object in mind, sends a corresponding signal and expects another example of the same class *according to his private categorization* in return (social coordination). In addition, if the initial object and the received one belong to the same class according to the external (natural) grouping, players receive a positive payoff (environmental pressure). Thus, the overall payoff is given by $r = w_S * s + w_N * n$, where s, n are payoffs for social and external alignment, respectively, and $w_S + w_N = 1$.

We investigate learning of categorizations for a 2-term/27-object game played by two communicating agents whose roles alternate within consecutive exchanges. We take Herrnstein reinforcement learning (HRL) (Herrnstein, 1970; Roth & Erev, 1995; Skyrms, 2008) and its smoothed version (SRL) (Barrett &

---

[1]The model mentioned here originally appears in a BA thesis by the first author (Oktaba, 2017).

Zollman, 2009), with parameters $\delta = 0.9$ (high forgetfulness) and $\lambda = 25$ (high reactivity), as conceivable models of learning. The algorithms are modified appropriately to incorporate both environmental and social rewarding systems.

We test each algorithm in three conditions: *social* ($w_N = 0.1$), *environmental* ($w_N = 0.9$), and *mixed* ($w_N = 0.5$). Exemplary results are shown in Figure 1.
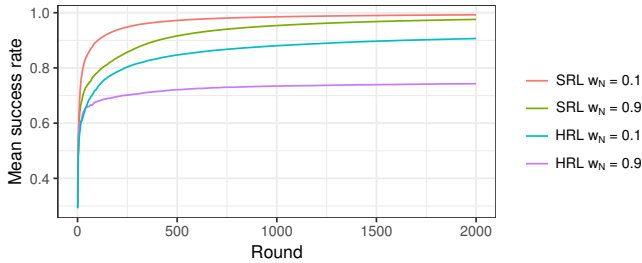

Figure 1. Mean success rates for HRL and SRL, social vs environmental condition, based on 100 simulations run across 2000 rounds. The mean success rate is calculated from the last 50 interactions.

After 2000 rounds, social condition yields the best success rates (HRL 95%, SRL 99%). Mixed condition results in lower values (HRL 81%, SRL 98%). The lowest success is obtained in environmental condition (HRL 72%, SRL 97%).

Prima facie, a plausible reading of the suggestions made by Wallentin and Frith (2008) is that the need to be identified as members of one group may run counter to environmental pressures. Our analysis does not support this interpretation. On the contrary, it turns out that higher social pressures support adaptation to a stable environment. This conclusion seems valid for both population- and, to a lesser extent, individual-based dynamics.[2] Moreover, SRL simulations give further evidence that forgetting improves learning (Barrett & Zollman, 2009).

Another reading of the Wallentin and Frith proposal, which we left largely open, is to consider groups that are socially biased towards themselves and indifferent to social pressures upon between-group interaction, or groups that are mutually antagonized. It seems likely that such groups will develop shared languages but not on the between-group level, particularly in view of similar research on dialect formation (Blythe, Jones, & Renton, 2016).

---

[2]HRL provides an approximation of two interacting groups: probability weights may be viewed as capturing relative numbers of agents in a population adhering to a given association. Individual-based dynamics is better captured by SRL, especially in view of high values of forgetfulness and reactivity, reflecting severe memory constraints and eager processing, respectively (Christiansen & Chater, 2016).

Acknowledgements

References

Barrett, J., & Zollman, K. J. S. (2009). The role of forgetting in the evolution and learning of language. *Journal of Experimental & Theoretical Artificial Intelligence*, *21*(4), 293–309.

Barrett, J. A. (2007). Dynamic partitioning and the conventionality of kinds. *Philosophy of Science*, *74*(4), 527–546.

Blythe, R. A., Jones, A. H., & Renton, J. (2016). Spontaneous Dialect Formation In A Population Of Locally Aligning Agents. In S. G. Roberts, C. Cuskley, L. McCrohon, L. Barcel-Coblijn, O. Fehr, & T. Verhoef (Eds.), *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11)*. Online at http://evolang.org/neworleans/papers/19.html.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(05), 489–509.

Christiansen, M. H., & Chater, N. (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, e62.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*(2), 243–266.

Lewis, D. K. (1969). *Convention.* Cambridge, MA: Harvard University Press.

Oktaba, K. (2017). *Learning to Signal: Socio-environmental Constraints in Category Formation.* Bachelor's thesis, Institute of Philosophy, University of Warsaw, Poland. (`http://semanticsarchive.net/Archive/TA1OTk0Y/Constraints_in_Category_Formation_Oktaba_18.pdf`)

Roth, A. E., & Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, *8*(1), 164–212.

Skyrms, B. (2008). Signals. *Philosophy of Science*, *75*(5), 489–500.

Skyrms, B. (2010). *Signals: Evolution, learning, and information.* OUP Oxford.

Wallentin, M., & Frith, C. D. (2008). Language is shaped for social interactions, as well as by the brain. *Behavioral and Brain Sciences*, *31*(5), 536–537.