EMERGENCE OF COMMUNICATION WITH SELFISH AGENTS

MICHAEL NOUKHOVITCH[*1,2], TRAVIS LACROIX[3,2], ANGELIKI LAZARIDOU[4], and AARON COURVILLE[1,2]

[*]Corresponding Author: mnoukhov@gmail.com
[1]Université de Montréal, Montréal, Canada
[2]Mila, Québec Artificial Intelligence Institute, Montréal, Canada
[3]Department of Logic & Philosophy of Science, University of California, Irvine, USA
[4]Deepmind, London, United Kingdom

Previous work on the evolution of meaning has taken advantage of the signalling-game framework—initially proposed by Lewis (1969) and extended to a dynamic setting by Skyrms (2014/1996, 2010). In the basic sender-receiver game (sometimes called a *referential* in the machine learning community), signals are used by an agent to disambiguate from a number of possible referents. There are two agents, called the sender and receiver. The sender observes some state and sends a message to the receiver; the receiver observes the message, but cannot observe the state of the world directly, and she subsequently chooses an action which determines payoffs for both agents.

A common assumption for this model is that the agents' interests are perfectly aligned, but more realistic assumptions entail that agents can have, at least partial, conflicts of interest. For example, vervet monkeys (and other mammalian species) employ simple systems of communication, but when and whether an individual sends a signal in a given context may be subject to 'audience effects'—individuals are more likely to produce alarm calls while in proximity to others than when they are alone, and they are more likely to produce alarm calls when they are in the presence of kin than when they are in the presence of non-kin (Cheney & Seyfarth, 2018). Previous work in game theory—both in the setting of economics and biology—has sought to determine whether information transfer is possible when the players have conflicting goals. For example, Skyrms (2010) considers a small number of cases where the players' interests are imperfectly aligned; Crawford and Sobel (1982) provide a more general treatment of the entailments of divergent interests; and Wagner (2012, 2014) shows that meaning can even be conveyed in a zero-sum game, though the resultant dynamics will be chaotic. When signalling is *costly*, communication may be stable—this applies equally well to economics (Spence, 1973) as it does to biology (Zahavi, 1975). In this case, a cost for send-

ing a signal can help stabilise the possibility of honest signalling.[1] As for costless cheap talk with an existing protocol, the general take away has been that it can theoretically lead to Pareto optimal behaviour, but this is not guaranteed (Farrell & Rabin, 1996). Furthermore, these results do not take into account situations where the protocol must be learned in the first place. Many of these results depend significantly upon the game, and modelling assumptions, in question. Godfrey-Smith and Martínez (2013) analyse static signalling games to see whether common interest is a predictor of communicative viability; they show that it is possible for communication to persist in games that can be characterised as having low levels of common interest. Martínez and Godfrey-Smith (2016) complement this work with a dynamic analysis of signalling with conflict of interest using the *replicator dynamic* (Taylor & Jonker, 1978).

To study the dynamic emergence of language in situations of partial conflict, we introduce a modified sender-receiver game with a parameter that smoothly defines the game between fully cooperative (agents share a reward) and fully competitive (zero-sum). The reward of the players depends on successful communication between the sender and receiver; but, in the presence of conflict of interests, neither agent should be fully informative or trustworthy in communication. We use computer agents to play the game and train them using deep reinforcement learning to selfishly optimise their own rewards and in doing so force them to learn to communicate from scratch (Foerster, Assael, Freitas, & Whiteson, 2016; Havrylov & Titov, 2017; Lazaridou, Peysakhovich, & Baroni, 2017).

For different levels of cooperation defined by our parameter, we do a thorough search over agent configurations to try to learn effective communication with our selfish learning rule. We find that, contrary to current literature in machine learning,[2] communication always *stably* emerges between selfish agents in games that are more cooperative than competitive, without any special learning rules. In games that are more competitive, we find that basic learning rules emerge communication unstably, inefficiently, and with chaotic dynamics as the act of honest communicating is no longer a strictly dominant strategy and more resembles cooperation in a prisoner's dilemma. We test a more complex learning rule that imbues agents with a form of theory-of-mind about their opponent (Foerster et al., 2018) and find that it improves stability and efficacy of communication and allows agents to effectively learn to communicate *even* in highly competitive scenarios. Thus, we propose that communication emerges naturally in games that are more cooperative than competitive. For more competitive 2-player games, we propose three properties that are beneficial to the efficacy and stabilisation of emergent

[1] See also Grafen (1990b, 1990a), Maynard Smith and Harper (2003), Zollman, Bergstrom, and Huttegger (2012). Lachmann, Szamado, and Bergstrom (2001) highlight that in some cases signals need only be costly outside of equilibrium.

[2] See Cao et al. (2018), Jaques et al. (2018)

communication: when the game is iterative or played with the same opponent, when the game is general-sum, and when agents are imbued with theory-of-mind.

References

Cao, K., Lazaridou, A., Lanctot, M., Leibo, J. Z., Tuyls, K., & Clark, S. (2018). Emergent communication through negotiation. *CoRR*, *abs/1804.03980*.

Cheney, D. L., & Seyfarth, R. M. (2018). Flexible usage and social function in primate vocalizations. *Proceedings of the National Academy of Sciences*, *15*(9), 1974–1979.

Crawford, V. P., & Sobel, J. (1982). Strategic information transmission. *Econometrica*, *50*(6), 1431–1451.

Farrell, J., & Rabin, M. (1996). Cheap talk. *The Journal of Economic Perspectives*, *10*(3), 103–118.

Foerster, J. N., Assael, Y. M., Freitas, N. de, & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. In *Advances in neural information processing systems.*

Foerster, J. N., Chen, R. Y., Al-Shedivat, M., Whiteson, S., Abbeel, P., & Mordatch, I. (2018). Learning with opponent-learning awareness. In *International conference on autonomous agents and multiagent systems.*

Godfrey-Smith, P., & Martínez, M. (2013). Communication and common interest. *PLoS Computational Biology*, *9*(11), 1–6.

Grafen, A. (1990a). Biological signals as handicaps. *Journal of Theoretical Biology*, *144*, 517–546.

Grafen, A. (1990b). Sexual selection unhandicapped by the fisher process. *Journal of Theoretical Biology*, *144*, 473–516.

Havrylov, S., & Titov, I. (2017). Emergence of language with multi-agent games: Learning to communicate with sequences of symbols. In *Advances in neural information processing systems* (pp. 2149–2159).

Jaques, N., Lazaridou, A., Hughes, E., Gülehre aglar, Ortega, P. A., Strouse, D., Leibo, J. Z., & Freitas, N. de. (2018). Social influence as intrinsic motivation for multi-agent deep reinforcement learning. In *International conference on machine learning.*

Lachmann, M., Szamado, S., & Bergstrom, C. T. (2001). Cost and conflict in animal signals and human language. *Proceedings of the National Academy of Sciences*, *98*(23), 13189–13194.

Lazaridou, A., Peysakhovich, A., & Baroni, M. (2017). Multi-agent cooperation and the emergence of (natural) language. In *International conference on learning representations.*

Lewis, D. (1969). *Convention: A philosophical study.* Oxford: Wiley-Blackwell.

Martínez, M., & Godfrey-Smith, P. (2016). Common interest and signaling games: a dynamic analysis. *Philosophy of Science*, *83*(3), 371–392.

Maynard Smith, J., & Harper, D. (2003). *Animal signals.* Oxford: Oxford Uni-

versity Press.
Skyrms, B. (2010). *Signals: Evolution, learning, & information.* Oxford: Oxford University Press.
Skyrms, B. (2014/1996). *Evolution of the social contract.* Cambridge: Cambridge University Press.
Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374.
Taylor, P., & Jonker, L. (1978). Evolutionarily stable strategies and game dynamics. *Mathematical Biosciences*, *40*, 145–156.
Wagner, E. O. (2012). Deterministic chaos and the evolution of meaning. *British Journal for the Philosophy of Science*, *63*, 547–575.
Wagner, E. O. (2014). Conventional semantic meaning in signalling games with conflicting interests. *British Journal for the Philosophy of Science*, *66*(4), 751–773.
Zahavi, A. (1975). Mate selection: A selection for a handicap. *Journal of Theoretical Biology*, *53*(1), 205–214.
Zollman, K. J. S., Bergstrom, C. T., & Huttegger, S. M. (2012). Between cheap and costly signals: The evolution of partially honest communication. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1750), 20121878.