Signalling signalhood in machine learning agents

Abstract for online workshop, **Machine Learning & the Evolution of Language**

Edward Hughes
DeepMind, London, UK
edwardhughes@google.com
http://edwardhughes.io

Abhinav Gupta
MILA, Montreal, QC, Canada
abhinavg@nyu.edu
https://mila.quebec/en/person/abhinav-gupta/

Ekaterina I. Tolstaya
Waymo LLC, New York, NY, USA
eig@waymo.com
http://katetolstaya.com/

Thom Scott-Phillips
Central European University, Budapest, Hungary
thom.scottphillips@gmail.com
https://thomscottphillips.com/

Background. Experimental language evolution investigates how pairs or groups interacting individuals create communicative conventions, and how those conventions develop some of the properties characteristic of natural languages. Here, we report ongoing work aiming to replicate a key task in this literature, commonly known as the Embodied Communication Game (Scott-Phillips et al., 2009; Kouwenhoven et al., 2022)—but with artificial agents rather than human participants. The distinctive feature of the Embodied Communication Game is that there is no *a priori* difference between communicative and non-communicative behaviour. This poses a boot-strapping challenge that may be important for artificial agents.

Methods & Results. We characterize the Embodied Communication Game as a decentralized partially observable Markov decision process, and implement it as a multi-agent reinforcement learning environment. Training state-of-the-art reinforcement learning agents in one-shot episodes via self-play or population-play produces a fixed default color strategy that is not only not communicative, but also cannot generalize to new co-players in the 'other-play' setting (Hu et al. 2020). In few-shot episodes, which is a memory-based meta-learning paradigm (Duan et al. 2016), agents find a local optimum of maximizing the number of rounds played, but do not achieve communication.

Discussion. Comparison between these results and existing experimental results with human participants reveals deep and serious challenges for replicating human communicative competence in artificial agents. Speculatively, we suggest these tasks could be performed by agents who have (i) goals with respect to others' internal ('mental') states, and (ii) models of others' goals and the means by which those goals might be satisfied. Some limited progress has been made in this direction using Bayesian approaches (e.g. Ho et al., 2021), but such work is in its infancy and there is an enormous amount to be done. Whether these capacities can emerge via multi-agent reinforcement learning is presently unknown.

References

Duan, Y., Schulman, J., Chen, X., Bartlett, P. L., Sutskever, I., & Abbeel, P. (2016). Rl \$^ 2\$: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint arXiv:1611.02779*.

Ho, M. K., Cushman, F., Littman, M. L., & Austerweil, J. L. (2021). Communication in action: Planning and interpreting communicative demonstrations. *Journal of Experimental Psychology: General, 150*(11), 2246-2272.

Hu, H., Lerer, A., Peysakhovich, A., & Foerster, J. (2020, November). “Other-Play” for Zero-Shot Coordination. In *International Conference on Machine Learning* (pp. 4399-4410). PMLR.

Kouwenhoven, T., de Kleijn, R., Raaijmakers, S., & Verhoef, T. (2022). Need for structure and the emergence of communication. In: *Proceedings of the Annual Meeting of the Cognitive Science Society, 44*.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition, 113*(2), 226-233.