The spontaneous emergence of discrete and compositional messages

Anonymous ACL submission

Abstract

blah blah blah

1 Introduction

In a signalling game, artificial agents communicate to achieve a common goal: a sender sees some piece of information and produces a message, this message is then sent to a receiver that must take some action. If the action is appropriate, the whole communication stream, and in particular the choice of the message, is reinforced. For instance, in a referential game, sender and receiver see a set of objects, and the sender must send a message to the receiver, so that the receiver can pick up the right object, as determined in advance for the sender, but unbeknownst to the receiver.

2 Function Games

We here introduce a general communication game setting, which we call Function Games. Our games contain three basic components: (i) a set of contexts C, (i) a set of actions A, (ii) a family of functions F, from contexts to actions. One play of a Function Game game runs as follows:

1. Nature chooses $f \in F$ and a context $c \in C$.
2. Sender sees the context c and $f(c)$. I like f(c) here, but f is a bit more appropriate. What do you all think?
3. Sender sends a message m to Receiver.
4. Receiver sees *a possibly different* context c' and the message m and chooses an action a'.
5. Both are 'rewarded' iff $a' = f(c')$.

Two concrete interpretations will be helpful in illustrating the various components.

Generalized referential games. A reference game is one in which Sender tries to get Receiver to pick the correct object out of a given set (Skyrms, 2010; Lazaridou et al., 2017, 2018; Havrylov and Titov, 2017; Chaabouni et al., 2019). Here, contexts are sets of objects (i.e. an $m \times n$ matrix, with m objects represented by n features). Normally (though we will drop this assumption later), $c' = \texttt{shuffled}(c)$: Sender and Receiver see the same objects, but in a different arrangement. Actions are the objects, and the functions $f \in F$ are *choice functions*: $f(c) \in c$ for every context c.

Belief update games. Contexts can represent possible belief states for the agents. Letting $A = C$, the functions will then be 'belief update' functions, representing e.g. how to update an agent's beliefs in the light of learning a new piece of information. What should we cite here? Something from dynamic semantics?

3 Experiment

3.1 Model

3.2 Game Parameters

- strict vs. non-strict context
  - num objects for non-strict
- equal vs. not equal
- object size (num properties)
- latent space (msg) dimension [didn't vary this]

4 Results

4.1 Communicative success

4.2 Discrete signals

4.3 Compositionality

5 Discussion

6 Conclusion

(Steinert-Threlkeld, 2019)

References

Rahma Chaabouni, Eugene Kharitonov, Emmanuel Dupoux, and Marco Baroni. 2019. Anti-efficient encoding in emergent communication. In *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*.

Serhii Havrylov and Ivan Titov. 2017. Emergence of Language with Multi-agent Games: Learning to Communicate with Sequences of Symbols. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS 2017)*.

Angeliki Lazaridou, Karl Moritz Hermann, Karl Tuyls, and Stephen Clark. 2018. Emergence of Linguistic Communication from Referential Games with Symbolic and Pixel Input. In *International Conference of Learning Representations (ICLR 2018)*.

Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. 2017. Multi-Agent Cooperation and the Emergence of (Natural) Language. In *International Conference of Learning Representations (ICLR2017)*.

Brian Skyrms. 2010. *Signals: Evolution, Learning, and Information*. Oxford University Press.

Shane Steinert-Threlkeld. 2019. Paying Attention to Function Words. In *Emergent Communication Workshop @ NeurIPS 2018*.