

Multi-agent communication: What works, what doesn't and what's next

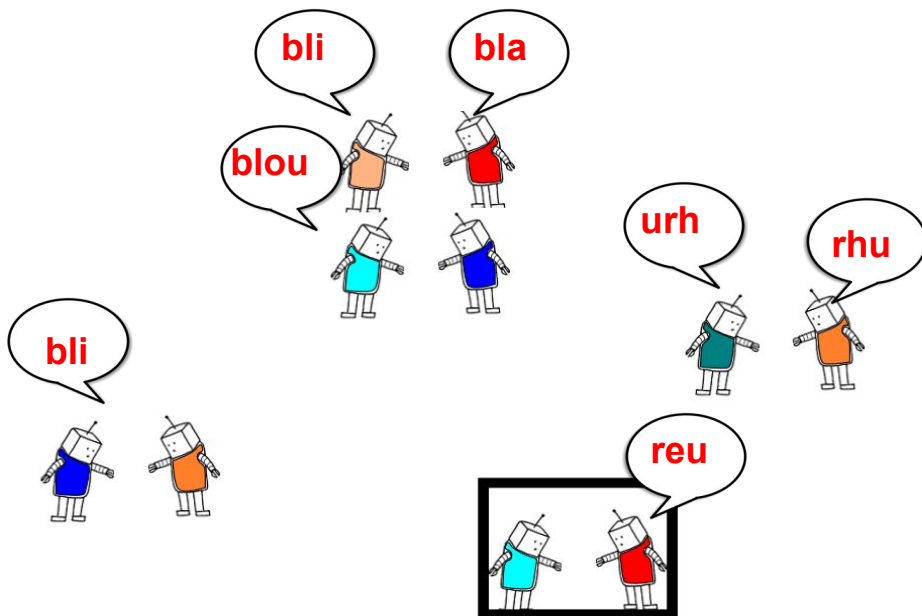
Angeliki Lazaridou

September 19th

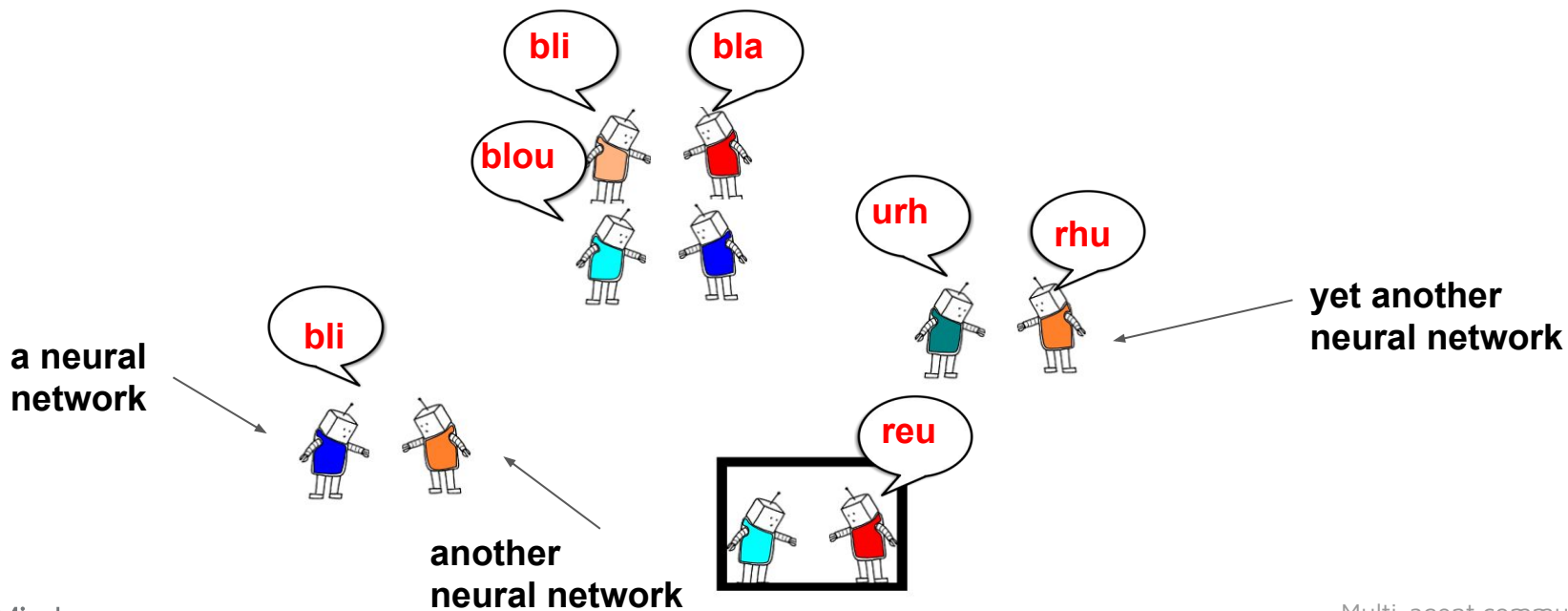
AθNLP



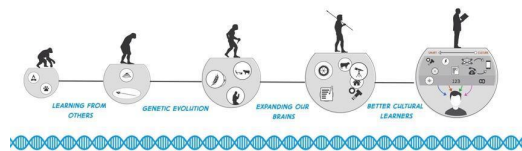
How does an effective communication system arise among a collection of initially noncommunicating individuals?



How does an effective communication system arise among a collection of initially noncommunicating individuals?

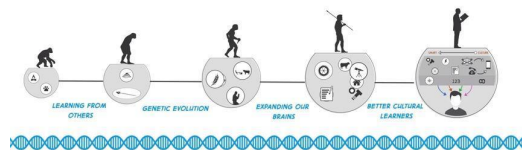


Different facets of multi-agent communication

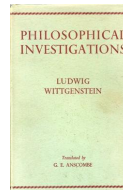
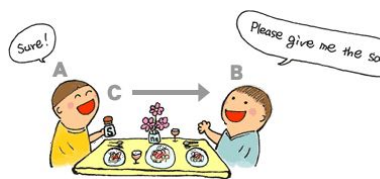


Experimental framework for **language emergence and evolution**

Different facets of multi-agent communication

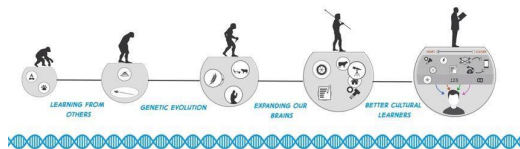


Experimental framework for **language emergence and evolution**



Alternative paradigm to **interactive language learning**

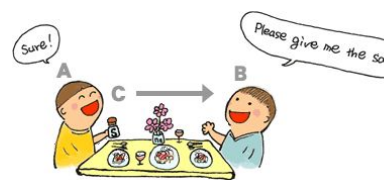
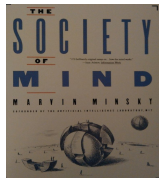
Different facets of multi-agent communication



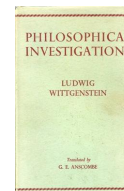
Experimental framework for **language emergence and evolution**



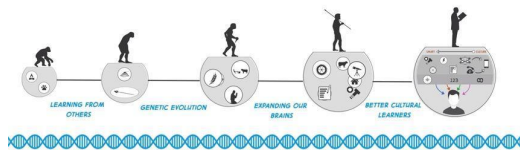
Facilitates **knowledge transfer** and **co-ordination** among agents



Alternative paradigm to **interactive language learning**



Different facets of multi-agent communication

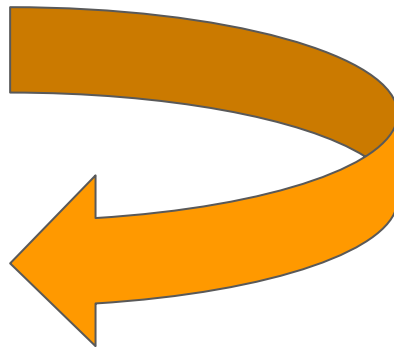
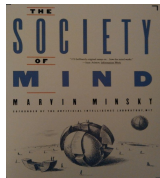


Experimental framework for **language emergence and evolution**

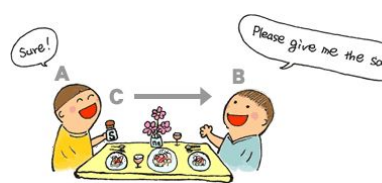
interested in scientific questions



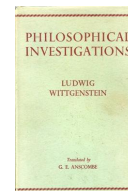
Facilitates **knowledge transfer** and **co-ordination** among agents



interested in engineering questions

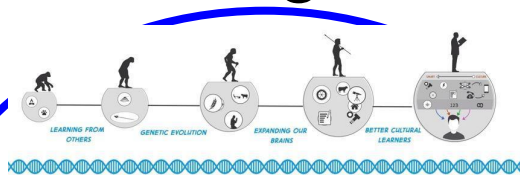


Alternative paradigm to **interactive language learning**



Different facets of multi-agent communication

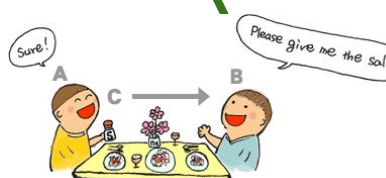
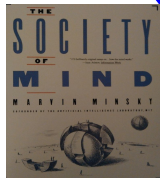
[^]
interconnected



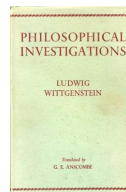
Experimental framework for **language emergence and evolution**



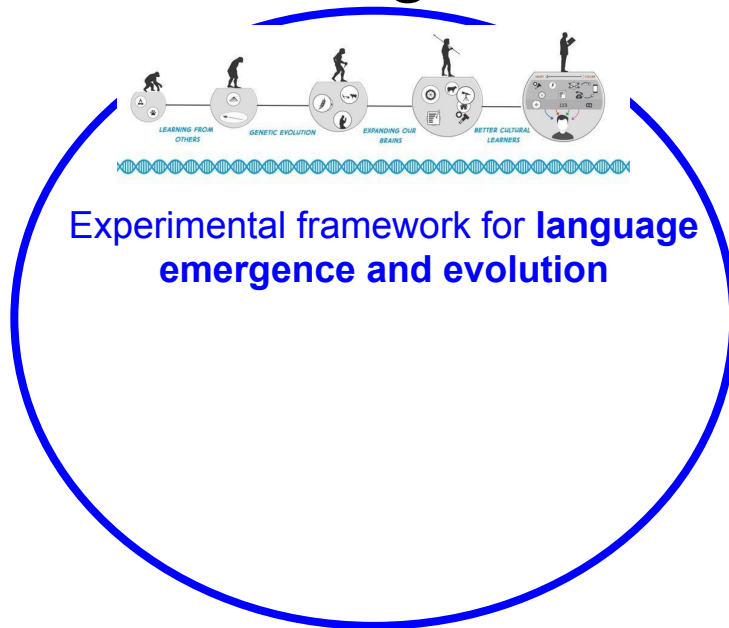
Facilitates **knowledge transfer** and **co-ordination** among agents



Alternative paradigm to **interactive language learning**



Different facets of multi-agent communication



The mystery (and many theories) of origins of language

- Chomsky
- Pinker & Bloom
- Jackendoff
- Deacon
- Hurford
- Tomasello
-
-

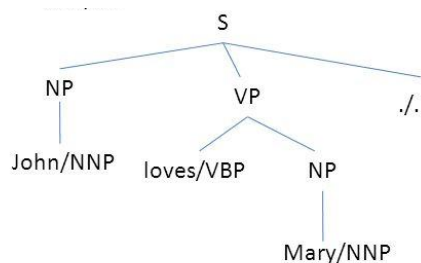


Human communication

The mystery (and theories) of emergence of structure

Compositionality of language

The meaning of the whole is a combination of the meaning of its parts and the way they combine



S => NP VP .
NP => NNP
VP => VBP NP
NNP => John
NNP => Mary
VBP => loves
. => .















- Perhaps the most **distinguished** property of human and communication
- Making use of **finite** means to create **infinite** meanings
 - Allows humans to communicate in a productive way, creating novel utterances on the fly and as needed

Human communication

Case study: Holistic vs compositional protocols

from Kirby et al. (2015)

emerged protocols

	pihino		kapa		newhomo
	nemone		gakho		kamone
	piga		wuwele		gaku
	kawake		nepi		hokako

grammar that explains the protocols

$S : 02 \rightarrow$ kapa

$S : 03 \rightarrow$ nepi

$S : 12 \rightarrow$ gaku













$S : 13 \rightarrow$ pihino

holistic protocols

Holistic vs compositional protocols

from Kirby et al. (2015)

emerged protocols

	pihino		kapa		newhomo
	nemone		gakho		kamone
	piga		wuwele		gaku
	kawake		nepi		hokako

grammar that explains the protocols













$S : 02 \rightarrow$ kapa

$S : 03 \rightarrow$ nepi

$S : 12 \rightarrow$ gaku

$S : 13 \rightarrow$ pihino

holistic protocols

	ege-wawu		mega		gamene-wawu
	ege-wawa		mega-wawa		gamene-wawa
	ege-wuwu		mega-wuwu		gamene-wuwu
	ege		wulagi		gamane

$S \rightarrow A B$

$A : 0 \rightarrow$ ege

$A : 1 \rightarrow$ mega

$B : 2 \rightarrow$ wawu

$B : 3 \rightarrow$ wawa

compositional protocols

(most papers don't find more than 1 level of recursivity)

Emergence of communication in the lab

- Communication of humans in the lab **without** pre-established linguistics conventions
- Very often communicating **not using language** (so as not to bias communication with linguistic compositionality), but rather drawings or sounds
- Source of inspiration of computer simulations to follow



Published online 2011 Feb 17. Prepublished online 2010 Sep 7.
doi: [10.3389/fnhum.2011.00011](https://doi.org/10.3389/fnhum.2011.00011)

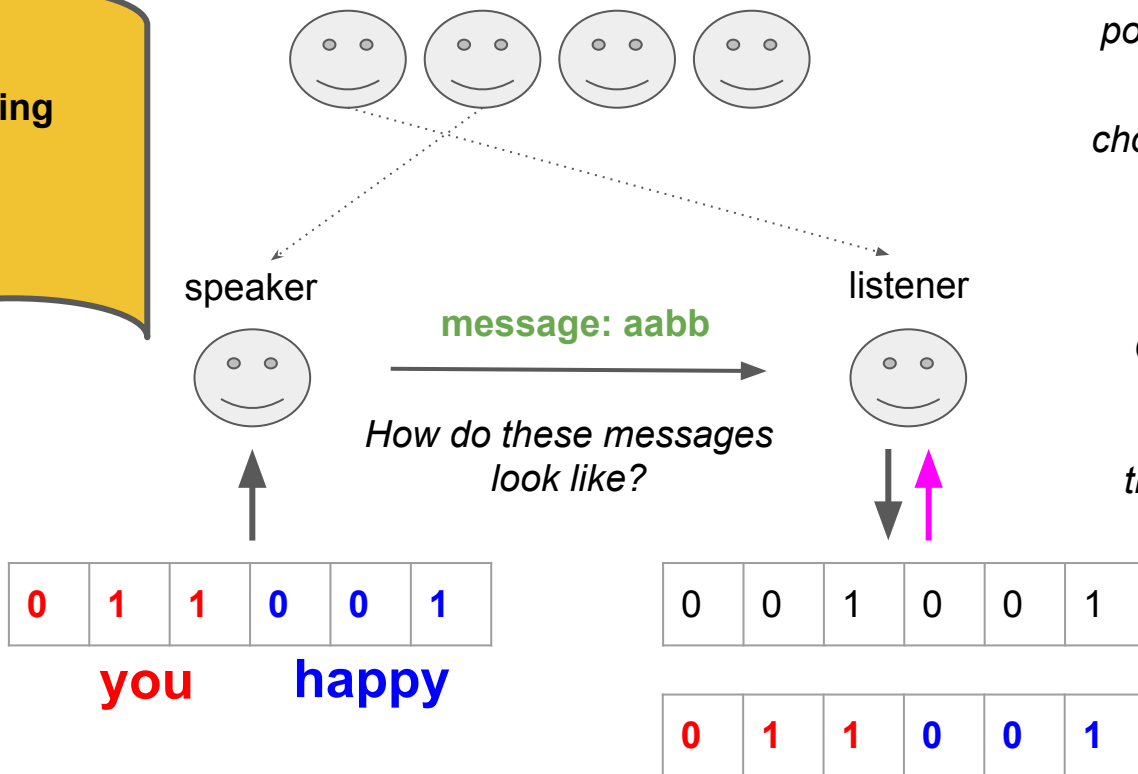
Experimental Semiotics: A Review

[Bruno Galantucci](#)^{1,2,*†} and [Simon Garrod](#)^{3,4,†}

Computational simulations: Emergent communication in **cultural transmission**

Observational Learning

Hare and Elman 1995
Batali 1998, 2002
Kirby
inter alia



population of agents

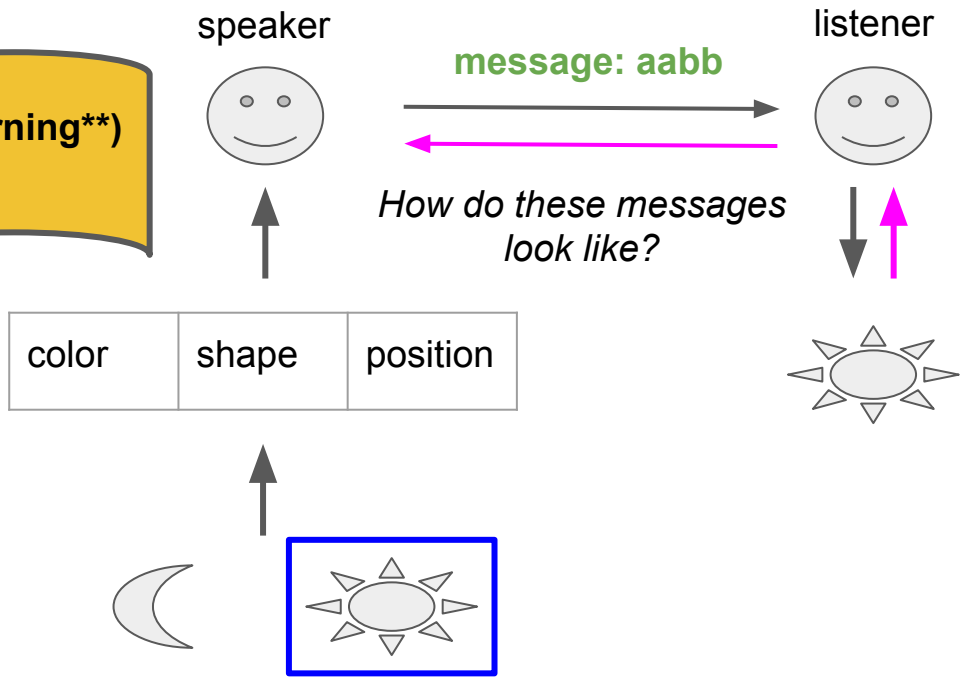
choose 2 agents from the population

Learning:

*Only the listener is being trained to correctly interpret the **message** of the speaker*

Computational simulations: Emergent communication in **language games**

(Reinforcement Learning**)
Luc Steels



Learning: Both speaker and listener are trained. The speaker has to describe the **target object** with a **message** that helps the listener pick the right object.

If you want to know more....

Progress in the Simulation of Emergent Communication and Language

Kyle Wagner¹, James A. Reggia², Juan Uriagereka³, Gerald S. Wilkinson⁴

¹Sparta, Inc. ²Department of Computer Science, University of Maryland, College Park

*³Department of Linguistics, University of Maryland, College Park ⁴Department of Biology,
University of Maryland, College Park*

Adaptive Behaviour, 2003

What works?

- Very important work laying the groundwork for multi-agent communication
- Asking interesting and important scientific questions
- Analyses and hypotheses approaches that are going to influence the future (our present)

Challenges: Beyond “simple” stimuli



has_legs

meows

has_whiskers

has_tail

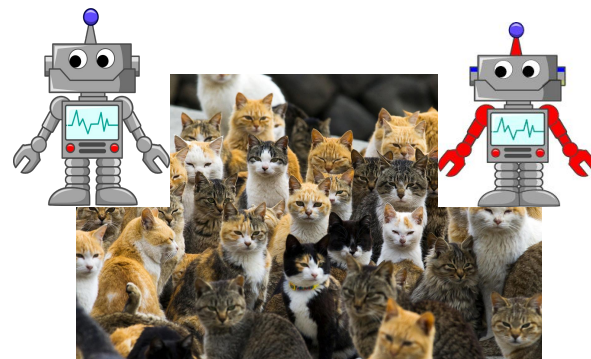
What this research is about

Challenges: Beyond “simple” stimuli



has_legs
meows
has_whiskers
has_tail

What this research is about



What people think it's all about

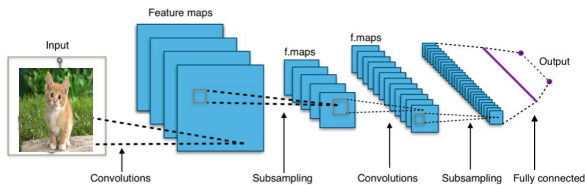
Will conclusions and findings transfer to
real world???**

Challenges: Beyond “simple” stimuli

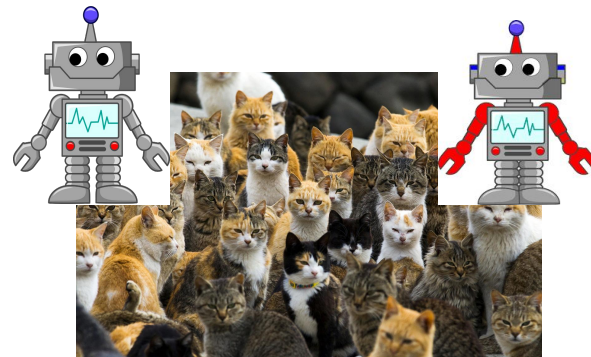


What this research is about

IMAGENET

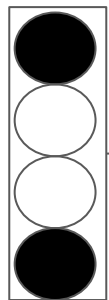


Perhaps, we can find a middle ground



What people think it's all about

Symbolic data



A cute grey cat

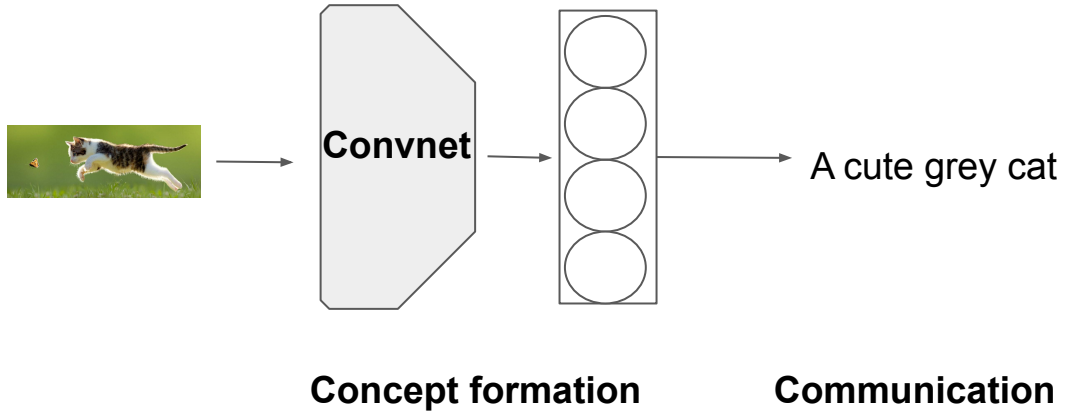
*Very nice representation, chunked into pieces already, a lot of information, object is a combination.
Concepts are there, we just need to communicate about them!!!*



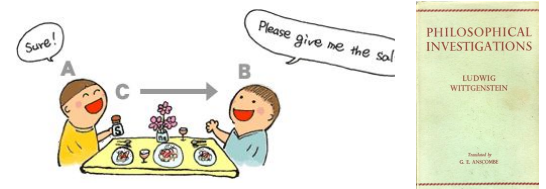
Symbolic data



Pixel data



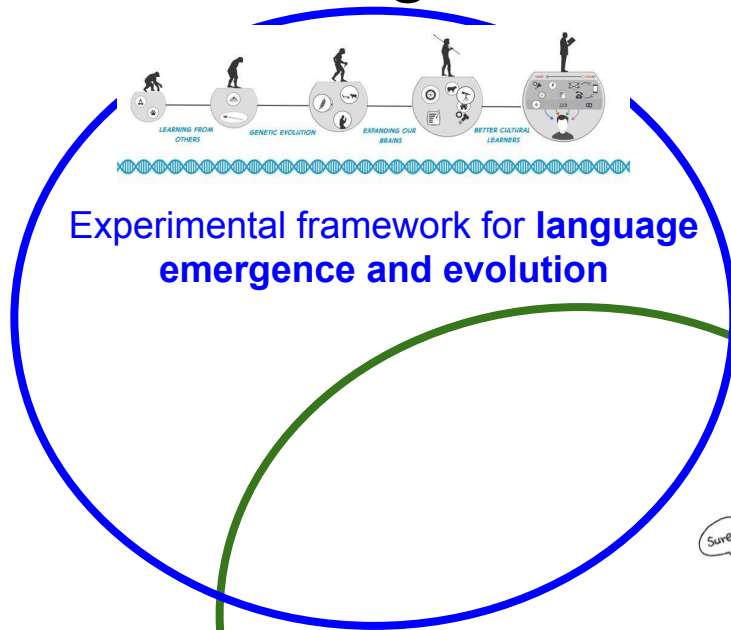
Different facets of multi-agent communication



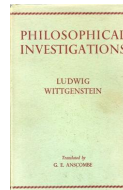
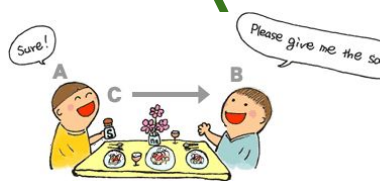
Alternative paradigm to
interactive language learning

Different facets of multi-agent communication

[^]
interconnected



Paradigm inspired and shaped from **previous-era research** but using modern experimental setup (Deep RL + vision)



Alternative paradigm to **interactive language learning**

Towards functional language learning

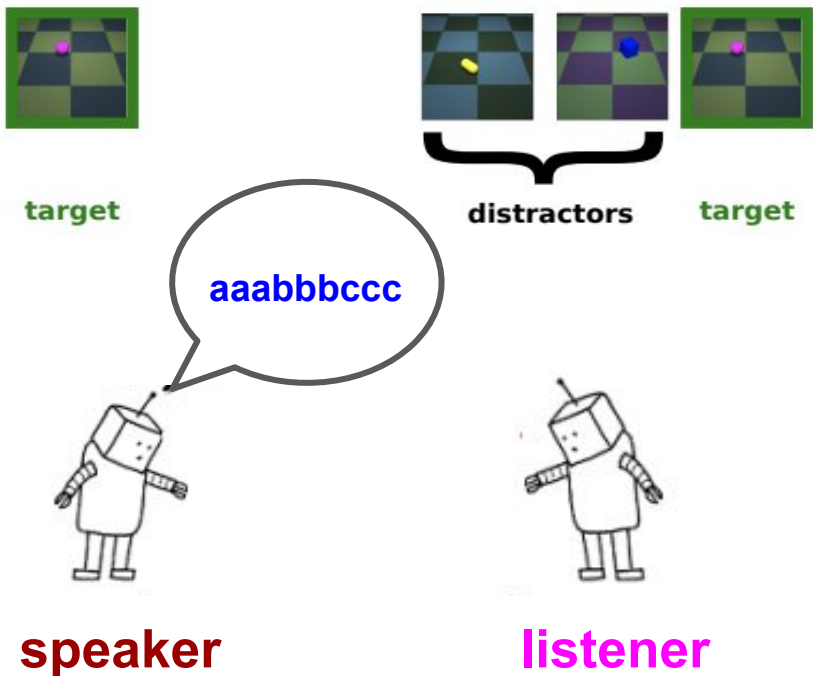
Language research is getting interested in more functional ways of training agents
(agents that have intentions and goals and communicate to achieve them vs agents that just do pattern learning like image classification, language modeling)

- Interested in emergent communication that can act as **scaffolding** to natural language or as a **pre-training** mechanism
- Questions regarding degree of **compositionality** of emergent communication and **interpretability** are central

Recent work on (cooperative) emergent communication

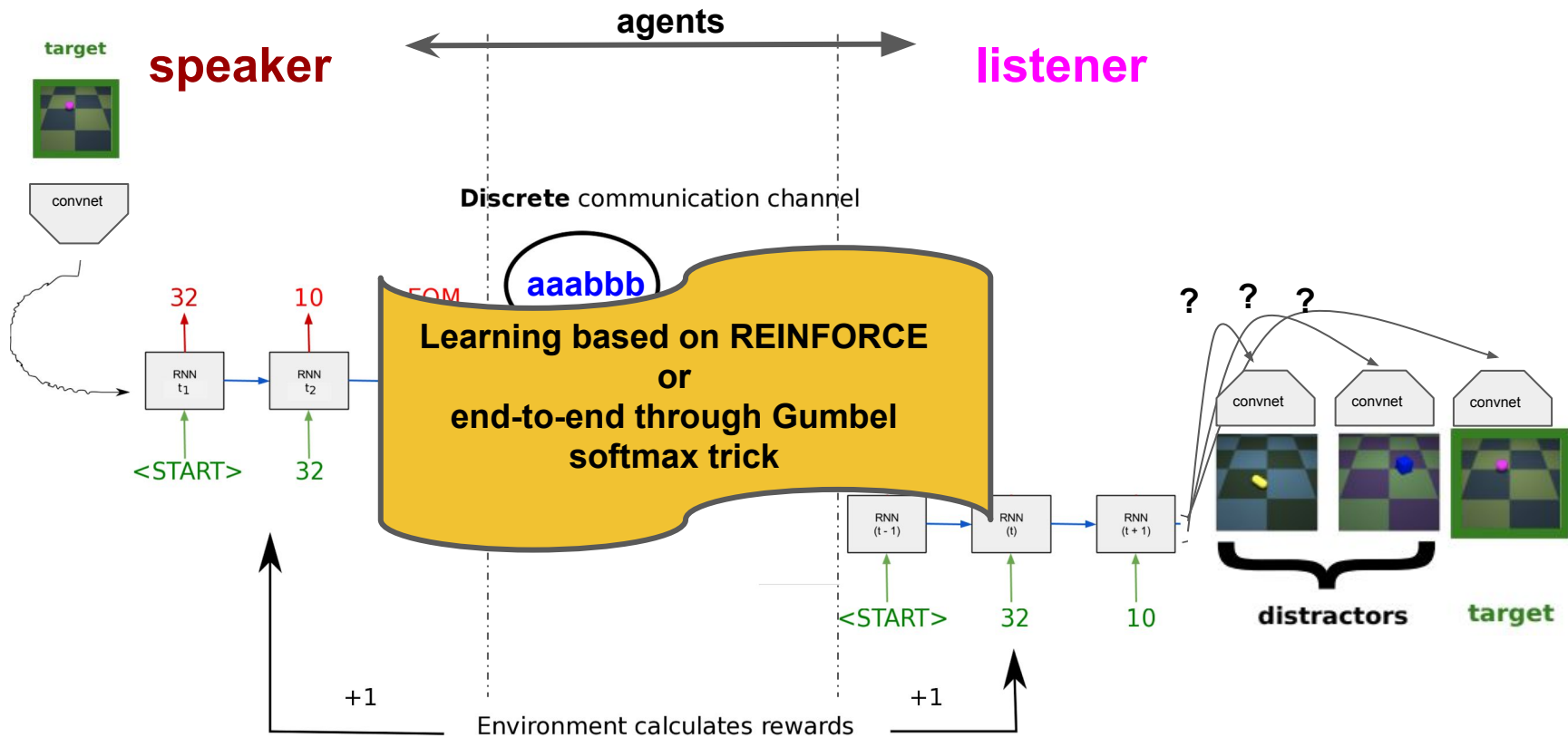
- Referential-ish (language) games on visual stimuli
 - Lazaridou et al. (2016); Havrylov and Titov, (2017); Evtimova et al. (2017); Lazaridou et al. (2017); Choi, Lazaridou et al. (2017); Bouchacourt and Baroni (2018); Graesser et al. (2018); Zhang et al. (2019)
- Dialogue on symbolic stimuli
 - Kottur et al. (2017); de Vries et al. (2018)
- Question asking on visual stimuli
 - Jorge et al. (2016)
- Co-operative tasks in visual worlds
 - Bogin et al. (2018); Das et al. (2018)

Emergent communication in referential games



- **Speaker** is presented with a **target** image
- The **listener** is presented with a collection of images, one of them being the **target** and needs to identify it.
- The **speaker** emits a discrete **communication** message, a single or a sequence of symbols
- The **listener** makes a choice of which one is the target.
- The meaning of the symbols is **emergent**. At the beginning, there is no association between symbols and meaning. It emerges throughout the game

Emergent communication in referential games (II)



What works?

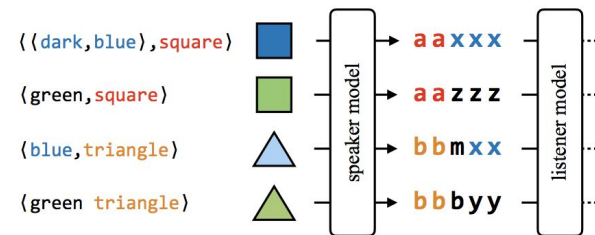
- Improvement from symbolic small data of past research to more realistic data
 - Larger datasets, often consisting of embodied environments or visual stimuli
- Novel and more realistic modeling
 - Agents with attention, equipped with ConvNets for visual processing, recurrent models of sequences
- Even with all the complexity, learning is happening!
 - The deep RL toolkit seems to be a good choice for answering these questions
 - We can start scaling up the research on language evolution

Reminder

- Interested in emergent communication that can act as **scaffolding** to natural language or as a **pre-training** mechanism
- Questions regarding degree of **compositionality** of emergent communication and **interpretability** are central

Challenges: Measuring structure of emergent protocols

- What's the **structure** (if any) in the messages: aaabbb, cccbbb???
- Zero-shot to novel stimuli, productivity [Lazaridou et al. (2017)]
- *Isomorphism* between emergent and natural language, e.g., emergent linguistic space is a rotation of the natural language space [Brighton and Kirby (2006); Lazaridou et al. (2017)]
- **Tree Reconstruction Error**: deconstructing messages into a set of *additive/concatenative* operations [Andreas (2019)]



Challenges: Interfacing emergent language to natural language

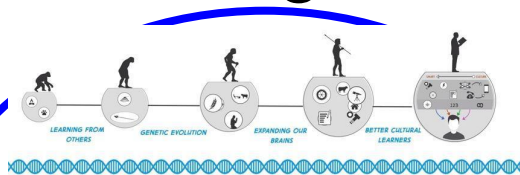
- Emergent communication is emergent: aaabbb, bbccabab??
- Assuming **parallel stimuli**
 - Multi-task learning: REINFORCE for communication game and cross-entropy for supervised natural language [Lazaridou et al. (2017); Lee et al., 2018]
 - Post-training: IBM-1 model between messages and environment concepts [Bogin et al., 2018]
- Not assuming parallel stimuli
 - Language model regularizer so that emergent languages has low perplexity under a language model trained on natural language [Havrylov and Titov., 2017]

Challenges: Emergent Languages are ad-hoc and not generic

- Emergent languages are compositional only if the experimental configuration (e.g., **number of distractors**) is imposing the right bias [Lazaridou et al. (2017)]
- Compositional languages emerge only with the right biases on the agents' model (e.g., **limited working memory**) [Kottur et al. (2017)]
- Consistent with recent results on disentanglement research random seeds and hyperparameters matter more than the model choice [Locatello et al., 2019]
- No free-lunch -- if we want compositional languages, **we need biases!**
 - Moving from two-agent interactions to **populations of agents** [Tieleman, Lazaridou et al. (under submission)]
 - Mixing referential games (for expressivity) with cultural transmission (for learnability) [ILM, Kirby,]

Different facets of multi-agent communication

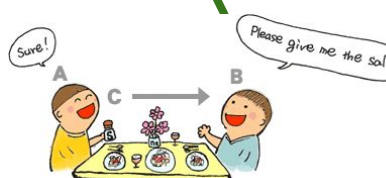
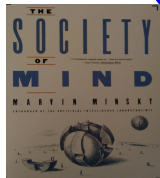
[^]
interconnected



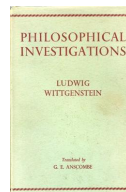
Experimental framework for **language emergence and evolution**



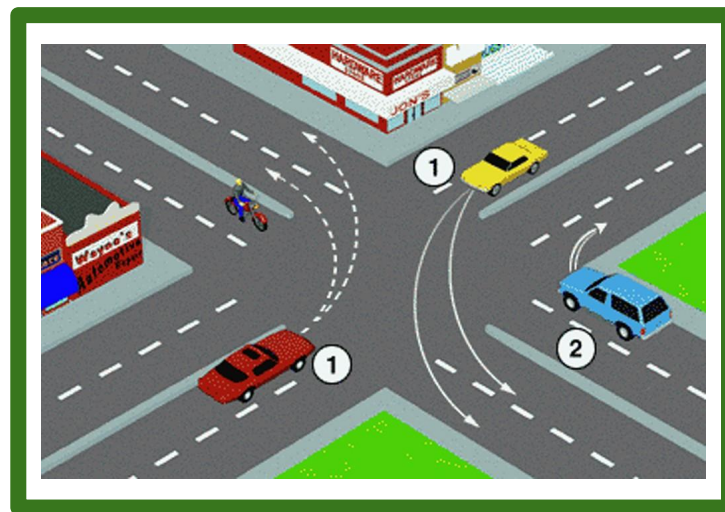
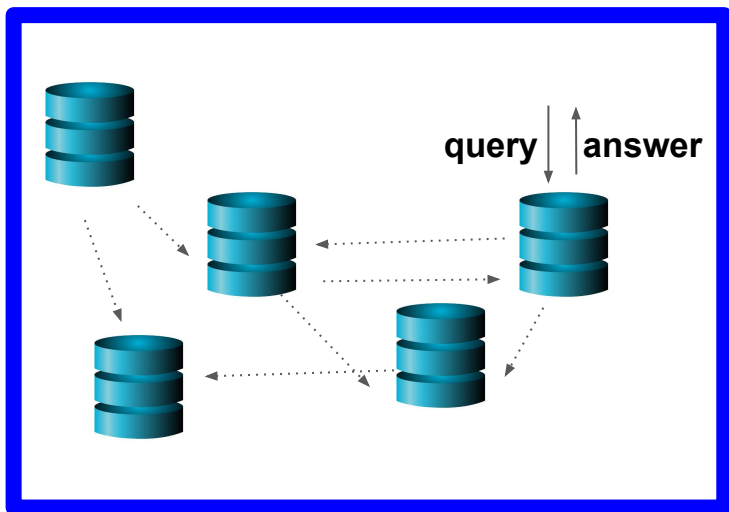
Facilitates **knowledge transfer** and **co-ordination** among agents



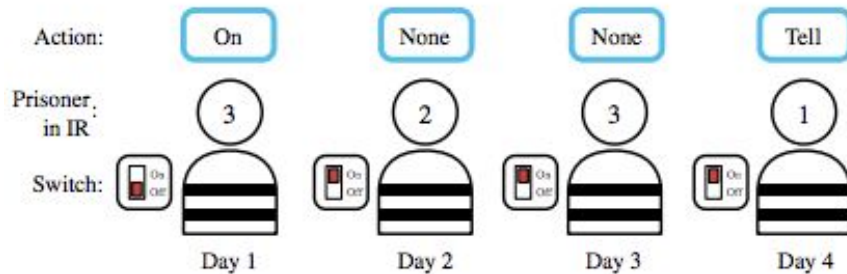
Alternative paradigm to **interactive language learning**



- Multiple agents that exist in an environment and need to solve a task
- Communication is important in facilitating **distributed intelligence** and **coordination**

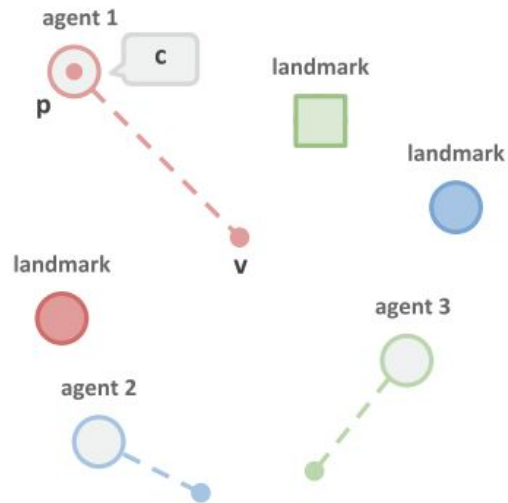


- Agents that can move and talk vs **agents that only talk**
- Visual environments vs **symbolic environments?**
- **Co-operative agents** vs semi cooperative agents?



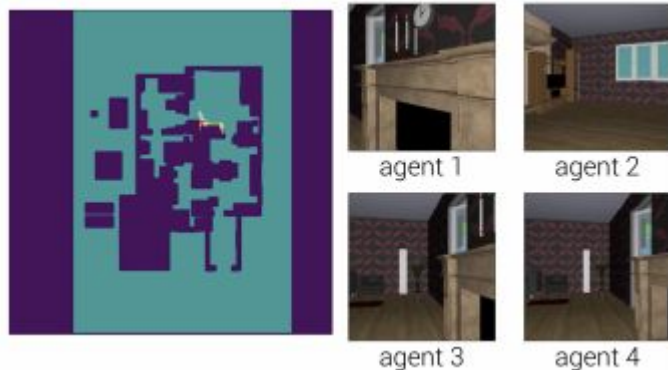
[Foerster et al. (2016)]

- **Agents that can move and talk** vs agents that only talk
- Visual environments vs **symbolic environments**?
- **Co-operative agents** vs semi cooperative agents?



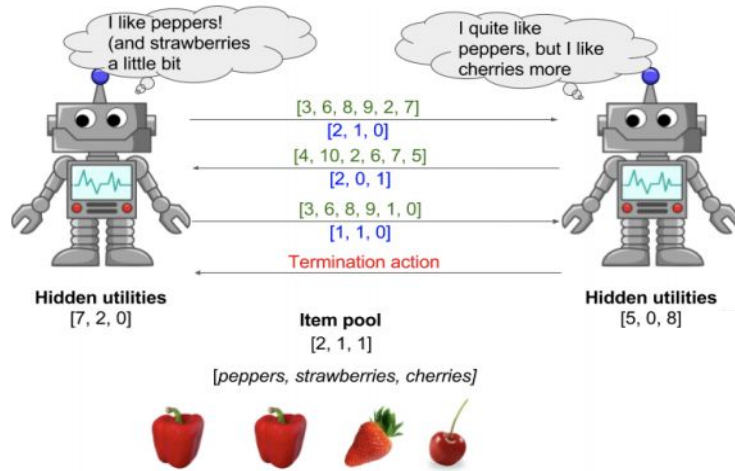
[Mordatch and Abiel et al. (2018)]

- **Agents that can move and talk** vs agents that only talk
- **Visual environments** vs symbolic environments?
- **Co-operative agents** vs semi cooperative agents?



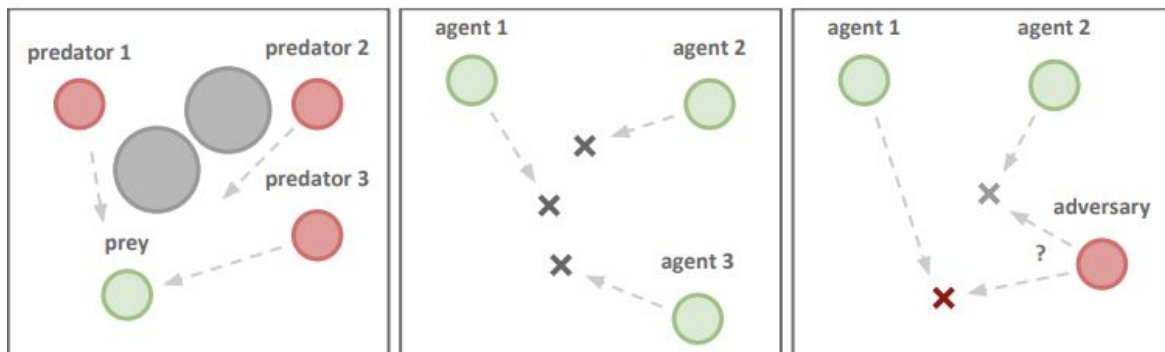
[Das et al. (2019)]

- Agents that can move and talk vs **agents that only talk**
- Visual environments vs **symbolic environments?**
- Co-operative agents vs **semi cooperative agents?**



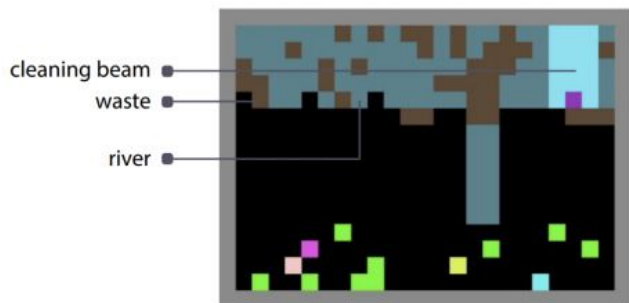
[Cao, Lazaridou et al. (2018)]

- **Agents that can move and talk** vs agents that only talk
- Visual environments vs **symbolic environments**?
- Co-operative agents vs **semi cooperative agents**?



[Lowe et al. (2018)]

- **Agents that can move and talk** vs agents that only talk
- **Visual environments** vs symbolic environments?
- Co-operative agents vs **semi cooperative agents**?

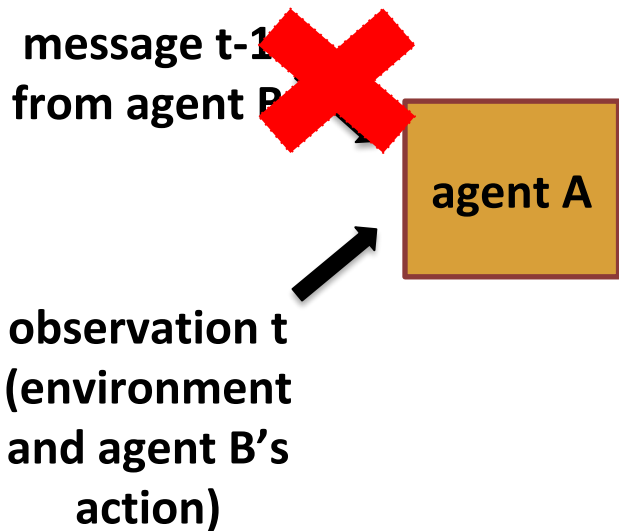


[Jaques, Lazaridou et al. (2019)]

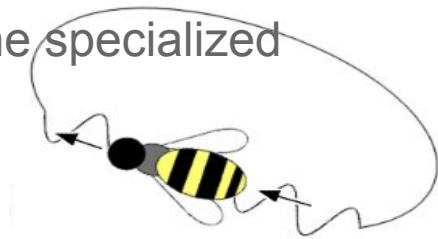


[Singh, Jain et al. (2019)]

Challenges: Joint exploration problem



- Agents A and B are both learning at the same time
 - They are both learning to **talk** and **listen** at the same time
- At the beginning of the training, the messages are meaningless (**random**), but learners are trying to interpret them
- It's easier to communicate on the action space, like **bee dancing** rather than on the extra communication channel.
- All in all, it's easier to **ignore** the specialized communication channel!



Challenges: Do agents really communicate?

- **Suboptimal** use of communication channel
- We need measures of communication: Agents appear to be communicating, but there is no useful information being transmitted, limited **positive signaling** and **positive listening** [Lowe, et al. (2019)]
- Can we do better? **Human communication** can provide inspiration for useful inductive biases
 - E.g., force agents to adhere to Gricean maxims [Eccles, Lazaridou et al. (NeurIPS 2019)]

Challenges: What are we learning, really???

- Messages are at best sequence of discrete symbols, at worse continuous vectors
- Although it might not always be necessary to interpret the messages, it's always good to have an understanding of what's going on
- Either start from language early on, or post-training experiments for testing different hypothesis (e.g., are agents communicating about beliefs vs communicating about object coordinates)

Key takeaways (or, my 2 cents :-))

- **Multi-agent communication needs biases as well**

- Human-centric intelligence is key inspiration for a lot of computation (CNNs, genetic algorithms). Linguistic communication can provide a lot of inspiration (e.g., size of linguistic community, gricean maxims in communication)

- **Metrics, metrics, metrics!**

- We need to be able to measure what we are learning, e.g., how structured are the protocols, what concepts do they communicate, how much communication is really needed?

- **Interfacing to natural language**

- Not only for interpretability and debugging of agents but also for wider applicability of this research to applications like dialogue, human-agent communication or even transfer learning

Are you new to the topic?

- Skim through the papers in these slides (give me 1 day to add proper references :))
- [EGG toolkit](#) by Facebook AI Research
- Emergent Communication [workshop](#) at NeurIPS (deadline on Friday!)
- Review on multi-agent communication under way, stay tuned!