# Emergent communication in cooperative multi-agent environments

Irina Barykina

Outline

- Why communication is important?

- What might be wrong with statistics-based NLP methods?

- Situated goal-driven language learning paradigm.

- Basics of reinforcement Q learning.

- Application of Q learning to emergent languages.

- When does syntax emerges?

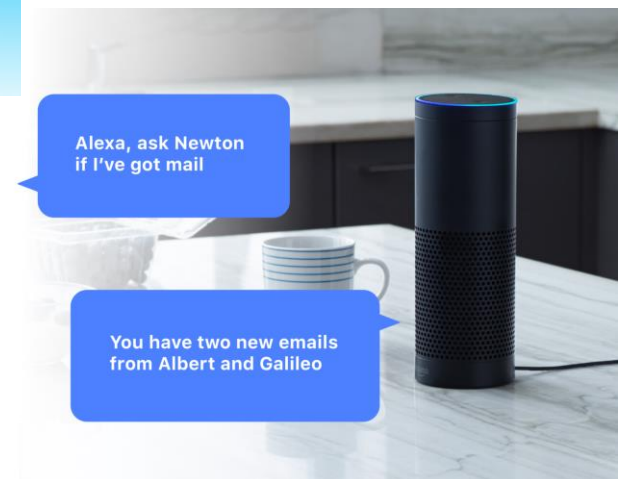- Pros and cons of approach

Why communication is important?
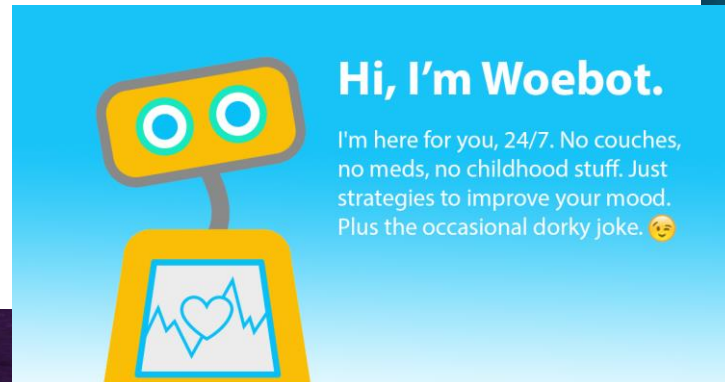
[softbankrobotics.com]
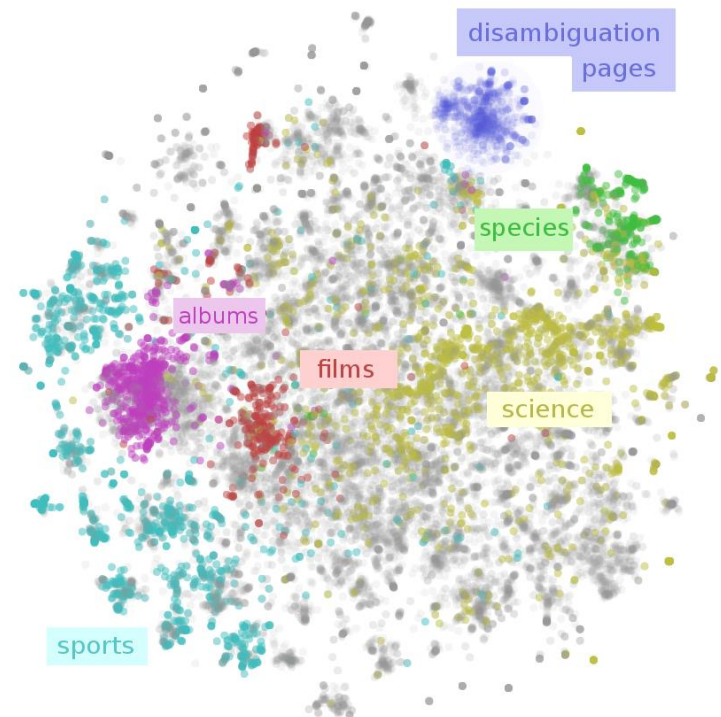

[woebot.io]

[wow.gamepedia.com]

[blog.newtonhq.com]

Natural language processing methods

- Capture statistical and structural patterns in provided corpora

- Hard to evaluate

- Rise the question: can AI reach a real language understanding using these methods?

[t-SNE map of Wikipedia, taken from colah.github.io]

Chinese Room Argument by Searle (1980)

- Can AI understand language if it "seats in the box"?

- "Internalist" vs "Externalist" perspective to grounding problem

- Embodied language theory
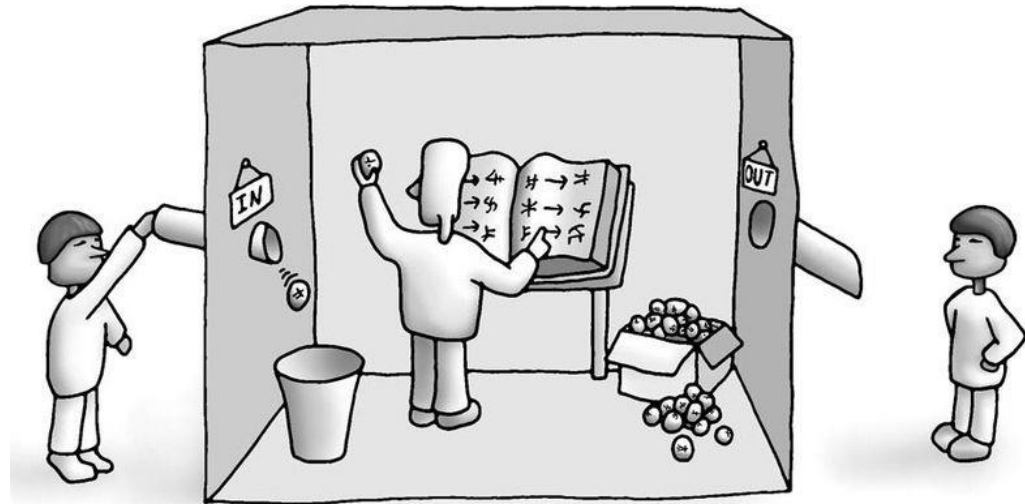
[commons.wikimedia.org]

Utilitarian language definition, Gauthier (2016)

- Situated, goal-driven paradigm

- Language as a tool for non-linguistical goals achieving

- Implicit evaluation

- If human is involved => agent learns human language

Basics of reinforcement learning

- Aims to maximize return

$$R_t = \sum_{n=t}^{T} (\gamma^{n-t} r_n)$$

- Finds optimal policy that maps states to actions

- Introduces notion of Q function

$$Q^*(s, a) = \max_{\pi} E[R_t | s_t = s, a_t = a, \pi]$$

- Can be solved by iterative updates

$$Q_{i+1}(s, a) = E[r + \gamma \max_{a'} Q_i(s', a') | s, a]$$

Deep Q networks (DQN)

- Introduced for playing Atari games,
  Mnih et al (2013)

- Use neural networks as Q function
  approximators

  $$L_i(\theta_i) = E_{s,a \sim \rho(.)}[(y_i - Q(s,a; \theta_i))^2]$$

- Use the same network for
  calculation of target values

  $$y_i = E[r + \gamma \max_{a'} Q(s', a'; \theta_{i-1}) | s, a]$$

DQN for emergent languages

- Each agent chooses actions and produces messages independently

- Agents share reward signal

- Separate Q-networks (for actions and messages) reduce dimensionality

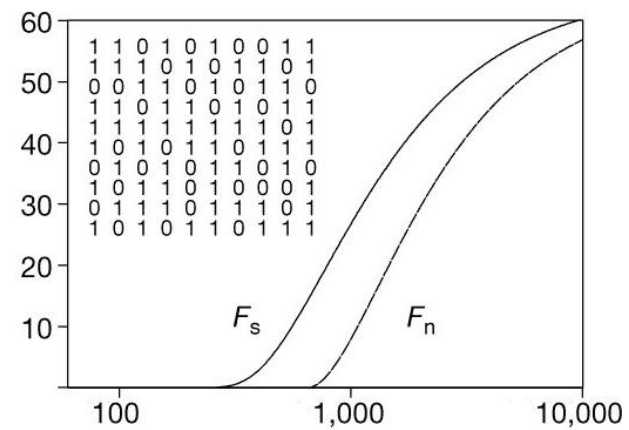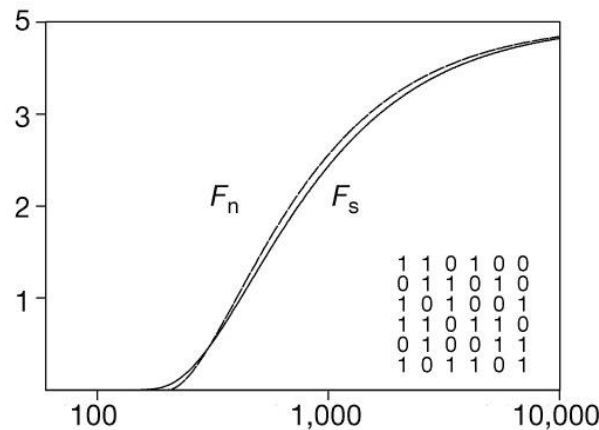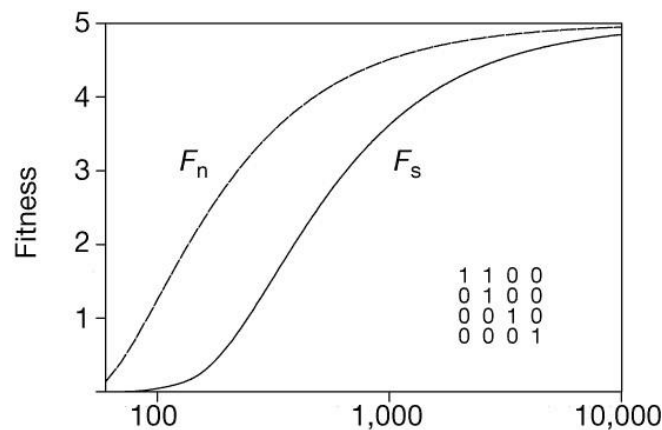- Independent Q-learning introduce instability
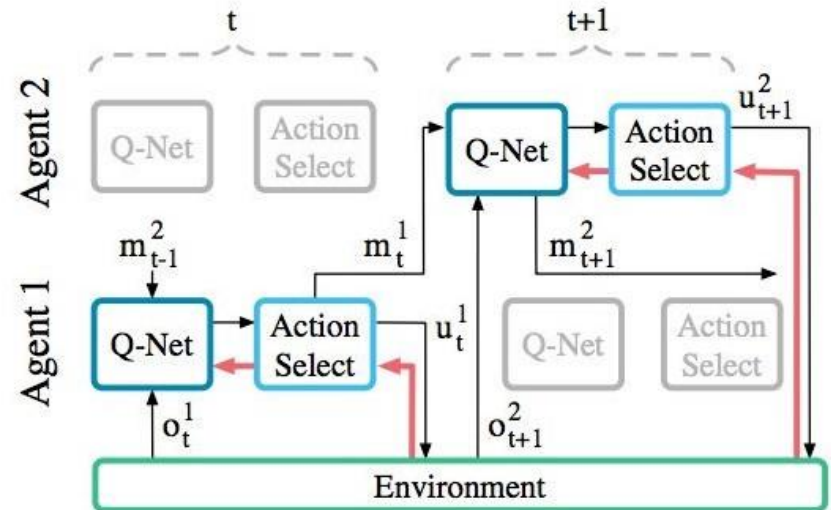
(a) RIAL - RL based communication

[Foerster (2016)]

When does syntax appear, Nowak et al.(2000)?

- If number of entities we need to reference > learning capacity,

- $n > \frac{3q}{pq_s}$, where n = amount of verbs and nouns, p = coefficient of world compositionality, $\frac{q}{q_s}$ = shows how hard to memorize syntactic expression

Can syntax emerge?

- Studied by Mordatch et al. (2017)

- Large vocabulary size => one-to-one mapping between concepts and words

- Small vocabulary size => local minima with conflated concepts

- Penalize large vocabulary sizes in "rich get richer" fashion.

$$p(c_k) = \frac{n_k}{\alpha + n - 1}$$

$$r_c = \sum_{i,t,k} 1\left[c_i^t = c_k\right] \log p(c_k)$$

Example: Environment

- Homogeneous agents with the same policy, observation and action spaces

- Shared reward signal

- Non-linguistic grounded goals, ex:"agent 2 go to blue landmark"

- Continuous space, discrete time
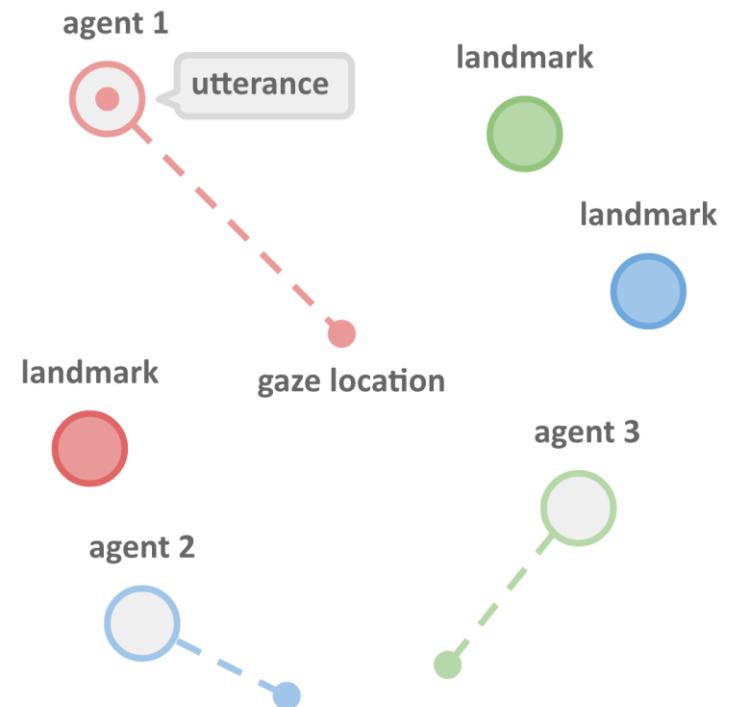
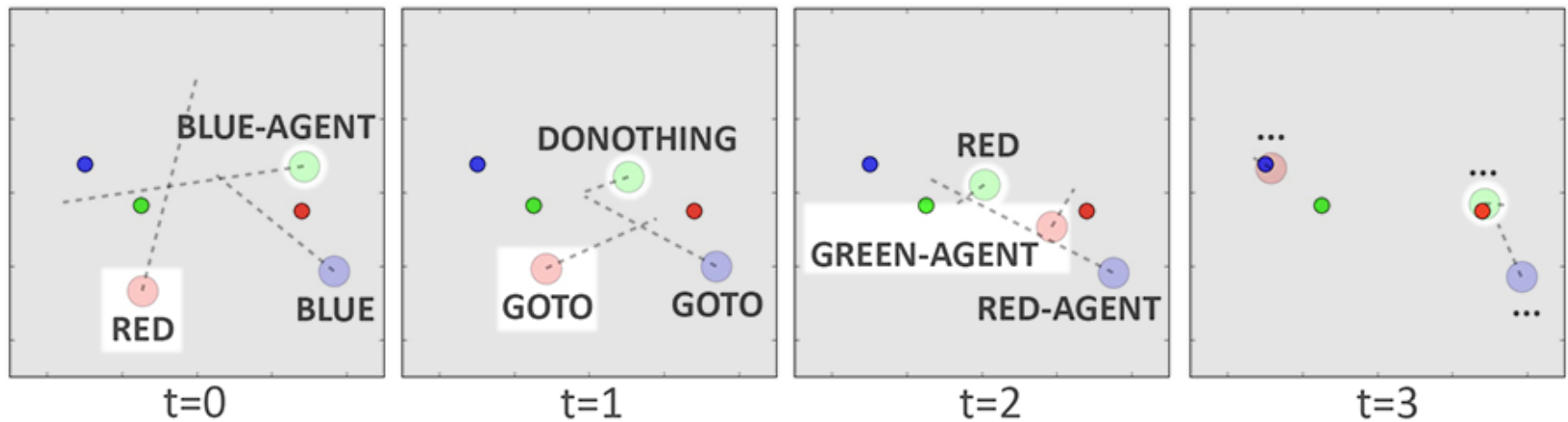[Environment configuration, Mordatch et al. (2017)]

Example: Approach

- Agents utter discrete symbols *c* at every time step

- Symbols from vocabulary do not have predefined meanings

- Agents have private memory bank *m*
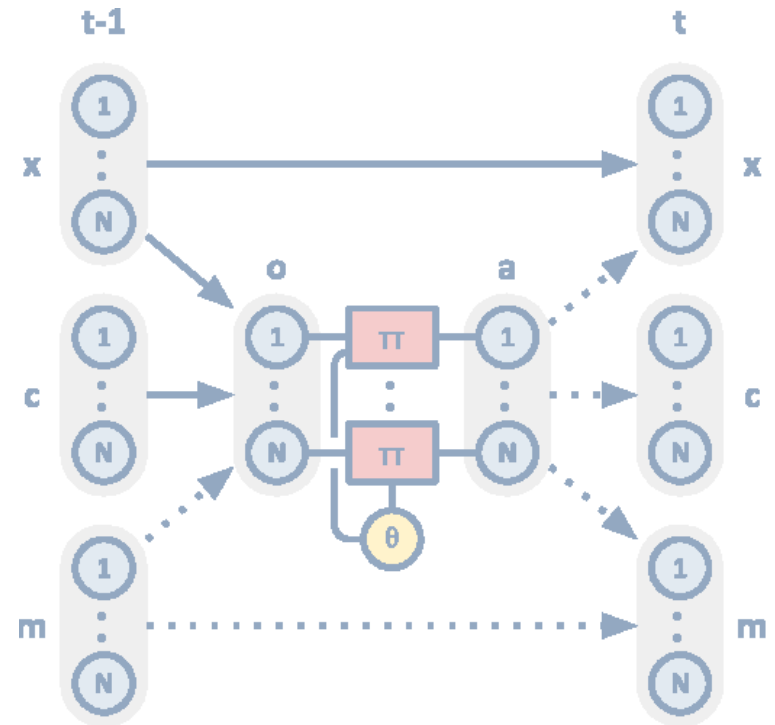
[Transition dynamics from time *t-1* to *t*.
*S*olid lines show al-to-all dependencies.
Mordatch et al. (2017)]

Video demonstration, Mordatch (2017)

Pros

- Biologically plausible grounded language

- Naturally emergent syntax

- Efficient communication

Cons

- Hard to control or shape language properties

- Hard to interpret and analyze

- Increase input dimensionality and training time

Conclusion

- Statistic-based NLP methods rise question about language understanding.

- Language derives its meaning from use, Wittgenstein (1953).

- Reinforcement learning allows to evolve grounded goal-driven language.

- Syntax appears similarly in human and artificial languages.

Thank you for your attention! Any questions?

References

[1]. Picture on the first slide is taken from: ssr.seas.harvard.edu

[2]. Buccino, G., Colagè, I., Gobbi, N., & Bonaccorso, G. (2016). Grounding meaning in experience: A broad perspective on embodied language. *Neuroscience & Biobehavioral Reviews*, *69*, 69-78.

[3]. Searle, J. R. (1980). Minds, brains, and programs. Behavioral and brain sciences, 3(3).

[4]. Mordatch, I., & Abbeel, P. (2017). Emergence of grounded compositional language in multi-agent populations. arXiv preprint arXiv:1703.04908

[5]. Gauthier, J., & Mordatch, I. (2016). A paradigm for situated and goal-driven language learning. *arXiv preprint arXiv:1610.03585*

[6]. Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., & Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*

References (cont)

[7]. Foerster, J., Assael, I. A., de Freitas, N., & Whiteson, S. (2016). Learning to communicate with deep multi-agent reinforcement learning. In *Advances in Neural Information Processing Systems.*

[8]. Nowak, M. A., Plotkin, J. B., & Jansen, V. A. (2000). The evolution of syntactic communication. *Nature*, *404*(6777), 495.