Complexity reduction in the negotiation of new linguistic conventions

William Schueller[1,2], Vittorio Loreto[3,4,5], Pierre-Yves Oudeyer[1]

[1]INRIA Bordeaux Sud-Ouest, [2]University of Bordeaux, [3]Sapienza University of Rome, [4]Sony CSL, Paris, [5]Complexity Science Hub, Vienna

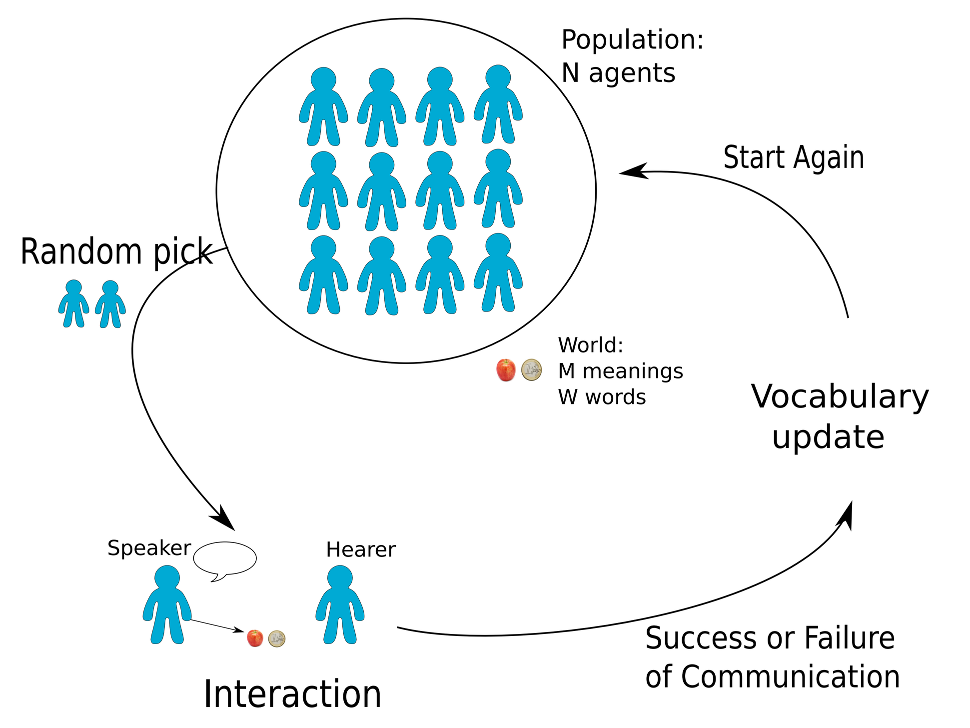
Introduction

How do we agree with each other on the new words to use for new concepts?

In the process of **collectively inventing new words for new concepts** in a population, conflicts can quickly become numerous, in the form of synonymy and homonymy. Remembering all of them could cost too much memory, and remembering too few may slow down the overall process.

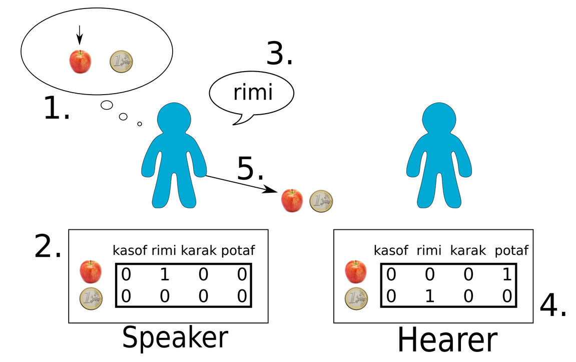
Is there an **efficient behavior** that could help balance the two?

A computational model: The Naming Game [1]



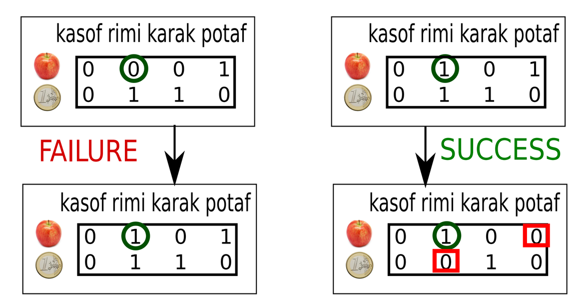
Evolution towards a **global pattern** (shared lexicon) through repeated **local interactions**

Interaction process



1. Speaker chooses a topic
2. Speaker checks its vocabulary to find or invent an associated word
3. Speaker utters the word
4. Hearer guesses the intended meaning
5. Speaker indicates the intended meaning.

Vocabulary update



Success (Hearer interpreted correctly the word as the topic): both hearer and speaker remove synonyms/homonyms conflicting with the word-meaning association used in the interaction.

Failure (Hearer interpreted the word not as the topic): the word-meaning association used by the speaker is added to the hearer's vocabulary. The speaker adds it as well if it was just invented.

Active Topic Choice

Opposed to Random Topic Choice, where all meanings are always equiprobable as topic [2]

Balancing two behaviors:

- **Exploration:** Invention of new conventions by selecting meanings not associated to any word. Necessary to build a complete lexicon, but might introduce synonyms and homonyms in the system.
- **Exploitation:** Spread known conventions by selecting meanings already having a word. Limits efficiently homonymy and synonymy.

Those options are balanced using a **measure of confidence** on their lexicon being shared by the rest of the population. Our measure of confidence here is the **LAPS measure**.

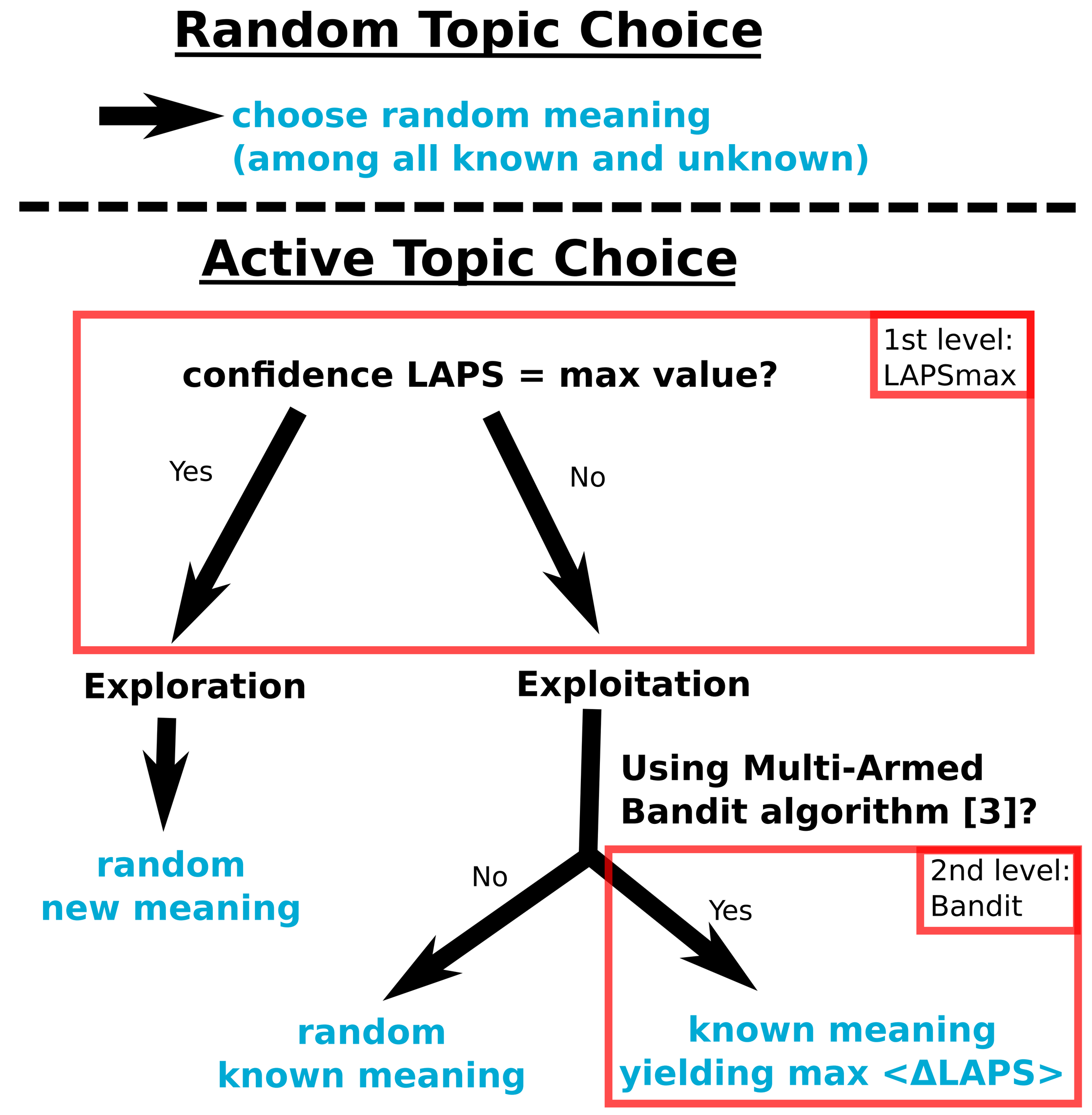
LAPS: a principled measure

Local **A**pproximated **P**robability of **S**uccess

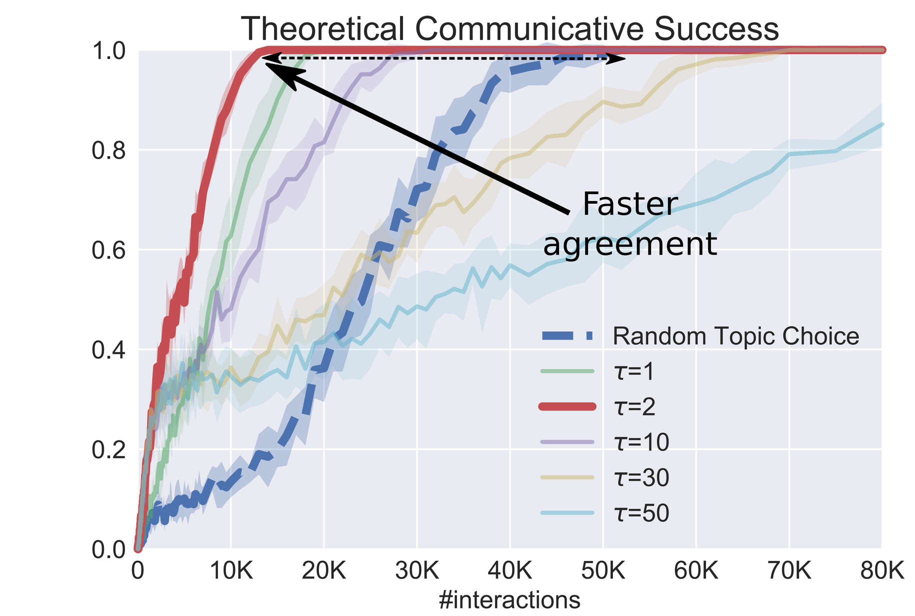
- An agent uses its past interactions as **information samples** on the population state.
- From this information, reconstruct an approximated **average lexicon** of the population, using a **sliding window** $\mathcal{T}$ to take into account only the recent past.(*)
- Estimate the **probability of having a successful communication** considering this average lexicon.

*: Each meaning has its own sliding window

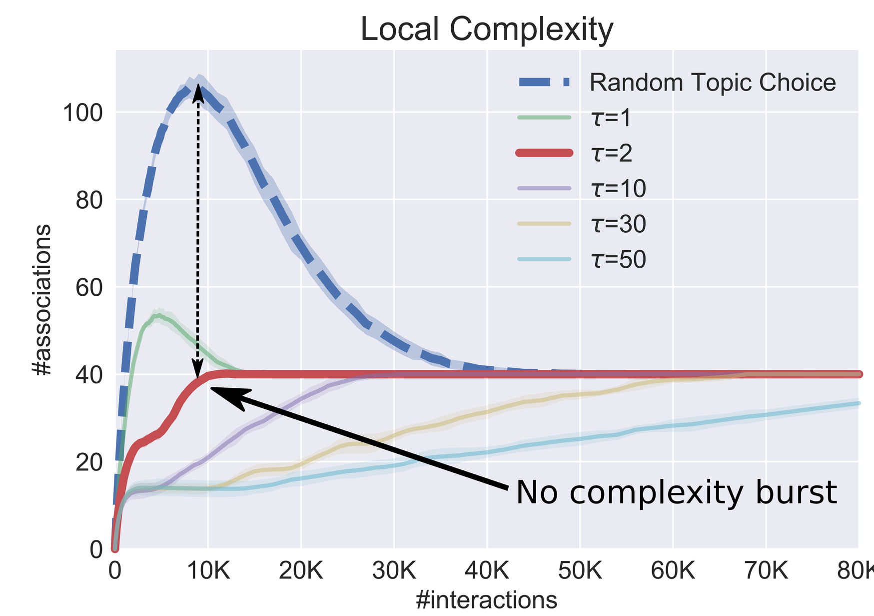
Pseudo-algorithm



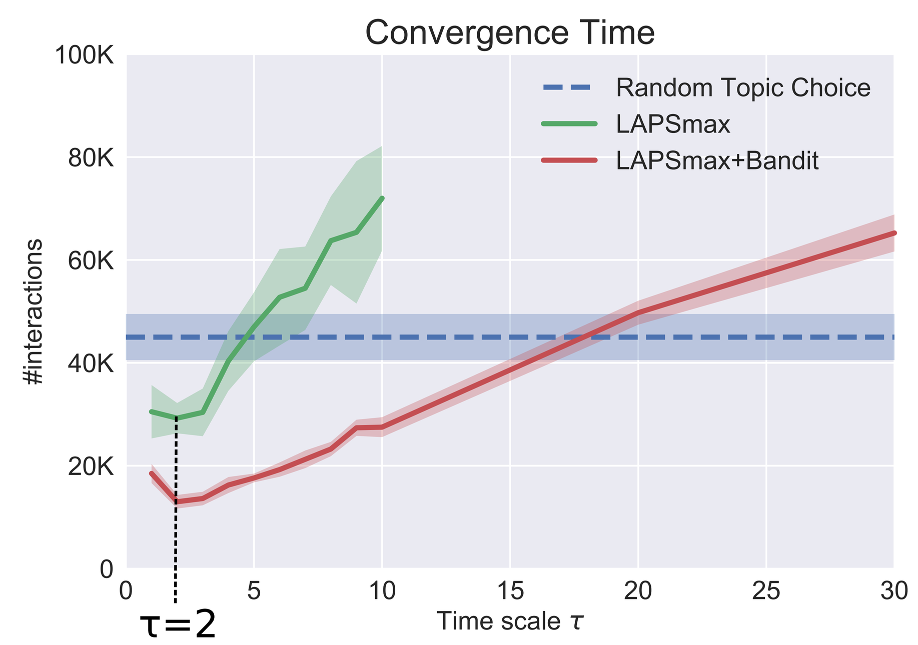
Results: Active vs. Random Topic Choice



Dynamics of the **probability of having a successful communication** between the agents, in the LAPSmax+Bandit condition, for different values of τ, compared to Random Topic Choice. 40 meanings, 40 words, 40 agents. Averaged over 8 trials.



Dynamics of **the average number of associations** in the lexicon of the agents, in the LAPSmax+Bandit condition, for different values of τ, compared to Random Topic Choice. 40 meanings, 40 words, 40 agents. Averaged over 8 trials.



Number of interactions needed **to reach full agreement** on the lexicon, as a function of τ, compared to Random Topic Choice. 40 meanings, 40 words, 40 agents. Averaged over 8 trials.

- Faster convergence to a shared lexicon
- No complexity burst for $\tau \geq 2$
- Already efficient without the bandit: both parts of the algorithm are useful
- $\tau = 2$ optimal: only limited memory needed, few recent past interactions are enough to reconstruct a meaningful average vocabulary of the population.

References and resources

[1] Vittorio Loreto, Andrea Baronchelli, Animesh Mukherjee, Andrea Puglisi, and Francesca Tria. Statistical physics of language dynamics, 2011.

[2] William Schueller and Pierre-Yves Oudeyer. Active control of complexity growth in naming games: Hearer's choice. In *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANGX11)*, 2016.

[3] Sébastien Bubeck, Nicolo Cesa-Bianchi, et al. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. *Foundations and Trends® in Machine Learning*, 5(1):1–122, 2012.

https://github.com/wschuell/notebooks_cogsci2018: Re-run the experiment