# Growing Opportunities to Grow: Toward Open-Ended Multi-Agent Communication Learning

Marina Dubova

Indiana University, Bloomington, IN, 47408
mdubova@iu.edu

## Abstract

Humans have developed a great variety of complex communicative systems (languages) without any centralized assistance. Therefore, evolution of human communication has often been modeled as a result of distributed learning among agents which are reinforced for successfully transmitting information to each other. These models, however, face two major challenges: 1) even in most successful cases, the agents can only develop a very small number of communicative conventions, whereas humans managed to successfully agree upon thousands of words; 2) after groups of artificial agents converge on a set of communicative conventions, they have no incentive to improve or expand it, whereas the development of human languages is open-ended. Here, I show how these two challenges could be resolved by dynamically changing the problem that the agents are learning to solve with communication. I suggest that the communicative problem that starts small and gradually increases in difficulty as the agents agree upon new communicative conventions is essential for achieving tractable evolution of rich communicative systems in decentralized multi-agent communities.

## Introduction

Human communities have arrived at thousands of communicative conventions, and continue to "invent" new words up until today. The impressive complexity of human languages presents a fundamental challenge: how could humans develop such complex communicative systems from scratch? The existing models of language evolution provide insights into how groups of individuals can converge on a small set of shared communicative conventions, and how these conventions, once developed, change over time (e.g. Skyrms 2010; Reali and Griffiths 2010). However, the mechanisms which could allow the groups of learning agents to develop communicative systems with fairly large amounts of conventions have been understudied.

In preliminary modeling experiments, I show that communicative problem that adaptively increases in difficulty, rather than constantly demanding the desired end-state of communication learning, allows the groups of agents to gradually agree upon large numbers of communicative conventions, while always keeping space to learn more.

## Method

### Communication Game

The model consists of a group of agents performing two-agent communicative interactions, which follow the communication game setup introduced by Lewis (1969). At each timestep, two agents from the group (size=$G$) are randomly chosen to participate in a communicative interaction. On each interaction, one agent is randomly assigned a speaker role, and the other is assigned a listener role. Then, the game proceeds as follows: 1) the world randomly chooses one of $N$ discrete states (probability distribution of the states is uniform); 2) the speaker perceives the state, and chooses one of $V$ discrete messages to transmit; 3) the listener receives the signal and chooses one of $A$ discrete actions to take; 4) if the listener's chosen action corresponds to the initial state of the world, both agents receive a positive reward ($+1$); otherwise, the reward is negative ($-1$).

### Agents

Each agent independently learns from its own experience according to the Roth-Erev reinforcement learning algorithm (Roth and Erev, 1995). Thus, each agent stores the cumulative reward history for all states and actions in two tables: world states$\times$messages (as talker) and messages$\times$actions (as listener). On each interaction, the agents choose actions proportional to their accumulated reward for a received state with softmax smoothing.

### Dynamic and Static Communication Problems

I compare the static and dynamically changing communication game conditions. The static game condition represents the desirable end-point of learning: e.g., if the goal is to develop a communicative system with 100 conventions, the interactions between agents will involve a world with 100 discrete states. For simulations with the static communicative problems, the number of messages available to speakers was always set to 100, and the number of listeners' actions always corresponded to the amount of world states.

The dynamically changing communicative problem increases in difficulty whenever the agents' average reward in

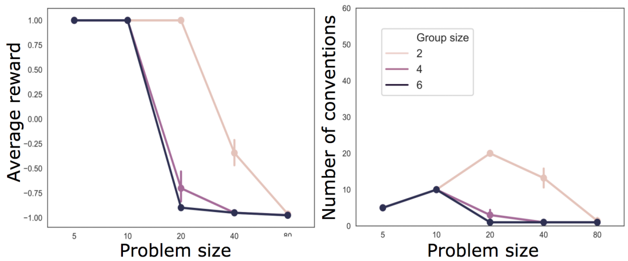


Figure 1: **Static communication game**: results after 300000 communicative interactions (each condition was simulated 5 times). Left: average reward for the last 10000 interactions. Right: average number of conventions that the group agreed upon for the last 10000 interactions. X-axis shows five world state size conditions for the simulations; the hue represents size of the group.

the last $100$ communicative interactions is larger than $0.95$ (no more often than after every $100$ interactions). I tested two types of dynamic communicative problem:

1. **Dynamic world**. The number of perceived world states starts with $2$ states and grows one state at a time when the "learning success" condition is met.
2. **Dynamic actions in dynamic world**. Both the number of perceived world states and the amount of listeners' actions start at $2$ and synchronously grow when the "learning success" condition is met.

The number of speakers' available messages was set to $100$ in all simulations, and the number of listeners' actions in the "dynamic world" condition was set to $40$.

## Results

Simulations demonstrated the limitations of multi-agent communication learning in static problem conditions. The groups of 2, 4, and 6 agents did not agree on a single communicative convention when the communicative problem involved 80, 40, and 20 world states respectively (Fig.1).

Preliminary results show that the groups solving the dynamically changing communicative problems overcame some limitations found in the static learning conditions (Fig.2). After the same amount of communicative interactions, the agents learning with the changing number of actions and world states converged on more conventions than their counterparts in the static condition, and continued to learn up until the very end of the fixed simulation time. The groups of $4$ and $6$ agents in the dynamic problem condition continued to show a steady increase in the number of communicative conventions even for the communicative problem difficulty (e.g. 20 and 30 states) when no learning at all was achieved in the static scenario. The group size still had a drastic effect on the speed of multi-agent learning.

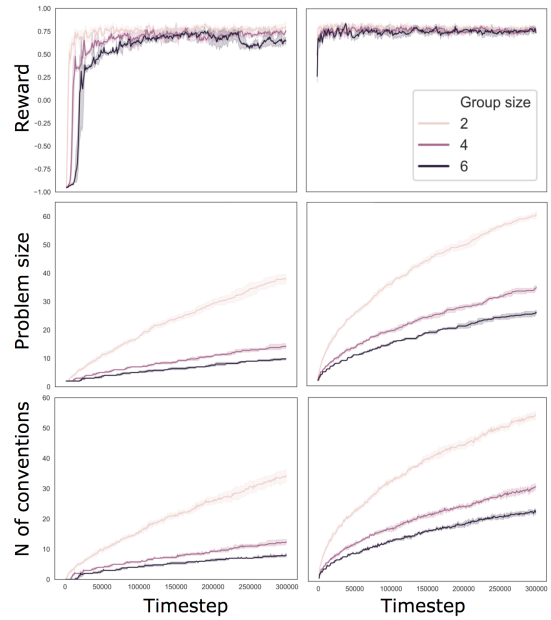


Figure 2: **Dynamic communication game**: results after 300000 communicative interactions. Left column: dynamic world; right column: dynamic actions in dynamic world. Top: rolling average reward for 1000 interactions over time; middle: rolling average number of world states over time; bottom: rolling average number of conventions that the group agreed upon. Hue represents the group size.

## Discussion

Results suggest that starting small and gradually increasing the complexity of the communicative problem for the group of agents allows them to continuously master large amounts communicative conventions. Agents learning in the dynamic problem condition continued to agree on new communicative conventions up until the very end of the experiment, whereas the agents learning to solve the static problem got stuck on a particular number of conventions and even completely failed to learn when the problem was too difficult. Future experiments need to clarify the limitations of learning in different dynamic game conditions.

The communicative problem that changes when it is being solved could result from a variety of adaptive processes which typically accompany human learning. For example, human children and adults often focus on learning a limited set of items and gradually move attention to new challenges after solving the current ones (Kidd and Hayden, 2015). Moreover, human perceptual processing rapidly changes when new labelled categories are being learned (Lupyan et al., 2020). Potential "inter-individual" causes of the communicative problem adaptivity include cultural evolution of a group (e.g. the number of items that the individuals wants to exchange, degree of labor specialization) which may be facilitated by the group's communicative success.

## References


Kidd, C. and Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, 88(3):449–460.

Lewis, D. K. (1969). Convention: A philosophical study.

Lupyan, G., Rahman, R. A., Boroditsky, L., and Clark, A. (2020). Effects of language on visual perception. *Trends in cognitive sciences*.

Reali, F. and Griffiths, T. L. (2010). Words as alleles: connecting language evolution with bayesian learners to models of genetic drift. *Proceedings of the Royal Society B: Biological Sciences*, 277(1680):429–436.

Roth, A. E. and Erev, I. (1995). Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term. *Games and economic behavior*, 8(1):164–212.

Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford University Press.