Quantifier Learning
An Agent-based Coordination Model

Dariusz Kalociński[†], Marcin Mostowski[†], Nina Gierasimczuk[‡]

[†] Institute of Philosophy, University of Warsaw

[‡] Institute for Logic, Language and Computation, University of Amsterdam

{d.kalocinski,m.mostowski}@uw.edu.pl, nina.gierasimczuk@gmail.com

UNIVERSITY OF WARSAW

UNIVERSITY OF AMSTERDAM

Abstract

We consider the problem of learning the meaning of natural language expressions. In contrast to traditional settings, in which agents infer prescribed meanings from observations, we focus on an algorithm for the coordination of meaning among many agents. We do not assume any external correctness criterion. Similar research on colour categorization can be found in [4]. We propose an agent-based iterative algorithm for coordinating the semantics of upward monotone proportional quantifiers. We consider coordination for two agents. We define three models: a simple one, one with authority parameter, and one combining authorities and more complex conversation patterns. We represent the models in terms of Markov chains and study the influence of authorities and conversation patterns on coordination. We observe a mathematical connection between the possibility of convergence and specific levels of agents authority and complexity of communication patterns. We mention the possibility of extending the model to cover the parameter of spatial separation.

Introduction

We restrict to upward monotone proportional quantifiers of type (1). Natural language quantifier interpretable in this way is *Most*. We identify the meaning of a quantified sentence with algorithm for checking the sentence truth value in finite models. Reduced fractions from $[0,1]$ (Farey fractions) serve as rough approximations of meanings of such quantifiers. For a quantifier expression Q with meaning p/q, a finite model (U,R) with $U \neq \emptyset$ and $R \subseteq U$: the sentence $QxR(x)$ is true in (U,R) iff $|R|/|U| > p/q$. The only information necessary to calculate the truth value is the proportion $|R|/|U|$. Thus, we identify relevant finite models with rational numbers from $[0,1]$.

We model the coordination of meaning of a quantifier expression Q via communication. We think of agents as equipped with simple linguistic constructions, relevant for uttering sentences of the form $QxR(x)$.

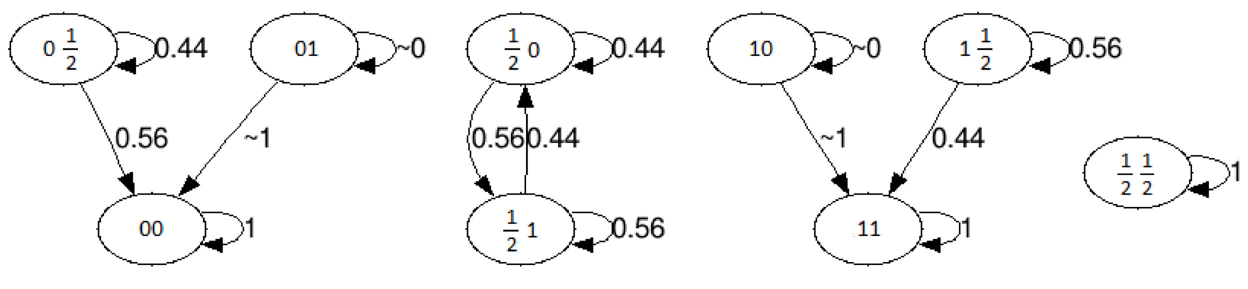

Figure 1: Markov chain for Model IIA, authorities $w_1 > w_2$, $H = \{0, \frac{1}{2}, 1\}$ and $X \sim B(50, 0.5)$.

Model Parameters

n — number of agents, $A = \{1, 2, \ldots, n\}$ is referred to as the population.

k — $H = F_k$ is the space of hypotheses. F_k denotes the set of irreducible fractions between 0 and 1 (inclusively) whose denominators do not exceed k.

X — a random variable with an associate probability function P. X assumes values in $[0,1]$ and approximates the contexts (environments) in which agents communicate. Random deviates of X are to be interpreted as proportions $|R|/|U|$.

Coordination Process

Agents are equipped with semantics $s : A \to H$. Each agent knows his semantics, not semantics of others. Coordination evolves in discrete steps. Each step comprises of two stages: **communication** and **adjustment**. During **communication** each agent gets to know how some of the other agents evaluate quantified sentences in some situations. During **adjustment** each agent tries to change his current semantics so as to maximize communicative success in a situation encountered in the communication stage. After that s gets updated and we get back to communication and adjustment stage once again, and so on.

Communication Conversation patterns for the present stage are generated. A conversation pattern consists of two agents a, b and a topic $r \in [0,1]$. Agents communicate according to the generated patterns. In a conversation (a,b,r), a communicates to b the truth value of '$r > s(a)$' and b does the same towards a with '$r > s(b)$'. We say a conversation (a,b,r) is successful, if the communicated truth values are equal.

Adjustment An agent-based coordination mechanism is performed simultaneously by all agents. The goal of each agent $a \in A$ is to become more successful in situations as the one encountered in the communication stage. Success of an agent in a given situation is measured by the the number of successful conversations he participated in. The idea of coordination is as follows: choose at random the simplest semantic hypothesis from those which guarantee maximal success in a given situation. Simplicity is understood as possessing small denominator. The simplicity criterion is introduced to mimic our natural preference for simple solutions.

Algorithm 1 Agent-based coordination mechanism

Agent: current hypothesis $h_0 \in \mathcal{H}$
Input: $b_1 b_2 \ldots b_m$ – complete list of interlocutors from communication stage
$r_1 r_2 \ldots r_m$ – corresponding topics of conversations
Output: (possibly new) hypothesis from H

for all $h \in \mathcal{H}$ **do** ▷ each h is assigned a score
  $z_1 z_2 \ldots z_m :=$ a binary vector, where $z_i = 1$, if $r_i > s(b_i) \Leftrightarrow r_i > h$, and $z_i = 0$ otherwise
  $score(h) := \begin{cases} \Sigma_{i=1}^{m}(z_i) + 1 & \text{if } h_0 = h \\ \Sigma_{i=1}^{m}(z_i) & \text{otherwise} \end{cases}$
end for
$M := \{h \in H : \forall h' \; score(h) \geq score(h')\}$
return random element from $S(M)$ ▷ $M \supseteq S(M) =$ fractions with the smallest denominators

Model I

Fix the parameters, $n = 2$. At each stage, agents perform one conversation on a single topic. We represent Model I by the Markov chain on $S = F_k \times F_k$. A state $s = s_1 s_2 \in S$ refers to the situation in which agents 1 and 2 understand Q as s_1 and s_2, respectively. For any $s, s' \in S$, the transition probability $p_{ss'}$ describes chances that a population changes its semantics from s to s' during one stage of coordination.

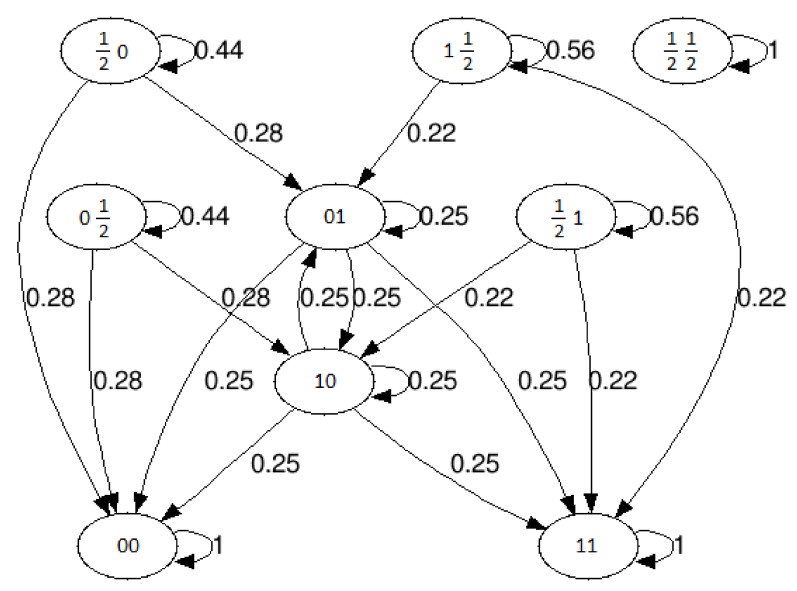

Figure 2: Markov chain for Model I, $H = \{0, \frac{1}{2}, 1\}$ and $X \sim B(50, 0.5)$.

By the Markov representation, we observe that only the existential quantifier or the trivial (always false) quantifier *more than everything* may emerge, unless initial semantics is a constant function $s : A \to \{u\}$, for some $u \in H$, $0 < u < 1$. This observation is not favourable for Model I, as it does not explain how more complex semantics could emerge.

Model IIA: Authorities

Model IIA has one new parameter—authority function $w : A \to \mathbb{R}_+$. We only modify the scoring procedure. Let us take $a \in A$, and let w_0 be a's authority and $w_1 w_2 \ldots w_m$ authorities of a's interlocutors. We set $score(h) = \Sigma_{i=1}^{m}(z_i \cdot w_i)$. Additionally, if h is the same as a's current hypothesis, we add w_0 to the final score of h.

For a constant non-zero authority function, Model IIA resolves into Model I. However, differentiated authorities facilitate coordination. Consider authorities $w_1 > w_2$ (see Fig. 1). It is more probable in Model IIA than in Model I to change from 01 to 00. Moreover, we cannot have the following cycles: $01, 10, 01, 10, \ldots$, while they may occur in Model I. In Model IIA, if an agent with the greatest authority starts with 0, then we cannot stabilize on anything else than 00 (similarly for 1). In Model I, it does not matter whether agent starts with 0 or 1—she can always change to 0 or 1. If an agent with the greatest authority starts with something other than 0 and 1, then the semantics diverge forever. This effect is partially due to the simplicity criterion that tells agents to choose among the simplest hypotheses, but another reason for this is a very low complexity of communication patterns.

Model IIB: Multiple Topics

In Model IIB communication patterns are more complex. At each stage, two topics are generated and each agent communicates with every other agent about the two topics. We observe that an agent may choose more complex semantics only if her interlocutor possesses complex semantics and has greater authority. *Most* is achievable in such a model.

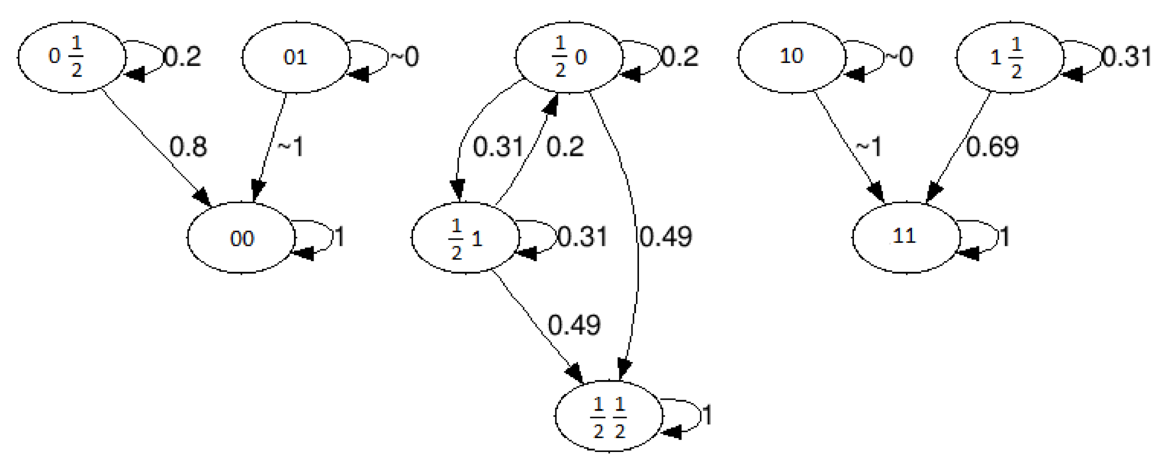

Figure 3: Markov chain for Model IIB, authorities $w_1 > w_2$, $H = \{0, \frac{1}{2}, 1\}$ and $X \sim B(50, 0.5)$.

We hypothesize that if the topics of real-life conversations obey a rule similar to normal distribution, then our coordination model explains why the quantifier *Most* emerged in natural language.

Conclusions

It turns out that authority functions significantly affect the behaviour of the population. Differentiated authority functions are propitious for coordination and the quality of communication, whereas equality among agents makes the coordination more difficult and the communication less successful. This observation extends to larger populations. Moreover, higher complexity of communication patterns may lead to the emergence of more complex semantics.

Forthcoming Research

The authority has less impact on others when communication is less frequent. Empirical experiments reveal a strong negative correlation between the physical distance and the frequency of communication (viz. Allen's Curve, [1]). The next step of this research is to take into account larger populations and to account for the distance factor.

References

[1] T. J. Allen. *Managing the flow of technology: technology transfer and the dissemination of technological information within the R&D organization.* MIT Press, Cambridge, MA, 1977.

[2] N. Gierasimczuk. The problem of learning the semantics of quantifiers. In *TBiLLC'05: 6th International Tbilisi Symposium on Logic, Language, and Computation. Revised Selected Papers*, Volume 4363 of *Lecture Notes in Artificial Intelligence*, pages 117–126. Springer, 2007.

[3] M. Mostowski and D. Wojtyniak. Computational complexity of the semantics of some natural language constructions. *Annals of Pure and Applied Logic*, 127(1-3):219–227, 2004.

[4] L. Steels and T. Belpaeme. Coordinating perceptually grounded categories through language: A case study for colour. *Behav Brain Sci*, 28(4):469–529, 2005.