Quantifier Learning: An Agent-based Coordination Model

(Extended Abstract)

Dariusz Kalociński
Institute of Philosophy
University of Warsaw
d.kalocinski@uw.edu.pl

Nina Gierasimczuk
Institute for Logic, Language and Computation
University of Amsterdam
nina.gierasimczuk@gmail.com

Marcin Mostowski
Institute of Philosophy
University of Warsaw
Institute of Philosophy
Jagiellonian University
m.mostowski@uw.edu.pl

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. We propose an agent-based iterative algorithm for coordinating the semantics of upward monotone proportional quantifiers. We describe simple instances of our model in terms of Markov chains. We observe a mathematical connection between the possibility of convergence and specific levels of agents authority and complexity of communication patterns. We discuss the possibility of extending the model to cover the parameter of spatial separation.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*coherence and coordination*; I.2.0 [**Artificial Intelligence**]: General—*philosophical foundations*

Keywords

Multi-agent learning; coordination; semantics; quantifiers; natural language

1. INTRODUCTION

Semantics acquisition is traditionally modelled as convergence to an external, prescribed meaning on the basis of observations. The existence of such 'true' semantics is however conditional on the prior emergence of meanings through communication between many agents. Even the 'true' meanings change and are rarely completely shared within a population. Nevertheless, humans *somehow* adjust the repertoire of meanings and become successful in communicating. Completely shared unique semantics is not necessary for this purpose. We study a possible agent-based mechanism for distributed coordination of meanings.

Learnability of quantifiers has been studied in the single-agent context (see e.g. [2]). The idea of semantic change is mentioned, in a relevant context, in [3]. Semantic coordination methods have been used before in the context of colour categorization [4].

2. COORDINATION MECHANISM

We identify meanings of expressions with algorithms for checking the expression's truth value in finite models. During coordination each agent may reassign an algorithm to a given expression. We focus on mathematically usable quantitative expressions, generalized quantifiers. For simplicity, we restrict to upward monotone proportional quantifiers of type (1). The meanings of quantifiers are identified with completely reduced fractions from $[0, 1]$, called Farey fractions, which serve as rough approximations of algorithms. Given a quantifier expression Q, identified with p/q, and a finite model (U, R) with the universe $U \neq \emptyset$ and the property $R \subseteq U$, the sentence $QxR(x)$ is true in (U, R) iff $|R|/|U| > p/q$. The computation procedure is clear from the definition. The most prominent example of a natural language quantifier interpretable in this way is *Most*. Note that the only information necessary to calculate the truth value is the proportion $|R|/|U|$, where $|\cdot|$ stands for cardinality. Thus, we identify relevant finite models with rational numbers from $[0, 1]$.

Below we model the coordination of meaning of a quantifier expression Q via communication. We think of agents as equipped with a quantifier Q and simple linguistic construction, relevant for uttering sentences of the form $QxR(x)$.[1] Let F_k denote the set of irreducible fractions between 0 and 1 (inclusively) whose denominators do not exceed k. Parameters of the model are n, k, X, where n, k are positive integers and X is a random variable with an associate probability function P. $H = F_k$ is the space of hypotheses which agents may choose from. $A = \{1, 2, \ldots, n\}$ is referred to as the population. X assumes values in $[0, 1]$ and approximates the contexts (environments) in which agents communicate.

The coordination proceeds in stages. At any given stage, each agent $a \in A$ associates with Q a meaning $s(a) \in H$. 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 '$r > s(b)$'. Next, an agent-based coordination mechanism is performed simultaneously by all agents. Let us take an agent $a \in A$, and let $b_1 b_2 \ldots b_m$

[1]We assume that this type of statements are the only ones allowed in communication among agents.

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

and $r_1 r_2 \ldots r_m$ be the complete lists of a's interlocutors and corresponding topics of conversations. Agent a's goal now is to adjust her current semantics. Each $h \in H$ is assigned a score, corresponding to the degree of agreement that a would attain, if h were her new semantics. Let us fix $h \in H$. Generate a binary vector $z_1 z_2 \ldots z_m$ where $z_i = 1$, if $r_i > s(b_i) \Leftrightarrow r_i > h$, and $z_i = 0$ otherwise. Intuitively, z_i indicates whether the truth value communicated by b_i on r_i is the same as the truth value that a would communicate given the semantics h. Eventually, h is assigned a score $score(h) = \Sigma_{i=1}^{m} z_i$. Additionally, if h is the same as a's current hypothesis, the final score of h is increased by 1. Finally, we set $M := \{h \in H : \forall h' \; score(h) \geq score(h')\}$. Let $S(M)$ denote the set of all Farey fractions from M with the smallest denominators.[2] A random element from $S(M)$ is chosen as a's current hypothesis—this is the semantics that a shall use in the next stage of the coordination.

3. COORDINATION MODELS

As the first step 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. Below we sum up the preliminary findings for each of the models. We represent the models in terms of Markov chains. We study the influence of authorities and conversation patterns on coordination.

Let us start with a simple Model I. We 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. 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 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 $w_1 > w_2$. It is four times more probable in Model IIA than in Model I that the population changes from 01 to 00. Moreover, population cannot change from 01 to 10 (or from 10 to 01). Thus, 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 the semantics cannot stabilize on anything else than 00 (similarly for 1). In Model I, it does not matter whether agents 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.

In Model IIB the communication pattern is 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.

Moreover, 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.

4. CONCLUSIONS AND OUTLOOK

Markov analysis of simple models reveals several features of coordination mechanisms. 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.

The authority may have less impact on others when communication is less frequent. It is worth noting that empirical experiments have already revealed 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.

5. ACKNOWLEDGMENTS

Dariusz Kalociński is supported by the Polish National Science Centre grant 2014/13/N/HS1/02048. Nina Gierasimczuk is funded by an Innovational Research Incentives Scheme Veni grant 275-20-043, Netherlands Organisation for Scientific Research (NWO). Marcin Mostowski is supported by Polish National Science Centre grant 2013/11/B/HS1/04168.

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.

[2] This choice is driven by the preference of simple solutions.