Lexicon acquisition in an uncertain world

Andrew D. M. Smith and Paul Vogt
LEC, Linguistics, University of Edinburgh; ILK, Computational Linguistics, Tilburg University
andrew@ling.ed.ac.uk

One of the distinctive features of human language is its use of arbitrary symbols to convey meanings from one person to another. In this paper, we focus on the problem, famously described by Quine (1960), of how learners learn the meanings of words, when they cannot receive any explicit information about the association between the two elements. Without such information, learners must rely on some external information to provide clues to the intended meaning, such as the pragmatic context in which the word is presented.

We have previously developed a model of meaning creation and inference in which agents use a Bayesian learning strategy to learn the meanings of words by disambiguating potential meanings through the presentation of words in multiple contexts (Vogt, 2000; Smith, 2001); we now formalise this computational model so that we can make accurate predictions of the likely outcomes of future experiments. In particular, we present a mathematical model for predicting the time needed to learn an associative lexicon of a given size and a given level of referential uncertainty, based on the cross-situational statistical learning used in the computational model.

We quantify the number of communicative interactions which are necessary for one agent to learn a lexicon from another, given the degree of uncertainty, and show that our mathematical model compares well to our computational simulations of lexicon acquisition under similar conditions. Furthermore, the model predicts that successful learning will take place even with surprisingly large levels of uncertainty in the model.

We go on to compare the model to other cross-situational learning models (e.g. Siskind, 1996), and show how the model can be extended to take account of more realistic Zipfian distributions of word frequency. This allows us to explore whether the model can provide helpful predictions about the conditions under which children learn language. With such distributions, learning is clearly much harder, as it takes much longer for us to be sure that the learner has been exposed to the whole lexicon, and so the level of uncertainty must be reduced relative to the uniform model. Many psycholinguistic biases, indeed, have been proposed to account for this necessary reduction of referential uncertainty, and thereby for the speed with which children acquire their lexicons (Bloom, 2000); the model presented here will also provide a formal mechanism for exploring the relative effectiveness of these hypotheses.

References

Bloom, P. (2000). *How children learn the meanings of words.* Cambridge, MA: MIT Press.

Quine, W. v. O. (1960). *Word and object.* Cambridge, MA: MIT Press.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition, 61,* 39–91.

Smith, A. D. M. (2001). Establishing communication systems without explicit meaning transmission. In J. Kelemen & P. Sosík (Eds.), *Advances in Artificial Life: Proceedings of the* 6^{th} *European Conference on Artificial Life* (pp. 381–390). Heidelberg: Springer-Verlag.

Vogt, P. (2000). Bootstrapping grounded symbols by minimal autonomous robots. *Evolution of Communication, 4*(1), 89–118.