How the poverty of stimulus solves the poverty of stimulus

Willem Zuidema
jelle@arti.vub.ac.be
Artificial Intelligence Laboratory
Vrije Universiteit Brussel
Pleinlaan 2, B-1050, Brussel, Belgium

To be presented at "Evolution of Language: Fourth International Conference", Harvard University, March 2002.

The logical problem of language acquisition is the problem every child is facing when acquiring the grammar of its native language: it has too little evidence to uniquely determine which is the target grammar from the set of all grammars that would in principle be possible. The basic intuition that the grammars of natural languages are not learnable from available evidence is supported with a series of mathematical models. Here, we will focus our discussion on two such models: Gold (1967) and Nowak et al. (2001). These models are often interpreted as showing the necessity of strict, a-priori and innate constraints on the set of grammars a child needs to consider in the first place.

The underlying assumptions of these "proofs" of Universal Grammar have been controversial. Gold, for instance, shows that context-free grammars are not learnable, if one assumes that the grammars are deterministic, that there is no negative evidence and that learnability corresponds to "identification in the limit". Many have argued that it is unreasonable to assume that a child learns language in such a worst-case scenario (see e.g. Elman et al, 1996). Although we suspect that some of this criticism is in fact justified, we leave that issue as an empirical question. Instead, we study a computer model in which we make similar assumptions as in the mathematical models, evaluate the claims about Universal Grammar within that framework, and nevertheless arrive at very different conclusions.

Our model consists of an evolving population of language learners, that learn a grammar from their parents and get offspring proportional to the success in communicating with other individuals in their generation. The grammar induction procedure is fixed; it is inspired by Kirby (2000). The details of the grammatical formalism (context-free grammars) and the population structure are deliberately close to Gold (1967) and Nowak et al. (2001) respec-

tively. Surprisingly, we observe in the simulations successful acquisition of grammars that are unlearnable by Gold's criterion. Further, we observe grammatical coherence although many more grammars are allowed than Nowak et al. calculate as an upper bound.

The reason for these surprising results is that language acquisition is a very particular type of learning problem: it is a problem where the target of the learning process is itself the outcome of a learning process. That opens up the possibility of language itself to adapt to the language acquisition procedure of children. In such iterated learning situations (Kirby 2000), learners are only presented with targets that other learners have been able to learn, and most likely even with targets that are best learnable.

Isn't this Universal Grammar in disguise? Learnability is, consistent with Gold's proof, still achieved by constraining the set of targets; but, unlike in usual interpretations, these constraints are not strict (some grammars are better learnable than others, allowing for an infinite "Grammar Universe"), and they are not a-priori: they are the outcome of iterated learning. The poverty of stimulus is no longer a problem; instead, the ancestors' poverty is the solution to the child's.

References

Elman, J., Bates, E. et al. (1996). Rethinking innateness. MIT Press.

Gold, E. M. (1967). Language identification in the limit. Information and Control (now Information and Computation) 10, 447–474.

Kirby, S. (2000). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In: The Evolutionary Emergence of Language: Social function and the origins of linguistic form (Knight, C., Hurford, J. & Studdert-Kennedy, M., eds.). Cambridge University Press.

Nowak, M. A., Komarova, N. & Niyogi, P. (2001). Evolution of universal grammar. Science 291, 114–118.