At the Interface of Computational Learning Theory and Human Language Learning

Lisa Pearl
University of Maryland
UC Irvine Cognitive Sciences Colloquium
Mar 12, 2007

In this talk, I will investigate the mechanism of language acquisition given the boundary conditions provided by linguistic representation and the time course of acquisition. Exploration of the mechanism is vital once we consider the complexity of the system to be learned and the non-transparent relationship between the observable data and the underlying system. It is not enough to restrict the potential systems the learner could acquire, which can be done by defining a finite set of parameters the learner must set. Even supposing that the system is defined by n binary parameters, we must *still* explain how the learner converges on the correct system(s) out of the possible 2^n systems, using data that is often highly ambiguous and exception-filled (Clark, 1994).

The two case studies presented will use computational modeling, and will be embedded in a framework that conceptualizes language learning as three separate components, assuming that learning is the process of selecting the best-fit option given the available data. These components are (1) a defined hypothesis space, (2) a definition of the data used for learning (data intake), and (3) an algorithm that updates the learner's belief in the available hypotheses, based on data intake. Defining the learning components in this somewhat abstract manner allows us to apply the framework to a range of language learning problems and linguistic domains. In addition, we can combine discrete linguistic representations with probabilistic methods and so account for the gradualness and variation in learning that human children display.

One of the discoveries of this line of work is that filtering the data intake set can lead to acquisition success even when the learner is faced with a complex, noisy system. Moreover, acquisition success in some cases seems to require that the data intake be a filtered subset of the available input.

Computational modeling proves itself a very useful tool in addressing data intake filtering questions of this kind, since they would be very difficult to explore with standard experimental techniques. In addition, the results of computational modeling can generate predictions that can then be tested experimentally.