Challenging the limits of learning: Human mind vs. yardstick of a machine

January 19 2011

Although we're convinced that baby is brilliant when she mutters her first words, cognitive scientists have been conducting a decades-long debate about whether or not human beings actually "learn" language.

Most theoretical linguists, including the noted researcher Noam Chomsky, argue that people have little more than a "language organ" -- an inherent capacity for language that's activated during early childhood. On the other hand, researchers like Dr. Roni Katzir of Tel Aviv University's Department of Linguistics insist that what humans can actually learn is still an open question -- and he has built a computer program to try and find an answer.

"I have built a computer program that learns basic grammar using only the bare minimum of cognitive machinery -- the bare minimum that children might have -- to test the hypothesis that language can indeed be learned," says Dr. Katzir, a graduate of the Massachusetts Institute of Technology (where he took classes taught by Chomsky) and a former faculty member at Cornell University. His early results suggest that the process of language acquisition might be much more active than the majority of linguists have assumed up until now.

Dr. Katzir's work was recently presented at a Cornell University workshop, where researchers from fields in linguistics, psychology, and computer science gathered to discuss learning processes.

A math model in mind

Able to learn basic grammar, the [computer program](#) relies on no preconceived assumptions about language or how it might be learned. Still in its early stages of development, the program helps Dr. Katzir explore the limits of learning -- what kinds of information can a complex cognitive system like the human mind acquire and then store at the unconscious level? Do people "learn" language, and if so, can a computer be made to learn the same way?

Using a type of machine learning known as "unsupervised learning," Dr. Katzir has programmed his computer to "learn" simple grammar on its own. The program sees raw data and conducts a random search to find the best way to characterize what it sees.

The computer looks for the simplest description of the data using a criterion known as Minimum Description Length. "The process of human learning is similar to the way computers compress files: it searches for recognizable patterns in the data. Let's say, for instance, that you want to describe a string of 1,000 letters. You can be very naïve and list all the letters in order, or you can start to notice patterns -- maybe every other character is a vowel -- and use that information to give a more compact description. Once you understand something better, you can describe it more efficiently," he says.

Artificial intelligence for answering machines

His early results point to the conclusion that the computer, modeling the human mind, is indeed able to "learn" -- that language acquisition need not be limited to choosing from a finite series of possibilities.

While it's primarily theoretical, Dr. Katzir's research may have

applications in technologies such as voice dialogue systems: a computer that, on its own, can better understand what callers are looking for. A more advanced version of Dr. Katzir's program might learn natural [language](#) grammar and be able to process data received in a realistic setting, reflecting the manner in which humans actually talk.

The results of the research might also be applied to study how we learn to "read" visual images, and may be able to teach a robot how to reconstruct a three-dimensional space from a two-dimensional image and describe what it sees. Dr. Katzir plans to pursue this line of research with engineering colleagues at Tel Aviv University and abroad.

"Many linguists today assume that there are severe limits on what is learnable," Dr. Katzir says. "I take a much more optimistic view about those limitations and the capacity of humans to learn."

Provided by Tel Aviv University