The Reasoning Revolution Revealed

Hanna M. Wallach

Ten years ago, the technology industry was in the middle of an information revolution. We could store, process, and share far more information than ever before, and were working out what these new technology capabilities meant for society. Now, storing and processing information is easy, and scientists are facing a new set of questions. How can we learn from, and reason about, the vast quantities of information now available to us? Until recently we used computer databases to store data, but left understanding data as a job for humans. Humans are very good at reasoning about data, even when the available information is highly uncertain. Humans are still much better at reasoning than computers, but machine reasoning is becoming more and more important, for practical and financial reasons. Machines can work faster and for longer hours than humans, and make fewer mistakes.

Most of us have only heard of "Bayesian" reasoning in the last few years, perhaps first in the context of spam filtering. But Bayesian reasoning was actually invented over 250 years ago. The cornerstone of the Bayesian framework is the idea that probabilities can be used to describe degrees of belief. Advocates of Bayesian reasoning argue that in order to carry out useful reasoning about data, computers need to do what humans do, and start by making assumptions. These assumptions are encoded mathematically as 'priors', which describe how much we believe that each possible state of affairs is really true before seeing any data. Bayesian inference provides a framework that combines these prior beliefs with actual observations, enabling us to adjust our beliefs as we see more data.

Until recently, many people questioned whether the Bayesian framework made sense, or was useful for real problems. But thanks to recent improvements in the understanding of Bayesian inference outside the statistics community, and big increases in computing power, Bayesian probability is now hailed as the key to machine reasoning. It has already been quietly adopted in many systems, enabling computers to carry out reasoning that was previously done by humans. Some of the first areas to adopt the Bayesian framework include medicine, where Bayesian methods have been used to identify the parts of the brain responsible for particular tasks, and the media, where articles, TV programs, radio broadcasts are now automatically categorised using Bayesian techniques to bring us the most relevant information about the topics we're most interested in.

You can bet that when he drew up his "doctrine of chances" in the 1700s, Reverend Thomas Bayes, the father of the Bayesian framework, never realised that he would be responsible for a new generation of machine-based reasoning. But those of us currently experiencing this reasoning revolution are extremely grateful.