Hierarchical Bayesian Networks

In 1999 Joshua Tenenbaum graduated at MIT with a dissertation on "A Bayesian Framework for Concept Learning". In 2001 he and his student Thomas Griffiths at Stanford University developed their "hierarchical Bayesian models" for inductive generalization at various levels of abstraction, i.e. learning higher-level concepts ("Structure Learning in Human Causal Induction", 2001). These are directed graphical models (like Pearl's Bayesian nets and unlike Boltzmann machines) implemented in conjunction with the Markov Chain Monte Carlo procedure.

The hierarchical Bayesian framework was later refined by Tai-Sing Lee of Carnegie Mellon University ("Hierarchical Bayesian Inference In The Visual Cortex", 2003). These studies were also the basis for the widely-publicized "Hierarchical Temporal Memory" model of the startup Numenta, founded in 2005 in Silicon Valley by Jeff Hawkins, Dileep George and Donna Dubinsky; yet another path to get to the same paradigm: hierarchical Bayesian belief networks. The startup was charged with developing the theory presented by Silicon Valley entrepreneur Jeff Hawkins in his book "On Intelligence" (2004), ambitiously subtitled "How a new understanding of the brain will lead to the creation of truly intelligent machines".

In 2008 Alison Gopnik proposed the "Theory Theory" of how children learn new concepts, the idea being that children use the same approach that a scientist uses to develop a scientific theory. Tenenbaum's hierarchical Bayesian models were immediately identified as a plausible mathematical tool to mimic what happens in the child's mind. The neural network is also unable to do much with the concept that it learns. We generally use the concepts that we learn, we do it right away and we do it naturally. We can imagine a whole universe of relationships between the new concepts and other concepts, and we can explain those relationships. More importantly, we can take action based on the new concept. Joshua Tenenbaum, Charles Kemp (now at CMU) and Thomas Griffiths (now at UC Berkeley) wrote ("How to Grow a Mind", 2011) in which they

argued that hierarchical Bayesian models underlie all of our cognitive life. In 2014 Tenenbaum, working with Brenden Lake of New York University and Ruslan Salakhutdinov of the University of Toronto, used Bayesian reasoning instead of neural networks and devised a program that learns in a more human-like fashion, although only in the narrow domain of handwritten characters ("Human-level Concept Learning through Probabilistic Program Induction", 2014).

There is also a trend to think of transfer learning and multi-task learning (the two kinds of learning that are common among humans but so hard to realize in algorithms) as forms of generalization, after Tenenbaum and Griffiths at Stanford University viewed Bayesian inference as the path to generalization ("Generalization, Similarity, and Bayesian Inference", 2001), following the ideas of Stanford psychologist Roger Shepard ("Toward a Universal Law of Generalization for Psychological Science", 1987).