TECHNICAL REPORT
R-368
October 2010

DOI:10.1145/1831407.1831430

Technical Perspective Belief Propagation

By Yair Weiss and Judea Pearl

WHEN A PAIR of nuclear-powered Russian submarines was reported patrolling off the eastern seaboard of the U.S. last summer, Pentagon officials expressed wariness over the Kremlin's motivations. At the same time, these officials emphasized their confidence in the U.S. Navy's tracking capabilities: "We've known where they were," a senior Defense Department official told the *New York Times*, "and we're not concerned about our ability to track the subs."

While the official did not divulge the methods used by the Navy to track submarines, the *Times* added that such tracking "can be done from aircraft, ships, underwater sensors, or other submarines." But the article failed to mention perhaps the most important part of modern tracking technology—the algorithm that fuses different measurements at different times. Nearly every modern tracking system is based on the seminal work of Rudolf Kalman[1] who developed the optimal fusion algorithm for linear dynamics under Gaussian noise. This algorithm, now known simply as the "Kalman filter" is used in a remarkably broad range of real-world applications—from patient monitoring to spaceship navigation. But in the 50 years since Kalman first published his algorithm, it has become apparent that the problem it addresses is a special case of a much more general problem.

This general problem, known as "Bayesian inference in graphical models," is defined on a graph where the nodes denote random variables and edges encode direct probabilistic dependencies. Each node has access to a noisy measurement about its state. In the case of tracking a submarine, the tth node will represent the location of a submarine at time t, and edges will connect node t to node $t+1$ in a temporal chain, representing the fact that a submarine's current location is highly dependent on its location in the previous time. Kalman's algorithm allows one to efficiently compute the optimal estimate of the submarine's location, given *all the measurements*. It assumes the probabilistic dependencies are Gaussian and the graph is a temporal chain.

The generalization of Kalman's algorithm to arbitrary graphical models is called "belief propagation"[2] and it originated in the late 1970s after Judea Pearl read a paper by the cognitive psychologist David Rumelhart on how children comprehend text.[3] Rumelhart presented compelling evidence that text comprehension must be, first, a collaborative computation among a vast number of autonomous, neural-like modules, each doing an extremely simple and repetitive task and, second, that some kind of friendly "handshaking" must take place between top-down and bottom-up modes of inference, for example, the meaning of a sentence helps disambiguate a word while, at the same time, recognizing a word helps disambiguate a sentence. This disambiguation is similar to what happens in a Kalman filter (where measurements at one time can disambiguate measurements at another time), but the dependency structure is certainly not a temporal chain.

Not caring much about generality, Pearl pieced together the simplest structure he could think of (that is, a tree) and tried to see if anything useful can be computed by assigning each variable a simple processor, forced to communicate only with its neighbors. This gave rise to the tree-propagation algorithm and, a year later, to belief propagation on poly-trees, which supported bi-directional inferences and interactions known as "explaining-away."

Although several algorithms were later developed to perform Bayesian updating in general, "loopy" structures, the prospects of achieving such updating by local message passing process remained elusive. Out of total frustration, yet still convinced that such algorithms must guide many of our cognitive abilities, Pearl imagined a "shortsighted" algorithm that totally ignores the loopy structure of the graph and propagates messages as if each module is situated in a polytree environment. He then assigned as a homework exercise[2] the task of evaluating the extent to which this uninformed algorithm could serve as an approximation to the exact Bayesian inference problem. This "homework exercise" was partially solved by different researchers in the last decade and loopy belief propagation is now used successfully in applications ranging from satellite communication to driver assistance.

The success of loopy belief propagation, however, has been limited to discrete state spaces. In the following paper, Sudderth et al. provide an elegant algorithm that handles continuous variables. Unlike the Kalman filter, it does not require the probabilistic dependencies to be Gaussian, relying instead on stochastic algorithms known as "Monte-Carlo" algorithms. An extension to Kalman filters called "particle filters" also uses Monte-Carlo algorithms, but the authors provide an algorithm that can work with any dependency structure, not just a temporal chain. They show how their algorithm successfully solves some important "loopy" problems in computer-vision and sensor networks. One only wonders if the future such algorithms will be used to solve the really difficult problems—figuring out the Kremlin's intent from partial, noisy observations, or reading text as children do.

References

1. Kalman, R.E. A new approach to linear filtering and prediction problems. *Transactions of the ASME—Journal of Basic Engineering 82*, Series D (1960) 35-45.
2. Pearl, J. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufmann, 1988.
3. Rumelhart, D. Toward and interactive model of reading. Technical Report CHIP-56, Center for Human Information Processing, University of California, San Diego, 1976.

Yair Weiss (yweiss@cs.huji.ac.il) is a professor in the School of Computer Science and Engineering at The Hebrew University of Jerusalem, Israel.

Judea Pearl (Judea@cs.ucla.edu) is a professor of computer science and director of the Cognitive Systems Laboratory at University of California, Los Angeles.

© 2010 ACM 0001-0782/10/1000 $10.00