INFERENCE AND LEARNING IN COMPLEX STOCHASTIC PROCESSES

Xavier Boyen, Ph.D.

Stanford University, 2003

Adviser: Daphne Koller

Systems with stochastic dynamics and complex structures are pervasive in the real world. Examples abound in a variety of domains, and include the flow of wealth in the stock market, and the ecology of evolving ecosystems. Such systems are composed of many variables that tend to interact in a structured fashion.

The ability to perform inference in such systems is key to many applications, such as tracking or learning based on observed behavior over time. The main obstacle to scalability is the need to reason with probability distributions, which quickly become unmanageable with more than a few variables. In the static case, Bayesian Networks are known to provide an appropriate solution for structured systems. In the dynamic case, inference is almost always intractable, even for nicely structured processes. Intuitively, this is because correlations propagate over time from one variable to another, resulting in all variables becoming fully entangled.

These correlations are shown to be typically weak, and, for inference purposes, to have a limited lifetime. These phenomena are exploited in an efficient approximate inference algorithm for stochastic processes represented as Dynamic Bayesian Networks (DBNs), by factoring out provably weak correlations. Under certain assumptions, it is shown that, even if approximate inference goes on forever, the total approximation error remains small and bounded on expectation. Novel notions of "weak" and "sparse" interaction are introduced, to capture quantitative aspects of the interaction strength between subprocesses. These notions are useful both to refine the estimation of inference errors, and to guide the design of approximations. Experimental performance and accuracy evaluations are provided for tracking complex systems borrowed from real life.

With regard to learning the structure and parameters of partially observable processes, it is shown how approximate inference can greatly accelerate structural and parametric learning, with minimal accuracy penalty. An additional temporal approximation based on similar ideas is also shown to produce high-performance online learners. Finally, it is suggested how to discover hidden state variables simply by identifying unexplained temporal correlations. Experimental results are given in which DBN structures and parameters for several real-life stochastic processes are learned from raw time series data.

Approved for publication:

By:

_____

For the Computer Science Department