

Statistical Learning of Probabilistic BDDs

Taisuke Sato

Department of Computer Science
Graduate School of Information Science and Engineering
Tokyo Institute of Technology
`sato@mi.cs.titech.ac.jp`

Abstract. BDDs (binary decision diagrams) are a graphical data structure which compactly represents Boolean functions and has been extensively used in VLSI CAD. In this talk, I will describe how BDDs are applied to machine learning, or more concretely how they are used in probability computation and statistical parameter learning for probabilistic modeling. The basic idea is that of PPC (propositionalized probabilistic computation) which considers “ $X = x$ ”, a random variable X taking a value x , as a probabilistic proposition. A complex (discrete) probabilistic model such as a Bayesian network is representable by way of PPC as a probabilistic Boolean formula, thus by a probabilistic BDD.

We developed an efficient sum-product probability computation algorithm on BDDs and also derived a parameter learning algorithm called the BDD-EM algorithm for maximum likelihood estimation in statistics. We applied the BDD-EM algorithm to the problem of finding a most probable hypothesis about a metabolic pathway in systems biology.