Toward Provably Correct Feature Selection in Arbitrary Domains

Dimitris Margaritis

- Optimal Feature Selection for X ⇔ Finding Markov Boundary (min Markov Blanket) of X
  - Pr(X | all vars) = Pr(X | boundary of X)
- Many algos for "normal" probability distributions → GS, IAMB & variants, MMMB etc
- Don't work for corner cases (e.g., parity functions in discrete domains); these contain higher-order interaction (involving more vars) and are hard to detect
- We give 2 algorithms and a basic theorem for provably correct, approximate algorithms for such cases as well: 1 deterministic, 1 randomized; parameter called algo's "margin"