Learning Bayesian networks

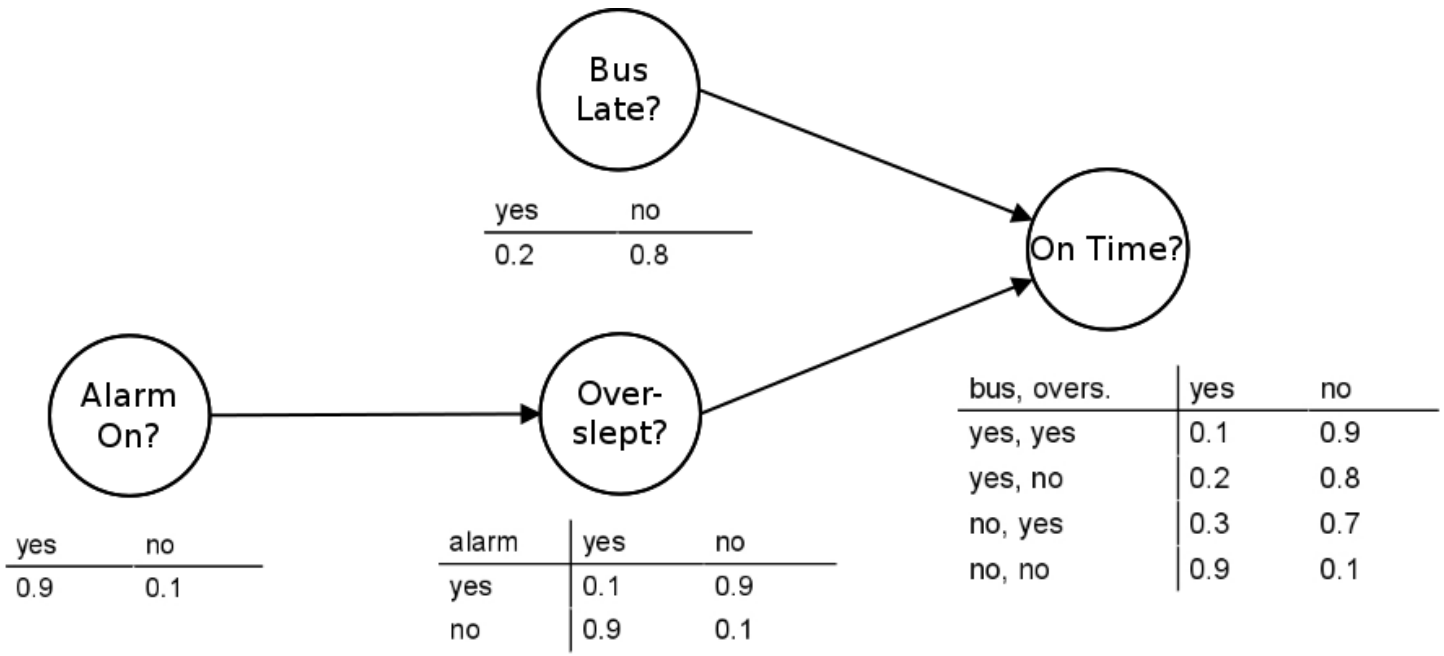
Pekka Parviainen

Department of Informatics
University of Bergen

NORA Kickoff Seminar
1.4.2019

Bayesian networks

- Representations of joint probability distributions
- Consist of:
  - The structure is a directed acyclic graph (DAG) that represents conditional independencies between variables
  - The local conditional probability distributions that are specified by parameters



Why are Bayesian networks useful?

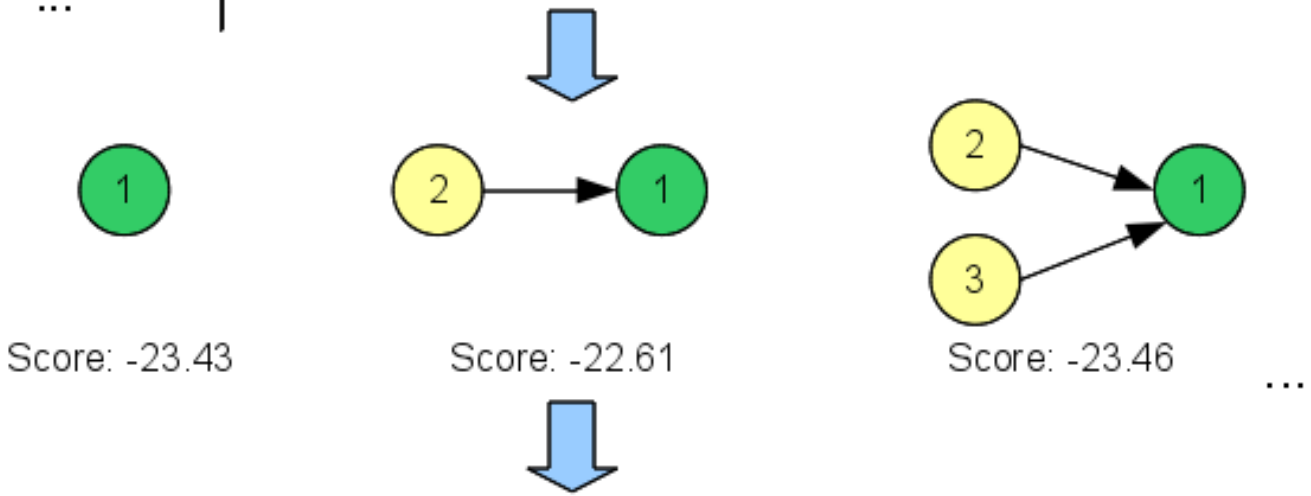
- Compact
  - Number of parameters grows with the maximum number of parents
- Flexible
  - It is possible to compute any conditional probability
  - This is called *inference*
- Interpretable
  - Human expert can evaluate semantics and properties of the network

Learning Bayesian networks

- Input: samples from a multivariate probability distribution
- Learning the parameters of a Bayesian network
  - Given the structure, learning parameters is straightforward
  - Just use your favorite statistical principle (maximum likelihood, Bayesian, . . .)
- Learning the structure of a Bayesian network
  - The structure determines the parameters that are needed.
  - Thus, one can first learn the structure and then the parameters.
  - NP-hard in general

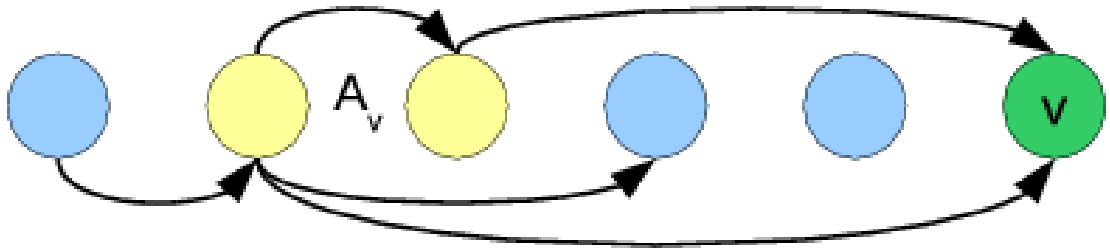
Score-based Structure Discovery

	Var. 1	Var. 2	Var. 3	Var. 4	...
Person A	1	1	2	1	
Person B	2	2	1	1	
Person C	1	2	2	2	
Person D	2	1	2	1	
...					



Research questions

- Exact learning of the structure
  - NP-hard, but doable for moderate-sized networks
- Learning Bayesian networks with bounded graph parameters
  - Motivation: Guarantee fast inference
- Heuristics with quality guarantees
  - Motivation: Exact algorithms are unlikely to scale-up into large networks



Thank you!