Informatics 2D: Reasoning and Agents

Alex Lascarides

School of Informatics
The University of Edinburgh

Lecture 23c: Efficient Representations of CPTs in BNs

Where are we?

- BNs are:
  1. A DAG that captures conditional independence
  2. CPTs for each variable: $\mathbf{P}(X|Parents(X))$
- BNs are a compact representation of JPDs, so can answer any query with a BN (more next time)
- Today: **Efficient representation of CPTs in BNs**

Efficient representation of conditional distributions

- Even the 2^k (k parents) conditioning cases that have to be provided require a great deal of experience and knowledge of the domain
- Arbitrary relationships are unlikely, often describable by **canonical distributions** that fit some standard pattern
- By specifying pattern by a few parameters we can save a lot of space!
- Simplest case: **deterministic node** that can be directly inferred from values of parents
- For example, logical or mathematical functions

Noisy-OR relationships

Generalisation of logical OR

- Any cause *can* make effect true, but won't *necessarily* (effect **inhibited**; $P(effect|cause) < 1$)
- Assumes all causes are listed (**leak node** can be used to cater for "miscellaneous" unlisted causes)
- Also assumes inhibitions are mutually conditionally independent
  - Whatever inhibits C_1 from making E true is independent of what inhibits C_2 from making E true.
- So E is *false* only if each of its *true* parents are inhibited and we can compute this likelihood from product of probabilities for each individual cause inhibiting E.
- How does this help?

Example of Noisy-OR

- *Fever* is caused by *Cold*, *Flu* or *Malaria* and that's all (!!)
- Inhibitions of *Cold*, *Flu* and *Malaira* are mutually conditionally independent
- Likelihood that *Cold* is inhibited from causing *Fever* is $P(\neg fever|cold, \neg flu, \neg malaria)$ (similarly for other causes)
- Individual inhibition probabilities:

$$P(\neg fever|cold, \neg flu, \neg malaria) = 0.6$$
$$P(\neg fever|\neg cold, flu, \neg malaria) = 0.2$$
$$P(\neg fever|\neg cold, \neg flu, malaria) = 0.1$$

- Inhibitions mutually independent, so:

$$P(\neg fever|cold, flu, \neg malaria) = P(\neg fever|cold, \neg flu, \neg malaria)P(\neg fever|\neg cold, flu, \neg malaria)$$

Noisy-OR relationships

- We can construct entire CPT from this information

Cold	*Flu*	*Malaria*	$P(Fever)$	$P(\neg Fever)$
F	F	F	0.0	1.0
F	F	T	0.9	**0.1**
F	T	F	0.8	**0.2**
F	T	T	0.98	$0.02 = 0.2 \times 0.1$
T	F	F	0.4	**0.6**
T	F	T	0.94	$0.06 = 0.6 \times 0.1$
T	T	F	0.88	$0.12 = 0.6 \times 0.2$
T	T	T	0.988	$0.012 = 0.6 \times 0.2 \times 0.1$

- Encodes CPT with k instead of 2^k values!

BNs with continuous variables

- Often variables range over continuous domains
- **Discretisation** one possible solution but often leads to inaccuracy or requires a lot of discrete values
- Other solution: use of standard families of probability distributions specified in terms of a few parameters
- Example: normal/Gaussian distribution $N(\mu,\sigma^2)(x)$ defined in terms of mean μ and variance σ^2 (needs just two parameters)
- **Hybrid Bayesian Networks** use mixture of discrete and continuous variables (special methods to deal with links between different types – not discussed here)

Summary

- Introduced Bayesian Networks as a structured way of reasoning under uncertainty using probabilities and independence
- Defined their semantics in terms of JPD representation, and conditional independence statements
- Gave numerical and topological interpretation of semantics
- Talked about issues of efficient representation of CPTs
- Discussed continuous variables and hybrid networks
- Next time: **Exact Inference in Bayesian Networks**