Informatics 2D: Reasoning and Agents

Alex Lascarides

School of informatics

The University of Edinburgh

Lecture 22a: Joint Probability Distributions (JPDs)

Where are we?

So far ...

- Introduced basics of decision theory
  (probability theory + utility)
- Talked about random variables, probability distributions
- Introduced basic probability notation and axioms

Today ...

- **Probabilities and Bayes' Rule**

Inference with joint probability distributions

- Last time we talked about joint probability distributions (JPDs) but didn't present a method for **probabilistic inference** using them
- Problem: Given some observed evidence and a query proposition, how can we compute the **posterior probability** of that proposition?
- We will first discuss a simple method using a JPD as "knowledge base"
- Although not very useful in practice, it helps us to discuss interesting issues along the way

Example

- Domain consisting only of Boolean variables *Toothache*, *Cavity* and *Catch* (steel probe catches in tooth)
- Consider the following JPD:

	toothache		$\neg$*toothache*	
	catch	$\neg$*catch*	*catch*	$\neg$*catch*
cavity	0.108	0.012	0.072	0.008
$\neg$*cavity*	0.016	0.064	0.144	0.576

- Probabilities (table entries) sum to 1
- We can compute probability of any proposition, e.g. $P(catch \vee cavity) =$ $0.108 + 0.016 + 0.072 + 0.144 + 0.012 + 0.008 = 0.36$

Marginalisation, conditioning & normalisation

- Extracting distribution of subset of variables is called **marginalisation**: $\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z}} \mathbf{P}(\mathbf{Y}, \mathbf{z})$
- Example:

$$
\begin{aligned}
P(cavity) = {} & P(cavity, toothache, catch) + P(cavity, toothache, \neg catch) \\
& + P(cavity, \neg toothache, catch) + P(cavity, \neg toothache, \neg catch) \\
& = 0.108 + 0.012 + 0.072 + 0.008 = 0.2
\end{aligned}
$$

- **Conditioning** – variant using the product rule:

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z}} \mathbf{P}(\mathbf{Y}|\mathbf{z})P(\mathbf{z})$$

Marginalisation, conditioning & normalisation

- Computing conditional probabilities:

$$P(cavity|toothache) = \frac{P(cavity \wedge toothache)}{P(toothache)}$$
$$= \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$

- **Normalisation** ensures probabilities sum to 1, normalisation constants often denoted by α
- Example:

$$\mathbf{P}(Cavity|toothache) = \alpha\mathbf{P}(Cavity, toothache)$$
$$= \alpha[\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]$$
$$= \alpha[\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle] = \alpha\langle 0.12, 0.08\rangle = \langle 0.6, 0.4\rangle$$

A general inference procedure

- Let X be a query variable (e.g. *Cavity*), $\mathbf{E}$ set of evidence variables (e.g. $\{Toothache\}$) and $\mathbf{e}$ their observed values, $\mathbf{Y}$ remaining unobserved variables
- Query evaluation: $\mathbf{P}(X|\mathbf{e}) = \alpha\mathbf{P}(X,\mathbf{e}) = \alpha\sum_{\mathbf{y}}\mathbf{P}(X,\mathbf{e},\mathbf{y})$
- Note that X, $\mathbf{E}$, and $\mathbf{Y}$ constitute complete set of variables, i.e. $\mathbf{P}(x,\mathbf{e},\mathbf{y})$ simply a subset of probabilities from the JPD
- For every value x_i of X, sum over all values of every variable in $\mathbf{Y}$ and normalise the resulting probability vector
- Only theoretically relevant, it requires $O(2^n)$ steps (and entries) for n Boolean variables
- Basically, all methods we will talk about deal with tackling this problem!

Summary

- You can use a JPD to answer any query
  - Marginalisation
  - Normalisation