Informatics 2D: Reasoning and Agents

Alex Lascarides

School of Informatics
The University of Edinburgh

Lecture 22b: Independence and Bayes' Rule

Where are we?

Last time. . .

- Used JPDs to answer queries
- Today: **Independence and Bayes' Rule**

Independence

- Suppose we extend our example with the variable *Weather*
- What is the relationship between old and new JPD?
- Can compute $P(toothache, catch, cavity, Weather = cloudy)$ as:

$$P(Weather = cloudy | toothache, catch, cavity) P(toothache, catch, cavity)$$

- And since the weather does not depend on dental stuff, we expect that

$$P(Weather = cloudy | toothache, catch, cavity) = P(Weather = cloudy)$$

- So

$$P(toothache, catch, cavity, Weather = cloudy) = P(Weather = cloudy) P(toothache, catch, cavity)$$

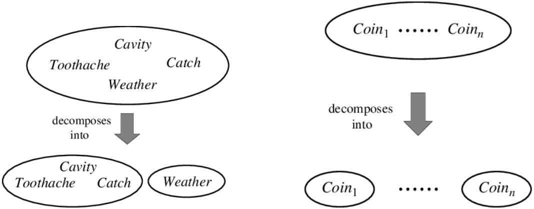
- One 8-element and one 4-element table rather than a 32-table!

Independence

- This is called **independence**, usually written as

$$\mathbf{P}(X|Y) = \mathbf{P}(X) \text{ or } \mathbf{P}(Y|X) = \mathbf{P}(Y) \text{ or } \mathbf{P}(X,Y) = \mathbf{P}(X)\mathbf{P}(Y)$$

- Depends on domain knowledge; can factor distributions



- Such independence assumptions can help to dramatically reduce complexity
- Independence assumptions are sometimes *necessary* even when not entirely justified, so as to make probabilistic reasoning in the domain practical (more later).

Bayes' rule

- **Bayes' rule** is derived by writing the product rule in two forms and equating them:

$$\left.\begin{array}{l} P(a \wedge b) = P(a|b)P(b) \\ P(a \wedge b) = P(b|a)P(a) \end{array}\right\} \Rightarrow P(b|a) = \frac{P(a|b)P(b)}{P(a)}$$

- General case for multivaried variables using background evidence **e**:

$$\mathbf{P}(Y|X,\mathbf{e}) = \frac{\mathbf{P}(X|Y,\mathbf{e})\mathbf{P}(Y|\mathbf{e})}{\mathbf{P}(X|\mathbf{e})}$$

- Useful because often we have good estimates for three terms on the right and are interested in the fourth

Applying Bayes' rule

- Example: meningitis causes stiff neck with 50%, probability of meningitis (m) 1/50000, probability of stiff neck (s) 1/20

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = \frac{\frac{1}{2} \times \frac{1}{50000}}{\frac{1}{20}} = \frac{1}{5000}$$

- Previously, we were able to avoid calculating probability of evidence ($P(s)$) by using normalisation
- With Bayes' rule: $\mathbf{P}(M|s) = \alpha \langle P(s|m)P(m), P(s|\neg m)P(\neg m) \rangle$
- Usefulness of this depends on whether $P(s|\neg m)$ is easier to calculate than $P(s)$
- Obvious question: why would conditional probability be available in one direction and not in the other?
- Diagnostic knowledge (from symptoms to causes) is often fragile
  (e.g. $P(m|s)$ will go up if $P(m)$ goes up due to epidemic)

Summary

- Independence is critical for making probabilistic reasoning tractable.
  - Sometimes you must assume independence that's false to achieve practical reasoning
- Bayes Rule critical for optimising use of available data for answering your queries.
- Next time: **Combining evidence to answer queries**