Independence and Conditional Independence

Frederik Mallmann-Trenn
6CCS3AIN

Inference by enumeration

- We saw that with our joint distribution table

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

  we can calculate any probability
- Obvious problems:
  1. Worst-case time complexity $O(d^n)$ where d is the largest arity
  2. Space complexity $O(d^n)$ to store the joint distribution
  3. How to find the numbers for $O(d^n)$ entries???
- These problems effectively stopped the use of probability in AI until the mid 80s

Computational efficiency

- To get efficient probabilistic computations, we need two things.
  1. (Conditional) independence.
  2. Bayes rule.
- Will cover these now, in order.

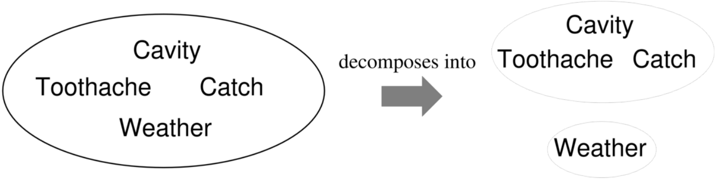
Independence

- A and B are independent iff

$$\mathbf{P}(A, B) = \mathbf{P}(A)\mathbf{P}(B)$$

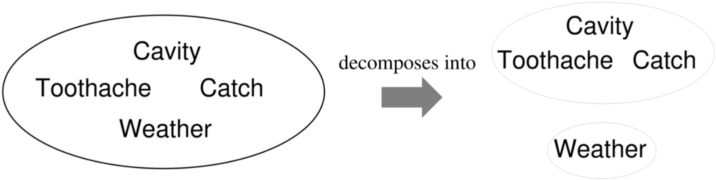
- Why is this interesting?
- Can help with the size of the problem.

Independence



- $\mathbf{P}(Toothache, Catch, Cavity, Weather)$
  $= \mathbf{P}(Toothache, Catch, Cavity) \odot \mathbf{P}(Weather)$
- If you store all values naively, this requires $2 \cdot 2 \cdot 2 \cdot 4 = 32$ entries.
- You can do it in 31 by leaving one entry empty. This is possible since you know that the probabilities add up to 1.
- Using the independence, you can even reduce the 31 values to 10:

Independence



- $\mathbf{P}(Toothache, Catch, Cavity, Weather)$
  $= \mathbf{P}(Toothache, Catch, Cavity) \odot \mathbf{P}(Weather)$
- If you store all values naively, this requires $2 \cdot 2 \cdot 2 \cdot 4 = 32$ entries.
- You can do it in 31 by leaving one entry empty. This is possible since you know that the probabilities add up to 1.
- Using the independence, you can even reduce the 31 values to 10:
- You store 7 for the 3 dependent variables and 3 for weather which is independent (note that we're using the probabilities adding up to 1 trick twice here)
- For n independent biased coins, $2^n \to n$

Independence

- Absolute independence powerful but rare
- Dentistry is a large field with hundreds of variables, none of which are independent. What to do?
- Conditional independence

Conditional independence

- $\mathbf{P}(Toothache, Cavity, Catch)$ has:
- Three binary variables.
- Thus 2^3 entries in the joint probability table.
- But these sum to 1.
- So $2^3 - 1$ independent entries
- That's 7 independent entries

Conditional independence

- But, wait! There's more!
- If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:

$$P(catch|toothache, cavity) = P(catch|cavity) \tag{1}$$

- The same independence holds if I haven't got a cavity:

$$P(catch|toothache, \neg cavity) = P(catch|\neg cavity) \tag{2}$$

- $Catch$ is conditionally independent of $Toothache$ given $Cavity$

$$\mathbf{P}(Catch|Toothache, Cavity) = \mathbf{P}(Catch|Cavity)$$

Conditional independence

- Equivalent statements:

$$\begin{aligned}\mathbf{P}(Toothache|Catch, Cavity) &= \mathbf{P}(Toothache|Cavity)\\ \mathbf{P}(Toothache, Catch|Cavity) &= \mathbf{P}(Toothache|Cavity) \odot \mathbf{P}(Catch|Cavity)\end{aligned}$$

Conditional independence

- Write out full joint distribution using chain rule:

$$\begin{aligned}
&\mathbf{P}(Toothache, Catch, Cavity)\\
&= \mathbf{P}(Toothache|Catch, Cavity) \odot \mathbf{P}(Catch, Cavity)\\
&= \mathbf{P}(Toothache|Catch, Cavity) \odot \mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)\\
&= \mathbf{P}(Toothache|Cavity) \odot \mathbf{P}(Catch|Cavity) \odot \mathbf{P}(Cavity)
\end{aligned}$$

- 2 + 2 + 1 = 5 independent numbers
- Equations 1 and 2 remove 2.

Conditional independence

- In most cases, the use of conditional independence reduces the size of the representation of the joint distribution from exponential in n to (close to) linear in n.
- Conditional independence is our most basic and robust form of knowledge about uncertain environments.
- Can often make conditional independence statements when little else is known.