

Probabilistic Models

- Models describe how (a portion of) the world works
- Models are always simplifications
 - May not account for every variable
 - May not account for all interactions between variables

Probability

Independence

- $\forall x, y : \mathbb{P}(x, y) = \mathbb{P}(x)\mathbb{P}(y)$
 - $\forall x, y : \mathbb{P}(x|y) = \mathbb{P}(x)$
- Also true if $\forall x, y : \mathbb{P}(x) = \mathbb{P}(x|y)$

$$X \perp\!\!\!\perp Y$$

- Written as:
- Independent actions can be considered individually; do not need to keep a massive table with independent probabilities
- Pretty rare (especially in models)

Conditional Independence

- Variables can become independent if you know one of the conditions is true

- **P(Toothache, Cavity, Catch)**

- If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:

- $P(+catch \mid +toothache, +cavity) = P(+catch \mid +cavity)$

- The same independence holds if I don't have a cavity:

- $P(+catch \mid +toothache, -cavity) = P(+catch \mid -cavity)$

- Catch is *conditionally independent* of Toothache given Cavity:

- $P(Catch \mid Toothache, Cavity) = P(Catch \mid Cavity)$

- **Equivalent statements:**

- $P(Toothache \mid Catch, Cavity) = P(Toothache \mid Cavity)$

- $P(Toothache, Catch \mid Cavity) = P(Toothache \mid Cavity) P(Catch \mid Cavity)$

- One can be derived from the other easily

- eg:

- $\forall x, y, z : \mathbb{P}(x, y|z) = \mathbb{P}(x|z)\mathbb{P}(y|z)$



- $\forall x, y, z : \mathbb{P}(x|z, y) = \mathbb{P}(x, z)$

▪ **Chain rule:**

$$P(X_1, X_2, \dots, X_n) = P(X_1)P(X_2|X_1)P(X_3|X_1, X_2) \dots$$

▪ **Trivial decomposition:**

$$T \perp\!\!\!\perp U \mid R$$

$$P(\text{Traffic}, \text{Rain}, \text{Umbrella}) =$$

$$P(\text{Rain})P(\text{Traffic}|\text{Rain})P(\text{Umbrella}|\text{Rain}, \text{Traffic})$$

▪ **With assumption of conditional independence:**

$$P(\text{Traffic}, \text{Rain}, \text{Umbrella}) =$$

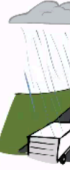
$$P(\text{Rain})P(\text{Traffic}|\text{Rain})P(\text{Umbrella}|\text{Rain})$$

- Chain Rule with conditional independence:
 - Note the necessity of the conditional independence assumption
 - Can be represented with Bayes' nets

Bayes' Nets

- A technique for describing complex joint distributions (models) using simple, local distributions (conditional probabilities)
- Good because sometimes the tables get too big
- Hard to learn full table empirically about tables with more than a few variables
- Type of graph
 - Nodes: variables (with domains)
 - * Can be assigned (observed) or unassigned (unobserved)
 - Arcs: interactions
 - * Similar to CSP constraints
 - * Indicate 'direct influence' between variables
 - * Formally: encode conditional independence
- Given a set of nodes, one per variable X , it is a directed, acyclic graph with a conditional distribution for each node
 - A collection of distributions over X , one for each combination of parents' values: $\mathbb{P}(X|a_1, \dots, a_n)$
 - CPT: conditional probability table
- To see what probability a BN gives to a full assignment, multiply all the relevant conditionals together:

$$\mathbb{P}(x_1, x_2, \dots, x_n) = \prod_{i=1}^n \mathbb{P}(x_i | \text{parents}(X_i))$$
 - Good because it assumes conditional independence
-



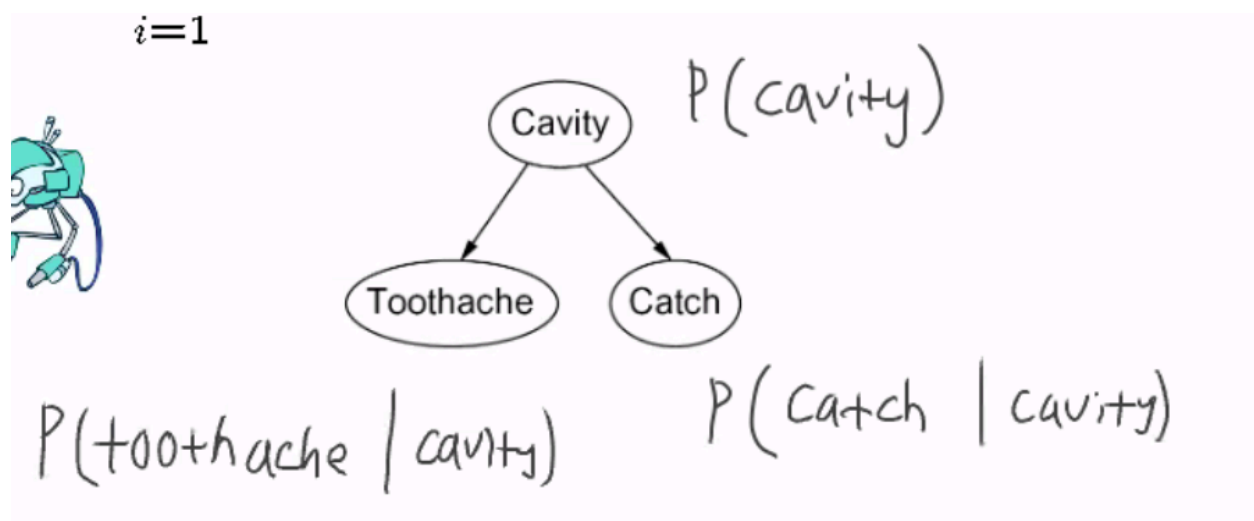


Figure 1: Screenshot_2023-10-05_at_6.19.03_PM.png