CS 486 — Lecture 11: Independence and Introduction to Bayesian Networks

# 1 Independence

- Unconditional independence: $X, Y$ are unconditionally independent iff:

$$P(X|Y) = P(X)$$
$$P(Y|X) = P(Y)$$
$$P(X \wedge Y) = P(X)P(Y)$$

  Learning $Y$ does not influence our beliefs about $X$!

- Conditional independence: $X$ and $Y$ are conditionally independent given $Z$ if:

$$P(X|Y \wedge Z) = P(X|Z)$$
$$P(Y|X \wedge Z) = P(Y|Z)$$
$$P(Y \wedge X|Z) = P(Y|Z)P(X|Z)$$

  Learning $Y$ does not influence $X$ if we already know $Z$.

- In general, to specify the joint distribution for $n$ RV, you need $2^n - 1$ probabilities. This is as:

$$P(A \wedge B \wedge C) = P(A)P(B|A)P(C|A \wedge B)$$

  This means we need $P(A), P(B|A), P(B|\neg A), P(C|A \wedge B), P(C|A \wedge \neg B), P(C|\neg A \wedge B), P(C|\neg A \wedge \neg B)$.

- However, if they are all independent, then we only need $n$ probabilities. As:

$$P(A \wedge B \wedge C) = P(A)P(B)P(C)$$

- Lastly, if $A$ and $B$ are conditionally independent given $C$. We would need 5 probabilities.

- We can figure this out by drawing dependency graphs (ie: in the previous example, $C$ is needed for $A$ and $B$, so we need $P(C), P(A|C), P(A|\neg C), P(B), P(B|\neg C)$.

# 2 Bayes Net

- We can compute any probability using a joint distribution, but they quickly grow as the number of variables increases, and they are unnatural and tedious to specify all probabilities.

- A Bayesian Network is a compact version of the joint distribution, taking advantage of the unconditional and conditional independence among the variables.

- For example, a 6 node joint distro would need $2^6 - 1$ probabilities — the Bayesian network for the same given problem could use only 12!