Importance of independence

- Conditional independence assertions allow chain rule to be simplified:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P(X_i \mid X_{i-1}, \ldots, X_1)$$

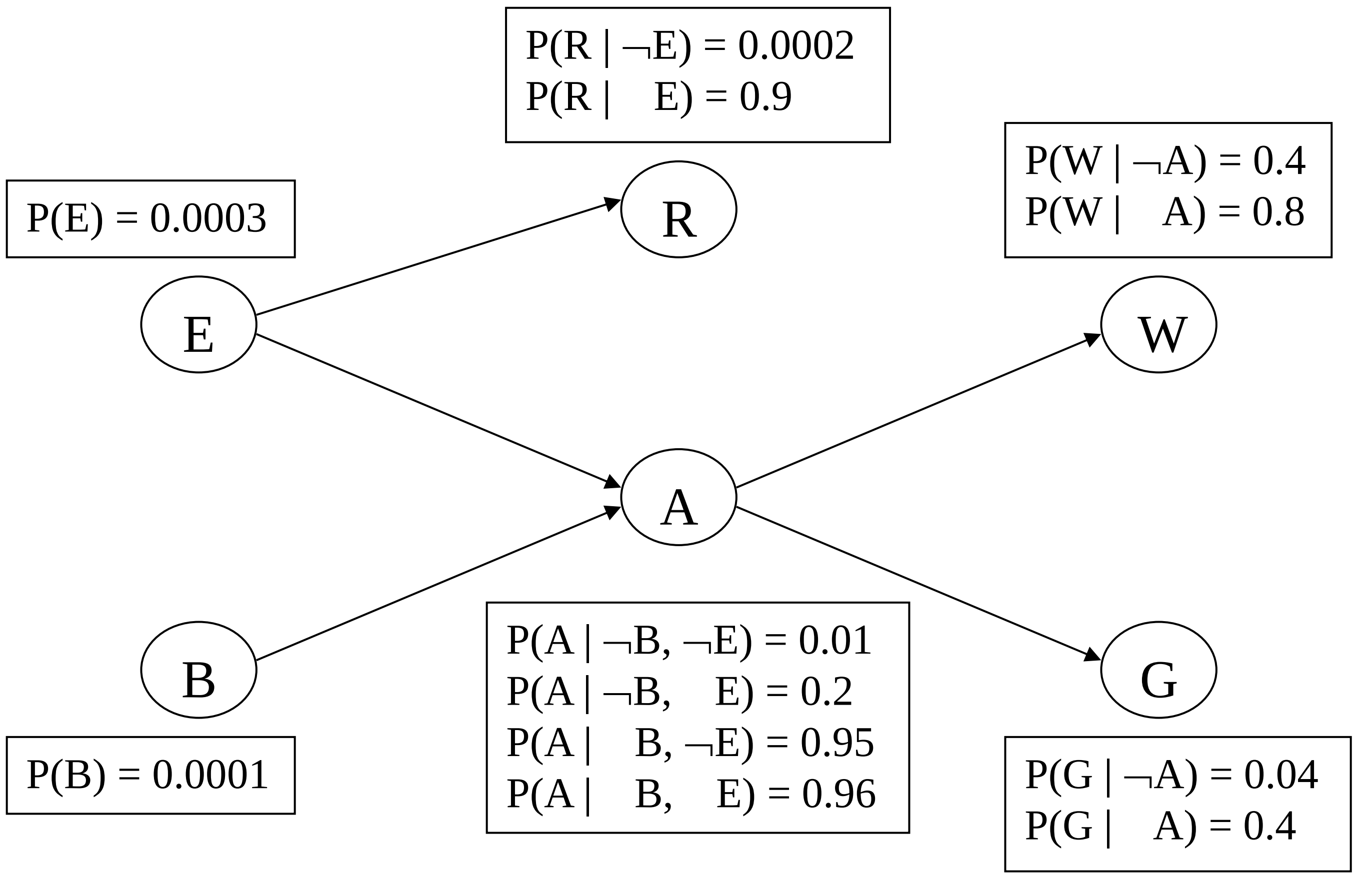
  - reduce number of probabilities that need to be specified
  - simplify computation of probabilistic queries

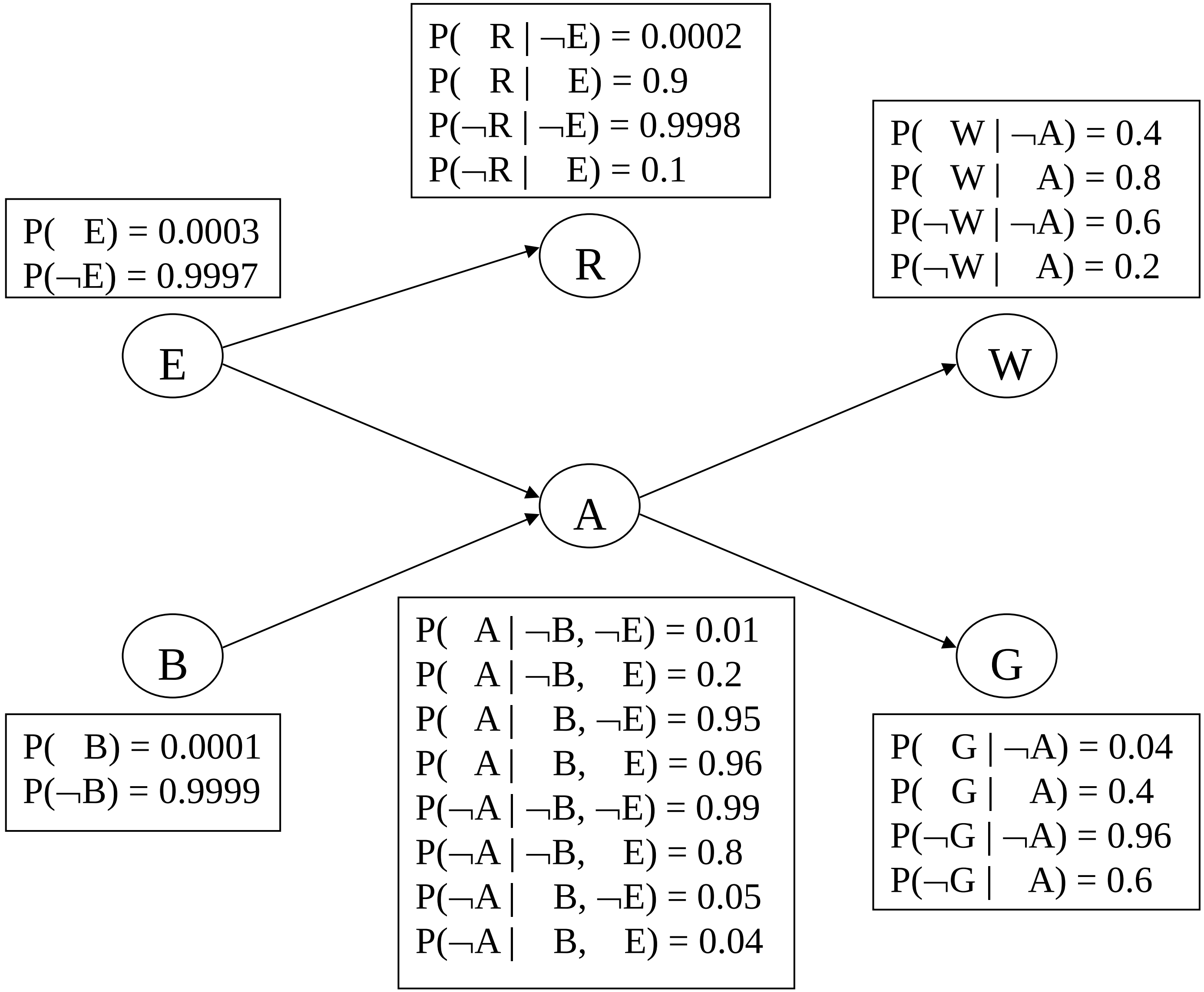
Belief networks

- A belief network is a directed acyclic graph (DAG) where:
  - nodes are random variables
  - directed arcs connect pairs of nodes
    - intuitive meaning: if arc $X \rightarrow Y$, X has a direct influence on Y
  - each node has a conditional probability table specifying the effects the parents have on the node

Belief network for Holmes example

Belief network for Holmes example





P(R | ¬E) = 0.0002
P(R | E) = 0.9
P(¬R | ¬E) = 0.9998
P(¬R | E) = 0.1
P(W | ¬A) = 0.4
P(W | A) = 0.8
P(¬W | ¬A) = 0.6
P(¬W | A) = 0.2
P(E) = 0.0003
P(¬E) = 0.9997
R
E
W
A
B
G
P(A | ¬B, ¬E) = 0.01
P(A | ¬B, E) = 0.2
P(A | B, ¬E) = 0.95
P(A | B, E) = 0.96
P(¬A | ¬B, ¬E) = 0.99
P(¬A | ¬B, E) = 0.8
P(¬A | B, ¬E) = 0.05
P(¬A | B, E) = 0.04
P(B) = 0.0001
P(¬B) = 0.9999
P(G | ¬A) = 0.04
P(G | A) = 0.4
P(¬G | ¬A) = 0.96
P(¬G | A) = 0.6

Semantics of belief networks

Two ways to understand belief networks

1. As a representation of the joint probability distribution
2. As an encoding of conditional independence assumptions

1. Representation of joint probability distribution

- Every entry in the joint probability distribution can be calculated from the network:

$$P(X_1, \ldots, X_n) = \prod P(X_i \mid \text{Parents}(X_i)) \qquad (1)$$

where each X_i can be negated or not negated

- From this, any query can be answered:

P(Query | Evidence)

2. Encoding of conditional independence assumptions

- Every entry in the joint probability distribution can be calculated from the network:

$$P(X_1, \ldots, X_n) = \prod P(X_i \mid \text{Parents}(X_i)) \qquad (1)$$

where each X_i can be negated or not negated

- Contrast with chain rule:

$$P(X_1, \ldots, X_n) = \prod P(X_i \mid X_{i-1}, \ldots, X_1) \qquad (2)$$

Encoding of conditional independence assumptions

$$\mathrm{P}(X_1, \ldots, X_n) = \Pi\ \mathrm{P}(X_i \mid \mathrm{Parents}(X_i)) \qquad (1)$$
$$\mathrm{P}(X_1, \ldots, X_n) = \Pi\ \mathrm{P}(X_i \mid X_{i-1}, \ldots, X_1) \qquad (2)$$

- Want resulting joint probability distribution to be a good representation of the domain
- Equation 1 is a <u>correct</u> representation of a domain only if each node is conditionally independent of its predecessors (in the node ordering), given its parents