# Review: probability

**Random variables**: sunshine $S \in \{0, 1\}$, rain $R \in \{0, 1\}$

**Joint distribution**:

$\mathbb{P}(S, R) =$

| $s$ | $r$ | $\mathbb{P}(S = s, R = r)$ |
|---|---|---|
| 0 | 0 | 0.20 |
| 0 | 1 | 0.08 |
| 1 | 0 | 0.70 |
| 1 | 1 | 0.02 |

**Marginal distribution**:

$\mathbb{P}(S) =$

| $s$ | $\mathbb{P}(S = s)$ |
|---|---|
| 0 | 0.28 |
| 1 | 0.72 |

(aggregate rows)

# Review: probability

**Random variables**: sunshine $S \in \{0, 1\}$, rain $R \in \{0, 1\}$

**Joint distribution**:

$\mathbb{P}(S, R) =$

| $s$ | $r$ | $\mathbb{P}(S = s, R = r)$ |
|---|---|---|
| 0 | 0 | 0.20 |
| 0 | 1 | 0.08 |
| 1 | 0 | 0.70 |
| 1 | 1 | 0.02 |

**Marginal distribution**:

$\mathbb{P}(S) =$

| $s$ | $\mathbb{P}(S = s)$ |
|---|---|
| 0 | 0.28 |
| 1 | 0.72 |

(aggregate rows)

**Conditional distribution**:

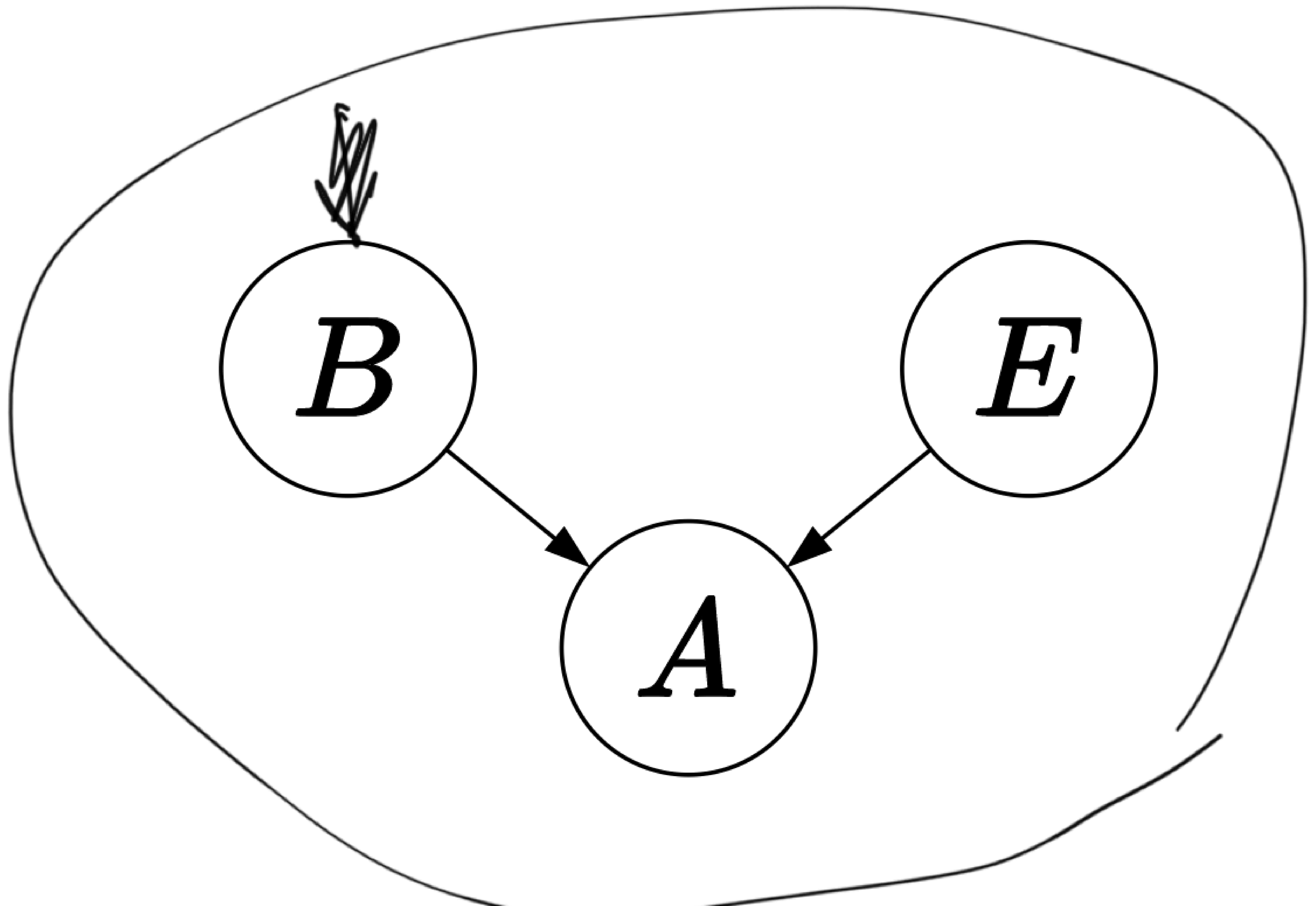
$\mathbb{P}(S \mid R = 1) =$

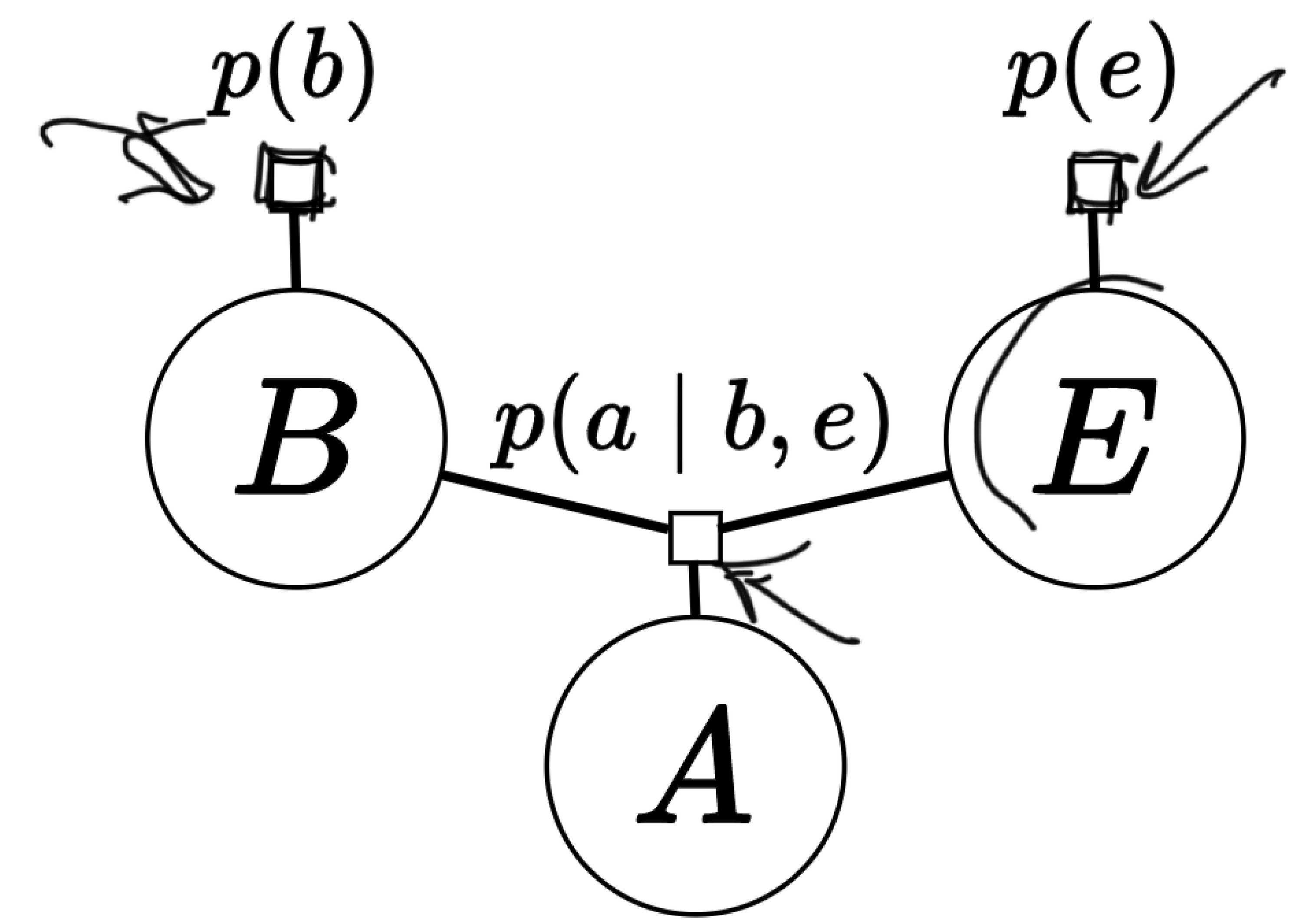
| $s$ | $\mathbb{P}(S = s \mid R = 1)$ |
|---|---|
| 0 | 0.8 |
| 1 | 0.2 |

(select rows, normalize)

divide by $0.08 + 0.02$

# Bayesian network (alarm)



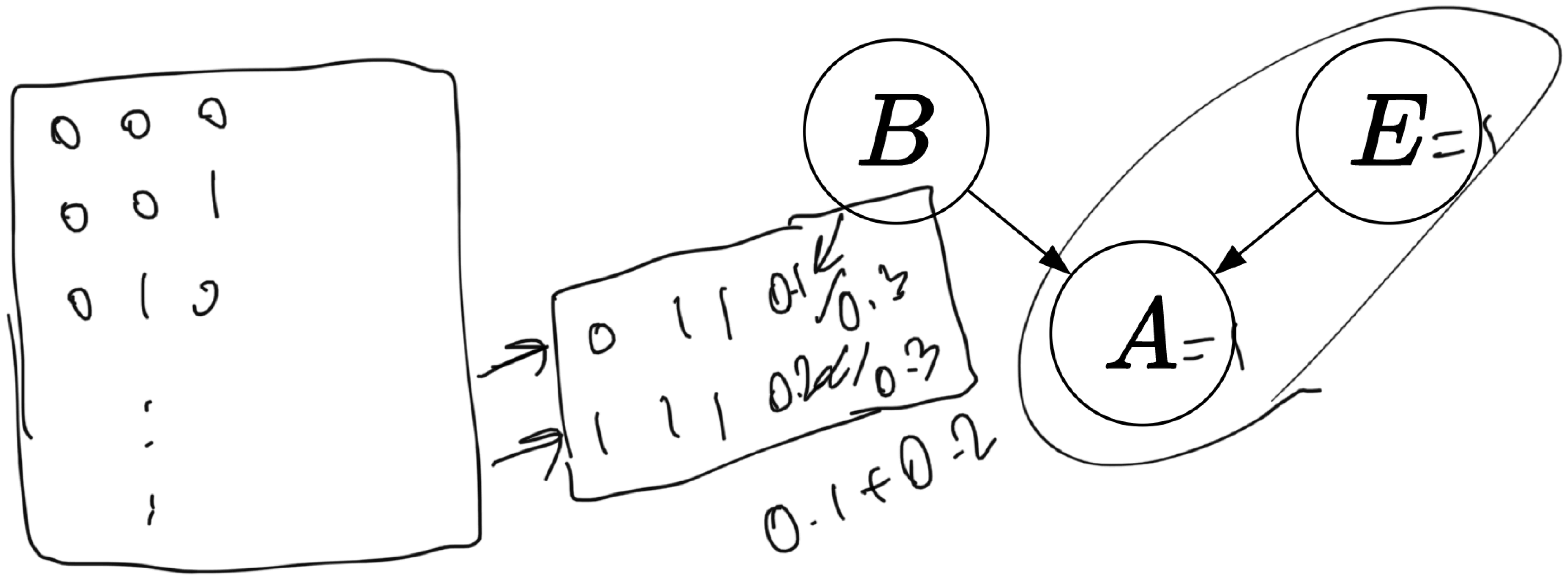




$$\mathbb{P}(B = b, E = e, A = a) = p(b)p(e)p(a \mid b, e)$$

Bayesian networks are a special case of factor graphs!

probability = weight

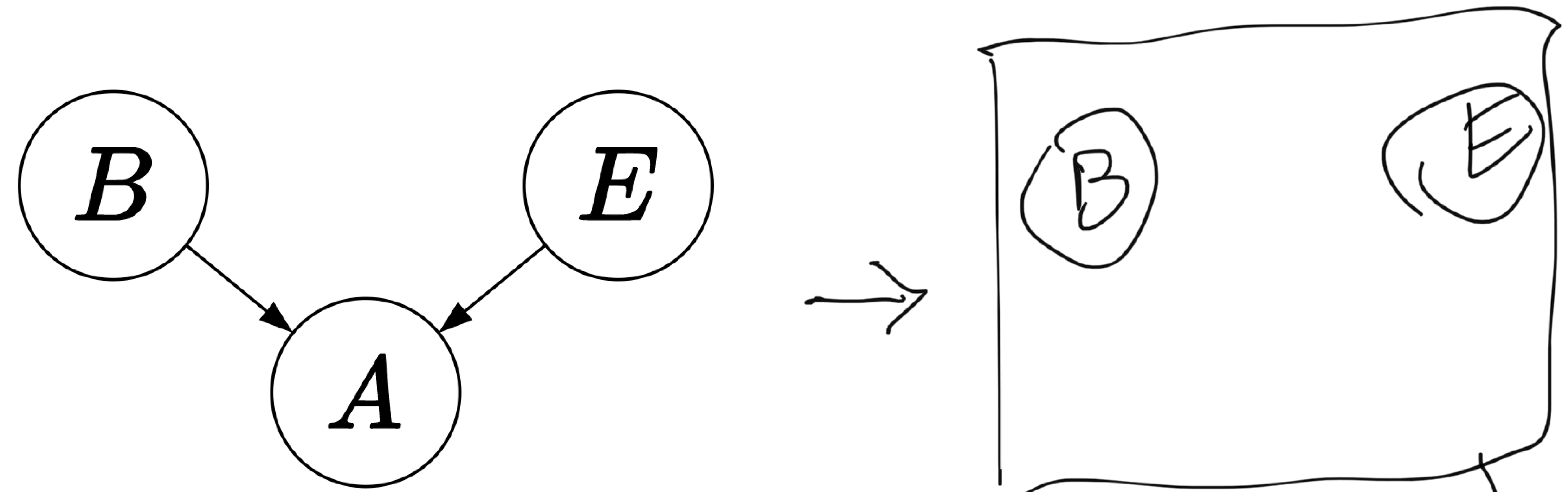
# Explaining away




**Key idea: explaining away**

Suppose two causes positively influence an effect. Conditioned on the effect, conditioning on one cause reduces the probability of the other cause.

# Consistency of sub-Bayesian networks




A short calculation:

$$\mathbb{P}(B = b, E = e) = \textstyle\sum_a \mathbb{P}(B = b, E = e, A = a)$$

$$\stackrel{\text{def}}{=} \textstyle\sum_a p(b)p(e)p(a \mid b, e)$$

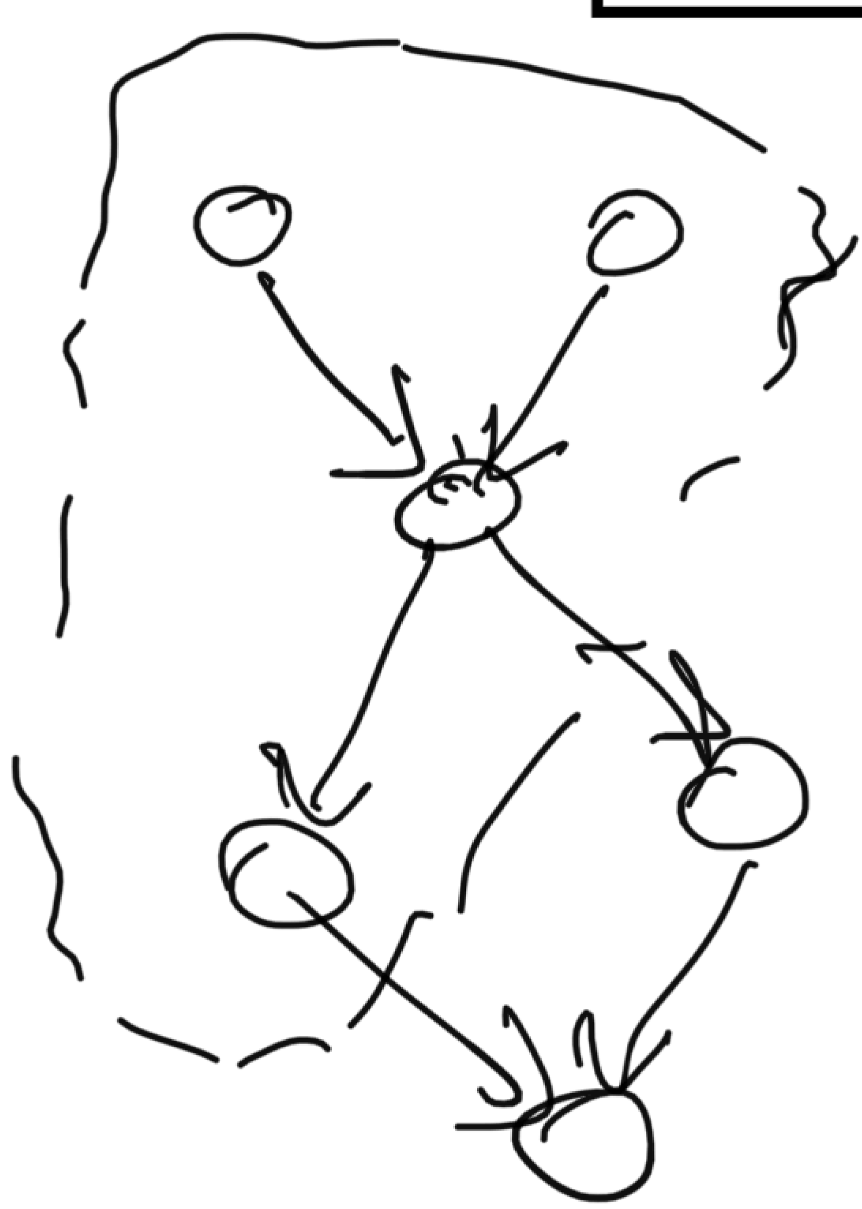
$$= p(b)p(e) \textstyle\sum_a p(a \mid b, e)$$

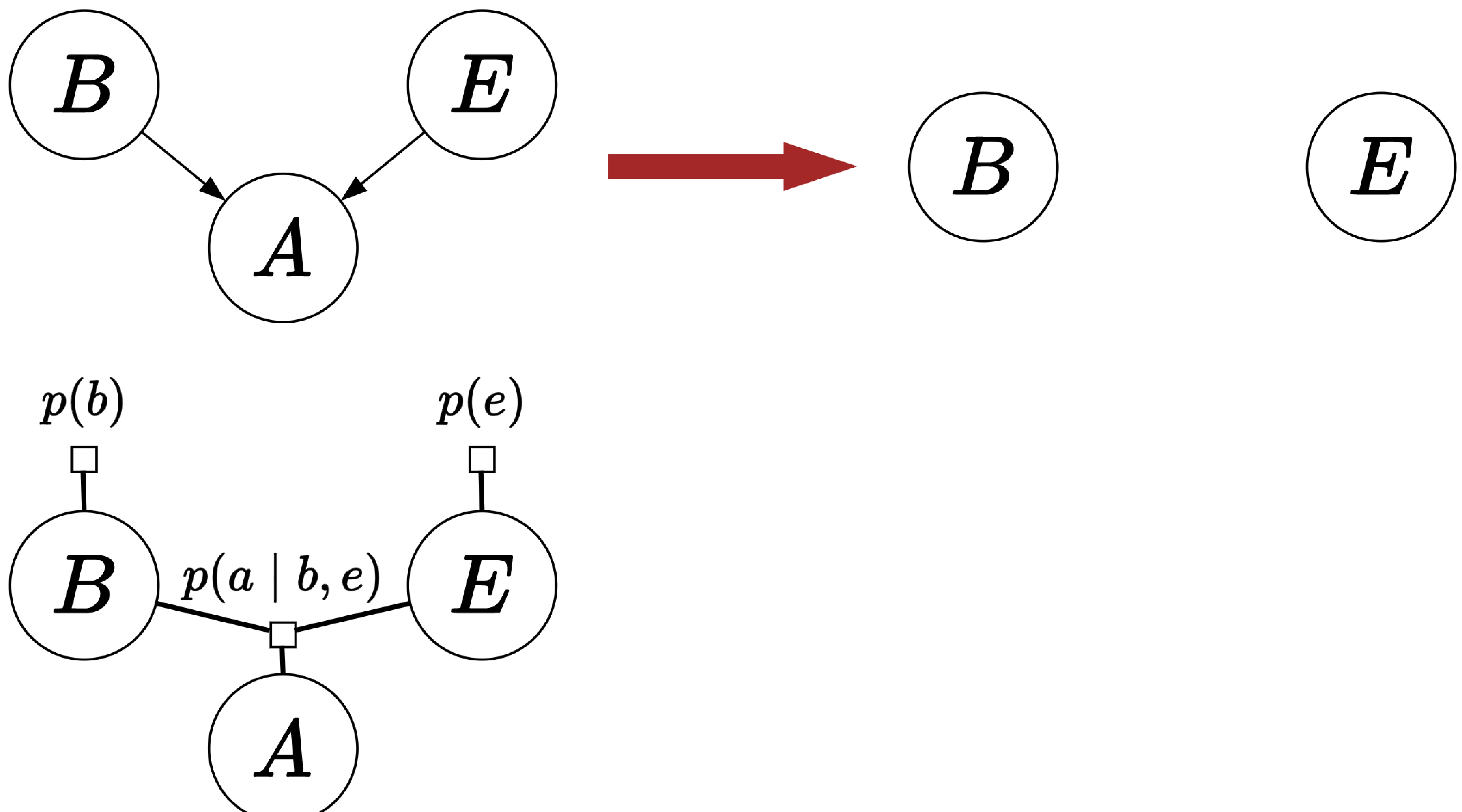
$$= p(b)p(e)$$

# Consistency of sub-Bayesian networks

**Key idea: marginalization**

Marginalization of a leaf node yields a Bayesian network without the node.

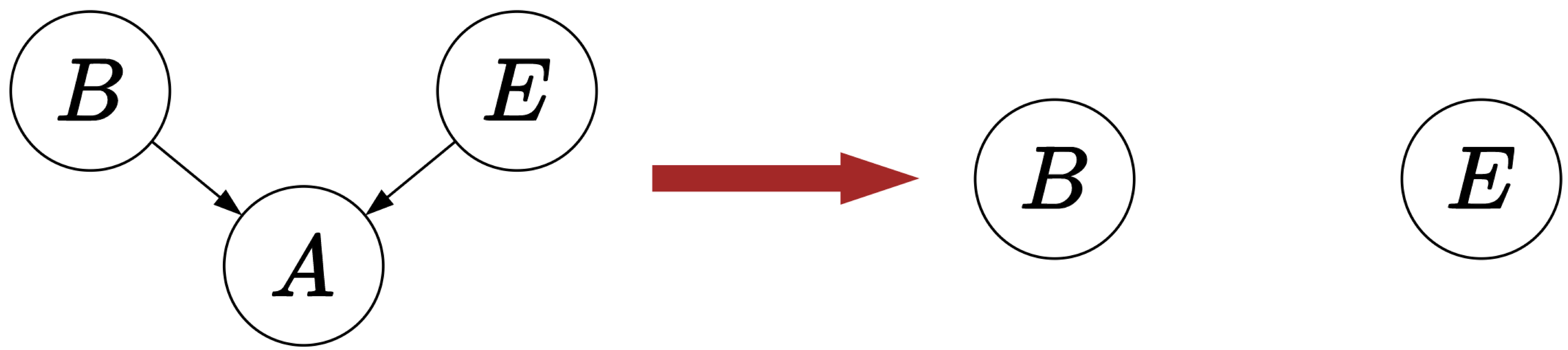


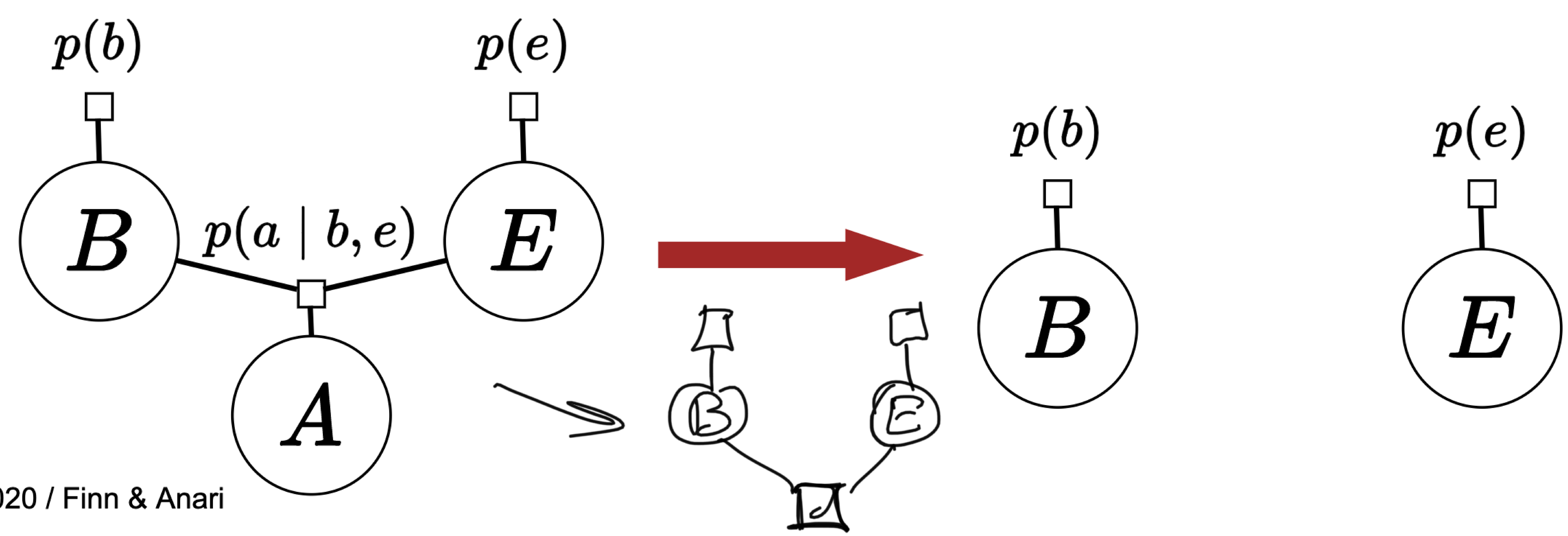


# Consistency of sub-Bayesian networks

**Key idea: marginalization**

Marginalization of a leaf node yields a Bayesian network without the node.

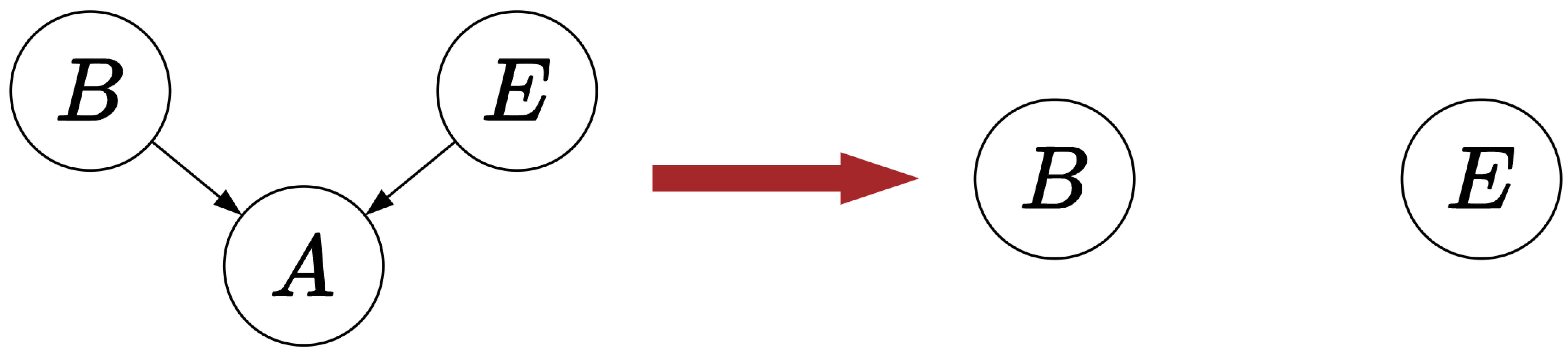


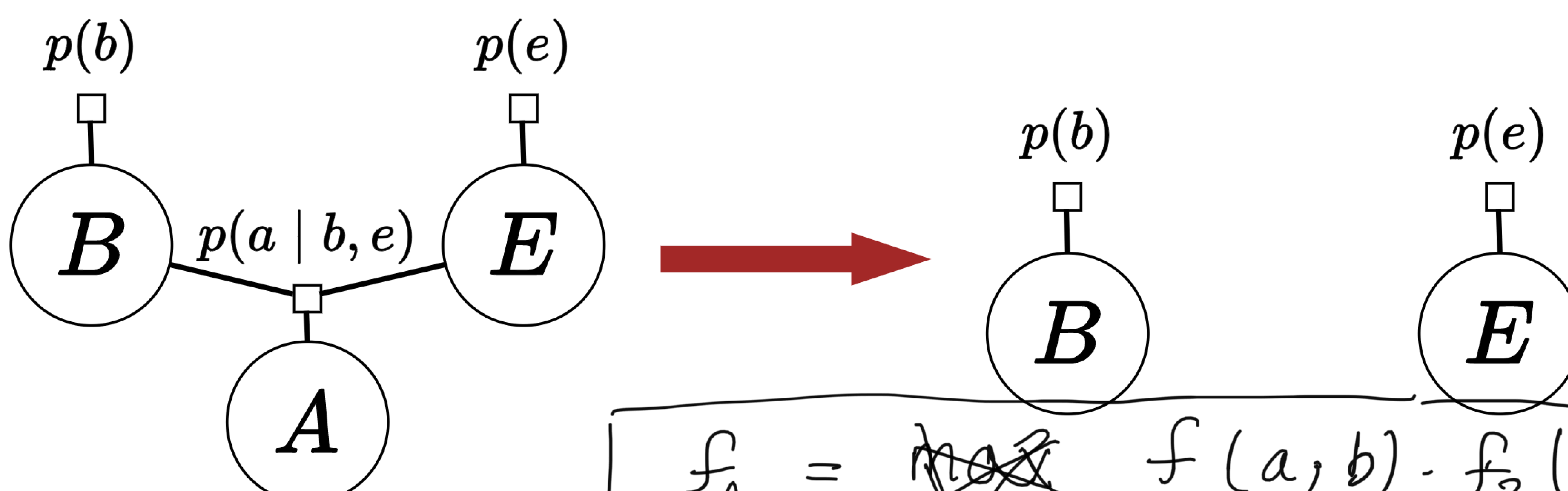



# Consistency of sub-Bayesian networks

**Key idea: marginalization**

Marginalization of a leaf node yields a Bayesian network without the node.




$$f_A = \cancel{\max_a} \sum_a f_1(a, b) \cdot f_2(a, c)$$

# Probabilistic program: example

**Probabilistic program: object tracking**

$$X_0 = (0, 0)$$

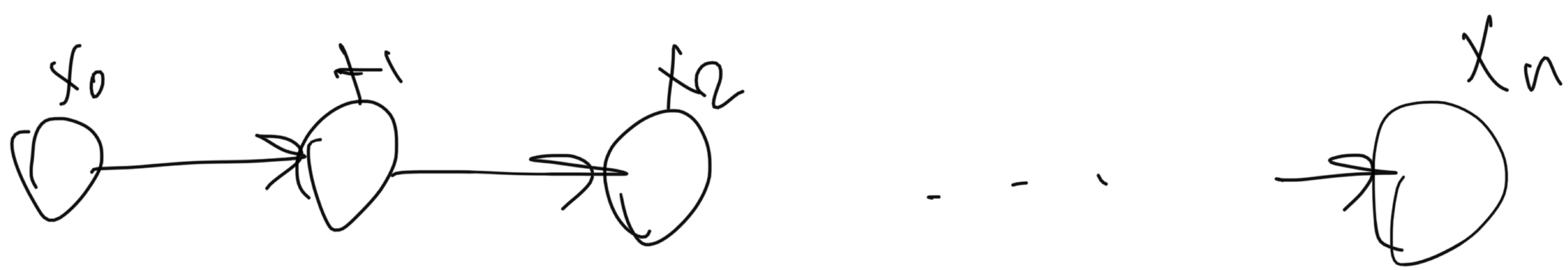
For each time step $i = 1, \ldots, n$:

With probability $\alpha$:

$X_i = X_{i-1} + (1, 0)$ [go right]

With probability $1 - \alpha$:

$X_i = X_{i-1} + (0, 1)$ [go down]



# Review: probabilistic inference

**Input**

Bayesian network: $\mathbb{P}(X_1, \dots, X_n)$

Evidence: $E = e$ where $E \subseteq X$ is subset of variables

Query: $Q \subseteq X$ is subset of variables

**Output**

$$\mathbb{P}(Q \mid E = e)$$

$\equiv$

$P(Q=q \mid E=e)$ for all $q$

# Example: Markov model

1. Eliminate all non-ancestors

2.

$$X_1 \to X_2 \to X_3 \to \cancel{X_4}$$

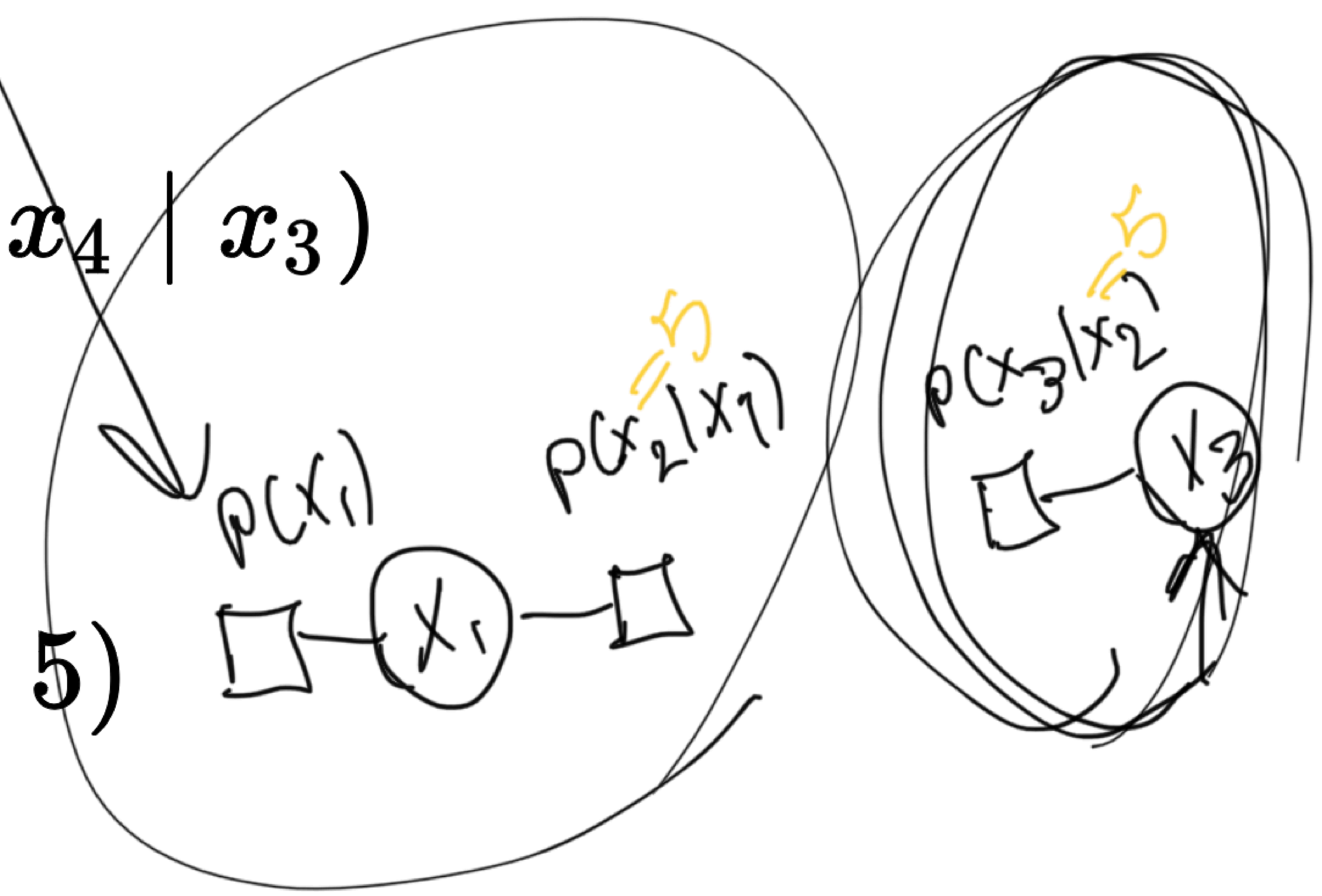
Query: $\mathbb{P}(\underline{X_3} = x_3 \mid \underline{X_2} = 5)$ for all $x_3$

Tedious way:

$$\propto \sum_{x_1, x_4} p(x_1) p(x_2 = 5 \mid x_1) p(x_3 \mid x_2 = 5) p(x_4 \mid x_3)$$

$$\propto \left( \sum_{x_1} p(x_1) p(x_2 = 5 \mid x_1) \right) p(x_3 \mid x_2 = 5)$$

$$\propto p(x_3 \mid x_2 = 5)$$




Fast way:

[whiteboard]

# General strategy

Query:

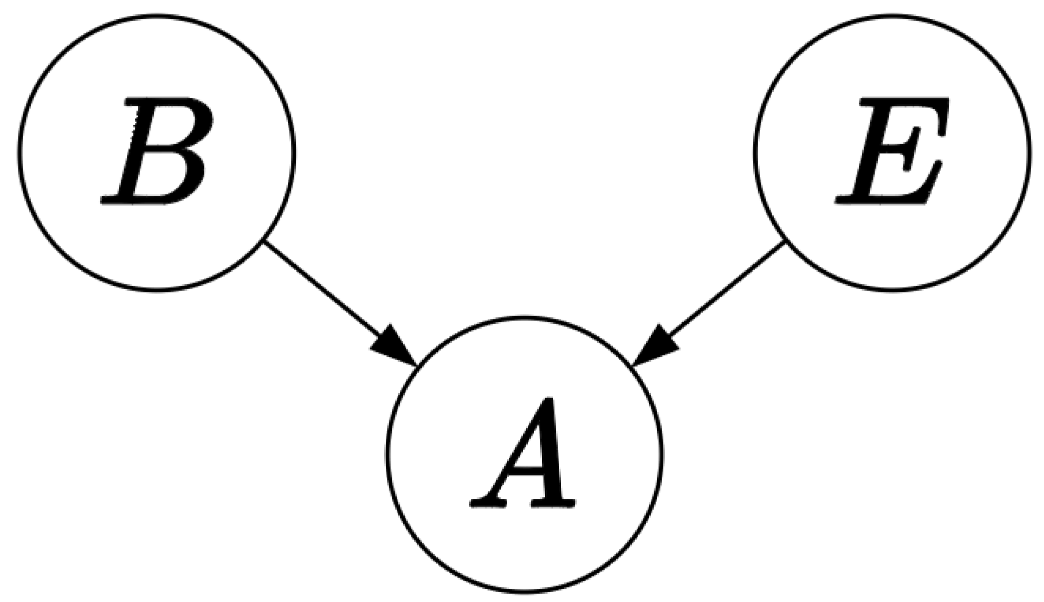
$$\mathbb{P}(Q \mid E = e)$$

**Algorithm: general probabilistic inference strategy**

- Remove (marginalize) variables that are not ancestors of $Q$ or $E$ (sub-Bayesian network).
- Convert Bayesian network to factor graph.
- Condition on $E = e$ (shade nodes + disconnect).
- Remove (marginalize) nodes disconnected from $Q$.


$$\sum_{a \in Domain(A)} f_1 \cdot f_2$$

# Example: alarm




| $b$ | $p(b)$ |
|---|---|
| 1 | $\epsilon$ |
| 0 | $1-\epsilon$ |

| $e$ | $p(e)$ |
|---|---|
| 1 | $\epsilon$ |
| 0 | $1-\epsilon$ |

| $b$ | $e$ | $a$ | $p(a \mid b, e)$ |
|---|---|---|---|
| 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 |

[whiteboard]


Query: $\mathbb{P}(B)$

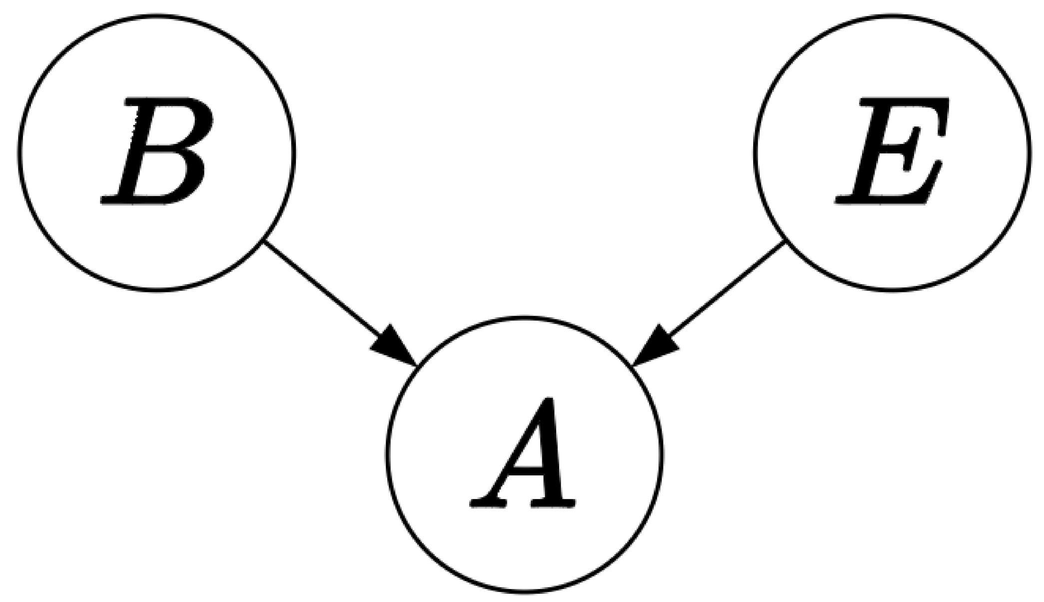
- Marginalize out $A, E$

Query: $\mathbb{P}(B \mid A = 1)$

- Condition on $A = 1$

# Example: alarm

$$p(b)\left(\sum_e p(e)\, p(a=1 \mid b,e)\right)$$

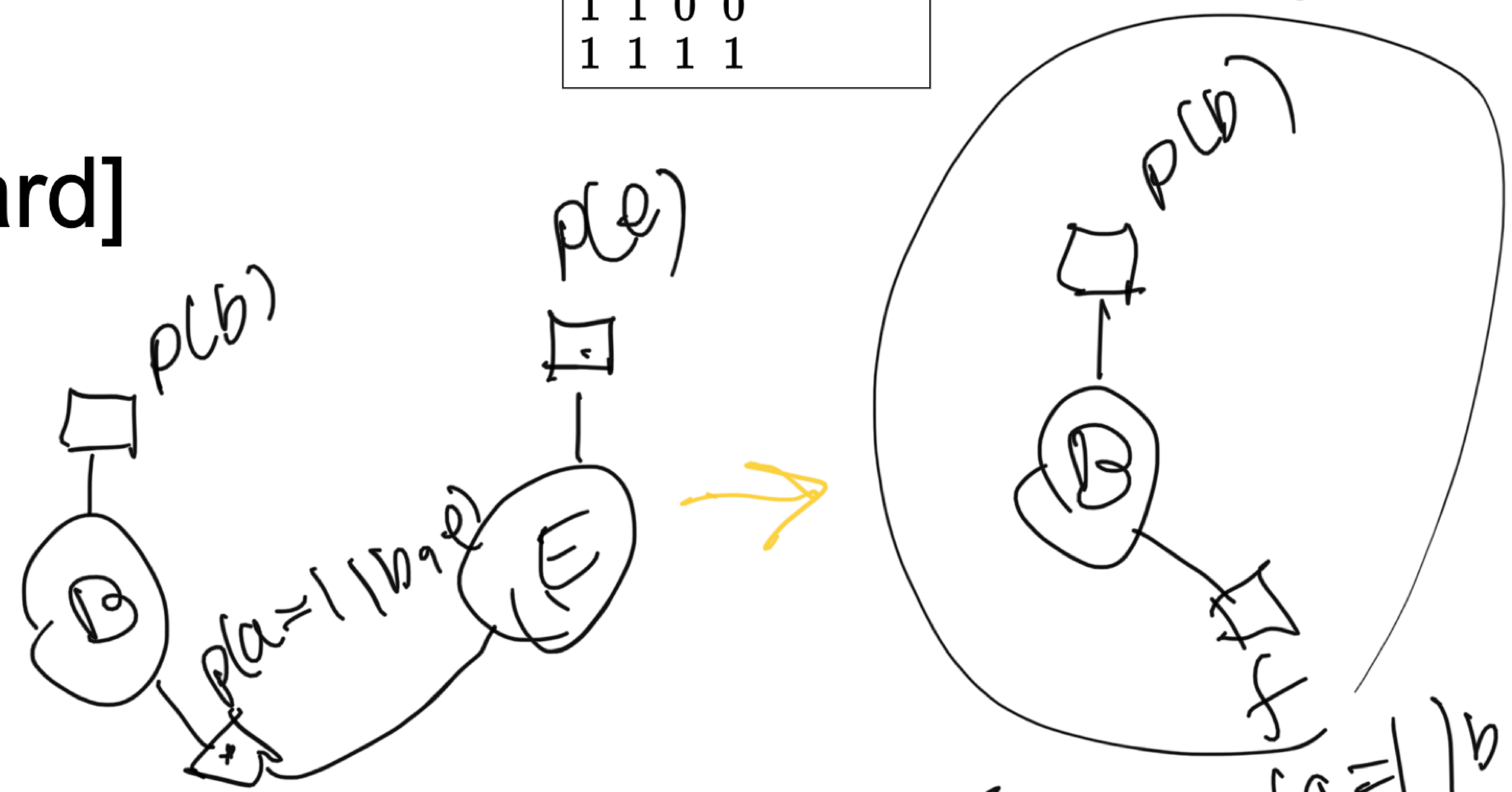



| $b$ | $p(b)$ |
|---|---|
| 1 | $\epsilon$ |
| 0 | $1-\epsilon$ |

| $e$ | $p(e)$ |
|---|---|
| 1 | $\epsilon$ |
| 0 | $1-\epsilon$ |

| $b$ | $e$ | $a$ | $p(a \mid b, e)$ |
|---|---|---|---|
| 0 | 0 | 0 | 1 |
| 0 | 0 | 1 | 0 |
| 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 1 |
| 1 | 0 | 0 | 0 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 0 |
| 1 | 1 | 1 | 1 |

[whiteboard]




Query: $\mathbb{P}(B)$

- Marginalize out $A, E$

Query: $\mathbb{P}(B \mid A = 1)$ ←

- Condition on $A = 1$

$$f(b) = \sum_e p(e) \cdot p(a=1 \mid b,e)$$