

Fundamentals of Artificial Intelligence

Chapter 13: Quantifying Uncertainty

Roberto Sebastiani

DISI, Università di Trento, Italy – roberto.sebastiani@unitn.it

https://disi.unitn.it/rseba/DIDATTICA/fai_2025/

Teaching assistant:

Paolo Morettin, paolo.morettin@unitn.it, <https://paolomorettin.github.io/>

M.S. Course “Artificial Intelligence Systems”, academic year 2024-2025

Last update: Friday 12th September, 2025, 18:35

Copyright notice: Most examples and images displayed in the slides of this course are taken from [\[Russell & Norvig, “Artificial Intelligence, a Modern Approach”, 3rd ed., Pearson\]](#), including explicitly figures from the above-mentioned book, so that their copyright is detained by the authors. A few other material (text, figures, examples) is authored by (in alphabetical order): [Pieter Abbeel](#), [Bonnie J. Dorr](#), [Anca Dragan](#), [Dan Klein](#), [Nikita Kitaev](#), [Tom Lenaerts](#), [Michela Milano](#), [Dana Nau](#), [Maria Simi](#), who detain its copyright.

These slides cannot be displayed in public without the permission of the author.

Outline

- 1 Acting Under Uncertainty
- 2 Basics on Probability
- 3 Probabilistic Inference via Enumeration
- 4 Independence and Conditional Independence
- 5 Applying Bayes' Rule
- 6 An Example: The Wumpus World Revisited

Outline

- 1 Acting Under Uncertainty
- 2 Basics on Probability
- 3 Probabilistic Inference via Enumeration
- 4 Independence and Conditional Independence
- 5 Applying Bayes' Rule
- 6 An Example: The Wumpus World Revisited

Acting Under Uncertainty

- Agents often make decisions based on incomplete information
 - partial observability
 - nondeterministic actions
- Partial solution (see previous chapters): maintain **belief states**
 - represent the set of all **possible world states** the agent might be in
 - generating a **contingency plan** handling every possible eventuality
- Several drawbacks:
 - must consider every possible explanation for the observation (even very-unlikely ones)
⇒ impossibly complex belief-states
 - contingent plans handling every eventuality grow arbitrarily large
 - sometimes there is no plan that is **guaranteed** to achieve the goal
- Agent's knowledge cannot guarantee a successful outcome ...
... but can provide some **degree of belief (likelihood)** on it
- **A rational decision depends on both the relative importance of (sub)goals and the likelihood that they will be achieved**
- **Probability theory** offers a clean way to quantify likelihood

Acting Under Uncertainty: Example

Automated taxi to Airport

- Goal: deliver a passenger to the airport on time
- Action A_t : leave for airport t minutes before flight
 - How can we be sure that A_{90} will succeed?
- Too many sources of uncertainty:
 - partial observability (ex: road state, other drivers' plans, etc.)
 - uncertainty in action outcome (ex: flat tire, etc.)
 - noisy sensors (ex: unreliable traffic reports)
 - complexity of modelling and predicting traffic

⇒ With purely-logical approach it is difficult to anticipate everything that can go wrong

- risks falsehood: " A_{25} will get me there on time" or
- leads to conclusions that are too weak for decision making:
" A_{25} will get me there on time if there's no accident on the bridge , and it doesn't rain and my tires remain intact , and..."
- Over-cautious choices are not rational solutions either
 - ex: A_{1440} causes staying overnight at the airport

Acting Under Uncertainty: Example (2)

A medical diagnosis

- Given the symptoms (toothache) infer the cause (cavity)
- How to encode this relation in logic?
 - diagnostic rules:
 - $Toothache \rightarrow Cavity$ (wrong)
 - $Toothache \rightarrow (Cavity \vee GumProblem \vee Abscess \vee \dots)$
(too many possible causes, some very unlikely)
 - causal rules:
 - $Cavity \rightarrow Toothache$ (wrong)
 - $(Cavity \wedge \dots) \rightarrow Toothache$ (many possible (con)causes)
- Problems in specifying the correct logical rules:
 - **Complexity**: too many possible antecedents or consequents
 - **Theoretical ignorance**: no complete theory for the domain
 - **Practical ignorance**: no complete knowledge of the patient

Summarizing Uncertainty

- Probability allows to summarize the uncertainty on effects of
 - laziness: failure to enumerate exceptions, qualifications, etc.
 - ignorance: lack of relevant facts, initial conditions, etc.
- Probability can be derived from
 - statistical data (ex: 80% of toothache patients so far had cavities)
 - some knowledge (ex: 80% of toothache patients has cavities)
 - their combination thereof
- Probability statements are made with respect to a state of knowledge (aka evidence), not with respect to the real world
 - e.g., “The probability that the patient has a cavity, given that she has a toothache, is 0.8”:
 $P(\text{HasCavity}(\text{patient}) \mid \text{hasToothAche}(\text{patient})) = 0.8$
- Probabilities of propositions change with new evidence:
 - “The probability that the patient has a cavity, given that she has a toothache and a history of gum disease, is 0.4”:
 $P(\text{HasCavity}(\text{patient}) \mid \text{hasToothAche}(\text{patient}) \wedge \text{HistoryOfGum}(\text{patient})) = 0.4$

Making Decisions Under Uncertainty

- Ex: Suppose I believe:

$P(A_{25} \text{ gets me there on time} \mid \dots) = 0.04$

$P(A_{90} \text{ gets me there on time} \mid \dots) = 0.70$

$P(A_{120} \text{ gets me there on time} \mid \dots) = 0.95$

$P(A_{1440} \text{ gets me there on time} \mid \dots) = 0.9999$

Which action to choose?

⇒ Depends on tradeoffs among preferences:

- missing flight vs. costs (airport cuisine, sleep overnight in airport)
- When there are conflicting goals the agent may express preferences among them by means of a **utility function**.
- Utilities are combined with probabilities in the **general theory of rational decisions**, aka **decision theory**:
Decision theory = Probability theory + Utility theory
- **Maximum Expected Utility (MEU)**: an agent is rational if and only if it chooses the action that yields the maximum expected utility, averaged over all the possible outcomes of the action.

Outline

- 1 Acting Under Uncertainty
- 2 Basics on Probability**
- 3 Probabilistic Inference via Enumeration
- 4 Independence and Conditional Independence
- 5 Applying Bayes' Rule
- 6 An Example: The Wumpus World Revisited

Probabilities Basics: an AI-sh Introduction

- **Probabilistic assertions**: state how likely possible worlds are
- **Sample space Ω** : the set of all possible worlds
 - $\omega \in \Omega$ is a **possible world** (aka **sample point** or **atomic event**)
 - ex: **the dice roll** (1,4)
 - the possible worlds are **mutually exclusive** and **exhaustive**
 - ex: **the 36 possible outcomes** of rolling two dice: (1,1), (1,2), ...
- **A probability model (aka probability space)** is a sample space with an assignment $P(\omega)$ for every $\omega \in \Omega$ s.t.
 - $0 \leq P(\omega) \leq 1$, for every $\omega \in \Omega$
 - $\sum_{\omega \in \Omega} P(\omega) = 1$
- Ex: 1-die roll: $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$
- An **Event A** is any subset of Ω , s.t. $P(A) = \sum_{\omega \in A} P(\omega)$
 - events can be described by **propositions** in some formal language
 - ex: $P(\text{Total} = 11) = P(5, 6) + P(6, 5) = 1/36 + 1/36 = 1/18$
 - ex: $P(\text{doubles}) = P(1, 1) + P(2, 2) + \dots + P(6, 6) = 6/36 = 1/6$

Random Variables

- Factored representation of possible worlds: sets of $\langle \text{variable}, \text{value} \rangle$ pairs
- Variables in probability theory: **Random variables**
 - **domain**: the set of possible values a variable can take on
ex: **Die**: $\{1, 2, 3, 4, 5, 6\}$, **Weather**: $\{\text{sunny}, \text{rain}, \text{cloudy}, \text{snow}\}$, **Odd**: $\{\text{true}, \text{false}\}$
 - a r.v. can be seen as a function from sample points to the domain:
ex: $\text{Die}(\omega)$, $\text{Weather}(\omega)$,... (" ω ") typically omitted)
- **Probability Distribution** gives the probabilities of all the possible values of a random variable
 X : $P(X = x_i) \stackrel{\text{def}}{=} \sum_{\omega \in X(\omega)} P(\omega)$
 - ex: $P(\text{Odd} = \text{true}) = P(1) + P(3) + P(5) = 1/6 + 1/6 + 1/6 = 1/2$

Propositions and Probabilities

- We think a proposition a as the event A (set of sample points) where the proposition is true
 - odd is a propositional random variable of range $\{true, false\}$
 - notation: $a \iff "A = true"$ (e.g., $odd \iff "Odd = true"$)
 - Given Boolean random variables A and B :
 - a : set of sample points where $A(\omega) = true$
 - $\neg a$: set of sample points where $A(\omega) = false$
 - $a \wedge b$: set of sample points where $A(\omega) = true, B(\omega) = true$
- $\implies$ with Boolean random variables, sample points are PL models
- Proposition: disjunction of the sample points in which it is true
 - ex: $(a \vee b) \equiv (\neg a \wedge b) \vee (a \wedge \neg b) \vee (a \wedge b)$
- $\implies P(a \vee b) = P(\neg a \wedge b) + P(a \wedge \neg b) + P(a \wedge b)$
- Some derived facts:
 - $P(\neg a) = 1 - P(a)$
 - $P(a \vee b) = P(a) + P(b) - P(a \wedge b)$

Probability Distributions

- **Probability Distribution** gives the probabilities of all the possible values of a random variable

- ex: **Weather**: {*sunny*, *rain*, *cloudy*, *snow*}

$$\Rightarrow \mathbf{P}(\text{Weather}) = (0.6, 0.1, 0.29, 0.01) \iff \left\{ \begin{array}{ll} P(\text{Weather} = \text{sunny}) & = 0.6 \\ P(\text{Weather} = \text{rain}) & = 0.1 \\ P(\text{Weather} = \text{cloudy}) & = 0.29 \\ P(\text{Weather} = \text{snow}) & = 0.01 \end{array} \right\}$$

- normalized: their sum is 1

- **Joint Probability Distribution** for multiple variables

- gives the probability of every sample point

● ex: $\mathbf{P}(\text{Weather}, \text{Cavity}) =$

<i>Weather</i> =	<i>sunny</i>	<i>rain</i>	<i>cloudy</i>	<i>snow</i>
<i>Cavity</i> = <i>true</i>	0.144	0.02	0.016	0.02
<i>Cavity</i> = <i>false</i>	0.576	0.08	0.064	0.08

- Every event is a sum of sample points,

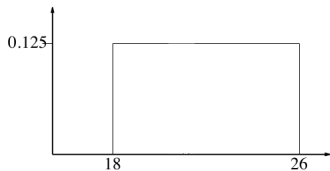
$\Rightarrow$ its probability is determined by the joint distribution

Probability for Continuous Variables

- Express continuous probability distributions:
 - density functions $f(x) \in [0, 1]$ s.t. $\int_{-\infty}^{+\infty} f(x)dx = 1$
- $P(x \in [a, b]) = \int_a^b f(x) dx$
 $\Rightarrow P(x \in [val, val]) = 0, P(x \in [-\infty, +\infty]) = 1$
 - ex: $P(x \in [20, 22]) = \int_{20}^{22} 0.125 dx = 0.25$
- Density: $P(x) = P(X = x) \stackrel{\text{def}}{=} \lim_{dx \rightarrow 0} P(X \in [x, x + dx])/dx$
 - ex: $P(20.1) = \lim_{dx \rightarrow 0} P(X \in [20.1, 20.1 + dx])/dx = 0.125$
 - note: $P(v) \neq P(x \in [v, v]) = 0$

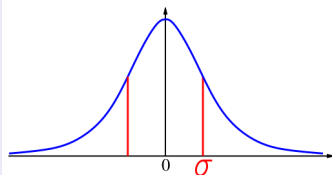
Uniform density between 18 and 26

$$f(x) = U[18, 26](x)$$



Gaussian density

$$P(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2}$$



Conditional Probabilities

- **Unconditional** or **prior probabilities** refer to degrees of belief in propositions **in the absence of any other information (evidence)**
 - ex: $P(\text{cavity}) = 0.2$, $P(\text{Total} = 11) = 1/18$, $P(\text{double}) = 1/6$
- **Conditional** or **posterior probabilities** refer to degrees of belief in proposition **a given some evidence b**: $P(a|b)$
 - **evidence**: information already revealed
 - ex: $P(\text{cavity}|\text{toothache}) = 0.6$: p. of a cavity given a toothache (assuming no other information is provided!)
 - ex: $P(\text{Total} = 11 | \text{die}_1 = 5) = 1/6$: p. of total 11 given first die is 5

⇒ restricts the set of possible worlds to those where the first die is 5
- Note: $P(a|\dots \wedge a) = 1$, $P(a|\dots \wedge \neg a) = 0$
 - ex: $P(\text{cavity}|\text{toothache} \wedge \text{cavity}) = 1$, $P(\text{cavity}|\text{toothache} \wedge \neg \text{cavity}) = 0$
- Less specific belief still valid after more evidence arrives
 - ex: $P(\text{cavity}) = 0.2$ holds even if $P(\text{cavity}|\text{toothache}) = 0.6$
- New evidence may be irrelevant, allowing for simplification
 - ex: $P(\text{cavity}|\text{toothache}, 49\text{ersWin}) = P(\text{cavity}|\text{toothache}) = 0.8$

Conditional Probabilities [cont.]

- **Conditional probability:** $P(a|b) \stackrel{\text{def}}{=} \frac{P(a \wedge b)}{P(b)}$, s.t. $P(b) > 0$
 - ex: $P(\text{Total} = 11 | \text{die}_1 = 5) = \frac{P(\text{Total}=11 \wedge \text{die}_1=5)}{P(\text{die}_1=5)} = \frac{1/6 \cdot 1/6}{1/6} = 1/6$
 - observing b restricts the possible worlds to those where b is true
- **Production rule:** $P(a \wedge b) = P(a|b) \cdot P(b) = P(b|a) \cdot P(a)$
- **Production rule for whole distributions:** $\mathbf{P}(X, Y) = \mathbf{P}(X|Y) \cdot \mathbf{P}(Y)$
 - ex: $\mathbf{P}(\text{Weather}, \text{Cavity}) = \mathbf{P}(\text{Weather}|\text{Cavity})\mathbf{P}(\text{Cavity})$, that is:
 $P(\text{sunny}, \text{cavity}) = P(\text{sunny}|\text{cavity})P(\text{cavity})$
...
 $P(\text{snow}, \neg\text{cavity}) = P(\text{snow}|\neg\text{cavity})P(\neg\text{cavity})$
 - a 4×2 set of equations, not matrix multiplication!
- **Chain rule** is derived by successive application of product rule:
$$\begin{aligned} & \mathbf{P}(X_1, \dots, X_n) \\ &= \mathbf{P}(X_1, \dots, X_{n-1})\mathbf{P}(X_n|X_1, \dots, X_{n-1}) \\ &= \mathbf{P}(X_1, \dots, X_{n-2})\mathbf{P}(X_{n-1}|X_1, \dots, X_{n-2})\mathbf{P}(X_n|X_1, \dots, X_{n-1}) \\ &= \dots \\ &= \mathbf{P}(X_1)\mathbf{P}(X_2|X_1)\dots\mathbf{P}(X_n|\dots X_1, \dots, X_{n-1}) \\ &= \prod_{i=1}^n \mathbf{P}(X_i|X_1, \dots, X_{i-1}) \end{aligned}$$

Logic vs. Probability

<i>Logic</i>	<i>Probability</i>
a	$P(a) = 1$
$\neg a$	$P(a) = 0$
$a \rightarrow b$	$P(b a) = 1$
$(a, a \rightarrow b)$	$P(a) = 1, P(b a) = 1$
b	$P(b) = 1$
$(\neg b, a \rightarrow b)$	$P(b) = 0, P(b a) = 1$
$\neg a$	$P(a) = 0$
$(a \rightarrow b, b \rightarrow c)$	$P(b a) = 1, P(c b) = 1$
$a \rightarrow c$	$P(c a) = 1$

(Courtesy of Maria Simi, UniPI)

Outline

- 1 Acting Under Uncertainty
- 2 Basics on Probability
- 3 Probabilistic Inference via Enumeration**
- 4 Independence and Conditional Independence
- 5 Applying Bayes' Rule
- 6 An Example: The Wumpus World Revisited

Probabilistic Inference via Enumeration

Basic Ideas

- Start with the joint distribution
- For any proposition φ , sum the atomic events where φ is true: $P(\varphi) = \sum_{\omega : \omega \models \varphi} P(\omega)$

Probabilistic Inference via Enumeration: Example

Example: Generic Inference

- Start with the joint distribution $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
- For any proposition φ , sum the atomic events where φ is true: $P(\varphi) = \sum_{\omega : \omega \models \varphi} P(\omega)$:
- Ex: $P(\textit{cavity} \vee \textit{toothache}) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

(© S. Russell & P. Norwig, AIMA)

Probabilistic Inference via Enumeration: Example

Example: Generic Inference

- Start with the joint distribution $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
- For any proposition φ , sum the atomic events where φ is true: $P(\varphi) = \sum_{\omega : \omega \models \varphi} P(\omega)$:
- Ex: $P(\textit{cavity} \vee \textit{toothache}) = 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

(© S. Russell & P. Norwig, AIMA)

Marginalization

- Start with the joint distribution (e.g., $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$)
- **Marginalization** (aka **summing out**):
sum up the probabilities for each possible value of the other variables:

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}, \mathbf{z})$$

$$\text{Ex: } \mathbf{P}(\textit{Toothache}) = \sum_{\mathbf{z} \in \{\textit{Catch}, \textit{Cavity}\}} \mathbf{P}(\textit{Toothache}, \mathbf{z})$$

- **Conditioning**: variant of marginalization, involving conditional probabilities instead of joint probabilities (using the product rule)

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}|\mathbf{z})P(\mathbf{z})$$

$$\text{Ex: } \mathbf{P}(\textit{Toothache}) = \sum_{\mathbf{z} \in \{\textit{Catch}, \textit{Cavity}\}} \mathbf{P}(\textit{Toothache}|\mathbf{z})P(\mathbf{z})$$

Marginalization: Example

- Start with the joint distribution $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
- Marginalization** (aka **summing out**):
sum up the probabilities for each possible value of the other variables:

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}, \mathbf{z})$$

$$\text{Ex: } \mathbf{P}(\textit{Toothache}) = \sum_{\mathbf{z} \in \{\textit{Catch}, \textit{Cavity}\}} \mathbf{P}(\textit{Toothache}, \mathbf{z})$$

$$P(\textit{toothache}) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$$

$$P(\neg \textit{toothache}) = 1 - P(\textit{toothache}) = 1 - 0.2 = 0.8$$

$$\Rightarrow \mathbf{P}(\textit{Toothache}) = \langle 0.2, 0.8 \rangle$$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

Marginalization: Example

- Start with the joint distribution $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
- Marginalization** (aka **summing out**):
sum up the probabilities for each possible value of the other variables:

$$\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}, \mathbf{z})$$

$$\text{Ex: } \mathbf{P}(\textit{Toothache}) = \sum_{\mathbf{z} \in \{\textit{Catch}, \textit{Cavity}\}} \mathbf{P}(\textit{Toothache}, \mathbf{z})$$

$$P(\textit{toothache}) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$$

$$P(\neg \textit{toothache}) = 1 - P(\textit{toothache}) = 1 - 0.2 = 0.8$$

$$\Rightarrow \mathbf{P}(\textit{Toothache}) = \langle 0.2, 0.8 \rangle$$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

Conditional Probability via Enumeration: Example

- Start with the joint distribution $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
- Conditional Probability:

$$\text{Ex: } P(\neg \textit{cavity} | \textit{toothache}) = \frac{P(\neg \textit{cavity} \wedge \textit{toothache})}{P(\textit{toothache})}$$

$$= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

$$\text{Ex: } P(\textit{cavity} | \textit{toothache}) = \frac{P(\textit{cavity} \wedge \textit{toothache})}{P(\textit{toothache})} = \dots = 0.6$$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

Conditional Probability via Enumeration: Example

- Start with the joint distribution $P(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
- Conditional Probability:

$$\text{Ex: } P(\neg \textit{cavity} | \textit{toothache}) = \frac{P(\neg \textit{cavity} \wedge \textit{toothache})}{P(\textit{toothache})}$$

$$= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

$$\text{Ex: } P(\textit{cavity} | \textit{toothache}) = \frac{P(\textit{cavity} \wedge \textit{toothache})}{P(\textit{toothache})} = \dots = 0.6$$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

(© S. Russell & P. Norwig, AIMA)

Conditional Probability via Enumeration: Example

- Start with the joint distribution $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
- Conditional Probability:

$$\text{Ex: } P(\neg \textit{cavity} | \textit{toothache}) = \frac{P(\neg \textit{cavity} \wedge \textit{toothache})}{P(\textit{toothache})}$$

$$= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

$$\text{Ex: } P(\textit{cavity} | \textit{toothache}) = \frac{P(\textit{cavity} \wedge \textit{toothache})}{P(\textit{toothache})} = \dots = 0.6$$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

Normalization

- Let $\mathbf{X}$ be all the variables. Typically, we want $\mathbf{P}(\mathbf{Y}|\mathbf{E} = \mathbf{e})$:
 - the conditional joint distribution of the query variables $\mathbf{Y}$
 - given specific values $\mathbf{e}$ for the evidence variables $\mathbf{E}$
 - let the hidden variables be $\mathbf{H} \stackrel{\text{def}}{=} \mathbf{X} \setminus (\mathbf{Y} \cup \mathbf{E})$
- The summation of joint entries is done by summing out the hidden variables:
$$\mathbf{P}(\mathbf{Y}|\mathbf{E} = \mathbf{e}) = \alpha \mathbf{P}(\mathbf{Y}, \mathbf{E} = \mathbf{e}) = \alpha \sum_{\mathbf{h} \in \mathbf{H}} \mathbf{P}(\mathbf{Y}, \mathbf{E} = \mathbf{e}, \mathbf{H} = \mathbf{h})$$

where $\alpha \stackrel{\text{def}}{=} 1/\mathbf{P}(\mathbf{E} = \mathbf{e})$ (Notice: different α 's for different values of $\mathbf{e}$!)

⇒ it is easy to compute α by normalization

 - note: the terms in the summation are joint entries, because $\mathbf{Y}, \mathbf{E}, \mathbf{H}$ together exhaust the set of random variables $\mathbf{X}$
- Idea: compute whole distribution on query variable by:
 - fixing evidence variables and summing over hidden variables
 - normalize the final distribution, so that $\sum \dots = 1$
- Complexity: $O(2^n)$, n number of propositions ⇒ impractical for large n 's

Common practice: deal with non-normalized distributions, normalize at the end of the process (see e.g. “Wumpus world” example at the end of this chapter)

Normalization: Example

- $\alpha \stackrel{\text{def}}{=} 1/P(\textit{toothache})$ (previous example) can be viewed as a normalization constant
- Idea: compute **whole distribution** on **query variable** by:
 - fixing **evidence variables** and summing over **hidden variables**
 - **normalize** the final distribution, so that $\sum \dots = 1$
- Ex:^a $\mathbf{P}(\textit{Cavity}|\textit{toothache}) = \alpha \mathbf{P}(\textit{Cavity}, \textit{toothache})$
 $= \alpha [\mathbf{P}(\textit{Cavity}, \textit{toothache}, \textit{catch}) + \mathbf{P}(\textit{Cavity}, \textit{toothache}, \neg \textit{catch})]$
 $= \alpha [\langle 0.108, 0.016 \rangle + \langle 0.012, 0.064 \rangle]$
 $= \alpha \langle 0.12, 0.08 \rangle = (\textit{normalization}) = \langle 0.6, 0.4 \rangle [\alpha = 5]$
 $\mathbf{P}(\textit{Cavity}|\neg \textit{toothache}) = \dots = \alpha \langle 0.08, 0.72 \rangle = \langle 0.1, 0.9 \rangle [\alpha = 1.25]$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

(© S. Russell & P. Norwig, AIMA)

^an.b.: here “Cavity” is a variable, “toothache” is a proposition (i.e. Toothache=true)

Normalization: Example

- $\alpha \stackrel{\text{def}}{=} 1/P(\textit{toothache})$ (previous example) can be viewed as a normalization constant
- Idea: compute **whole distribution** on **query variable** by:
 - fixing **evidence variables** and summing over **hidden variables**
 - **normalize** the final distribution, so that $\sum \dots = 1$
- Ex:^a $\mathbf{P}(\textit{Cavity}|\textit{toothache}) = \alpha \mathbf{P}(\textit{Cavity}, \textit{toothache})$
 $= \alpha [\mathbf{P}(\textit{Cavity}, \textit{toothache}, \textit{catch}) + \mathbf{P}(\textit{Cavity}, \textit{toothache}, \neg \textit{catch})]$
 $= \alpha [\langle 0.108, 0.016 \rangle + \langle 0.012, 0.064 \rangle]$
 $= \alpha \langle 0.12, 0.08 \rangle = (\textit{normalization}) = \langle 0.6, 0.4 \rangle [\alpha = 5]$
 $\mathbf{P}(\textit{Cavity}|\neg \textit{toothache}) = \dots = \alpha \langle 0.08, 0.72 \rangle = \langle 0.1, 0.9 \rangle [\alpha = 1.25]$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

(© S. Russell & P. Norwig, AIMA)

^an.b.: here “Cavity” is a variable, “toothache” is a proposition (i.e. Toothache=true)

Normalization: Example

- $\alpha \stackrel{\text{def}}{=} 1/P(\text{toothache})$ (previous example) can be viewed as a normalization constant
- Idea: compute **whole distribution** on **query variable** by:
 - fixing **evidence variables** and summing over **hidden variables**
 - **normalize** the final distribution, so that $\sum \dots = 1$
- Ex:^a $\mathbf{P}(\text{Cavity}|\text{toothache}) = \alpha \mathbf{P}(\text{Cavity}, \text{toothache})$
 $= \alpha [\mathbf{P}(\text{Cavity}, \text{toothache}, \text{catch}) + \mathbf{P}(\text{Cavity}, \text{toothache}, \neg \text{catch})]$
 $= \alpha [\langle 0.108, 0.016 \rangle + \langle 0.012, 0.064 \rangle]$
 $= \alpha \langle 0.12, 0.08 \rangle = (\text{normalization}) = \langle 0.6, 0.4 \rangle [\alpha = 5]$
 $\mathbf{P}(\text{Cavity}|\neg \text{toothache}) = \dots = \alpha \langle 0.08, 0.72 \rangle = \langle 0.1, 0.9 \rangle [\alpha = 1.25]$

	<i>toothache</i>		$\neg$ <i>toothache</i>	
	<i>catch</i>	$\neg$ <i>catch</i>	<i>catch</i>	$\neg$ <i>catch</i>
<i>cavity</i>	.108	.012	.072	.008
$\neg$ <i>cavity</i>	.016	.064	.144	.576

(© S. Russell & P. Norwig, AIMA)

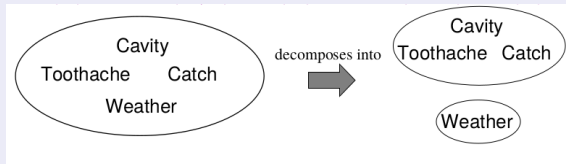
^an.b.: here “Cavity” is a variable, “toothache” is a proposition (i.e. Toothache=true)

Outline

- 1 Acting Under Uncertainty
- 2 Basics on Probability
- 3 Probabilistic Inference via Enumeration
- 4 Independence and Conditional Independence**
- 5 Applying Bayes' Rule
- 6 An Example: The Wumpus World Revisited

Independence

- Variables X and Y are **independent** iff $\mathbf{P}(X, Y) = \mathbf{P}(X)\mathbf{P}(Y)$
(equivalently, iff $\mathbf{P}(X|Y) = \mathbf{P}(X)$ and iff $\mathbf{P}(Y|X) = \mathbf{P}(Y)$)
 - ex: $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity}, \textit{Weather}) = \mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})\mathbf{P}(\textit{Weather})$
 $\Rightarrow$ e.g. $P(\textit{toothache}, \textit{catch}, \textit{cavity}, \textit{cloudy}) = P(\textit{toothache}, \textit{catch}, \textit{cavity})P(\textit{cloudy})$
 - ex: $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity}|\textit{Weather}) = \mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$
 $\Rightarrow$ e.g. $P(\textit{toothache}, \textit{catch}, \textit{cavity}|\textit{cloudy}) = P(\textit{toothache}, \textit{catch}, \textit{cavity})$
 - typically **based on domain knowledge**
- May drastically reduce the number of entries and computation
 $\Rightarrow$ ex: 32-element table decomposed into one 8-element and one 4-element table
- Unfortunately, absolute independence is quite rare



Conditional Independence

- Variables X and Y are **conditionally independent given Z** iff $\mathbf{P}(X, Y|Z) = \mathbf{P}(X|Z)\mathbf{P}(Y|Z)$ (equivalently, iff $\mathbf{P}(X|Y, Z) = \mathbf{P}(X|Z)$ and iff $\mathbf{P}(Y|X, Z) = \mathbf{P}(Y|Z)$)
 - Consider $\mathbf{P}(\textit{Toothache}, \textit{Cavity}, \textit{Catch})$
 - if I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache: $P(\textit{catch}|\textit{toothache}, \textit{cavity}) = P(\textit{catch}|\textit{cavity})$
 - the same independence holds if I haven't got a cavity:
 $P(\textit{catch}|\textit{toothache}, \neg\textit{cavity}) = P(\textit{catch}|\neg\textit{cavity})$
- ⇒ **Catch is conditionally independent of Toothache given Cavity:**
 $\mathbf{P}(\textit{Catch}|\textit{Toothache}, \textit{Cavity}) = \mathbf{P}(\textit{Catch}|\textit{Cavity})$
or, equivalently: $\mathbf{P}(\textit{Toothache}|\textit{Catch}, \textit{Cavity}) = \mathbf{P}(\textit{Toothache}|\textit{Cavity})$, or
 $\mathbf{P}(\textit{Toothache}, \textit{Catch}|\textit{Cavity}) = \mathbf{P}(\textit{Toothache}|\textit{Cavity})\mathbf{P}(\textit{Catch}|\textit{Cavity})$
- Hint: *Toothache* and *Catch* are two (mutually-independent) **effects** of the same **cause** *Cavity*

Conditional Independence [cont.]

- In many cases, the use of conditional independence reduces the size of the representation of the joint distribution dramatically

- even from exponential to linear!

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

- Ex:
$$= \mathbf{P}(\textit{Toothache} | \textit{Catch}, \textit{Cavity}) \mathbf{P}(\textit{Catch}, \textit{Cavity})$$
$$= \mathbf{P}(\textit{Toothache} | \textit{Catch}, \textit{Cavity}) \mathbf{P}(\textit{Catch} | \textit{Cavity}) \mathbf{P}(\textit{Cavity})$$
$$= \mathbf{P}(\textit{Toothache} | \textit{Cavity}) \mathbf{P}(\textit{Catch} | \textit{Cavity}) \mathbf{P}(\textit{Cavity})$$

⇒ Passes from 7 to $2+2+1=5$ independent numbers

- $\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$ contains 7 independent entries (the 8th can be obtained as $1 - \sum \dots$)
 - $\mathbf{P}(\textit{Toothache} | \textit{Cavity}), \mathbf{P}(\textit{Catch} | \textit{Cavity})$ contain 2 independent entries (2×2 matrix, each row sums to 1)
 - $\mathbf{P}(\textit{Cavity})$ contains 1 independent entry
- General Case: if one cause has n independent effects:
$$\mathbf{P}(\textit{Cause}, \textit{Effect}_1, \dots, \textit{Effect}_n) = \mathbf{P}(\textit{Cause}) \prod_i \mathbf{P}(\textit{Effect}_i | \textit{Cause})$$

⇒ reduces from $2^{n+1} - 1$ to $2n + 1$ independent entries

Exercise

Consider the joint probability distribution described in the table in previous section:

$P(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$

- Consider the example in previous slide:

$P(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$

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

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

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

- Compute separately the distributions $P(\textit{Toothache}|\textit{Catch}, \textit{Cavity})$, $P(\textit{Catch}|\textit{Cavity})$, $P(\textit{Cavity})$, $P(\textit{Toothache}|\textit{Cavity})$.

- Recompute $P(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$ in two ways:

- $P(\textit{Toothache}|\textit{Catch}, \textit{Cavity})P(\textit{Catch}|\textit{Cavity})P(\textit{Cavity})$

- $P(\textit{Toothache}|\textit{Cavity})P(\textit{Catch}|\textit{Cavity})P(\textit{Cavity})$

and compare the result with $P(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$

Outline

- 1 Acting Under Uncertainty
- 2 Basics on Probability
- 3 Probabilistic Inference via Enumeration
- 4 Independence and Conditional Independence
- 5 Applying Bayes' Rule**
- 6 An Example: The Wumpus World Revisited

Bayes' Rule

Bayes' Rule/Theorem/Law

- Bayes' rule: $P(a|b) = \frac{P(a \wedge b)}{P(b)} = \frac{P(b|a)P(a)}{P(b)}$
- In distribution form $P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)} = \alpha P(X|Y)P(Y)$
 - $\alpha \stackrel{\text{def}}{=} 1/P(X)$: normalization constant to make $P(Y|X)$ entries sum to 1 (different α 's for different values of X)
- A version conditionalized on some background evidence $\mathbf{e}$:

$$P(Y|X, \mathbf{e}) = \frac{P(X|Y, \mathbf{e})P(Y|\mathbf{e})}{P(X|\mathbf{e})}$$

Using Bayes' Rule: The Simple Case

- Used to assess **diagnostic probability** from **causal probability**:

$$P(\text{cause}|\text{effect}) = \frac{P(\text{effect}|\text{cause})P(\text{cause})}{P(\text{effect})}$$

- $P(\text{cause}|\text{effect})$ goes from effect to cause (**diagnostic** direction)
- $P(\text{effect}|\text{cause})$ goes from cause to effect (**causal** direction)

Example

- An expert doctor is likely to have **causal knowledge** ... $P(\text{symptoms}|\text{disease})$ (i.e., $P(\text{effect}|\text{cause})$)
... and needs producing **diagnostic knowledge** $P(\text{disease}|\text{symptoms})$ (i.e., $P(\text{cause}|\text{effect})$)
- Ex: let m be meningitis, s be stiff neck
 - $P(m) = 1/50000$, $P(s) = 0.01$ (prior knowledge, from statistics)
 - “meningitis causes to the patient a stiff neck in 70% of cases”: $P(s|m) = 0.7$ (doctor's experience)

$$\Rightarrow P(m|s) = \frac{P(s|m)P(m)}{P(s)} = \frac{0.7 \cdot 1/50000}{0.01} = 0.0014$$

Using Bayes' Rule: Combining Evidence

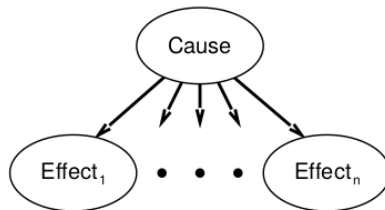
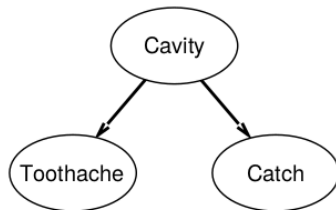
- A **naive Bayes model** is a probability model that assumes **the effects are conditionally independent, given the cause**

$$\Rightarrow \mathbf{P}(\text{Cause}, \text{Effect}_1, \dots, \text{Effect}_n) = \mathbf{P}(\text{Cause}) \prod_i \mathbf{P}(\text{Effect}_i | \text{Cause})$$

- total number of parameters is linear in n
- ex: $\mathbf{P}(\text{Cavity}, \text{Toothache}, \text{Catch}) = \mathbf{P}(\text{Cavity})\mathbf{P}(\text{Toothache} | \text{Cavity})\mathbf{P}(\text{Catch} | \text{Cavity})$

Q: How can we compute $\mathbf{P}(\text{Cause} | \text{Effect}_1, \dots, \text{Effect}_k)$?

- ex $\mathbf{P}(\text{Cavity} | \text{toothache} \wedge \text{catch})$?



Using Bayes' Rule: Combining Evidence [cont.]

Q: How can we compute $\mathbf{P}(\text{Cause}|\text{Effect}_1, \dots, \text{Effect}_k)$?

- ex: $\mathbf{P}(\text{Cavity}|\text{toothache} \wedge \text{catch})$?

A: Apply Bayes' Rule

$$\begin{aligned} & \mathbf{P}(\text{Cavity}|\text{toothache} \wedge \text{catch}) \\ &= \mathbf{P}(\text{toothache} \wedge \text{catch}|\text{Cavity})\mathbf{P}(\text{Cavity})/P(\text{toothache} \wedge \text{catch}) \\ &= \alpha \mathbf{P}(\text{toothache} \wedge \text{catch}|\text{Cavity})\mathbf{P}(\text{Cavity}) \\ &= \alpha \mathbf{P}(\text{toothache}|\text{Cavity})\mathbf{P}(\text{catch}|\text{Cavity})\mathbf{P}(\text{Cavity}) \end{aligned}$$

- $\alpha \stackrel{\text{def}}{=} 1/P(\text{toothache} \wedge \text{catch})$ not computed explicitly
 - General case: $\mathbf{P}(\text{Cause}|\text{Effect}_1, \dots, \text{Effect}_n) = \alpha \mathbf{P}(\text{Cause}) \prod_i \mathbf{P}(\text{Effect}_i|\text{Cause})$
 - $\alpha \stackrel{\text{def}}{=} 1/\mathbf{P}(\text{Effect}_1, \dots, \text{Effect}_n)$ not computed explicitly
(one α value for every value of $\text{Effect}_1, \dots, \text{Effect}_n$)
- $\Rightarrow$ reduces from $2^{n+1} - 1$ to $2n + 1$ independent entries

Outline

- 1 Acting Under Uncertainty
- 2 Basics on Probability
- 3 Probabilistic Inference via Enumeration
- 4 Independence and Conditional Independence
- 5 Applying Bayes' Rule
- 6 An Example: The Wumpus World Revisited**

An Example: The Wumpus World

A probability model of the Wumpus World

- Consider again the Wumpus World (restricted to pit detection)
- Evidence: no pit in (1,1), (1,2), (2,1), breezy in (1,2), (2,1)

Q. Given the evidence, what is the probability of having a pit in (1,3), (2,2) or (3,1)?

- Two groups of variables:
 - $P_{ij} = \text{true}$ iff $[i, j]$ contains a pit (“causes”)
 - $B_{ij} = \text{true}$ iff $[i, j]$ is breezy (“effects”, consider only $B_{1,1}, B_{1,2}, B_{2,1}$)

- Joint Distribution:

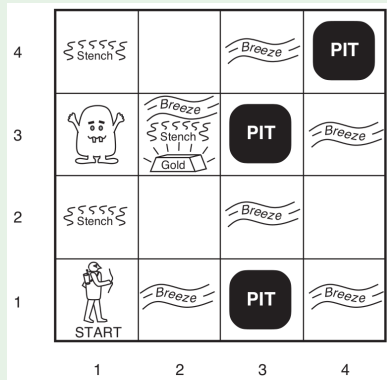
$$\mathbf{P}(P_{1,1}, \dots, P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$$

- Known facts (evidence):

$$b^* \stackrel{\text{def}}{=} \neg b_{1,1} \wedge b_{1,2} \wedge b_{2,1}$$

$$p^* \stackrel{\text{def}}{=} \neg p_{1,1} \wedge \neg p_{1,2} \wedge \neg p_{2,1}$$

- Queries: $\mathbf{P}(P_{1,3}|b^*, p^*)$? $\mathbf{P}(P_{2,2}|b^*, p^*)$?
($\mathbf{P}(P_{3,1}|b^*, p^*)$ symmetric)



An Example: The Wumpus World

A probability model of the Wumpus World

- Consider again the Wumpus World (restricted to pit detection)
- Evidence: no pit in (1,1), (1,2), (2,1), breezy in (1,2), (2,1)

Q. Given the evidence, what is the probability of having a pit in (1,3), (2,2) or (3,1)?

- Two groups of variables:
 - $P_{ij} = \text{true}$ iff $[i, j]$ contains a pit (“causes”)
 - $B_{ij} = \text{true}$ iff $[i, j]$ is breezy (“effects”, consider only $B_{1,1}, B_{1,2}, B_{2,1}$)
- Joint Distribution:
 $P(P_{1,1}, \dots, P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$
- Known facts (evidence):
 - $b^* \stackrel{\text{def}}{=} \neg b_{1,1} \wedge b_{1,2} \wedge b_{2,1}$
 - $p^* \stackrel{\text{def}}{=} \neg p_{1,1} \wedge \neg p_{1,2} \wedge \neg p_{2,1}$
- Queries: $P(P_{1,3}|b^*, p^*)$? $P(P_{2,2}|b^*, p^*)$?
($P(P_{3,1}|b^*, p^*)$ symmetric)

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 B OK	2,2	3,2	4,2
1,1 OK	2,1 B OK	3,1	4,1

An Example: The Wumpus World [cont.]

Specifying the probability model

Compute the joint distribution $\mathbf{P}(P_{1,1}, \dots, P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$:

- Apply the product rule to the joint distribution

$$\mathbf{P}(P_{1,1}, \dots, P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1}) = \mathbf{P}(B_{1,1}, B_{1,2}, B_{2,1} | P_{1,1}, \dots, P_{4,4}) \mathbf{P}(P_{1,1}, \dots, P_{4,4})$$

- $\mathbf{P}(B_{1,1}, B_{1,2}, B_{2,1} | P_{1,1}, \dots, P_{4,4})$: conditional probability distribution of a breeze configuration, given a pit configuration (deterministic)
 - 1 if one pit is adjacent to breeze,
 - 0 otherwise
- $\mathbf{P}(P_{1,1}, \dots, P_{4,4})$: prior probability of a pit configuration
 - assume pits are randomly placed with independent probability 0.2 in each square except in (1,1):
$$\mathbf{P}(P_{1,1}, \dots, P_{4,4}) = \prod_{i=1}^4 \prod_{j=1}^4 P(P_{i,j}), \text{ s.t. } P(P_{i,j}) = \begin{cases} 0.2 & \text{if } (i,j) \neq (1,1) \\ 0 & \text{otherwise} \end{cases}$$
 - e.g, if exactly three pits, except in (1,1):
$$\mathbf{P}(P_{1,1}, \dots, P_{4,4}) = 0.2^3 \cdot (1 - 0.2)^{15-3} \approx 0.00055$$

An Example: The Wumpus World [cont.]

Naive solution: Inference by enumeration

Case $P_{1,3}$:

- General form of query: $\mathbf{P}(\mathbf{Y}|\mathbf{E} = \mathbf{e}) = \alpha P(\mathbf{Y}, \mathbf{E} = \mathbf{e}) = \alpha \sum_{\mathbf{h}} \mathbf{P}(\mathbf{Y}, \mathbf{E} = \mathbf{e}, \mathbf{H} = \mathbf{h})$
 - $\mathbf{Y}$: query vars; $\mathbf{E}, \mathbf{e}$: evidence vars/values; $\mathbf{H}, \mathbf{h}$: hidden vars/values
- Our case: $\mathbf{P}(P_{1,3}|p^*, b^*)$, s.t. the evidence is
 - $b^* \stackrel{\text{def}}{=} \neg b_{1,1} \wedge b_{1,2} \wedge b_{2,1}$
 - $p^* \stackrel{\text{def}}{=} \neg p_{1,1} \wedge \neg p_{1,2} \wedge \neg p_{2,1}$
- Sum over hidden variables:
 $\mathbf{P}(P_{1,3}|p^*, b^*) =$
 $\alpha \sum_{\text{unknown}} \mathbf{P}(P_{1,3}|p^*, b^*, \text{unknown})$
 - unknown are all P_{ij} 's s.t.
 $(i, j) \notin \{(1, 1), (1, 2), (2, 1), (1, 3)\}$
 $\Rightarrow 2^{16-4} = 4096$ terms of the sum!
- Grows exponentially in the number of hidden variables $\mathbf{H}$!
 $\Rightarrow$ Inefficient
- Can we do better?

1,4	2,4	3,4	4,4
1,3 P_{13}	2,3	3,3	4,3
1,2 B OK	2,2	3,2	4,2
1,1 OK	2,1 B OK	3,1	4,1

An Example: The Wumpus World [cont.]

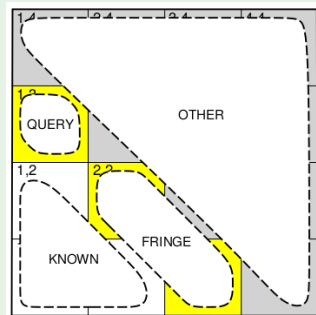
Smarter solution: Exploiting conditional independence

- Basic insight: Given the fringe squares (see below), b^* is conditionally independent of the other hidden squares

- $Unknown \stackrel{\text{def}}{=} Fringe \cup Other$

$$\Rightarrow \mathbf{P}(b^* | p^*, P_{1,3}, Unknown) \stackrel{\text{def}}{=} \mathbf{P}(b^* | p^*, P_{1,3}, Fringe, Others) = \mathbf{P}(b^* | p^*, P_{1,3}, Fringe)$$

- Next: manipulate the query into a form where this equation can be used



An Example: The Wumpus World [cont.]

$\mathbf{P}(p^*, b^*) = P(p^*, b^*)$ is scalar; use as a normalization constant

$$\mathbf{P}(P_{1,3}|p^*, b^*) = \mathbf{P}(P_{1,3}, p^*, b^*) / \underline{\mathbf{P}(p^*, b^*)} = \underline{\alpha} \mathbf{P}(P_{1,3}, p^*, b^*)$$

An Example: The Wumpus World [cont.]

Sum over the unknowns

$$\begin{aligned}\mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\ &= \alpha \sum_{\text{unknown}} \mathbf{P}(P_{1,3}, \text{unknown}, p^*, b^*)\end{aligned}$$

An Example: The Wumpus World [cont.]

Use the product rule

$$\begin{aligned}\mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, \underline{b^*}) \\ &= \alpha \sum_{unknown} \mathbf{P}(\underline{b^*} | P_{1,3}, p^*, unknown) \mathbf{P}(P_{1,3}, p^*, unknown)\end{aligned}$$

An Example: The Wumpus World [cont.]

Separate unknown into *fringe* and *other*

$$\begin{aligned}\mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\ &= \alpha \sum_{\text{unknown}} \mathbf{P}(P_{1,3}, \text{unknown}, p^*, b^*) \\ &= \alpha \sum_{\text{unknown}} \mathbf{P}(b^* | P_{1,3}, p^*, \text{unknown}) \mathbf{P}(P_{1,3}, p^*, \text{unknown}) \\ &= \alpha \sum_{\text{fringe}} \sum_{\text{other}} \mathbf{P}(b^* | p^*, P_{1,3}, \text{fringe}, \text{other}) \mathbf{P}(P_{1,3}, p^*, \text{fringe}, \text{other})\end{aligned}$$

An Example: The Wumpus World [cont.]

b^* is conditionally independent of *other* given *fringe*

$$\begin{aligned}\mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(b^* | P_{1,3}, p^*, unknown) \mathbf{P}(P_{1,3}, p^*, unknown) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^* | p^*, P_{1,3}, \underline{fringe}, other) \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^* | p^*, P_{1,3}, \underline{fringe}) \mathbf{P}(P_{1,3}, p^*, fringe, other)\end{aligned}$$

An Example: The Wumpus World [cont.]

Move $\mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe})$ outward

$$\begin{aligned}\mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\&= \alpha \sum_{\text{unknown}} \mathbf{P}(P_{1,3}, \text{unknown}, p^*, b^*) \\&= \alpha \sum_{\text{unknown}} \mathbf{P}(b^*|P_{1,3}, p^*, \text{unknown}) \mathbf{P}(P_{1,3}, p^*, \text{unknown}) \\&= \alpha \sum_{\text{fringe}} \sum_{\text{other}} \mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe}, \text{other}) \mathbf{P}(P_{1,3}, p^*, \text{fringe}, \text{other}) \\&= \alpha \sum_{\text{fringe}} \sum_{\text{other}} \frac{\mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe}) \mathbf{P}(P_{1,3}, p^*, \text{fringe}, \text{other})}{\mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe})} \\&= \alpha \sum_{\text{fringe}} \frac{\mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe})}{\mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe})} \sum_{\text{other}} \mathbf{P}(P_{1,3}, p^*, \text{fringe}, \text{other})\end{aligned}$$

An Example: The Wumpus World [cont.]

All of the pit locations are independent

$$\begin{aligned} \mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(b^*|P_{1,3}, p^*, unknown) \mathbf{P}(P_{1,3}, p^*, unknown) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^*|p^*, P_{1,3}, fringe, other) \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \sum_{other} \underline{\mathbf{P}(P_{1,3}, p^*, fringe, other)} \\ &= \alpha \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \sum_{other} \underline{\mathbf{P}(P_{1,3}) \mathbf{P}(p^*) \mathbf{P}(fringe) \mathbf{P}(other)} \end{aligned}$$

An Example: The Wumpus World [cont.]

Move $P(p^*)$, $\mathbf{P}(P_{1,3})$, and $P(\text{fringe})$ outward

$$\begin{aligned}\mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\&= \alpha \sum_{\text{unknown}} \mathbf{P}(P_{1,3}, \text{unknown}, p^*, b^*) \\&= \alpha \sum_{\text{unknown}} \mathbf{P}(b^*|P_{1,3}, p^*, \text{unknown}) \mathbf{P}(P_{1,3}, p^*, \text{unknown}) \\&= \alpha \sum_{\text{fringe}} \sum_{\text{other}} \mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe}, \text{other}) \mathbf{P}(P_{1,3}, p^*, \text{fringe}, \text{other}) \\&= \alpha \sum_{\text{fringe}} \sum_{\text{other}} \mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe}) \mathbf{P}(P_{1,3}, p^*, \text{fringe}, \text{other}) \\&= \alpha \sum_{\text{fringe}} \mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe}) \sum_{\text{other}} \mathbf{P}(P_{1,3}, p^*, \text{fringe}, \text{other}) \\&= \alpha \sum_{\text{fringe}} \mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe}) \sum_{\text{other}} \frac{\mathbf{P}(P_{1,3}) P(p^*) P(\text{fringe})}{P(\text{other})} P(\text{other}) \\&= \alpha \frac{P(p^*) \mathbf{P}(P_{1,3})}{\sum_{\text{fringe}}} \sum_{\text{fringe}} \mathbf{P}(b^*|p^*, P_{1,3}, \text{fringe}) \frac{P(\text{fringe})}{\sum_{\text{other}}} \sum_{\text{other}} P(\text{other})\end{aligned}$$

An Example: The Wumpus World [cont.]

Remove $\sum_{other} P(other)$ because it equals 1

$$\begin{aligned} \mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(b^*|P_{1,3}, p^*, unknown) \mathbf{P}(P_{1,3}, p^*, unknown) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^*|p^*, P_{1,3}, fringe, other) \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \sum_{other} \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \sum_{other} \mathbf{P}(P_{1,3}) P(p^*) P(fringe) P(other) \\ &= \alpha P(p^*) \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe) \underbrace{\sum_{other} P(other)}_{=1} \\ &= \alpha P(p^*) \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe) \end{aligned}$$

An Example: The Wumpus World [cont.]

$P(p^*)$ is scalar, so make it part of the normalization constant

$$\begin{aligned} \mathbf{P}(P_{1,3}|p^*, b^*) &= \mathbf{P}(P_{1,3}, p^*, b^*) / \mathbf{P}(p^*, b^*) = \alpha \mathbf{P}(P_{1,3}, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, b^*) \\ &= \alpha \sum_{unknown} \mathbf{P}(b^*|P_{1,3}, p^*, unknown) \mathbf{P}(P_{1,3}, p^*, unknown) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^*|p^*, P_{1,3}, fringe, other) \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \sum_{other} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \sum_{other} \mathbf{P}(P_{1,3}, p^*, fringe, other) \\ &= \alpha \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) \sum_{other} \mathbf{P}(P_{1,3}) P(p^*) P(fringe) P(other) \\ &= \alpha P(p^*) \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe) \sum_{other} P(other) \\ &= \underline{\alpha P(p^*)} \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe) \\ &= \underline{\alpha'} \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe) \end{aligned}$$

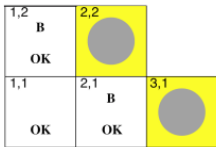
An Example: The Wumpus World [cont.]

- We have obtained: $\mathbf{P}(P_{1,3}|p^*, b^*) = \alpha' \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe)$
- We know that $\mathbf{P}(P_{1,3}) = \langle 0.2, 0.8 \rangle$ (see slide 38)
- We can compute the normalization coefficient α' afterwards
- $\sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe)$: only 4 possible fringes
- Start by rewriting as two separate equations:

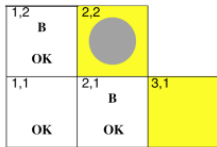
$$\mathbf{P}(p_{1,3}|p^*, b^*) = \alpha' P(p_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, p_{1,3}, fringe) P(fringe)$$

$$\mathbf{P}(\neg p_{1,3}|p^*, b^*) = \alpha' P(\neg p_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, \neg p_{1,3}, fringe) P(fringe)$$

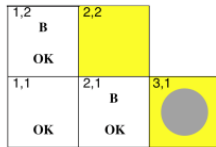
Four possible fringes:



$$0.2 \times 0.2 = 0.04$$



$$0.2 \times 0.8 = 0.16$$



$$0.8 \times 0.2 = 0.16$$



$$0.8 \times 0.8 = 0.64$$

An Example: The Wumpus World [cont.]

- Start by rewriting as two separate equations:

$$P(p_{1,3}|p^*, b^*) = \alpha' P(p_{1,3}) \sum_{fringe} P(b^*|p^*, p_{1,3}, fringe) P(fringe)$$

$$P(\neg p_{1,3}|p^*, b^*) = \alpha' P(\neg p_{1,3}) \sum_{fringe} P(b^*|p^*, \neg p_{1,3}, fringe) P(fringe)$$

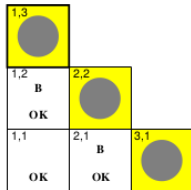
- For each of them, $P(b^*|...)$ is 1 if the breezes occur, 0 otherwise:

$$\sum_{fringe} P(b^*|p^*, p_{1,3}, fringe) P(fringe) = 1 \cdot 0.04 + 1 \cdot 0.16 + 1 \cdot 0.16 + 0 \cdot 0.64 = 0.36$$

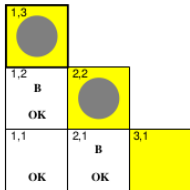
$$\sum_{fringe} P(b^*|p^*, \neg p_{1,3}, fringe) P(fringe) = 1 \cdot 0.04 + 1 \cdot 0.16 + 0 \cdot 0.16 + 0 \cdot 0.64 = 0.2$$

$$\Rightarrow P(P_{1,3}|p^*, b^*) = \alpha' P(P_{1,3}) \sum_{fringe} P(b^*|p^*, P_{1,3}, fringe) P(fringe)$$

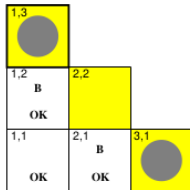
$$= \alpha' \langle 0.2, 0.8 \rangle \langle 0.36, 0.2 \rangle = \alpha' \langle 0.072, 0.16 \rangle = (\text{normalization, s.t. } \alpha' \approx 4.31) \approx \langle 0.31, 0.69 \rangle$$



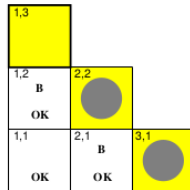
$$0.2 \times 0.2 = 0.04$$



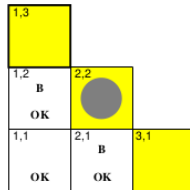
$$0.2 \times 0.8 = 0.16$$



$$0.8 \times 0.2 = 0.16$$



$$0.2 \times 0.2 = 0.04$$



$$0.2 \times 0.8 = 0.16$$

Exercise

Compute $\mathbf{P}(P_{2,2}|p^*, b^*)$ in the same way.