Fundamentals of Artificial Intelligence
Chapter 13: **Quantifying Uncertainty**

Roberto Sebastiani

DISI, Università di Trento, Italy – roberto.sebastiani@unitn.it
http://disi.unitn.it/rseba/DIDATTICA/fai_2021/

Teaching assistant: Mauro Dragoni – dragoni@fbk.eu
http://www.maurodragoni.com/teaching/fai/

M.S. Course "Artificial Intelligence Systems", academic year 2021-2022

Last update: Monday 13th December, 2021, 11:41

Outline

Outline

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)
    $\Longrightarrow$ 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

$\Longrightarrow$ 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)
    $Toothache \rightarrow (Cavity \lor GumProblem \lor Abscess \lor ...)$
    (too many possible causes, some very unlikely)
  - causal rules:
    $Cavity \rightarrow Toothache$ (wrong)
    $(Cavity \land ...) \rightarrow Toothache$ (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":
    $P(HasCavity(patient) \mid \text{hasToothAche(patient)}) = 0.8$
- 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":
    $P(HasCavity(patient) \mid hasToothAche(patient) \wedge HistoryOfGum(patient)) = 0.4$

Making Decisions Under Uncertainty

- Ex: Suppose I believe:
  $P(A_{25}$ gets me there on time $|\ldots) = 0.04$
  $P(A_{90}$ gets me there on time $|\ldots) = 0.70$
  $P(A_{120}$ gets me there on time $|\ldots) = 0.95$
  $P(A_{1440}$ gets me there on time $|\ldots) = 0.9999$
  Which action to choose?

$\Longrightarrow$ 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

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 $\omega \in \Omega$
  - $\Sigma_{\omega \in \Omega} P(\omega) = 1$
- 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. $P(A) = \Sigma_{\omega \in A} P(\omega)$
  - events can be described by propositions in some formal language
  - ex: $P(Total = 11) = P(5,6) + P(6,5) = 1/36 + 1/36 = 1/18$
  - ex: $P(doubles) = P(1,1) + P(2,2) + ... + P(6,6) = 6/36 = 1/6$

Random Variables

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

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: $a \iff$ "$A = true$"
- Given Boolean random variables A and B:
  - a: set of sample points where $A(\omega) = true$
  - $\neg a$: set of sample points where $\mathrm{A}(\omega) = false$
  - $a \land b$: set of sample points where $A(\omega) = true$, $B(\omega) = true$

$\Longrightarrow$ with Boolean random variables, sample points are PL models

- Proposition: disjunction of the sample points in which it is true
  - ex: $(a \lor b) \equiv (\neg a \land b) \lor (a \land \neg b) \lor (a \land b)$

  $\Longrightarrow$ $P(a \lor b) = P(\neg a \land b) + P(a \land \neg b) + P(a \land b)$
- Some derived facts:
  - $P(\neg a) = 1 - P(a)$
  - $P(a \lor b) = P(a) + P(b) - P(a \land b)$

Probability Distributions

- Probability Distribution gives the probabilities of all the possible values of a random variable
  - ex: Weather: $\{sunny, rain, cloudy, snow\}$

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

  - normalized: their sum is 1
- Joint Probability Distribution for multiple variables
  - gives the probability of every sample point
  - ex: $\mathbf{P}(Weather, Cavity) =$

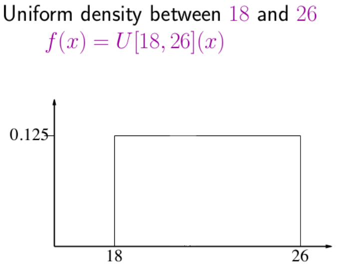
$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$	0.144	0.02	0.016	0.02
$Cavity = false$	0.576	0.08	0.064	0.08

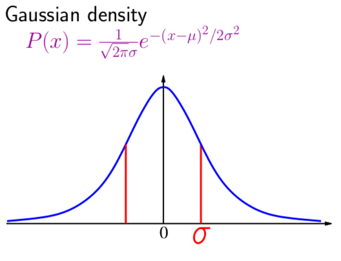
- Every event is a sum of sample points,
  $\Longrightarrow$ its probability is determined by the joint distribution

Probability for Continuous Variables

- Express continuous probability distributions:
  - density functions $f(x) \in [0,1]$ s.t $\int_{-\infty}^{+\infty} f(x)dx = 1$
- $P(x \in [a,b]) = \int_a^b f(x)\,dx$

  $\Longrightarrow$ $P(x \in [val, val]) = 0$, $P(x \in [-\infty, +\infty]) = 1$
  - ex: $P(x \in [20,22]) = \int_{20}^{22} 0.125\,dx = 0.25$
- Density: $P(x) = P(X = x) \stackrel{\text{def}}{=} \lim_{dx \mapsto 0} P(X \in [x, x+dx])/dx$
  - ex: $P(20.1) = \lim_{dx \mapsto 0} P(X \in [20.1, 20.1+dx])/dx = 0.125$
  - note: $P(v) \neq P(x \in [v,v]) = 0$





Conditional Probabilities

- Unconditional or prior probabilities refer to degrees of belief in propositions in the absence of any other information (evidence)
  - ex: $P(cavity) = 0.2$, $P(Total = 11) = 1/18$, $P(double) = 1/6$
- Conditional or posterior probabilities refer to degrees of belief in proposition a given some evidence b: P(a|b)
  - evidence: information already revealed
  - ex: $P(cavity|toothache) = 0.6$: p. of a cavity given a toothache (assuming no other information is provided!)
  - ex: $P(Total = 11|die_1 = 5) = 1/6$: p. of total 11 given first die is 5

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

Conditional Probabilities [cont.]

- Conditional probability: $P(a|b) \stackrel{\text{def}}{=} \frac{P(a \wedge b)}{P(b)}$, s.t. $P(b) > 0$
  - ex: $P(Total = 11|die_1 = 5) = \frac{P(Total=11 \wedge die_1=5)}{P(die_1=5)} = \frac{1/6 \cdot 1/6}{1/6} = 1/6$
  - 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: $\mathbf{P}(Weather, Cavity) = \mathbf{P}(Weather|Cavity)\mathbf{P}(Cavity)$, that is:

    $P(sunny, cavity) = P(sunny|cavity)P(cavity)$

    ...

    $P(snow, \neg cavity) = P(snow|\neg cavity)P(\neg cavity)$
  - a 4×2 set of equations, not matrix multiplication!
- Chain rule is derived by successive application of product rule:

$$\begin{aligned}
&\mathbf{P}(X_1, ..., X_n) \\
&= \mathbf{P}(X_1, ..., X_{n-1})\mathbf{P}(X_n|X_1, ..., X_{n-1}) \\
&= \mathbf{P}(X_1, ..., X_{n-2})\mathbf{P}(X_{n-1}|X_1, ..., X_{n-2})\mathbf{P}(X_n|X_1, ..., X_{n-1}) \\
&= ... \\
&= \prod_{i=1}^{n} \mathbf{P}(X_i|X_1, ..., X_{i-1})
\end{aligned}$$

Logic vs. Probability

Logic	Probability
a	$P(a) = 1$
$\neg a$	$P(a) = 0$
$a \rightarrow b$	$P(b \mid a) = 1$
$\dfrac{(a, a \rightarrow b)}{b}$	$\dfrac{P(a) = 1, P(b \mid a) = 1}{P(b) = 1}$
$\dfrac{(a \rightarrow b, b \rightarrow c)}{a \rightarrow c}$	$\dfrac{P(b \mid a) = 1, P(c \mid b) = 1}{P(c \mid a) = 1}$

- Proof of $P(b \mid a) = 1, P(c \mid b) = 1 \Longrightarrow P(c \mid a) = 1$
  - $P(b \mid a) = 1 \Longrightarrow P(\neg b, a) \stackrel{\text{def}}{=} P(\neg b \mid a)P(a) = 0$
  - $P(c \mid b) = 1 \Longrightarrow P(\neg c, b) \stackrel{\text{def}}{=} P(\neg c \mid b)P(b) = 0$
  - $P(\neg c, a) = P(\neg c, a, b) + P(\neg c, a, \neg b) \leq \underbrace{P(\neg c, b)}_{0} + \underbrace{P(a, \neg b)}_{0} = 0$
  - $P(\neg c \mid a) = P(\neg c, a)/P(a) = 0$
  - $P(c \mid a) = 1 - P(\neg c \mid a) = 1$

(Courtesy of Maria Simi, UniPI)

Outline

Probabilistic Inference via Enumeration

Basic Ideas

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

Probabilistic Inference via Enumeration: Example

Example: Generic Inference

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

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

Probabilistic Inference via Enumeration: Example

Example: Generic Inference

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

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	.108	.012	.072	.008
¬ *cavity*	.016	.064	.144	.576

(© S. Russell & P. Norwig, AIMA)

Marginalization

- Start with the joint distribution $\mathbf{P}(Toothache, Catch, Cavity)$
- Marginalization (aka summing out):
  sum up the probabilities for each possible value of the other variables:
  $\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}, \mathbf{z})$
  Ex: $\mathbf{P}(Toothache) = \sum_{\mathbf{z} \in \{Catch, Cavity\}} \mathbf{P}(Toothache, \mathbf{z})$
- Conditioning: variant of marginalization, involving conditional probabilities instead of joint probabilities (using the product rule)
  $\mathbf{P}(\mathbf{Y}) = \sum_{\mathbf{z} \in \mathbf{Z}} \mathbf{P}(\mathbf{Y}|\mathbf{z})P(\mathbf{z})$
  Ex: $\mathbf{P}(Toothache) = \sum_{\mathbf{z} \in \{Catch, Cavity\}} \mathbf{P}(Toothache|\mathbf{z})P(\mathbf{z})$

Marginalization: Example

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

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

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

  $$P(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$$
  $$P(\neg toothache) = 1 - P(toothache) = 1 - 0.2 = 0.8$$

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

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	.108	.012	.072	.008
¬ *cavity*	.016	.064	.144	.576

Marginalization: Example

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

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

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

  $P(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$
  $P(\neg toothache) = 1 - P(toothache) = 1 - 0.2 = 0.8$

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

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ cavity	**.016**	**.064**	**.144**	**.576**

Conditional Probability via Enumeration: Example

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

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

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

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

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

Conditional Probability via Enumeration: Example

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

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

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

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

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ cavity	**.016**	**.064**	**.144**	**.576**

Conditional Probability via Enumeration: Example

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

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

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

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

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.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:
  $\mathbf{P}(Y|\mathbf{E}=\mathbf{e}) = \alpha\mathbf{P}(Y, \mathbf{E}=\mathbf{e}) = \alpha\Sigma_{\mathbf{h}\in\mathbf{H}}\mathbf{P}(Y, \mathbf{E}=\mathbf{e}, \mathbf{H}=\mathbf{h})$

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

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

Normalization: Example

- $\alpha \stackrel{\text{def}}{=} 1/P(toothache)$ 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 $\sum ... = 1$

Ex:

$\mathbf{P}(Cavity|toothache) = \alpha\mathbf{P}(Cavity \wedge toothache)$
$= \alpha[\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]$
$= \alpha[\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle]$
$= \alpha\langle 0.12, 0.08\rangle = (normalization) = \langle 0.6, 0.4\rangle \; [\alpha{=}5]$

$\mathbf{P}(Cavity|\neg toothache) = ... = \alpha\langle 0.08, 0.72\rangle = \langle 0.1, 0.9\rangle[\alpha{=}1.25]$

	toothache		¬ toothache	
	catch	¬ catch	catch	¬ catch
cavity	.108	.012	.072	.008
¬ cavity	.016	.064	.144	.576

Normalization: Example

- $\alpha \stackrel{\text{def}}{=} 1/P(toothache)$ 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 $\sum \ldots = 1$

Ex:

$$\mathbf{P}(Cavity|toothache) = \alpha\mathbf{P}(Cavity \wedge toothache)$$
$$= \alpha[\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]$$
$$= \alpha[\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle]$$
$$= \alpha\langle 0.12, 0.08\rangle = (normalization) = \langle 0.6, 0.4\rangle\ [\alpha{=}5]$$

$$\mathbf{P}(Cavity|\neg toothache) = \ldots = \alpha\langle 0.08, 0.72\rangle = \langle 0.1, 0.9\rangle[\alpha{=}1.25]$$

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ cavity	**.016**	**.064**	**.144**	**.576**

Normalization: Example

- $\alpha \stackrel{\text{def}}{=} 1/P(toothache)$ 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 $\sum ... = 1$

Ex:

$$\mathbf{P}(Cavity|toothache) = \alpha\mathbf{P}(Cavity \land toothache)$$
$$= \alpha[\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]$$
$$= \alpha[\langle 0.108, 0.016\rangle + \langle 0.012, 0.064\rangle]$$
$$= \alpha\langle 0.12, 0.08\rangle = (normalization) = \langle 0.6, 0.4\rangle \; [\alpha{=}5]$$

$$\mathbf{P}(Cavity|\neg toothache) = ... = \alpha\langle 0.08, 0.72\rangle = \langle 0.1, 0.9\rangle [\alpha{=}1.25]$$

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ *cavity*	**.016**	**.064**	**.144**	**.576**

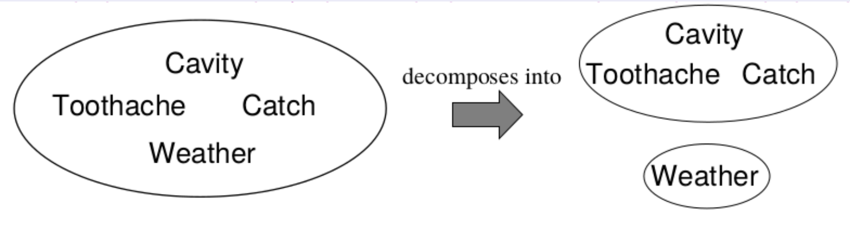
Outline

Independence

- Variables X and Y are independent iff $\mathbf{P}(X,Y) = \mathbf{P}(X)\mathbf{P}(Y)$ (or equivalently, iff $\mathbf{P}(X|Y) = \mathbf{P}(X)$ or $\mathbf{P}(Y|X) = \mathbf{P}(Y)$)
  - ex: $\mathbf{P}(Toothache, Catch, Cavity, Weather) = \mathbf{P}(Toothache, Catch, Cavity)\mathbf{P}(Weather)$

  $\Longrightarrow$ e.g. $P(toothache, catch, cavity, cloudy) = P(toothache, catch, cavity)P(cloudy)$
  - typically based on domain knowledge
- May drastically reduce the number of entries and computation

  $\Longrightarrow$ ex: 32-element table decomposed into one 8-element and one 4-element table
- Unfortunately, absolute independence is quite rare



(© S. Russell & P. Norwig, AIMA)

Conditional Independence

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

  $\Longrightarrow$ Catch is conditionally independent of Toothache given Cavity:
  $\mathbf{P}(Catch|Toothache, Cavity) = \mathbf{P}(Catch|Cavity)$
  or, equivalently: $\mathbf{P}(Toothache|Catch, Cavity) = \mathbf{P}(Toothache|Cavity)$, or
  $\mathbf{P}(Toothache, Catch|Cavity) = \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)$
- 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!
- Ex:
$$\begin{aligned}
&\mathbf{P}(Toothache, Catch, Cavity)\\
&= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch, Cavity)\\
&= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)\\
&= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)
\end{aligned}$$

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

- $\mathbf{P}(Toothache, Catch, Cavity)$ contains 7 independent entries (the 8th can be obtained as $1 - \sum ...$)
- $\mathbf{P}(Toothache|Cavity)$,$\mathbf{P}(Catch|Cavity)$ contain 2 independent entries (2×2 matrix, each row sums to 1)
- $\mathbf{P}(Cavity)$ contains 1 independent entry

- General Case: if one causes has n independent effects:
$\mathbf{P}(Cause, Effect_1, ..., Effect_n) = \mathbf{P}(Cause) \prod_i \mathbf{P}(Effect_i|Cause)$
$\Longrightarrow$ reduces from $2^{n+1} - 1$ to $2n + 1$ independent entries

Exercise

Consider the joint probability distribution described in the table in previous section (slide 20 onwards): $\mathbf{P}(Toothache, Catch, Cavity)$

- Consider the example in previous slide:
  $\mathbf{P}(Toothache, Catch, Cavity)$
  $= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch, Cavity)$
  $= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)$
  $= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)$
- Compute separately the distributions $\mathbf{P}(Toothache|Catch, Cavity)$, $\mathbf{P}(Catch|Cavity)$, $\mathbf{P}(Cavity)$, $\mathbf{P}(Toothache|Cavity)$.
- Recompute $\mathbf{P}(Toothache, Catch, Cavity)$ in two ways:
  - $\mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)$
  - $\mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)$

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

Outline

Bayes' Rule

Bayes' Rule/Theorem/Law

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

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

Using Bayes' Rule: The Simple Case

- Used to assess diagnostic probability from causal probability:

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

  - $P(cause|effect)$ goes from effect to cause (diagnostic direction)
  - $P(effect|cause)$ goes from cause to effect (causal direction)

Example

- An expert doctor is likely to have causal knowledge ... $P(symptoms|disease)$ (i.e., $P(effect|cause)$)

  ... and needs producing diagnostic knowledge $P(disease|symptoms)$ (i.e., $P(cause|effect)$)
- 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)

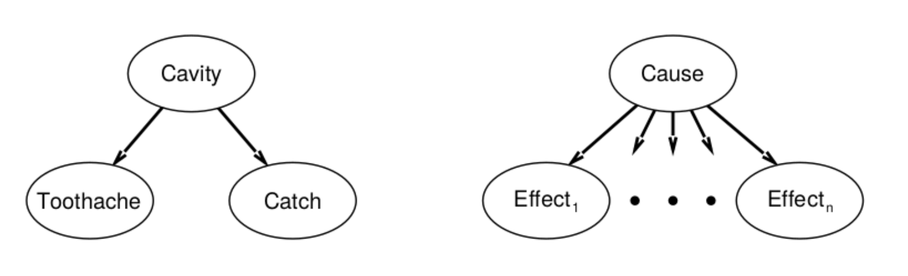
$$\Longrightarrow P(m|s) = \frac{P(s|m)P(m)}{P(s)} = \frac{0.7 \cdot 1/50000}{0.01} = 0.0014$$

Using Bayes' Rule: Combining Evidence

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

  $\Longrightarrow$ $\mathbf{P}(Cause, Effect_1, ..., Effect_n) = \mathbf{P}(Cause) \prod_i \mathbf{P}(Effect_i | Cause)$
  - total number of parameters is linear in n
  - ex: $\mathbf{P}(Cavity, Toothache, Catch) = \mathbf{P}(Cavity)\mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)$

Q: How can we compute $\mathbf{P}(Cause|Effect_1, ..., Effect_k)$?
  - ex $\mathbf{P}(Cavity|toothache \wedge catch)$?



(© S. Russell & P. Norwig, AIMA)

Using Bayes' Rule: Combining Evidence [cont.]

Q: How can we compute $\mathbf{P}(Cause|Effect_1, ..., Effect_k)$?

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

A: Apply Bayes' Rule

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

- $\alpha \stackrel{\text{def}}{=} 1/P(toothache \wedge catch)$ not computed explicitly

- General case: $\mathbf{P}(Cause|Effect_1, ..., Effect_n) = \alpha\mathbf{P}(Cause)\prod_i \mathbf{P}(Effect_i|Cause)$
  - $\alpha \stackrel{\text{def}}{=} 1/\mathbf{P}(Effect_1, ..., Effect_n)$ not computed explicitly (one α value for every value of $Effect_1, ..., Effect_n$)

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

Outline

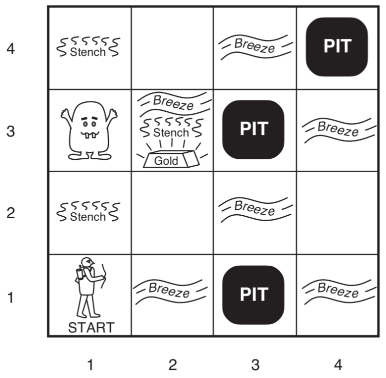
An Example: The Wumpus World

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:
  - $P_{ij} = true$ iff $[i, j]$ contains a pit ("causes")
  - $B_{ij} = true$ iff $[i, j]$ is breezy ("effects", consider only $B_{1,1}, B_{1,2}, B_{2,1}$)
- Joint Distribution: $\mathbf{P}(P_{1,1}, ..., P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$
- 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: $\mathbf{P}(P_{1,3}|b^*, p^*)$? $\mathbf{P}(P_{22}|b^*, p^*)$? ($\mathbf{P}(P_{3,1}|b^*, p^*)$ symmetric)



(© S. Russell & P. Norwig, AIMA)

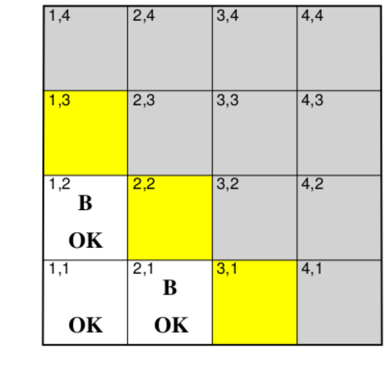
An Example: The Wumpus World

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:
  - $P_{ij} = true$ iff $[i, j]$ contains a pit ("causes")
  - $B_{ij} = true$ iff $[i, j]$ is breezy ("effects", consider only $B_{1,1}$, $B_{1,2}$, $B_{2,1}$)
- Joint Distribution: $\mathbf{P}(P_{1,1}, ..., P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$
- 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: $\mathbf{P}(P_{1,3}|b^*, p^*)$? $\mathbf{P}(P_{22}|b^*, p^*)$? ($\mathbf{P}(P_{3,1}|b^*, p^*)$ symmetric)



An Example: The Wumpus World [cont.]

Specifying the probability model

- Apply the product rule to the joint distribution $\mathbf{P}(P_{1,1}, ..., P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1}) =$
  $\mathbf{P}(B_{1,1}, B_{1,2}, B_{2,1} | P_{1,1}, ..., P_{4,4})\ \mathbf{P}(P_{1,1}, ..., P_{4,4})$
- $\mathbf{P}(B_{1,1}, B_{1,2}, B_{2,1} | P_{1,1}, ..., P_{4,4})$
  - 1 if one pit is adjacent to breeze,
  - 0 otherwise
- $\mathbf{P}(P_{1,1}, ..., P_{4,4})$: pits are placed randomly except in (1,1):
  $\mathbf{P}(P_{1,1}, ..., P_{4,4}) = \prod_{i=1}^{4} \prod_{j=1}^{4} P(P_{i,j})$

  $$P(P_{i,j}) = \begin{cases} 0.2 & \textit{if } (i,j) \neq (1,1)\} \\ 0 & \textit{otherwise} \end{cases}$$

  - ex: $\mathbf{P}(P_{1,1}, ..., P_{4,4}) = 0.2^3 \cdot 0.8^{15-3} \approx 0.00055$ if 3 pits

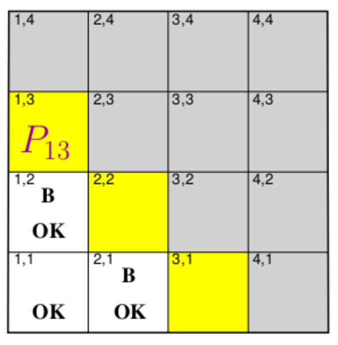
An Example: The Wumpus World [cont.]

Inference by enumeration

Case $P_{1,3}$:

- General form of query: $\mathbf{P}(\mathbf{Y}|\mathbf{E}=\mathbf{e}) = \alpha P(\mathbf{Y},\mathbf{E}=\mathbf{e}) = \alpha\sum_{\mathbf{h}}\mathbf{P}(\mathbf{Y},\mathbf{E}=\mathbf{e},\mathbf{H}=\mathbf{h})$
  - **Y**: query vars; **E**,**e**: evidence vars/values; **H**,**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:
  $\mathbf{P}(P_{1,3}|p^*, b^*) =$
  $\alpha\sum_{unknown}\mathbf{P}(P_{1,3}|p^*, b^*, unknown)$
  - *unknown* are all P_{ij}'s s.t. $(i,j) \notin \{(1,1),(1,2),(2,1),(1,3)\}$

  $\Longrightarrow 2^{16-4} = 4096$ terms of the sum!
- Grows exponentially in the number of hidden variables **H**!
  $\Longrightarrow$ Inefficient



(© S. Russell & P. Norwig, AIMA)

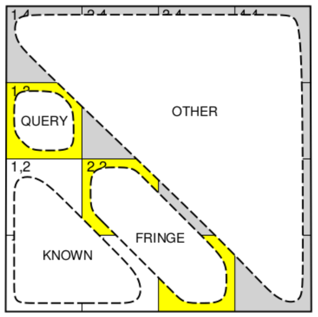
An Example: The Wumpus World [cont.]

Using conditional independence

- Basic insight: Given the fringe squares (see below), b^* is conditionally independent of the other hidden squares
  - $Unknown \stackrel{\text{def}}{=} Fringe \cup Other$

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

- 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

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

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

An Example: The Wumpus World [cont.]

Use the product rule

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

An Example: The Wumpus World [cont.]

Separate unknown into *fringe* and *other*

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

$$= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, b^*)$$

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

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

An Example: The Wumpus World [cont.]

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

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

An Example: The Wumpus World [cont.]

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

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

$$= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, b^*)$$

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

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

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

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

An Example: The Wumpus World [cont.]

All of the pit locations are independent

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

An Example: The Wumpus World [cont.]

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

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

An Example: The Wumpus World [cont.]

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

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

$$= \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, p^*, b^*)$$

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

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

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

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

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

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

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

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:

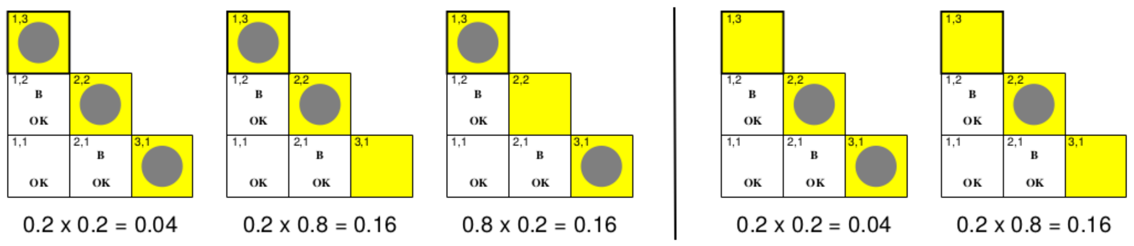
0.2 x 0.2 = 0.04 | 0.2 x 0.8 = 0.16 | 0.8 x 0.2 = 0.16 | 0.8 x 0.8 = 0.64

(© S. Russell & P. Norwig, AIMA)

An Example: The Wumpus World [cont.]

- Start by rewriting as two separate equations:
  $\mathbf{P}(\ \ p_{1,3}|p^*, b^*) = \alpha' P(\ \ p_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*,\ \ p_{1,3}, fringe) P(fringe)$
  $\mathbf{P}(\neg p_{1,3}|p^*, b^*) = \alpha' P(\neg p_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, \neg p_{1,3}, fringe) P(fringe)$
- For each of them, $P(b^*|...)$ is 1 if the breezes occur, 0 otherwise:
  $\sum_{fringe} \mathbf{P}(b^*|p^*,\ \ p_{1,3}, fringe) P(fringe) = 1 \cdot 0.04 + 1 \cdot 0.16 + 1 \cdot 0.16 + 0 \cdot 0.64 = 0.36$
  $\sum_{fringe} \mathbf{P}(b^*|p^*, \neg p_{1,3}, fringe) P(fringe) = 1 \cdot 0.04 + 1 \cdot 0.16 + 0 \cdot 0.16 + 0 \cdot 0.64 = 0.2$

$\Longrightarrow \mathbf{P}(P_{1,3}|p^*, b^*) = \alpha' \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b^*|p^*, P_{1,3}, fringe) P(fringe)$
$= \alpha' \langle 0.2, 0.8 \rangle \langle 0.36, 0.2 \rangle = \alpha' \langle 0.072, 0.16 \rangle = (normalization,\ s.t.\ \alpha' \approx 4.31) \approx \langle 0.31, 0.69 \rangle$



(© S. Russell & P. Norwig, AIMA)

Exercise

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