Chapter 13
Quantifying Uncertainty

CS5811 - Artificial Intelligence

Nilufer Onder
Department of Computer Science
Michigan Technological University

Outline

Probability basics

Syntax and semantics

Inference

Independence and Bayes' rule

Motivation

Uncertainty is everywhere. Consider the following proposition.

> A_t: *Leaving t minutes before the flight will get me to the airport.*

Problems:

1. partial observability (road state, other drivers' plans, etc.)
2. noisy sensors (radio traffic reports)
3. uncertainty in action outcomes (flat tire, etc.)
4. immense complexity of modelling and predicting traffic

Knowledge representation

Language	Main elements	Assignments
Propositional logic	facts	T, F, unknown
First-order logic	facts, objects, relations	T, F, unknown
Temporal logic	facts, objects, relations, times	T, F, unknown
Temporal CSPs	time points	time intervals
Fuzzy logic	set membership	degree of truth
Probability theory	facts	degree of belief

The first three do not represent uncertainty, while the last three do.

Probability

Probabilistic assertions summarize effects of

laziness: failure to enumerate exceptions, qualifications, etc.

ignorance: lack of relevant facts, initial conditions, etc.

Probabilities relate propositions to one's own state of knowledge.
They might be learned from past experience of similar situations.

e.g., $P(A_{25}) = 0.05$

Probabilities of propositions change with new evidence:

e.g., $P(A_{25} |$ no reported accidents$) = 0.06$

e.g., $P(A_{25} |$ no reported accidents, 5am$) = 0.15$

Probability basics

Begin with a set Ω called the *sample space*
A sample space is a set of possible outcomes

Each $\omega \in \Omega$ is a *sample point* (*possible world*, *atomic event*)
e.g., 6 possible rolls of a die: $\{1, 2, 3, 4, 5, 6\}$

Probability space or *probability model*:
Take a sample space Ω, and
assign a number *$P(\omega)$ (the probability of ω)*
to every atomic event $\omega \in \Omega$

Probability basics (cont'd)

A probability space must satisfy the following properties:

$0 \leq P(\omega) \leq 1$ for every $\omega \in \Omega$

$\sum_{\omega \in \Omega} P(\omega) = 1$

e.g., for rolling the die,
P (1) = P (2) = P (3) = P (4) = P (5) = P (6) = 1/6.

An *event* A is any subset of Ω
The probability of an event is defined as follows:

$P(A) = \sum_{\{\omega \in A\}} P(\omega)$

e.g., P(die roll < 4) =
P(1) + P(2) + P(3) = 1/6 + 1/6 + 1/6 = 1/2

Random variables

A *random variable* is a function from sample points to some range such as integers or Booleans.
We'll use capitalized words for random variables.

e.g., rolling the die:

$$Odd(\omega) = \begin{cases} \text{true} & \text{if } \omega \text{ is even,} \\ \text{false} & \text{otherwise} \end{cases}$$

A *probability distribution* gives a probability for every possible value.

If X is a random variable, then
$P(X = x_i) = \sum\{P(\omega) : X(\omega) = x_i\}$

e.g., P (Odd = true) = P (1) + P (3) + P (5) =
1/6 + 1/6 + 1/6 = 1/2
Note that we don't write Odd's argument ω here.

Propositions

Odd is a *Boolean* or *propositional* random variable:
its range is {true, false}

We'll use the corresponding lower-case word (in this case *odd*) for the event that a propositional random variable is true

e.g., $P(odd) = P(Odd = true) = 1/6$
$P(\neg odd) = P(Odd = false) = 5/6$

Boolean formula = disjunction of the sample points in which it is true

e.g., $(a \vee b) \equiv (\neg a \wedge b) \vee (a \wedge \neg b) \vee (a \wedge b)$
$\Rightarrow P(a \vee b) = P(\neg a \wedge b) + P(a \wedge \neg b) + P(a \wedge b)$

Syntax for propositions

Propositional or *Boolean* random variables
e.g., *Cavity* (do I have a cavity in one of my teeth?)
$Cavity = true$ is a proposition, also written *cavity*

Discrete random variables (finite or infinite)
e.g., *Weather* is one of $< sunny, rain, cloudy, snow >$
$Weather = rain$ is a proposition
Values must be exhaustive and mutually exclusive

Continuous random variables (bounded or unbounded)
e.g., $Temp = 21.6$; $Temp < 22.0$

Arbitrary Boolean combinations of basic propositions
e.g., $\neg cavity$ means $Cavity = false$

Probabilities of propositions
e.g., $P(cavity) = 0.1$ and $P(Weather = sunny) = 0.72$

Syntax for probability distributions

Represent a discrete probability distribution as a vector of probability values:

$\mathbb{P}(Weather) =< 0.72, 0.1, 0.08, 0.1 >$

The above is an ordered list representing the probabilities of *sunny*, *rain*, *cloudy*, and *snow*.
Probabilities of *sunny*, *rain*, *cloudy*, and *snow* must sum to 1 when the vector is *normalized*

If B is a Boolean random variable, then $P(B) =< P(b), P(\neg b) >$
e.g., if $P(cavity) = 0.1$ then
$P(Cavity = true) = 0.1$ and $\mathbb{P}(Cavity) =< 0.1, 0.9 >$

When the entries in the vector do not add up to 1, but represent the true ratios, the vector is preceded by a *normalizing constant*, α, e.g. $\mathbb{P}(Cavity) = \alpha < 0.01, 0.09 >$ where α is 10

Syntax for joint probability distributions

A *joint probability distribution* for a set of n random variables gives the probability of every atomic event on those variables,
i.e., every sample point

Represent it as an n-dimensional matrix,
e.g., $\mathbb{P}(Weather, Cavity)$ is a 4×2 matrix.
The entries contain propabilities for all possible combinations of *Weather* (4), and *Cavity* (2).

	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 question about a domain can be answered by the joint distribution because every event is a sum of sample points

Conditional probability

Prior (unconditional) probabilities refer to degrees of belief in the absence of any other information.

Posterior (conditional) probabilites refer to degrees of belief when we have some information, called *evidence*.

Consider drawing straws from a set of 1 long and 4 short straws, *long* refers to drawing a long straw, and *short* refers to drawing a short straw.

$P(long) = 0.2$
$P(long|short) = 0.25$
$P(long|long) = 0.0$
$P(long|short, short) = \frac{1}{3}$
$P(long|rain) = 0.2$

Conditional probability (cont'd)

$P(cavity|toothache) = 0.8$ means
the probability of *cavity* given that *toothache* is all we know
It does **not** mean "if *toothache* then 80% chance of *cavity*

Suppose we get more evidence, e.g., *cavity* is also given. Then
$P(cavity|toothache, cavity) = 1$
Note: the less specific belief remains valid, but is not always useful

New evidence may be irrelevant, allowing simplification, e.g.,
$P(cavity|toothache, 49ersWin) = P(cavity|toothache) =$0.8

Conditional distibutions are shown as vectors for all possible combinations of the evidence and query.
$\mathbb{P}(Cavity|Toothache)$ is a 2-element vector of 2-element vectors

$$< \underbrace{< 0.12, 0.08 >}_{toothache}, \underbrace{< 0.08, 0.72 >}_{\neg toothache} >$$

Conditional probability definitions

Definition of conditional probability:

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$

Product rule gives an alternative formulation and holds even if $P(b) = 0$

$$P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$$

A general version holds for an entire probability distribution, e.g.,

$$\mathbb{P}(Weather, Cavity) = \mathbb{P}(Weather|Cavity)\mathbb{P}(Cavity)$$

This is not matrix multiplication, it's a set of 4×2 equations:

$$\begin{array}{ll} P(sunny, cavity) = P(sunny|cavity)P(cavity) & P(sunny, \neg cavity) = P(sunny|\neg cavity)P(\neg cavity) \\ P(rain, cavity) = P(rain|cavity)P(cavity) & P(rain, \neg cavity) = P(rain|\neg cavity)P(\neg cavity) \\ P(cloudy, cavity) = P(cloudy|cavity)P(cavity) & P(cloudy, \neg cavity) = P(cloudy|\neg cavity)P(\neg cavity) \\ P(snow, cavity) = P(snow|cavity)P(cavity) & P(snow, \neg cavity) = P(snow|\neg cavity)P(\neg cavity) \end{array}$$

Chain rule

Chain rule is derived by successive applications of the product rule:

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

For example,

$$
\begin{aligned}
&\mathbb{P}(X_1, X_2, X_3, X_4) \\
&\quad = \mathbb{P}(X_1)\mathbb{P}(X_2|X_1)\mathbb{P}(X_3|X_1, X_2)\mathbb{P}(X_4|X_1, X_2, X_3) \\
&\quad = \mathbb{P}(X_4|X_3, X_2, X_1)\mathbb{P}(X_3|X_2, X_1)\mathbb{P}(X_2|X_1)\mathbb{P}(X_1)
\end{aligned}
$$

Inference by enumeration

The **Dentist Domain**:
What is the probability of a *cavity* given a *toothache*?
What is the probability of a *cavity* given the probe *catch*es?

We start with the joint distribution:

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

For any proposition q, add up the atomic events where it is true:

$$P(q) = \sum_{w:\, w \models q} P(w)$$

Computing the probability of a proposition

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

For any proposition q, add up the atomic events where it is true:

$$P(q) = \sum_{w:\, w \models q} P(w)$$

Red shows "the world" given what we know so far.
Green shows the (atomic) event we are interested in.

$$\begin{aligned} P(toothache) = & \; P(toothache, catch, cavity) + P(toothache, \neg catch, cavity) + \\ & \; P(toothache, catch, \neg cavity) + P(toothache, \neg catch, \neg cavity) \\ = & \; 0.108 + 0.012 + 0.016 + 0.064 = 0.2 \end{aligned}$$

Computing the probability of a logical sentence

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

$P(cavity \vee toothache)$
$= 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064$
$= 0.28$

Computing a conditional probability

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

Once *toothache* comes as evidence the world is restricted to those cells where *Toothache* is true as shown in red.

General idea:
Compute the distribution on the *query variable* (*Cavity*) (*Cavity*)
by fixing the *evidence variables* (*Toothache*) and
summing over all possible values of *hidden variables* (*Catch*, *Cavity*)

$$P(\neg cavity | toothache) = \frac{P(\neg cavity \land toothache)}{P(toothache)} = \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

Computing a conditional probability (cont'd)

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

General idea: Fix the *evidence variable* (*Toothache*) and sum over all possible values of *hidden variables* (*Catch* for the numerator, *Cavity* and *Catch* for the denominator)

$$P(Y=y|E=e) = \frac{P(Y{=}y,E{=}e)}{P(E{=}e)} = \frac{\sum_h P(Y{=}y,E{=}e,H{=}h)}{\sum_h P(E{=}e,H{=}h)}$$

$$P(\neg cav|tth) = \frac{P(\neg cav,tth)}{P(tth)} = \frac{\sum_h P(\neg cav,tth,H{=}h)}{\sum_h P(tth,H{=}h)}$$

$$= \frac{P(\neg cav,tth,cat)+P(\neg cav,tth,\neg cat)}{P(tth,cav,cat)+P(tth,cav,\neg cat)+P(tth,\neg cav,cat)+P(tth,\neg cav,\neg cat)}$$

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

Normalization

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

Recall that *events* are lower case, *random variables* are Capitalized
General idea: The denominator can be viewed as a *normalization constant* α

We take the probability distribution over the values of the hidden variables.

$\mathbb{P}(Cavity|toothache) = \alpha \mathbb{P}(Cavity, toothache)$
$= \alpha[\mathbb{P}(Cavity, toothache, catch) + \mathbb{P}(Cavity, toothache, \neg catch)]$
$= \alpha[< P(cavity, toothache, catch), P(\neg cavity, toothache, catch) > +$
$< P(cavity, toothache, \neg catch), P(\neg cavity, toothache, \neg catch) >]$
$= \alpha[< 0.108, 0.016 > + < 0.012, 0.064 >]$
$= \alpha[< 0.108 + 0.012, 0.016 + 0.64 >] = \alpha[< 0.12, 0.08 >]$
$=< 0.6, 0.4 >$ because the entries must add up to 1

Compute α from $\frac{1}{0.12+0.08}$

Inference by enumeration, summary

Let X be the set of all variables. Typically, we are interested in the posterior (conditional) joint distribution of the *query variables* Y given specific values e from the *evidence* variables E

Let the *hidden variables* be $H = X - Y - E$

Then the required summation of joint entries is done by summing out the hidden variables:

$$\mathbb{P}(Y|E=e) = \alpha \mathbb{P}(Y, E=e) = \alpha \sum_h \mathbb{P}(Y, E=e, H=h)$$

i.e., sum over every possible combination of values $h =< h_1, \ldots, h_n >$ of the hidden variables $H =< H_1, \ldots, H_n >$

The terms in the summation are joint entries because Y, E, and H together exhaust the set of random variables

Inference by enumeration, issues

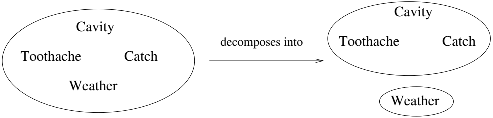
Consider that number of random variables is n, and d is the largest arity

- Worst case time complexity is $O(d^n)$
- Space complexity of $O(d^n)$, to store the entire joint distribution
- How to find the numbers for the $O(d^n)$ entries?

Independence

Random variables A and B are independent iff

$\mathbb{P}(A|B) = \mathbb{P}(A)$ or $\mathbb{P}(B|A) = \mathbb{P}(B)$ or $\mathbb{P}(A, B) = \mathbb{P}(A)\mathbb{P}(B)$



$\mathbb{P}(Toothache, Catch, Cavity, Weather)$
$= \mathbb{P}(Toothache, Catch, Cavity)\mathbb{P}(Weather)$

$2 \times 2 \times 2 \times 4 = 32$ entries reduced to $(2 \times 2 \times 2) + 4 = 12$ entries

For n independent biased coins, 2^n entries reduced to n

Absolute independence powerful but rare

E.g., dentistry is a large field with hundreds of variables, none of which are independent. What to do?

Conditional independence

Consider $\mathbb{P}(Toothache, Cavity, Catch)$

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:

$$P(catch|toothache, \neg cavity) = P(catch|\neg cavity)$$

Thus *Catch* is conditionally independent of *Toothache* given *Cavity*:

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

Or equivalently:

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

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

Conditional independence (cont'd)

Write out full joint distribution using chain rule:

$$\begin{aligned}
&\mathbb{P}(Toothache, Catch, Cavity) \\
&= \mathbb{P}(Toothache|Catch, Cavity)\mathbb{P}(Catch, Cavity) \\
&= \mathbb{P}(Toothache|Catch, Cavity)\mathbb{P}(Catch|Cavity)\mathbb{P}(Cavity) \\
&= \mathbb{P}(Toothache|Cavity)\mathbb{P}(Catch|Cavity)\mathbb{P}(Cavity)
\end{aligned}$$

In most cases, the use of conditional independence reduces the size of the representation of the joint distribution from exponential in n to linear in n.

Conditional independence is our most basic and robust from of knowledge about uncertain environments.

Bayes' rule

Product rule: $P(a \wedge b) = P(a|b)P(b) = P(b|a)P(a)$

Bayes' rule: $P(a|b) = \frac{P(b|a)P(a)}{P(b)}$

or in probability distribution form,

$$\mathbb{P}(Y|X) = \frac{\mathbb{P}(X|Y)\mathbb{P}(Y)}{\mathbb{P}(Y)} = \alpha\mathbb{P}(X|Y)\mathbb{P}(Y)$$

Useful for assessing *diagnostic* probability from *causal probability*:

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

Bayes' rule example

Useful for assessing *diagnostic* probability from *causal probability*:

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

E.g., let M be meningitis, S be stiff neck:

$$P(m|s) = \frac{P(s|m)P(m)}{P(s)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$$

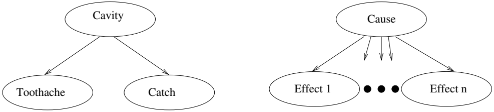
Note: posterior probability of meningitis is still very small

Bayes' rule and conditional independence

$\mathbb{P}(Cavity|toothache \wedge catch)$
$= \mathbb{P}(toothache \wedge catch|Cavity)\mathbb{P}(Cavity)/P(toothache \wedge catch)$
$= \alpha\mathbb{P}(toothache \wedge catch|Cavity)\mathbb{P}(Cavity)$
$= \alpha\mathbb{P}(toothache|Cavity)\mathbb{P}(catch|Cavity)\mathbb{P}(Cavity)$

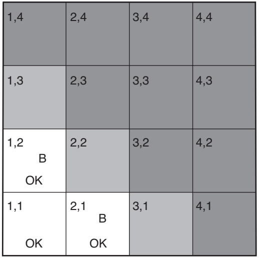
A *naive Bayes model* is a mathematical model that assumes the effects are conditionally independent, given the cause

$$\mathbb{P}(Cause, Effect_1, \ldots, Effect_n) = \mathbb{P}(Cause)\prod_i \mathbb{P}(Effect_i|Cause)$$



Naive Bayes model $\Rightarrow$ total number of parameters is linear in n

The wumpus world



The agent is navigating the wumpus world in search of gold.

The agent can perceive a breeze, a smell, or the gold.

Each cell has 0.2 probability of containing a pit.
Falling into a pit kills the agent.
The wumpus won't fall into a pit.

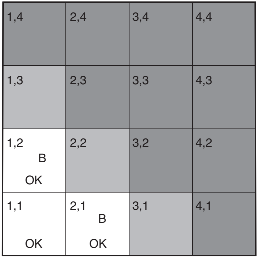
$P_{i,j} = true$ iff $[i,j]$ contains a pit.
$\forall i,j \; P(p_{i,j}) = 0.2$

Each pit causes a breeze in the adjacent cells.
$B_{i,j} = true$ iff $[i,j]$ is breezy.

There is one wumpus. Being in the same cell as the wumpus kills the agent. The cells adjacent to where the wumpus have a stench.

After finding a breeze in both [1,2] and [2,1], there is no safe place to explore.

Specifying the probability model for pits



The only breezes we care about are $B_{1,1}, B_{1,2}, B_{2,1}$. We can ignore the others.

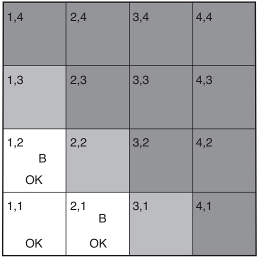
The full joint distribution is:
$\mathbb{P}(P_{1,1}, \ldots, P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$

Apply the product rule to get $P(\mathit{Effect}|\mathit{Cause})$:
$\mathbb{P}(B_{1,1}, B_{1,2}, B_{2,1}|P_{1,1}, \ldots, P_{4,4})\mathbb{P}(P_{1,1}, \ldots, P_{4,4})$

First term: 1 if pits are adjacent to breezes, 0 otherwise

Second term: Pits are placed independently. Calculate using probability 0.2 for each of the n pits. For example:
$P(p_{1,1}, \ldots, p_{4,4}) = 0.2^{16} \times 0.8^{0}$, as $n = 0$
$P(\neg p_{1,1}, \ldots, p_{4,4}) = 0.2^{15} \times 0.8^{1}$, as $n = 1$

Observations and query



We know the following facts (evidence):
$b = \neg b_{1,1} \wedge b_{1,2} \wedge b_{2,1}$
$known = \neg p_{1,1} \wedge \neg p_{1,2} \wedge \neg p_{2,1}$

The query is $\mathbb{P}(P_{1,3} | known, b)$

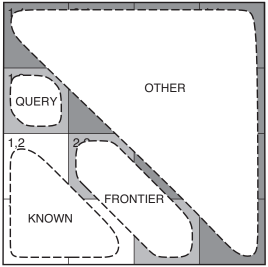
We need to sum over the hidden variables, so define $Unknown = P_{i,j}$s other than $P_{1,3}$ and $Known$

For inference by enumeration, we have
$\mathbb{P}(P_{1,3} | known, b) =$
$\alpha \sum_{unknown} \mathbb{P}(P_{1,3}, unknown, known, b)$

Exponential number of combinations based on the number of cells in *unknown*

Using conditional independence

Basic insight: Given the *frontier* squares, b is conditionally independent of the *other* hidden squares



Define $\mathit{Unknown} = \mathit{Frontier} \cup \mathit{Other}$

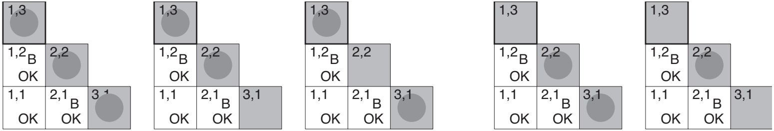
$$\mathbb{P}(b|P_{1,3}, \mathit{Known}, \mathit{Unknown})$$
$$= \mathbb{P}(b|P_{1,3}, \mathit{Known}, \mathit{Frontier}, \mathit{Other})$$
$$= \mathbb{P}(b|P_{1,3}, \mathit{Known}, \mathit{Frontier})$$

We want to manipulate the query into a form where we can use the above conditional independence.

Translating to use conditional independence

$$
\begin{aligned}
&\mathbb{P}(P_{1,3}|known, b)\\
&= \mathbb{P}(P_{1,3}, known, b)/\mathbb{P}(known, b)\\
&= \alpha\mathbb{P}(P_{1,3}, known, b)\\
&= \alpha \sum_{unknown} \mathbb{P}(P_{1,3}, known, b, unknown)\\
&= \alpha \sum_{unknown} \mathbb{P}(b|P_{1,3}, known, unknown)\mathbb{P}(P_{1,3}, known, unknown)\\
&= \alpha \sum_{frontier} \sum_{other} \mathbb{P}(b|P_{1,3}, known, frontier, other)\mathbb{P}(P_{1,3}, known, frontier, other)\\
&= \alpha \sum_{frontier} \sum_{other} \mathbb{P}(b|P_{1,3}, known, frontier)\mathbb{P}(P_{1,3}, known, frontier, other)\\
&= \alpha \sum_{frontier} \mathbb{P}(b|P_{1,3}, known, frontier) \sum_{other} \mathbb{P}(P_{1,3}, known, frontier, other)\\
&= \alpha \sum_{frontier} \mathbb{P}(b|P_{1,3}, known, frontier) \sum_{other} \mathbb{P}(P_{1,3})P(known)P(frontier)P(other)\\
&= \alpha P(known)\mathbb{P}(P_{1,3}) \sum_{frontier} \mathbb{P}(b|P_{1,3}, known, frontier) \sum_{other} P(frontier)P(other)\\
&= \alpha'\mathbb{P}(P_{1,3}) \sum_{frontier} \mathbb{P}(b|P_{1,3}, known, frontier) \sum_{other} P(frontier)P(other)\\
&= \alpha'\mathbb{P}(P_{1,3}) \sum_{frontier} \mathbb{P}(b|P_{1,3}, known, frontier)P(frontier) \sum_{other} P(other)\\
&= \alpha'\mathbb{P}(P_{1,3}) \sum_{frontier} \mathbb{P}(b|P_{1,3}, known, frontier)P(frontier)
\end{aligned}
$$

Results using conditional independence



0.2 x 0.2 = 0.04 0.2 x 0.8 = 0.16 0.8 x 0.2 = 0.16 0.2 x 0.2 = 0.04 0.2 x 0.8 = 0.16

$$\mathbb{P}(P_{1,3}|known, b) \approx < 0.31, 0.69 >$$

$$\mathbb{P}(P_{2,2}|known, b) \approx < 0.86, 0.14 >$$

Summary

Probability is a rigorous formalism for uncertain knowledge

Joint probability distribution specifies probability of every *atomic event*

Queries can be answered by *inference by enumeration* (summing over atomic events)

Can reduce combinatorial explosion using *independence* and *conditional independence*

Sources for the slides

- AIMA textbook (3^{rd} edition)
- Dana Nau's CMSC421 slides. 2010. http://www.cs.umd.edu/~nau/cmsc421/chapter13.pdf
- Mausam's CSL333 slides. 2014. http://www.cse.iitd.ac.in/~mausam/courses/csl333/spring2014/lectures/15-uncertmausam-15-uncertainty.pdf