COMP3411/9814: Artificial Intelligence

5a. Uncertainty

Alan Blair

School of Computer Science and Engineering

March 12, 2025

Outline

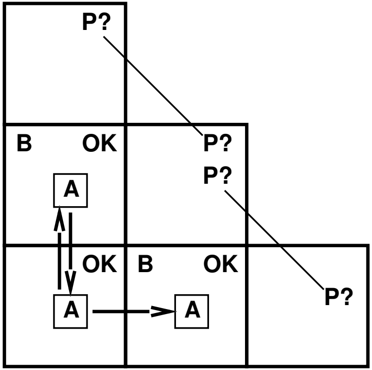
- Uncertainty
- Probability
- Bayes' Rule
- Joint Probability Distribution
- Conditional Independence

Uncertainty

In many situations, an AI agent has to choose an action based on incomplete information.

- ➛ stochastic environments (e.g. dice rolls in Backgammon)
- ➛ partial observability
  - → some aspects of environment hidden from agent
  - → robots can have noisy sensors, reporting quantities which differ from the “true” values

Uncertainty in the Wumpus World



In this situation no action is completely safe, because the agent does not know the location of the Pit(s).

Planning under Uncertainty

Let action A_t = leave for airport t minutes before flight.

Will A_t get me there on time? Problems:

- ➛ partial observability, noisy sensors
- ➛ uncertainty in action outcomes (flat tyre, etc.)
- ➛ immense complexity of modelling and predicting traffic

Hence a purely logical approach either

1) risks falsehood: "A_{30} will get me there on time", or

2) leads to conclusions that are too weak for decision making:

"A_{30} will get me there on time if there's no accident on the bridge and it doesn't rain and my tires remain intact etc. etc."

(A_{1440} might be safe but I'd have to stay overnight in the airport . . .)

Methods for handling Uncertainty

- **Default** or **Nonmonotonic** logic:
  - → Assume my car does not have a flat tire, etc.
  - → Assume A_{30} works unless contradicted by evidence
  - → Issues: What assumptions are reasonable?
    How to handle contradiction?

- **Probability**

  Given the available evidence,
  - → A_{30} will get me there on time with probability 0.04
  - → A_{90} will get me there on time with probability 0.70
  - → A_{120} will get me there on time with probability 0.95

Probabilistic Agents

We consider an Agent whose World Model consists not of a set of facts, but rather a set of **probabilities** of certain facts being true, or certain random variables taking particular values.

When the Agent makes an observation, it may update its World Model by adjusting these probabilities, based on what it has observed.

Making Decisions under Uncertainty

Suppose I believe the following:

$$\begin{aligned}
P(A_{30} \text{ gets me there on time}|\ldots) &= 0.04\\
P(A_{90} \text{ gets me there on time}|\ldots) &= 0.70\\
P(A_{120} \text{ gets me there on time}|\ldots) &= 0.95\\
P(A_{1440} \text{ gets me there on time}|\ldots) &= 0.9999
\end{aligned}$$

Which action to choose?

Depends on my **preferences** for missing flight vs. airport cuisine, etc.

Utility theory is used to represent and infer preferences

Decision theory = utility theory + probability theory

Probability Basics

Begin with a set Ω – the **sample space** (e.g. 6 possible rolls of a die)

Each $\omega \in \Omega$ is a **sample point** / **possible world** / **atomic event**

A **probability space** or **probability model** is a sample space with an assignment $P(\omega)$ for every $\omega \in \Omega$ such that

$$0 \leq P(\omega) \leq 1$$

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

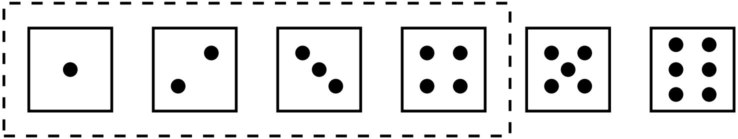
e.g. $P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{6}$.

Random Events

A **random event** A is any subset of Ω

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

e.g. $P(\text{die roll} < 5) = P(1) + P(2) + P(3) + P(4) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{2}{3}$



Random Variables

A **random variable** is a function from sample points to some range (e.g. the Reals or Booleans)

For example, $\texttt{Odd}(3) = \texttt{true}$

P induces a **probability distribution** for any random variable X:

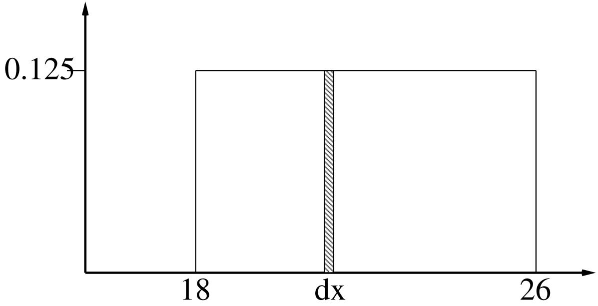
$$P(X = x_i) = \sum_{\{\omega : X(\omega) = x_i\}} P(\omega)$$

e.g., $P(\texttt{Odd} = \texttt{true}) = P(1) + P(3) + P(5) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$

Probability for Continuous Variables

For continuous variables, P is a **density**; it integrates to 1.

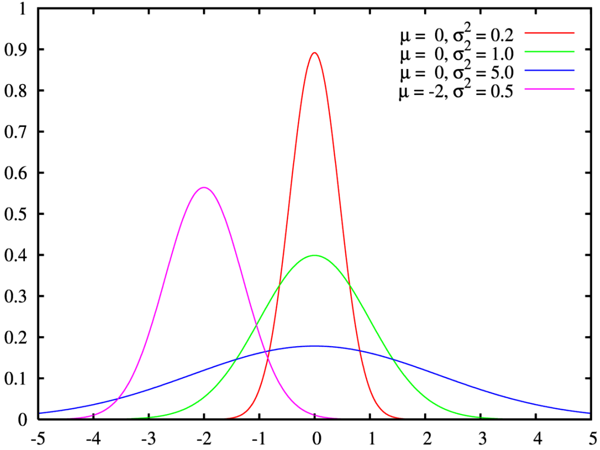
e.g. $P(X = x) = U[18, 26](x)$ = uniform density between 18 and 26



When we say $P(X = 20.5) = 0.125$, it really means

$$\lim_{dx \to 0} P(20.5 \leq X \leq 20.5 + dx)/dx = 0.125$$

Gaussian Distribution



$\mu =$ mean

$\sigma =$ standard deviation

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

Propositions

Think of a proposition as the event (set of sample points) where the proposition is true.

Given Boolean random variables A and B:

event $\quad a$ = set of sample points where $A(\omega) = \mathtt{true}$
event $\neg a$ = set of sample points where $A(\omega) = \mathtt{false}$
event $a \wedge b$ = points where $A(\omega) = \mathtt{true}$ and $B(\omega) = \mathtt{true}$

With Boolean variables, sample point = propositional logic model
e.g., $A = \mathtt{true}$, $B = \mathtt{false}$, or $a \wedge \neg b$.

Proposition = disjunction of atomic events 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 my tooth?)

(`Cavity = true`) is a proposition, also written `Cavity`

Discrete random variables (finite or infinite)

e.g., `Weather` is one of $\langle$`sunny,rain,cloudy,snow`$\rangle$

(`Weather = rain`) is a proposition

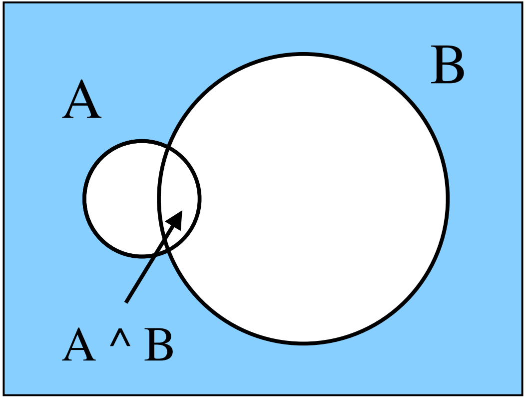
Values must be exhaustive and mutually exclusive.

Continuous random variables (bounded or unbounded)

e.g. `Temp` = 21.6; also allow, e.g. `Temp` < 22.0

Arbitrary Boolean combinations of basic propositions.

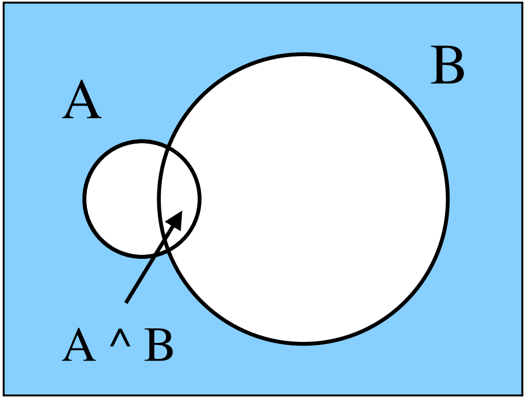
Probability and Logic



Logically related events must have related probabilities

For example, $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

Conditional Probability



If $P(B) \neq 0$, then the **conditional probability** of A given B is

$$P(A \,|\, B) = \frac{P(A \wedge B)}{P(B)}$$

Bayes' Rule

The formula for conditional probability can be manipulated to find a relationship when the two variables are swapped:
$P(a \wedge b) = P(a \,|\, b)P(b) = P(b \,|\, a)P(a)$

$$\rightarrow \textbf{Bayes' rule } P(a \,|\, b) = \frac{P(b \,|\, a)P(a)}{P(b)}$$

This is often useful for assessing the probability of an underlying **Cause** after an **Effect** has been observed:

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

Example: Medical Diagnosis

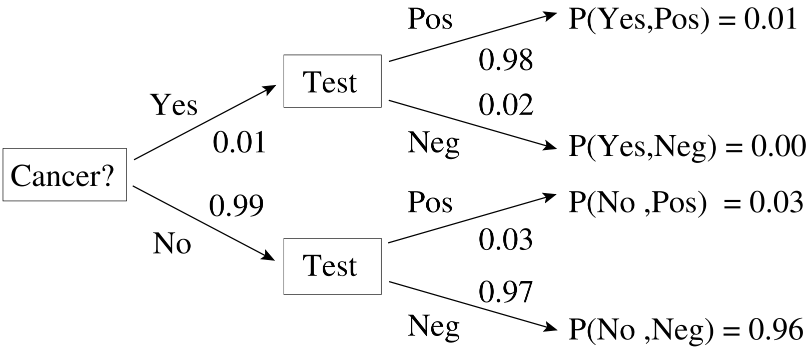
Question: Suppose we have a test for a type of cancer which occurs in 1% of patients. The test has a sensitivity of 98% and a specificity of 97%.
If a patient tests positive, what is the probability that they have the cancer?

Answer: There are two random variables: Cancer (true or false) and Test (positive or negative). The probability is called a **prior**, because it represents our estimate of the probability **before** we have done the test (or made some other observation).

The **sensitivity** and **specificity** are interpreted as follows:

$$P(\texttt{positive} \,|\, \texttt{cancer}) = 0.98, \quad \text{and} \quad P(\texttt{negative} \,|\, \neg\texttt{cancer}) = 0.97$$

Bayes' Rule for Medical Diagnosis



$$
\begin{aligned}
P(\texttt{cancer}\,|\,\texttt{positive}) &= \frac{P(\texttt{positive}\,|\,\texttt{cancer})P(\texttt{cancer})}{P(\texttt{positive})} \\
&= \frac{0.98 * 0.01}{0.98 * 0.01 + 0.03 * 0.99} = \frac{0.01}{0.01 + 0.03} = \frac{1}{4}
\end{aligned}
$$

Example: Light Bulb Defects

Question: You work for a lighting company which manufactures 60% of its light bulbs in Factory A and 40% in Factory B. One percent of the light bulbs from Factory A are defective, while two percent of those from Factory B are defective. If a random light bulb turns out to be defective, what is the probability that it was manufactured in Factory A?

Answer: There are two random variables: Factory (A or B) and Defect (Yes or No).

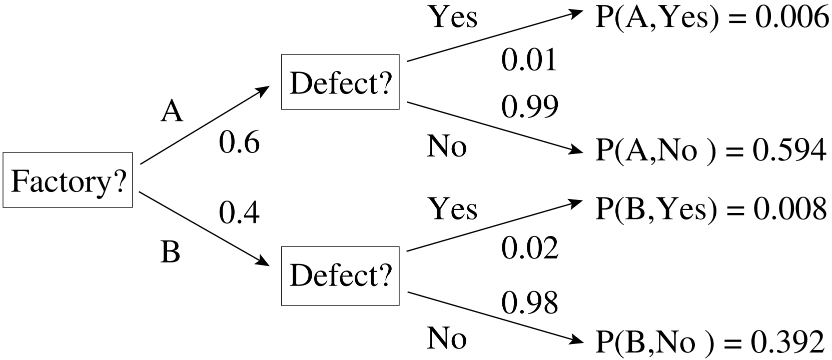
In this case, the prior is:

$$P(A) = 0.6, \qquad P(B) = 0.4$$

The conditional probabilities are:

$$P(\texttt{defect} \mid \texttt{A}) = 0.01, \quad \text{and} \quad P(\texttt{defect} \mid \texttt{B}) = 0.02$$

Bayes' Rule for Light Bulb Defects



$$
\begin{aligned}
P(\texttt{A} \mid \texttt{defect}) &= \frac{P(\texttt{defect} \mid \texttt{A})P(\texttt{A})}{P(\texttt{defect})} \\
&= \frac{0.01 * 0.6}{0.01 * 0.6 + 0.02 * 0.4} = \frac{0.006}{0.006 + 0.008} = \frac{3}{7}
\end{aligned}
$$

Prior Probability

Prior or **Unconditional** probabilities of propositions

e.g. $P(\texttt{Cavity} = \texttt{true}) = 0.1$ and $P(\texttt{Weather} = \texttt{sunny}) = 0.72$

correspond to belief **prior** to arrival of any (new) evidence.

Probability distribution gives values for all possible assignments:

$P(\texttt{Weather}) = \langle 0.72, 0.1, 0.08, 0.1\rangle$ (normalized, i.e., sums to 1)

Example: Tooth Decay

Assume you live in a community where, at any given time, 20% of people have a **cavity** in one of their teeth which needs a filling from the dentist.

$$P(\texttt{cavity}) = 0.2$$

If you have a toothache, suddenly you will think it is much more likely that you have a cavity, perhaps as high as 60%. We say that the **conditional probability** of cavity, given toothache, is 0.6, written as follows:

$$P(\texttt{cavity} \,|\, \texttt{toothache}) = 0.6$$

If you go to the dentist, they will use a small hook-shaped instrument called a probe, and check whether this probe can **catch** on the back of your tooth. If it does catch, this information will increase the probability that you have a cavity.

Joint Probability Distribution

We assume there is some underlying joint probability distribution over the three random variables Toothache, Cavity and Catch, which we can write in the form of a table:

	$toothache$		$\neg\, toothache$	
	$catch$	$\neg\, catch$	$catch$	$\neg\, catch$
$cavity$	**.108**	**.012**	**.072**	**.008**
$\neg\, cavity$	**.016**	**.064**	**.144**	**.576**

Note that the sum of the entries in the table is 1.0.

For any proposition ϕ, sum the atomic events where it is true:

$$P(\phi) = \sum_{\omega:\omega \models \phi} P(\omega)$$

Inference by Enumeration

Start with the joint distribution:

	$toothache$		$\neg\ toothache$	
	$catch$	$\neg\ catch$	$catch$	$\neg\ catch$
$cavity$	**.108**	**.012**	**.072**	**.008**
$\neg\ cavity$	**.016**	**.064**	**.144**	**.576**

For any proposition ϕ, sum the atomic events where it is true:

$P(\phi) = \sum_{\omega:\omega\models\phi} P(\omega)$

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

Inference by Enumeration

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

For any proposition ϕ, sum the atomic events where it is true:

$P(\phi) = \sum_{\omega:\omega\models\phi} P(\omega)$

$P(\texttt{cavity} \lor \texttt{toothache})$

$$= 0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28$$

Conditional Probability by Enumeration

	toothache		$\neg$ *toothache*	
	catch	$\neg$ *catch*	*catch*	$\neg$ *catch*
cavity	**.108**	**.012**	**.072**	**.008**
$\neg$ *cavity*	**.016**	**.064**	**.144**	**.576**

$$\begin{aligned} P(\neg\texttt{cavity} \mid \texttt{toothache}) &= \frac{P(\neg\texttt{cavity} \wedge \texttt{toothache})}{P(\texttt{toothache})} \\ &= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4 \end{aligned}$$

Independent Variables

Let's consider the joint probability distribution for Cavity and Weather.

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

Note that:

$$P(\texttt{cavity} \,|\, \texttt{Weather} = \texttt{sunny}) = \frac{0.144}{0.144 + 0.576} = 0.2 = P(\texttt{cavity})$$

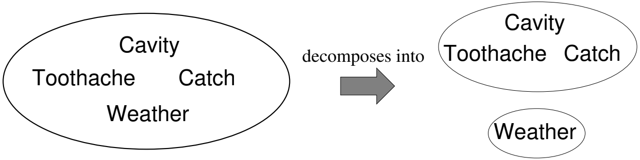
In other words, learning that the Weather is sunny has no effect on the probability of having a cavity (and the same for rain, cloudy and snow).

We say that `Cavity` and `Weather` are **independent** variables.

Independence

A and B are **independent** if

$P(A|B) = P(A)$ or $P(B|A) = P(B)$ or $P(A, B) = P(A)P(B)$



If variables not independent, would need 32 items in probability table.

Because Weather is independent of the other variables, only need two smaller tables, with a total of 8+4=12 items.

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

(Note: the number of free parameters is slightly less, because the values in each table must sum to 1).

Conditional independence

The variables Toothache, Cavity and Catch are not independent.
But, they do exhibit **conditional independence.**

If you have a cavity, the probability that the probe will catch is 0.9,
no matter whether you have a toothache or not.

If you don't have a cavity, the probability that the probe will catch is 0.2,
regardless of whether you have a toothache. In other words,

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

We say that `Catch` is **conditionally independent** of `Toothache` given `Cavity`.

Conditional independence

This conditional independence reduces the number of free parameters from 7 down to 5.

For larger problems with many variables, deducing this kind of conditional independence among the variables can reduce the number of free parameters substantially, and allow the Agent to maintain a simpler World Model.

Equivalent statements:

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

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

Bayes' Rule and Conditional Independence

$$\begin{aligned} & P(\texttt{cavity}, \texttt{toothache}, \texttt{catch}) \\ & \quad = \quad P(\texttt{toothache} \mid \texttt{catch}, \texttt{cavity}) P(\texttt{catch} \mid \texttt{cavity}) P(\texttt{cavity}) \\ & \quad = \quad P(\texttt{toothache}|\texttt{cavity}) P(\texttt{catch}|\texttt{cavity}) P(\texttt{cavity}) \end{aligned}$$

This is an example of a **Naive Bayes** model:

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

Total number of parameters is **linear** in n

Wumpus World

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

What is the probability of a Pit in (1,3) ? What about (2,2) ?

To answer this, we need a "prior" assumption about the placement of Pits.
We will assume a 20% chance of a Pit in each square at the beginning of the game (independent of what Pits are in the other squares).

Specifying the Probability Model

We will use $B_{i,j}$ to indicate a Breeze in square (i, j),
and $\mathrm{Pit}_{i,j}$ to indicate a Pit in square (i, j).
We use `known` to represent what we know, i.e.

$$B_{1,2} \wedge B_{2,1} \wedge \neg B_{1,1} \wedge \neg \mathrm{Pit}_{1,2} \wedge \neg \mathrm{Pit}_{2,1} \wedge \neg \mathrm{Pit}_{1,1}$$

We use `Unknown` to represent the joint probability of Pits in all the other squares, i.e.

$$P(\texttt{Unknown}) = P(\mathrm{Pit}_{1,4}, \ldots, \mathrm{Pit}_{4,1})$$

We divide `Unknown` into `Fringe` and `Other`, where

$$P(\texttt{Fringe}) = P(\mathrm{Pit}_{1,3}, \mathrm{Pit}_{2,2}, \mathrm{Pit}_{3,1})$$

and `Other` is all the other variables.

Manipulating Probabilities

$$P(\mathrm{Pit}_{1,3} | \texttt{known}) = \sum_{\texttt{unknown}} P(\mathrm{Pit}_{1,3}, \texttt{unknown} \,|\, \texttt{known})$$

$$= \sum_{\texttt{fringe}} \sum_{\texttt{other}} P(\mathrm{Pit}_{1,3}, \texttt{fringe}, \texttt{other} \,|\, \texttt{known})$$

$$= \sum_{\texttt{fringe}} \sum_{\texttt{other}} P(\mathrm{Pit}_{1,3} | \texttt{fringe}, \texttt{other}, \texttt{known})\, P(\texttt{fringe}, \texttt{other} \,|\, \texttt{known})$$

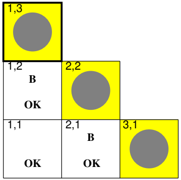
$$= \sum_{\texttt{fringe}} P(\mathrm{Pit}_{1,3} | \texttt{fringe}) \sum_{\texttt{other}} P(\texttt{fringe}, \texttt{other} \,|\, \texttt{known})$$

$$= \sum_{\texttt{fringe}} P(\mathrm{Pit}_{1,3} | \texttt{fringe}) \sum_{\texttt{other}} \frac{P(\texttt{known} \,|\, \texttt{fringe}, \texttt{other}) P(\texttt{fringe}, \texttt{other})}{P(\texttt{known})}$$

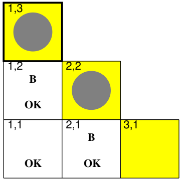
Note: have used the fact that $P_{1,3}$ is independent of `other` and `known`, given `fringe`.

Fringe Models

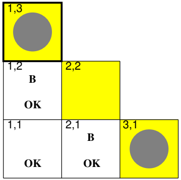
Let's denote by F the set of fringe models compatible with the known facts:



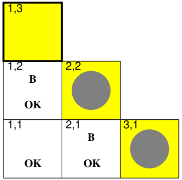
0.2 x 0.2 = 0.04



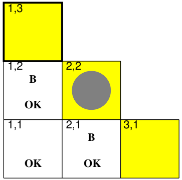
0.2 x 0.8 = 0.16



0.8 x 0.2 = 0.16



0.2 x 0.2 = 0.04



0.2 x 0.8 = 0.16

$P(\texttt{known}\,|\,\texttt{fringe},\texttt{other}) = 0$ outside F, so $P(\mathrm{Pit}_{1,3}|\,\texttt{known})$ reduces to:

$$\frac{\displaystyle\sum_{\texttt{fringe}\in F} P(\mathrm{Pit}_{1,3}|\,\texttt{fringe}) \sum_{\texttt{other}} P(\texttt{known}\,|\,\texttt{fringe},\texttt{other})P(\texttt{fringe},\texttt{other})}{P(\texttt{known})}$$

Note also that

$$P(\texttt{known}) = \sum_{\texttt{fringe}\in F}\sum_{\texttt{other}} P(\texttt{known}\,|\,\texttt{fringe},\texttt{other})P(\texttt{fringe},\texttt{other})$$

Using the Prior

Because of the prior, `other` and `fringe` become independent, and `known` becomes independent of `other`, given `fringe`.

$$P(\texttt{known} \,|\, \texttt{fringe}, \texttt{other}) = P(\texttt{known} \,|\, \texttt{fringe}) = 1, \quad \text{for } \texttt{fringe} \in F, \text{ so}$$

$$P(\texttt{known}) = \sum_{\texttt{fringe} \in F} P(\texttt{fringe}) = (0.2)^3 + 3 \times (0.2)^2(0.8) + (0.2)(0.8)^2$$

$$= 0.008 + 0.032 + 0.032 + 0.032 + 0.128 = 0.232$$

The numerator includes only those models for which $\text{Pit}_{1,3}$ is true, i.e.

$$P(\text{Pit}_{1,3} \,|\, \texttt{known}) = \frac{0.008 + 0.032 + 0.032}{0.232} = \frac{9}{29} \simeq 0.310$$

In a similar way,

$$P(\text{Pit}_{2,2} \,|\, \texttt{known}) = \frac{0.008 + 0.032 + 0.032 + 0.128}{0.232} = \frac{25}{29} \simeq 0.862$$

Summary

- Probability is a rigorous formalism for uncertain knowledge
- **Joint probability distribution** specifies probability of every **atomic event**
- Queries can be answered by summing over atomic events
- For nontrivial domains, we must find a way to reduce the joint distribution size
- **Independence** and **conditional independence** provide the tools