Introduction to AI

Chapter13
Uncertainty

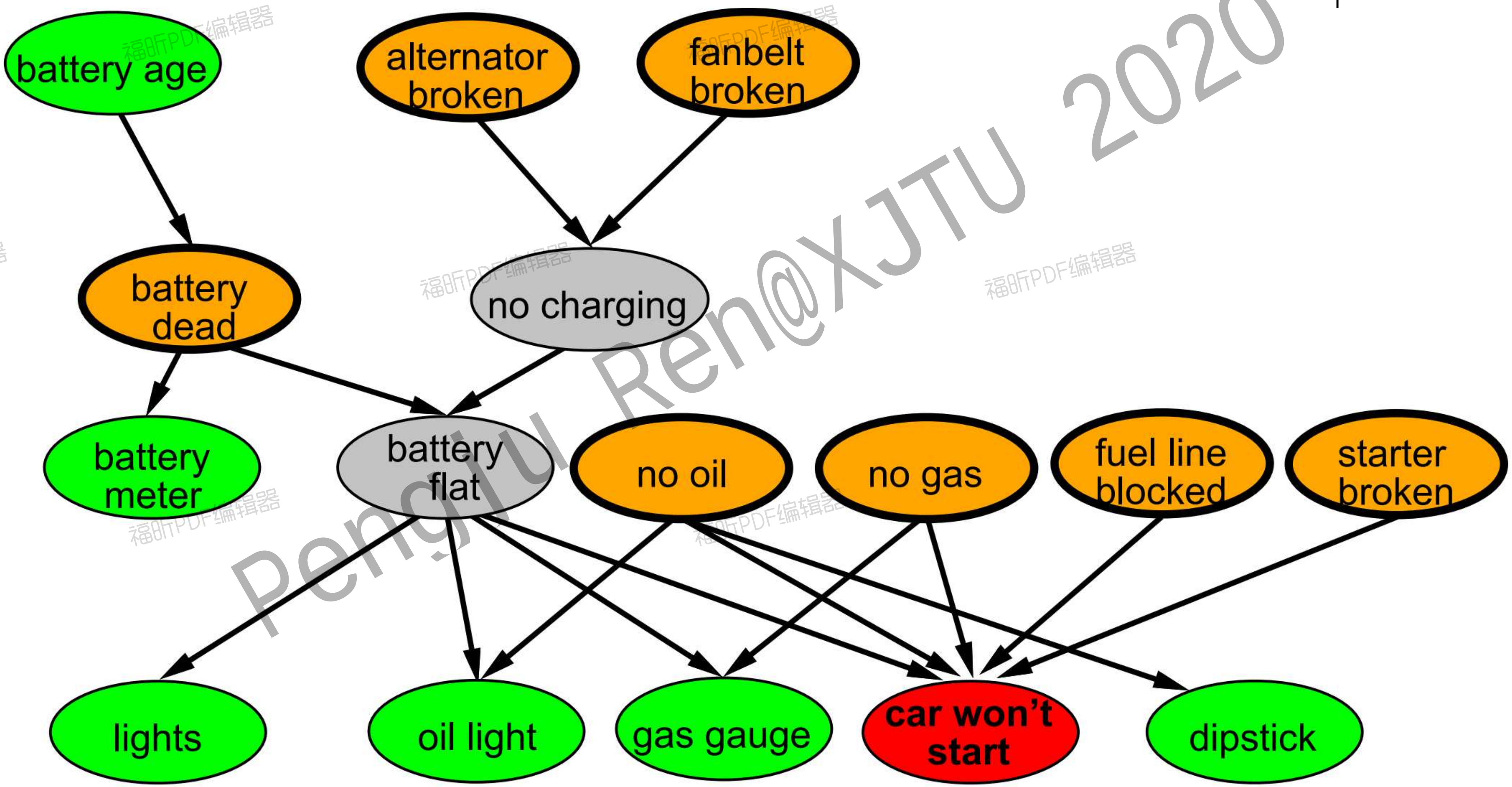
Pengju Ren@IAIR

Outline

- Uncertainty
- Probability
- Syntax and Semantics
- Inference
- Independence and Bayes’ Rule

机器人搬运实验

Example: Car diagnosis



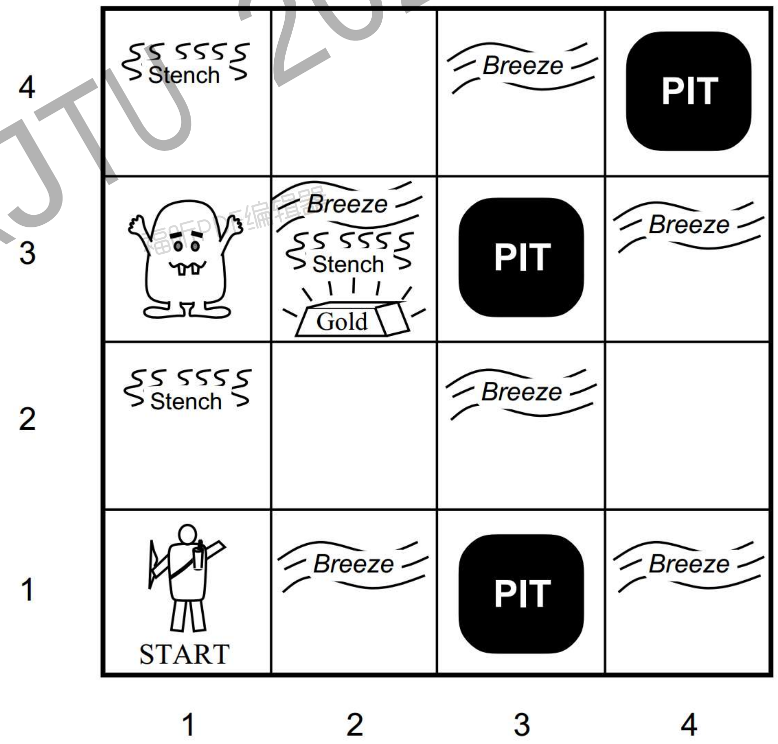
Wumpus World

Environment

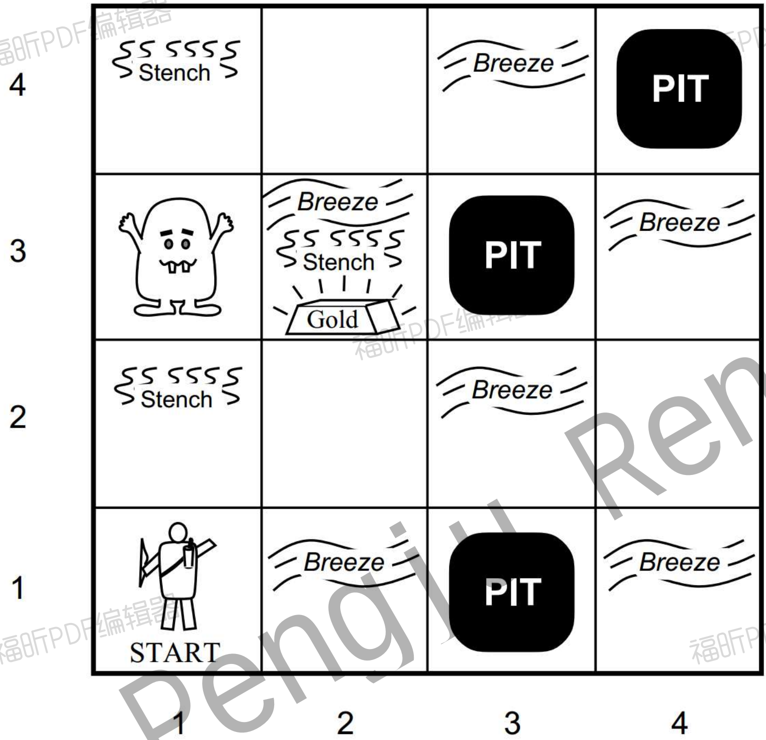
- Squares adjacent to wumpus are smelly
- Squares adjacent to pit are breezy
- Glitter iff gold is in the same square
- Shooting kills wumpus if you are facing it
- Shooting uses up the only arrow
- Grabbing picks up gold if in same square
- Releasing drops the gold in same square

Actuators Left turn, Right turn,
Forward, Grab, Release, Shoot

Sensors Breeze, Glitter, Smell



E.g. 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

$P_{i,j} = true$ iff [i,j] contains a pit

$B_{i,j} = true$ iff [i,j] is breezy

Include only $B_{1,1}, B_{1,2}, B_{2,1}$ in the probability model

Why your girlfriend is angry ?

1. Because she uncovered a leftover profile Weibo picture of you and your ex.
2. Because you didn't respond to her complaint the way she wanted you to.
3. Because she made something up in her head that she wanted you to do, and you didn't do it.
4. You peek other girls when you are hanging out.
5. She had a new hair style, but you didn’t notice that.
6. ….

Probability

Probabilistic assertions summarize effects of

laziness: failure to enumerate exceptions, qualifications, etc.

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

Subjective or Bayesian probability:

Probabilities relate propositions to one's own state of knowledge

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

These are not claims of a "probabilistic tendency" in the current situation (but might be learned from past experience of similar situations)

Probabilities of propositions change with new evidence:

e.g., p(A_{25} *|no reported accidents, 5 a.m.) = 0.15*

Making decisions under uncertainty

Suppose I believe the following:

p(A_{25} gets me there on time|. . .) = 0.04
p(A_{90} gets me there on time|. . .) = 0.70
p(A_{120} gets me there on time|. . .) = 0.95
p(A_{1440}gets me there on time|. . .) = 0.9999

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

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

Often in AI applications, the sample points are defined by the values of a set of random variables, i.e., the sample space is the Cartesian product of the ranges of the variables

With Boolean variables, sample point = propositional logic model

e.g., $A = true$, $B = false$, or $a \wedge \neg b$.

Proposition = disjunction of atomic events in which it is true

e.g., $(a \vee b) \equiv (\neg a \wedge b) \vee (a \wedge \neg b) \vee (a \wedge b)$

$\Rightarrow p(a \vee b) = p(\neg a \wedge b) + p(a \wedge \neg b) + p(a \wedge b)$

Prior probability

Prior or **unconditional probabilities** of propositions

e.g., $p(Cavity = true) = 0.1$ and *$p(Weather = sunny) = 0.72$*

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

Probability distribution gives values for all possible assignments:

$p(Weather) = (0.72, 0.1, 0.08, 0.1)$ (normalized, i.e., sums to 1)

Joint probability distribution for a set of r.v.s gives the probability of every atomic event on those r.v.s (i.e., every sample point)

$p(Weather, Cavity)$ = a 4×2 matrix of values:

$Weather =$	$sunny$	$rain$	$cloudy$	$snow$
$Cavity = true$	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 point

Probability basics

Begin with a set Ω—the sample space

e.g., 6 possible rolls of a die.

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

A probability space or probability model is a sample space with an assignment *p(ω)* for every *ω ∈ Ω* s.t.

$0 \le p(\omega) \le 1$

$\Sigma_\omega p(\omega) = 1$

e.g., p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = 1/6.

An event A is any subset of Ω

$$p(A) = \sum_{\{w \in A\}} p(w)$$

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

Probability basics

Marginalization: $p(x) = \sum_{Y} p(x, y = Y)$

Chain rule: $p(a, b) = p(a|b)p(b) = p(b|a)p(a)$

$$p(a, b, c) = p(a|b, c)p(b, c) \\ = p(a|b, c)p(b|c)p(c)$$

Bayes rule: $p(a|b) = \frac{p(a, b)}{p(b)} = \frac{p(b|a)p(a)}{p(b)} = \frac{p(b|a)p(a)}{\sum_a p(a, b)}$

Quiz 1

Suppose there are only two weather: Sunny and Rainy
And weather is only depended on pervious day's condition.

p(D1=Sunny) = 0.9
p(D2=Sunny|D1=Sunny) = 0.8
p(D2=Rainy|D1=Sunny) = ?
p(D2=Sunny|D1=Rainy)= 0.6
p(D2=Rainy|D1=Rainy)= ?

p(D2=Sunny)=? p(D3=Sunny)= ?

Quiz 2

C: Coronavirus;　+ or -: Test

$p(C)=0.01$, $p(+|C)=0.9$, $p(+|\neg C)=0.2$

$p(\neg C)=0.99$, $p(-|C)=0.1$, $p(-|\neg C)=0.8$

Joint Probabilities

$p(+,C)=0.009$

$p(-,C)=0.001$

$p(+,\neg C)=p(+|\neg C)*p(\neg C)=0.198$

$p(-,\neg C)=p(-|\neg C)*p(\neg C)=0.792$

Question

$p(C|+)=\ ?$

$p(\neg C\ |-)=\ ?$

Inference by enumeration

With the joint distribution:

	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:

$$p(toothache) = 0.108 + 0.012 + 0.016 + 0.064 = 0.2$$

Inference by enumeration

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:

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

Inference by enumeration

With the joint distribution:

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

Can also compute conditional probabilities:

$$p(\neg cavity|toothache) = \frac{p(\neg cavity, toothache)}{p(toothache)}$$

$$= \frac{0.072 + 0.008}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

$$p(cavity|toothache) = \frac{p(cavity, toothache)}{p(toothache)}$$

$$= \frac{0.108 + 0.012}{0.108 + 0.012 + 0.016 + 0.064} = 0.6$$

$$\mathrm{p}(\neg cavity|toothache) : \mathrm{p}(cavity|toothache) = <?:?>$$

Normalization

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	$\boxed{.108}$	$\boxed{.012}$	.072	.008
¬ cavity	$\boxed{.016}$	$\boxed{.064}$	.144	.576

Denominator can be viewed as a normalization constant α

$$\mathbf{P}(Cavity|toothache) = \alpha\mathbf{P}(Cavity, toothache)$$
$$= \alpha[\mathbf{P}(Cavity, toothache, catch) + \mathbf{P}(Cavity, toothache, \neg catch)]$$
$$= \alpha[< 0.108, 0.016 > + < 0.012, 0.064 >] = \alpha < 0.12, 0.08 > = < 0.6, 0.4 >$$

General idea: compute distribution on query variable by fixing evidence variables and summing over hidden variables

Inference by enumeration, contd.

Let X be all the variables. Typically, we want
the posterior joint distribution of the query variables Y
given specific values e for 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:

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

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

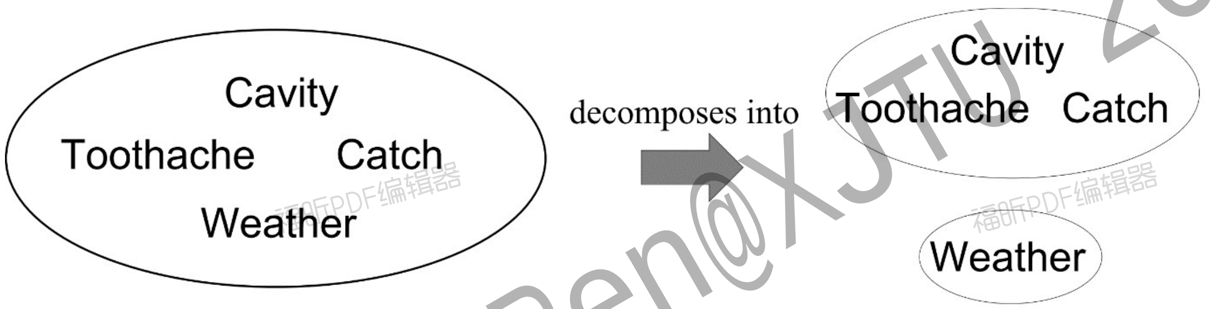
Obvious problems:

1) Worst-case time complexity $O(d^m)$ where d is the largest arity
2) Space complexity $O(d^n)$ to store the joint distribution

Independence

A and B are independent iff

$$P(A|B) = P(A) \qquad P(B|A) = P(B) \qquad P(A,B) = P(A)P(B)$$



Weather = {Sunny, Rainy, Cloudy, Snow}
Cavity, Toothache, Catch = {True, False}

32 entries reduced to 8 and 4; for n independent biased coins, 2^n → n
Absolute independence powerful but rare

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

Conditional independence

$P(Toothache, Cavity, Catch)$ **has** $2^3 - 1 = 7$ **independent entries**

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:

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

Catch is conditionally independent of Toothache given Cavity:

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

Equivalent statements:

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

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

Conditional independence

$P(MIT, \ Stanford, GPA)$ has $2^3 - 1 = 7$ independent entries

If I have a high GPA, the probability that I got MIT offer it doesn't depend on whether I already have a Stanford offer:

$$p(\text{mit}|\text{stanford}, \text{gpa}) = p(\text{mit}|gpa)$$

The same independence holds if I haven't got a high GPA:

$$p(\text{mit}|stanford, \neg\text{gpa}) = p(\text{mit}|\neg\text{gpa})$$

MIT is conditionally independent of Stanford given GPA:

$$P(MIT|Stanford, \ GPA) = P(MIT|GPA)$$

Equivalent statements:

$$P(Stanford|MIT, \ GPA) = P(Stanford|GPA)$$

$$P(MIT, Stanford|GPA) = P(MIT|GPA)P(Stanford|GPA)$$

Conditional independence contd.

Write out full joint distribution using chain rule:

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

I.e., 2 + 2 + 1 = 5 independent numbers (equations 1 and 2 remove 2)

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 form of knowledge about uncertain environments.

Conditional independence contd.

Write out full joint distribution using chain rule:

$$\begin{aligned} P(MIT, Stanford, GPA) &= P(MIT|Stanford, GPA)P(Stanford, GPA) \\ &= P(MIT|Stanford, GPA)P(Stanford/GPA)P(GPA) \\ &= P(MIT|GPA)P(Stanford/GPA)P(GPA) \end{aligned}$$

I.e., 2 + 2 + 1 = 5 independent numbers (equations 1 and 2 remove 2)

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 form of knowledge about uncertain environments.

Bayes' Rule

Product rule $p(a \wedge b) = p(a|b)p(b) = p(b|a)p(a)$

$$p(a|b) = \frac{p(b|a)p(a)}{p(b)}$$

Or in distribution form

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

Useful for assessing diagnostic probability from causal probability:

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

E.g., let M be Cancer, S be (+/-) test results:

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

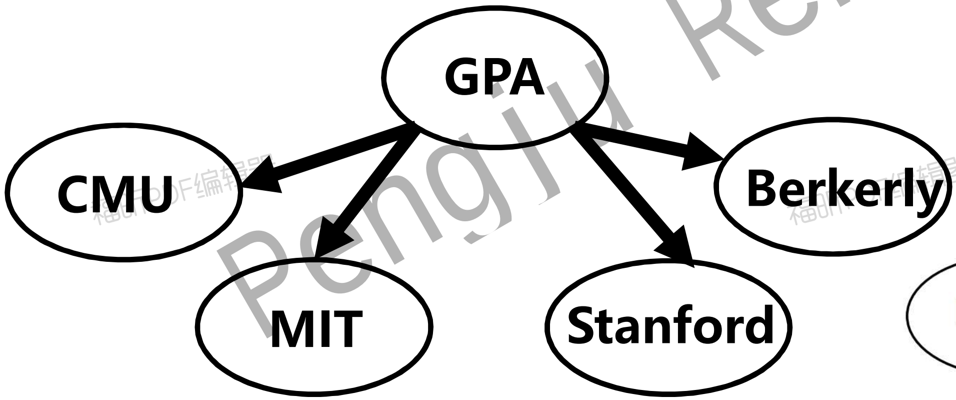
Note: posterior probability of cancer still very small!

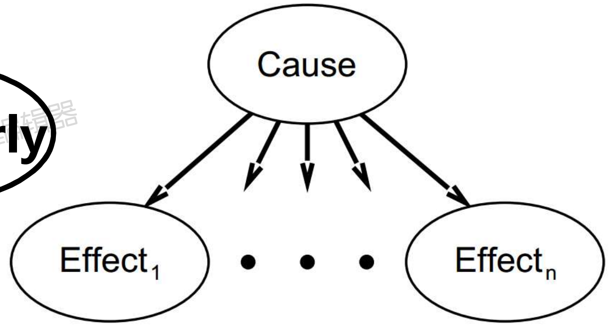
Bayes' Rule and conditional independence

$$
\begin{aligned}
P(MIT, Stanford, GPA) &= P(MIT \mid Stanford, GPA)P(Stanford, GPA) \\
&= P(MIT \mid Stanford, GPA)P(Stanford / GPA)P(GPA) \\
&= P(MIT \mid GPA)P(Stanford / GPA)P(GPA)
\end{aligned}
$$

This is an example of a naive Bayes model:

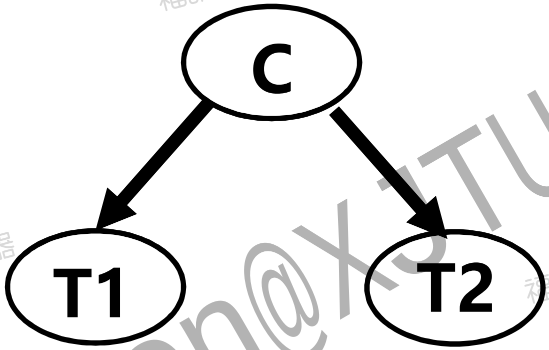
$$P(\mathbf{Cause}, Effect_1, \ldots, Effect_n) = P(Cause)\Pi_i P(Effect_i \mid Cause)$$





Total number of parameters is linear in n

Quiz 3



To detect Coronavirus, we run two conditional dependent Test: T1 and T2

$p(c)=0.01$, $P(+|c)=0.9$, $p(-|\neg c)=0.8$ **(+,- for T1 and T2 are the same)**

$$p(T2=+|T1=+) = ?$$

Answer Quiz 3

$$p(T2=+|T1=+)=\frac{p(T2=+,T1=+)}{p(T1=+)}=\frac{p(T2=+,T1=+,C)}{p(T1=+)}+\frac{p(T2=+,T1=\cdot}{p(T1=+)}$$

$$=\frac{p(T2=+,T1=+,C)}{p(T1=+,C)}*\frac{p(T1=+,C)}{p(T1=+)}+\frac{p(T2=+,T1=+,\neg C)}{p(T1=+,\neg C)}*\frac{p(T1=+,\neg (}{p(T1=+)}$$

$$=p(T2=+|T1=+,C)p(C|T1=+)+p(T2=+|T1=+,\neg C)p(\neg C|T1=+)$$

$$=p(T2=+|C)p(C|T1=+)+p(T2=+|\neg C)p(\neg C|T1=+)$$

$$p(C|+)=\frac{p(+|C)*p(C)}{p(+)}=\frac{p(+|C)*p(C)}{p(+|C)*p(C)+p(+|\neg C)*p(\neg C)}=\frac{0.01*0.9}{0.01*0.9+0.2*0.99}=0.043$$

$$p(\neg C|+)=\frac{p(+|\neg C)*p(\neg C)}{p(+)}=\frac{p(+|\neg C)*p(\neg C)}{p(+|C)*p(C)+p(+|\neg C)*p(\neg C)}=\frac{0.2*0.99}{0.01*0.9+0.2*0.99}=0.957$$

So， $p(T2=+|T1=+)=$ 0.9*0.043+0.2*0.957=0.0387+0.1914=0.23;

$$p(+)=\frac{p(+|C)p(C)}{p(C|+)}=\frac{0.9*0.01}{0.435}=0.2069$$

0.23/0.2069=1.112， Increase 11.2%

Answer Quiz 3

$$P(T2=+|T1=+)=\frac{P(T1=+,T2=+)}{P(T1=+)}=\frac{P(T1=+,T2=+,C)+P(T1=+,T2=+,\neg C)}{P(T1=+)}$$

$$=\frac{P(T1=+|C)P(T2=+|C)P(C)+P(T1=+|\neg C)P(T2=+|\neg C)P(\neg C)}{P(T1=+|C)P(C)+P(T1=+|\neg C)P(\neg C)}$$

$$=\frac{0.9*0.9*0.01+0.2*0.2*0.99}{0.8*0.01+0.2*0.99}=\frac{0.0477}{0.20691}=0.23$$

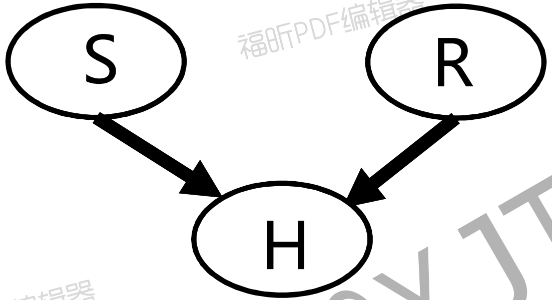
Homework (12am before next Class):

To detect Coronavirus, we run two conditional dependent Test: T1 and T2

$p(c)=0.01$, $P(+|c)=0.9$, $p(+|\neg c)=0.2$ **(+,- for T1 and T2 are the same)**

1) How many '+' results in a row, can determine $P(c) > 0.9$?
2) If $P(+|c)=0.9$, $P(+|\neg c)=0.1$, How many '+' results in a row, can determine $P(c)>0.9$?
3) If $P(+|c)=0.95$, $P(+|\neg c)=0.2$, How many '+' results in a row, can determine $P(c)>0.9$?

Quiz 4: Different type



S: The weather is Sunny;
R: I get promoted/Rise;
H: Happy

$p(S)=0.7$
$p(R)=0.01$

$p(H|S,R) = 1$
$p(H|\neg S, R) = 0.9$
$p(H|S,\neg R) = 0.7$
$p(H|\neg S, \neg R)=0.1$

Question 1: $p(R|S) = ?$

Question 2: $p(R|H,S) = ?$

Question 3: $p(R|H) = ?$

Question 4: $p(R|H, \neg S) = ?$

Answers: Different type

Q2:

$$p(R\mid H,S)=\frac{p(H,R,S)}{p(H,S)}=\frac{p(H\mid R,S)p(R,S)}{p(H,S,R)+p(H,S,\neg R)}$$

$$=\frac{p(H\mid R,S)p(R,S)/p(S)}{p(H,S,R)+p(H,S,\neg R)/p(S)}=\frac{p(H\mid R,S)p(R)}{p(H\mid S,R)p(R)+p(H\mid S,\neg R)/p(\neg R)}$$

$$=\frac{1*0.01}{1*0.01+0.7*0.99}=0.0142$$

Q3:

$$p(R\mid H)=\frac{p(R,H)}{p(H)}=\frac{p(R,H,\neg S)+p(R,H,S)}{p(H)}$$

$$=\frac{p(H\mid R,\neg S)p(R,\neg S)+p(H\mid R,S)p(R,S)}{\sum_{\substack{i=S,\neg S\\ j=R,\neg R}} p(H)p(H\mid i,j)}$$

$$=\frac{0.9*0.01*0.3+1*0.01*0.7}{1*0.7*0.01+0.9*0.3*0.01+0.7*0.7*0.99+0.1*0.3*0.99}=\frac{0.0027+0.007}{0.5245}=0.0185$$

Answers: Different type

Q4:

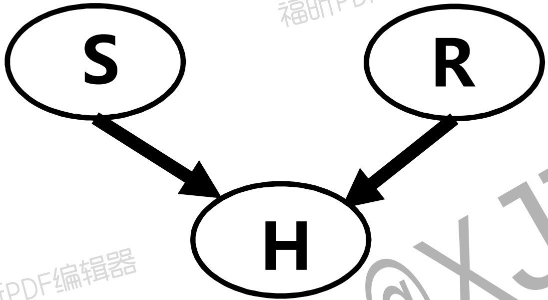
$$p(R \mid H, \neg S) = \frac{p(R, H, \neg S)}{p(H, \neg S)} = \frac{p(H \mid R, \neg S)\, p(R, \neg S)}{p(H, \neg S)}$$

$$= \frac{p(H \mid R, \neg S)\, p(R, \neg S)}{p(H, \neg S, R) + P(H, \neg S, \neg R)}$$

$$= \frac{0.9 * 0.01 * 0.3}{0.9 * 0.01 * 0.3 + p(H \mid \neg R, \neg S)\, p(\neg R, \neg S)}$$

$$= \frac{0.0027}{0.0027 + 0.0297} = 0.08333$$

Second Thoughts about Conditional Dependence



$p(R|S) = P(R)=0.01$

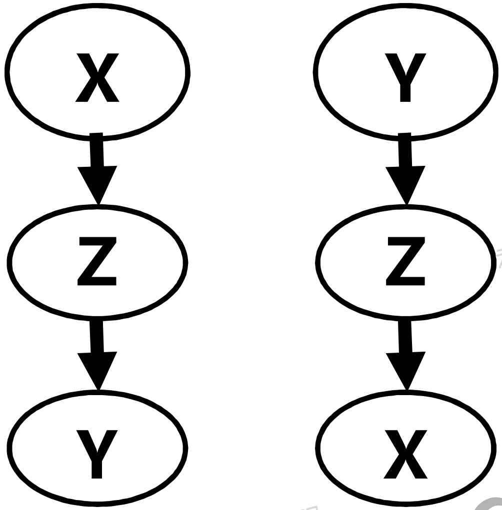
$p(R|H,S) = 0.0142$

$p(R|H) = 0.0185$

$p(R|H,\neg S) = 0.00833$

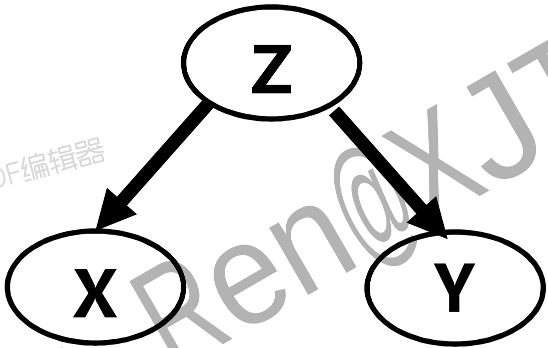
Independence does not imply conditional independence.

Four Relationship "D-separation"



cascade

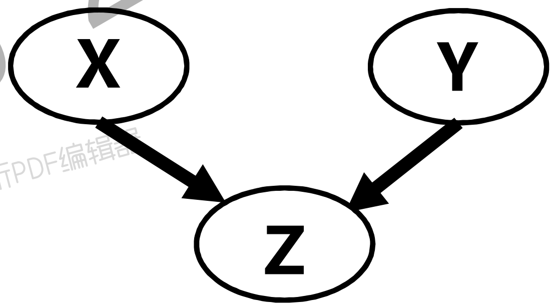
$$p(x,y,z) = p(x)p(z|x)p(y|z)$$



Common Parent

$$p(x,y,z) = p(x|z)p(y|z)p(z)$$

$$p(x,y|z) = p(x|z)p(y|z)$$



V-Structure

$$p(x,y,z) = p(x)p(y)p(z|x,y)$$

$$p(x,y|z) = \frac{p(x)p(y)p(z|x,y)}{p(z)}$$

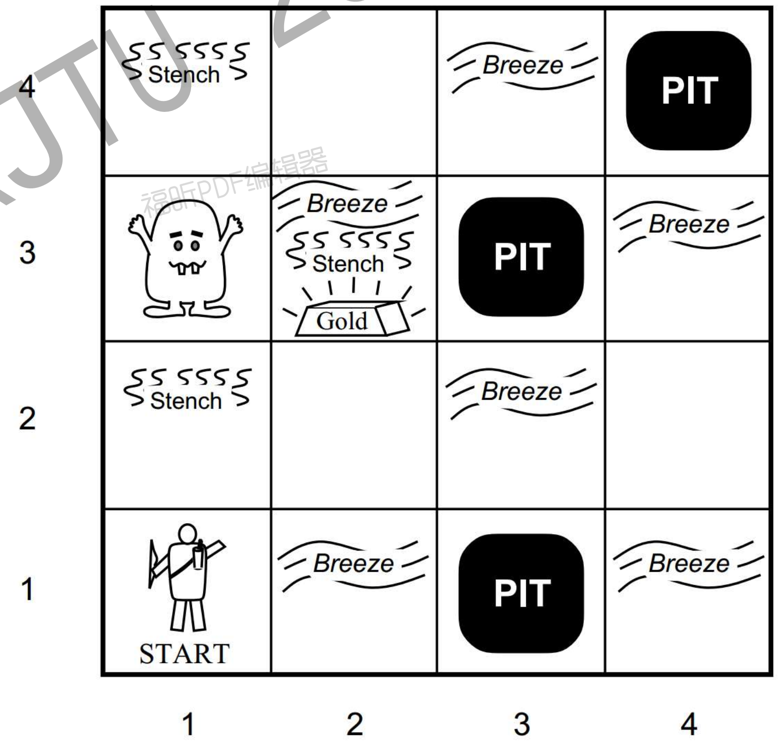
Back to the Wumpus World

Environment

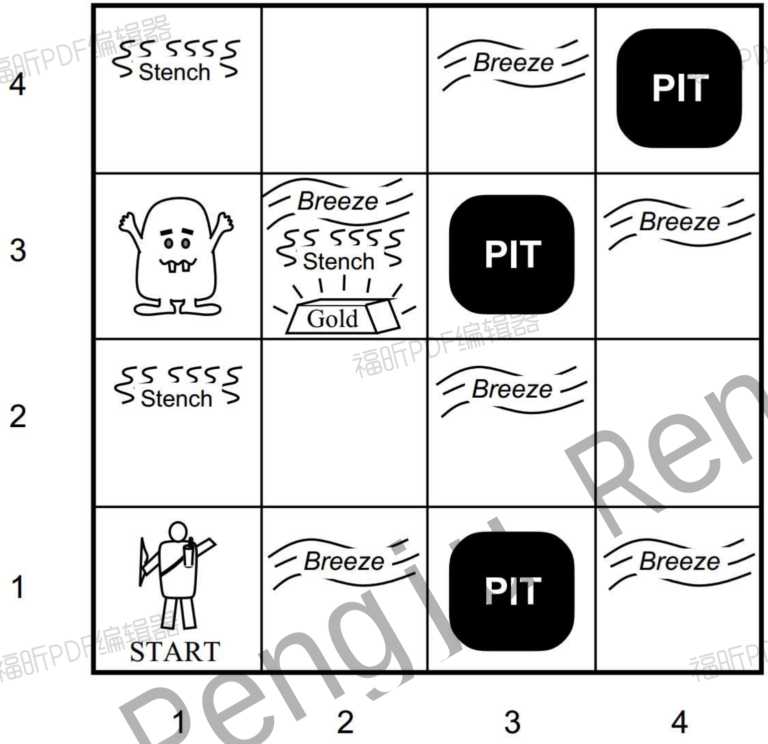
- Squares adjacent to wumpus are smelly
- Squares adjacent to pit are breezy
- Glitter iff gold is in the same square
- Shooting kills wumpus if you are facing it
- Shooting uses up the only arrow
- Grabbing picks up gold if in same square
- Releasing drops the gold in same square

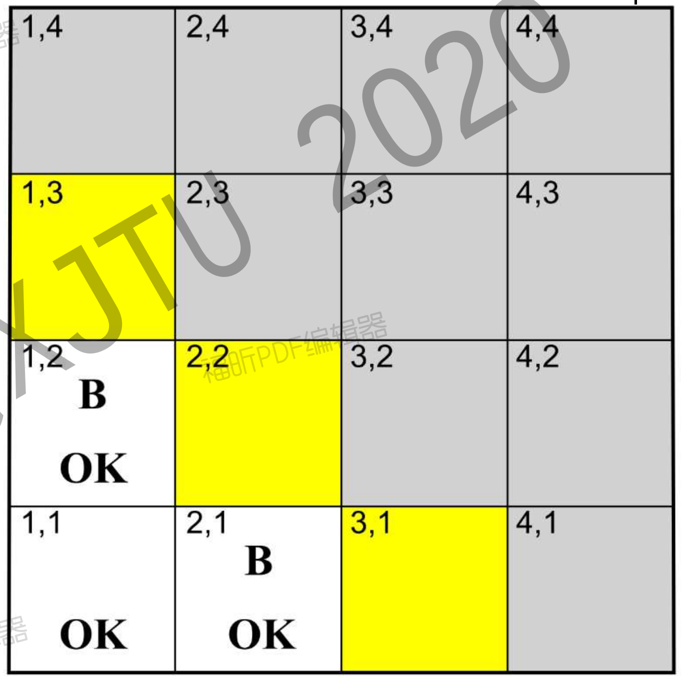
Actuators Left turn, Right turn, Forward, Grab, Release, Shoot

Sensors Breeze, Glitter, Smell



E.g. Wumpus World





$P_{i,j} = true$ iff [i,j] contains a pit

$B_{i,j} = true$ iff [i,j] is breezy

Include only $B_{1,1}, B_{1,2}, B_{2,1}$ in the probability model

Position 1,2 and 2,1 both feel breezy. Which is the better next move ?

Specifying the probability model

The full joint distribution is $p(p_{1,1},\ldots,p_{4,4},B_{1,1},B_{1,2},B_{2,1})$

Apply product rule: $p(B_{1,1},B_{1,2},B_{2,1}|p_{1,1},\ldots,p_{4,4})p(p_{1,1},\ldots,p_{4,4})$

(Do it this way to get $\boldsymbol{P(Effect|Cause)}$ **)**

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

Second term: pits are placed randomly, probability 0.2 per square:

$$p(p_{1,1},\ldots,p_{4,4}) = \prod_{i,j=1,1}^{4,4} p(p_{i,j}) = 0.2^n \times 0.8^{16-n}$$

Observations and query

We know the following facts:

$$b = \neg b_{1,1} \wedge b_{1,2} \wedge b_{2,1}$$
$$known = \neg p_{1,1} \wedge \neg p_{1,2} \wedge \neg p_{2,1}$$

Query is $\mathbf{P}(P_{1,3} | known, b)$

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

Define $Unknown = P_{i,j}$ other than $P_{1,3}$ and $Known$

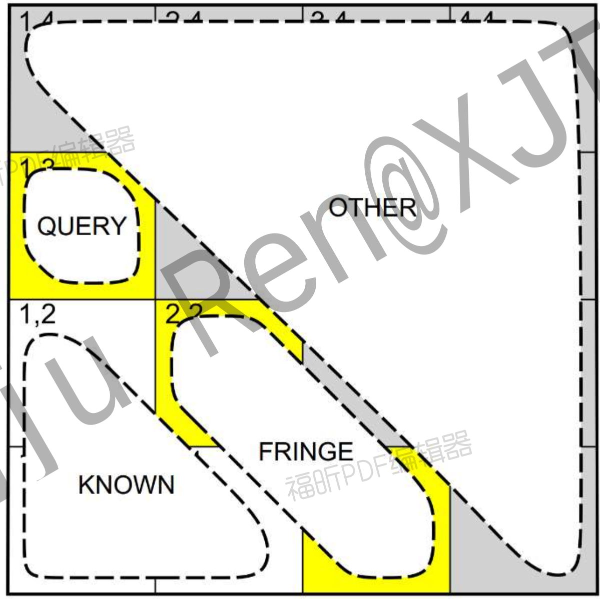
For inference by enumeration, we have

$$\mathbf{P}(P_{1,3} | Known, b) = \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, known, b)$$

Grows exponentially with number of squares!

Using conditional independence

Basic insight: observations are conditionally independent of other hidden squares given neighbouring hidden squares



Define $Unknown = Fringe \cup Other$

$$\mathrm{P}(b|P_{1,3}, Known, Unknown) = \mathrm{P}(b|P_{1,3}, Known, Fringe)$$

Manipulate query into a form where we can use this!

Using conditional independence contd.

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

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

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

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

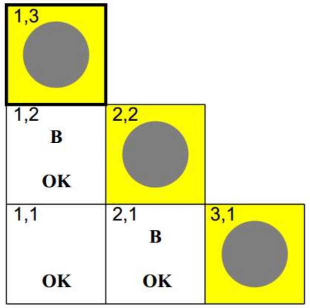
$$= \alpha \sum_{fringe} \mathbf{P}(b|known,P_{1,3},fringe)\sum_{other} \mathbf{P}(P_{1,3},known,fringe,other)$$

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

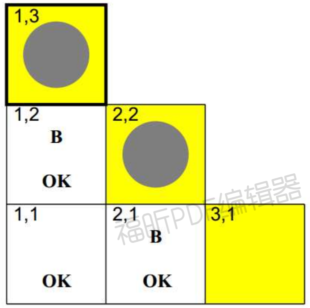
$$= \alpha P(known)\mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b|known,P_{1,3},fringe)P(fringe)\sum_{other} P(other)$$

$$= \alpha' \mathbf{P}(P_{1,3}) \sum_{fringe} \mathbf{P}(b|known,P_{1,3},fringe)P(fringe)$$

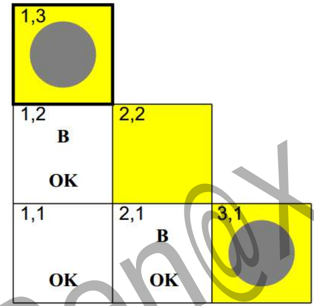
Using conditional independence contd.



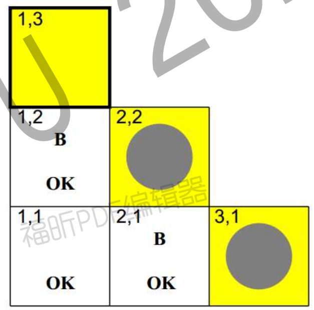
0.2 x 0.2 = 0.04



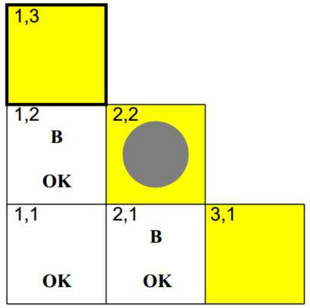
0.2 x 0.8 = 0.16



0.8 x 0.2 = 0.16



0.2 x 0.2 = 0.04



0.2 x 0.8 = 0.16

$$P(P_{1,3}|known, b) = \alpha' \langle 0.2(0.04 + 0.16 + 0.16), 0.8\,(0.04+0.16) \rangle \approx \langle 0.31,\ \ 0.69 \rangle$$

$$P(P_{2,2}|known, b) \approx \langle 0.86,\ \ 0.14 \rangle$$

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 size
- Independence and conditional independence provide the tools