Artificial Intelligence

Lecture 6：Probabilistic Reasoning

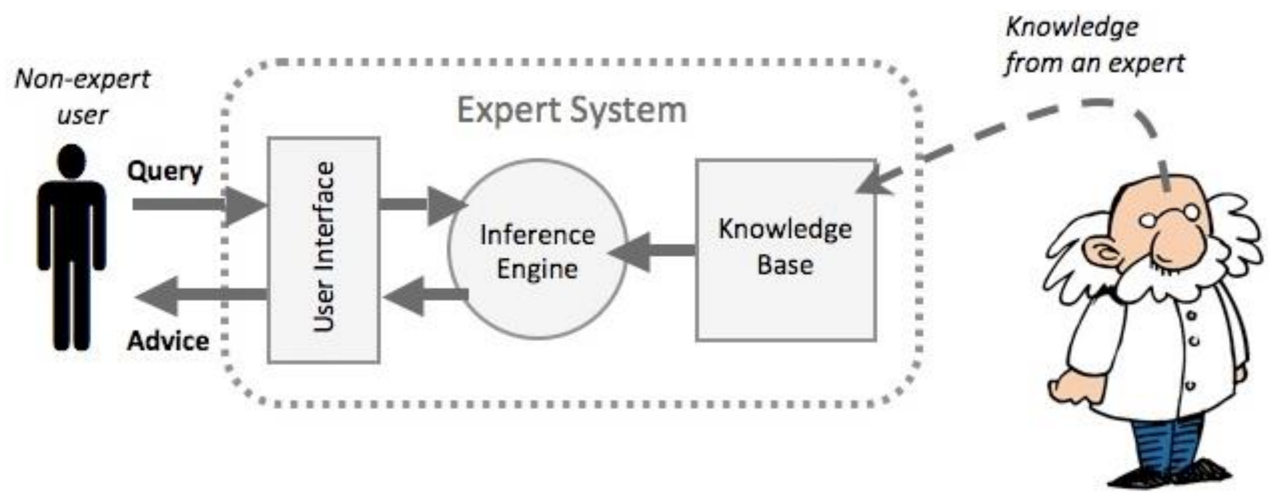
Xiaojin Gong

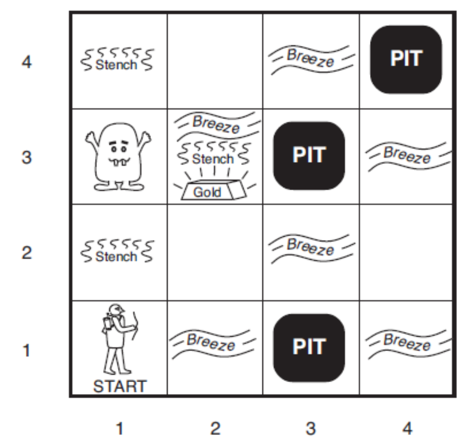
2022-03-27

Credits: AI Course in Berkeley & JHU

Review

- Knowledge and reasoning
  - Knowledge base
  - Inference
- Propositional logic
  - Inference: model checking, resolution, forward chaining





Review

Language	Ontological Commitment	Epistemological Commitment
Propositional logic	facts	true/false/unknown
First-order logic	facts, objects, relations	true/false/unknown
Temporal logic	facts, objects, relations, times	true/false/unknown
Probability theory	facts	degree of belief
Fuzzy logic	facts + degree of truth	known interval value

Higher-order logic:

relations and functions operate not only on objects,
but also on relations and functions

Outline

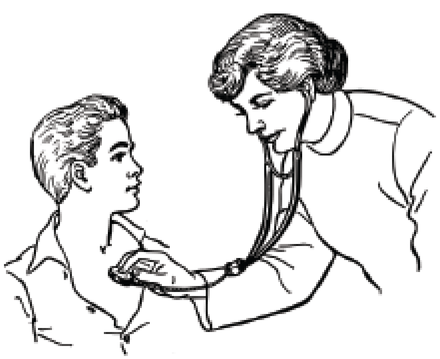
- Probability
  - Random variables
  - Joint and marginal distributions
  - Conditional distribution
  - Product rule, chain rule, Bayes' rule
  - Inference
  - Independence and conditional independence

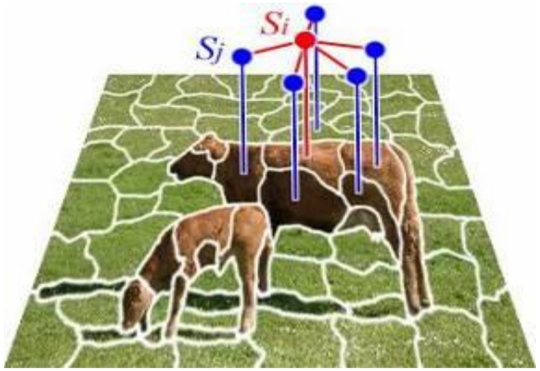
Uncertainty

- Task environments:
  - Partially observable
  - Non-deterministic

Uncertainty

- General situation:
  - Observed variables (evidence): Agent knows certain things about the state of the world (e.g., sensor readings or symptoms)
  - Unobserved variables: Agent needs to reason about other aspects (e.g. where an object is or what disease is present)
  - Model: Agent knows something about how the known variables relate to the unknown variables
- Probabilistic reasoning gives us a framework for managing our beliefs and knowledge





Making Decision Under Uncertainty

- Example:

$$
\begin{aligned}
P(A_{25} \text{ gets me there on time}|\ldots) &= 0.04 \\
P(A_{90} \text{ gets me there on time}|\ldots) &= 0.70 \\
P(A_{120} \text{ gets me there on time}|\ldots) &= 0.95 \\
P(A_{1440} \text{ gets me there on time}|\ldots) &= 0.9999
\end{aligned}
$$

- 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 = probability theory + utility theory

Rational Decision

```
function DT-AGENT(percept) returns an action
    persistent: belief_state, probabilistic beliefs about the current state of the world
                action, the agent's action

    update belief_state based on action and percept
    calculate outcome probabilities for actions,
        given action descriptions and current belief_state
    select action with highest expected utility
        given probabilities of outcomes and utility information
    return action
```

- Decision theory = probability theory + utility theory
  - An agent is rational iff it chooses the action that yields the highest expected utility, averaged over all the possible outcomes of the action.
  - A.k.a the principle of maximum expected utility.

Random Variable

- A random variable is some aspect of the world about which we have uncertainty, e.g.
  - R = Is it raining? {true, false}
  - T = Is it hot or cold? {hot, cold}
  - D = How long will it take to drive to airport? $[0, \infty)$

- We denote random variables with capital letters

Probability Distribution

- Associate a probability with each value

$$\forall x \;\; P(X = x) \geq 0, \;\; \text{and} \;\; \sum_{x} P(X = x) = 1$$

- Temperature

$P(T)$

T	P
hot	0.5
cold	0.5

- Weather

$P(W)$

W	P
sun	0.6
rain	0.4

Probability for Discrete Variables

- Unobserved random variables have distributions
- A distribution is a TABLE of probabilities of values
  - Temperature
  - Weather

$$P(T)$$

T	P
hot	0.5
cold	0.5

$$P(W)$$

W	P
sun	0.6
rain	0.4

Shorthand notation:

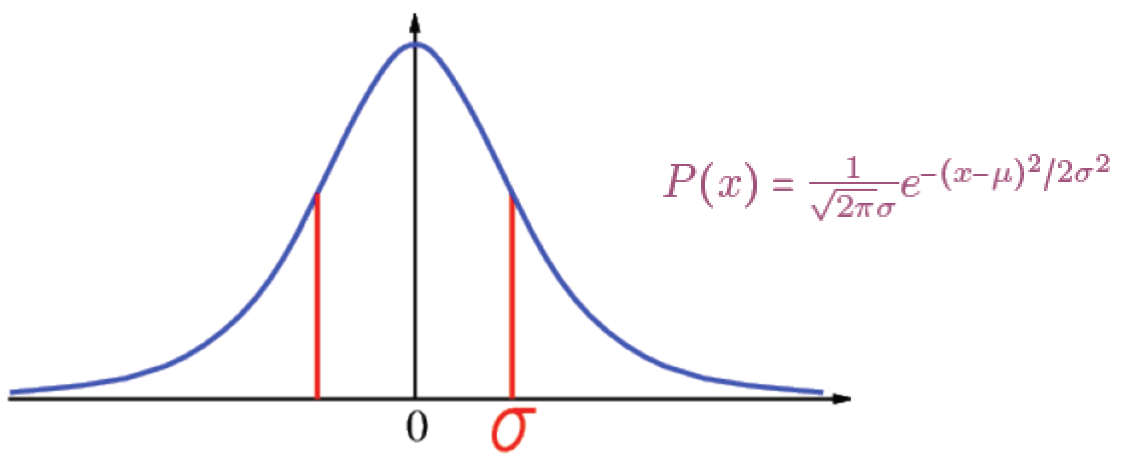
$$P(hot) = P(T = hot),$$
$$P(cold) = P(T = cold),$$
$$P(rain) = P(W = rain),$$
$$\ldots$$

OK *if* all domain entries are unique

Probability for Continuous Variables

- Probability density function

$$P(x) = \lim_{dx \to 0} \frac{P(x \le X \le x + dx)}{dx}$$



Joint Probability Distribution

- A joint probability distribution over a set of random variables: $X_1, \cdots, X_n$ specifies a real number of each assignment (or outcome):
  - Denote: $P(X_1 = x_1, \cdots, X_n = x_n) = P(x_1, \cdots, x_n)$

$$P(x_1, \cdots, x_n) \geq 0, \qquad and \quad \sum_{x_1, \cdots, x_n} P(x_1, \cdots, x_n) = 1$$

$P(T)$

T	P
hot	0.5
cold	0.5

$P(W)$

W	P
sun	0.6
rain	0.4

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Probabilistic Models

- A probabilistic model is a joint distribution over a set of random variables
- Probabilistic models:
  - Random variables with domains
  - Assignments are called outcomes
  - Joint distributions: say whether assignments (outcomes) are likely
  - Normalized: sum to 1.0
  - Ideally: only certain variables directly interact

Events

- An event is a set E of outcomes

$$P(E) = \sum_{(x_1,\cdots,x_n)\in E} P(x_1,\cdots,x_n)$$

- From a joint distribution, we can calculate the probability of any event:
  - Probability that it's hot AND sunny?
  - Probability that it's hot?
  - Probability that it's hot OR sunny?

$P(T,W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

- Typically, the events we care about are partial assignments, like P(T=hot)

Marginal Distributions

- Marginal distributions are sub-tables which eliminate variables
- Marginalization (summing out): Combine collapsed rows by adding

$P(T)$

T	P
hot	0.5
cold	0.5

$$P(T) = \sum_{W} P(T, W)$$

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P(W)$

W	P
sun	0.6
rain	0.4

$$P(W) = \sum_{T} P(T, W)$$

Conditional Probabilities

- Conditional distributions are probability distributions over some variables given fixed values of others

$$P(W = s|T = c) = \frac{P(W = s, T = c)}{P(T = c)} = \frac{0.2}{0.5} = 0.4$$

- Joint probability:

$$P(s, c) = P(s|c)P(c)$$

$P(T)$

T	P
hot	0.5
cold	0.5

$P(W)$

W	P
sun	0.6
rain	0.4

$\rightarrow$

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

Conditional Probability Distribution

- Normalization:

$$P(W = s|T = c) = \frac{P(W = s, T = c)}{P(T = c)}$$

$$= \frac{P(T = c, W = s)}{\sum_W P(T = c, W)}$$

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

SELECT the joint probabilities matching the evidence

$P(c, W)$

T	W	P
cold	sun	0.2
cold	rain	0.3

NORMALIZE the selection (make it sum to one)

$P(W|T = c)$

W	P
sun	0.4
rain	0.6

Independence

- Two variables are independent ($X \perp Y$) in a joint distribution if:

$$P(X, Y) = P(X)P(Y)$$

$$\forall x, y \ \ P(x, y) = P(x)P(y)$$

$P(T)$

T	P
hot	0.5
cold	0.5

$P_1(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

$P_2(T, W)$

T	W	P
hot	sun	0.3
hot	rain	0.2
cold	sun	0.3
cold	rain	0.2

$P(W)$

W	P
sun	0.6
rain	0.4

Conditional Independence

- X is conditionally independent of Y given Z ($X \perp Y|Z$) iff:

$$\forall x, y, z \qquad P(x, y|z) = P(x|z)P(y|z)$$

- Conditional independence is our most basic and robust form of knowledge about uncertain environments.
- Unconditional independence very rare.

Probabilistic Inference

- Probabilistic inference: compute a desired probability from other known probabilities
- We generally compute conditional probabilities
  - P(on time | no reported accidents) = 0.9
  - These represents the agent's beliefs given the evidence
- Probabilities change with new evidence:
  - P(on time | no accidents, 5 a.m.) = 0.95
  - P(on time | no accidents, 5 a.m., raining) = 0.80
  - Observing new evidence causes beliefs to be updated

Inference by Enumeration

- General case:
  - Evidence variables: $E_1 \cdots E_k = e_1 \cdots e_k$
  - Query variable: Q
  - Hidden variables: $H_1 \cdots H_k$

$\left.\begin{array}{l} E_1 \cdots E_k = e_1 \cdots e_k \\ Q \\ H_1 \cdots H_k \end{array}\right\} X_1 \cdots X_n$

All variables

- We want: $P(Q|e_1 \cdots e_k)$
- Solution:
  - Step 1: Select the entries consistent with the evidence
  - Step 2: Sum out H to get joint of Query and evidence
  - Step 3: Normalize

Inference by Enumeration

- P(sun)?

- P(sun | winter)?

- P(sun | winter, hot)?

S	T	W	P
summer	hot	sun	0.30
summer	hot	rain	0.05
summer	cold	sun	0.10
summer	cold	rain	0.05
winter	hot	sun	0.10
winter	hot	rain	0.05
winter	cold	sun	0.15
winter	cold	rain	0.20

Inference by Enumeration

- Worst-case time complexity $O(d^n)$

- Space complexity $O(d^n)$ to store the joint distribution

The Product Rule

- Infer the joint probability from conditional distributions

$$P(x, y) = P(y)P(x|y)$$

$P(W)$

R	P
sun	0.8
rain	0.2

$P(D|W)$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

$P(D, W)$

D	W	P
wet	sun	
dry	sun	
wet	rain	
dry	rain	

The Chain Rule

- The joint distribution can be written as an incremental product of conditional distributions:

$$P(x_1, x_2, x_3) = P(x_1)P(x_2|x_1)P(x_3|x_1, x_2)$$

$$P(x_1, x_2, \dots, x_n) = \prod P(x_i|x_1 \dots x_{i-1})$$

Bayes' Rule

- Two ways to factor a joint distribution:

$$P(x,y) = P(y)P(x|y) = P(x)P(y|x)$$

$$\Rightarrow P(x|y) = \frac{P(y|x)}{P(y)} P(x)$$

$$\Rightarrow \boxed{P(x|y)} \propto \boxed{P(y|x)}\boxed{P(x)}$$

Posterior Likelihood Prior

Bayes' Rule

$$P(x|y) = \frac{P(y|x)}{P(y)} P(x)$$

- Why is this at all helpful?
  - Build one conditional from its reverse
  - Often one conditional is tricky but the other one is simple
  - Foundation of many systems
- In the running for most important AI equation!

Inference with Bayes' Rule

- Example:
  - Diagnostic probability from causal probability:

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

- Example:
  - Given:

$P(W)$

R	P
sun	0.8
rain	0.2

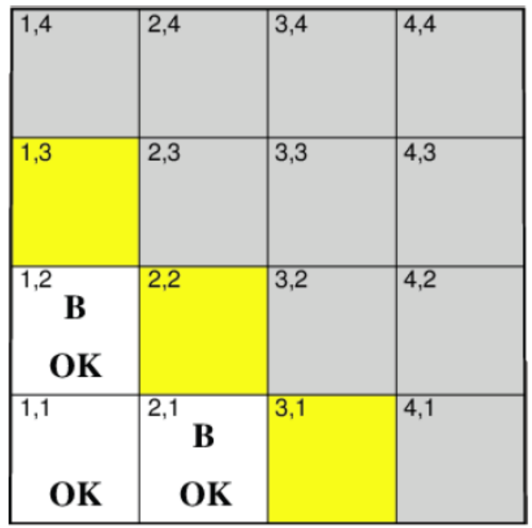
$P(D|W)$

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

  - What is P(W | dry) ?

The Wumpus World

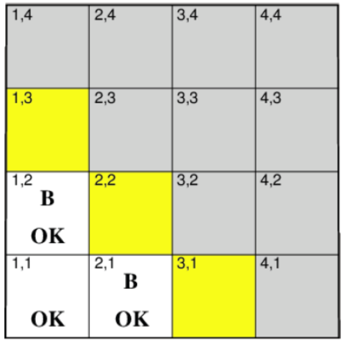
-



- $P_{ij} = true$ iff $[i, j]$ contains a pit
- $B_{ij} = true$ iff $[i, j]$ is breezy

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

The Wumpus World



- Specifying the Probability Model
  - The full joint distribution is $\mathbf{P}(P_{1,1}, \ldots, P_{4,4}, B_{1,1}, B_{1,2}, B_{2,1})$
  - Apply product rule: $\mathbf{P}(B_{1,1}, B_{1,2}, B_{2,1} \,|\, P_{1,1}, \ldots, P_{4,4})\mathbf{P}(P_{1,1}, \ldots, P_{4,4})$

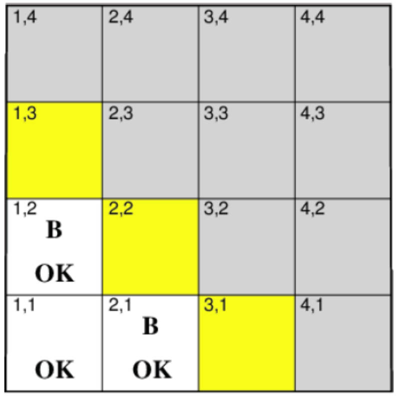
    This gives us: $P(Effect|Cause)$
  - First term: 1 if pits are adjacent to breezes, 0 otherwise
  - Second term: pits are placed randomly, probability 0.2 per square:

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

for n pits.

The Wumpus World

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

- Define $Unknown$ = P_{ij}s other than $P_{1,3}$ and $Known$

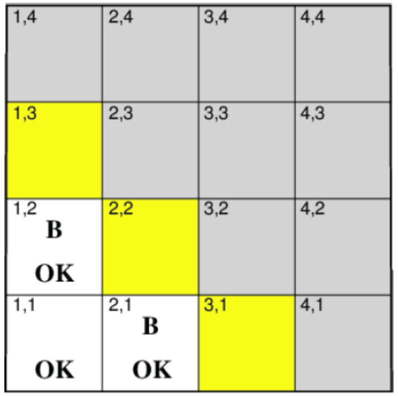
- For inference by enumeration, we have

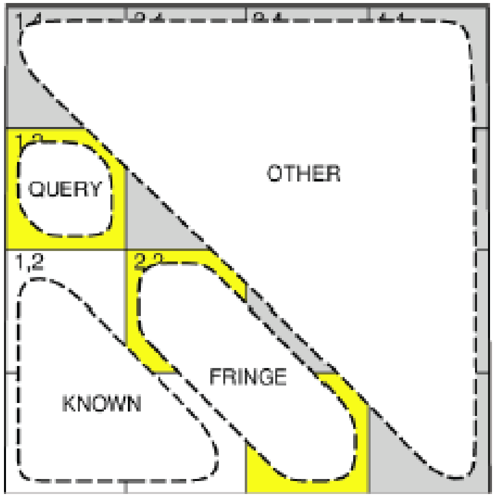
$$\mathbf{P}(P_{1,3} | known, b) = \alpha \sum_{unknown} \mathbf{P}(P_{1,3}, unknown, known, b)$$

- Grows exponentially with number of squares!

The Wumpus World

- Using Conditional Independence
  - Basic insight: observations are conditionally independent of other hidden squares given neighbouring hidden squares



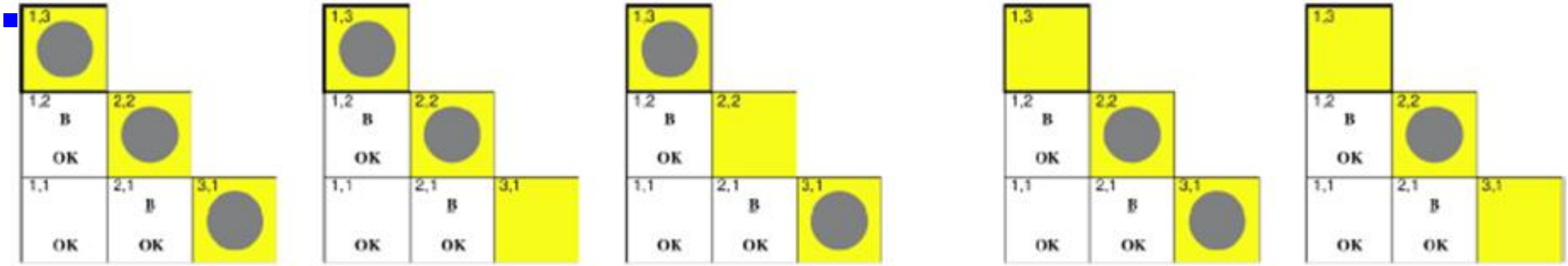


  - Define $Unknown = Fringe \cup Other$
    $\mathbf{P}(b|P_{1,3}, Known, Unknown) = \mathbf{P}(b|P_{1,3}, Known, Fringe)$
  - Manipulate query into a form where we can use this!

The Wumpus World

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

The Wumpus World



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

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

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

$$\approx \langle 0.31, 0.69 \rangle$$

Assignments

- Reading assignment:
  - Ch. 13