Uncertain knowledge

- Russell and Norvig, Chapters 7 + 13

Assignment 1

- Up on the website
- Will discuss the assignment next week during the lecture after the mid-semester exam

Mid Semester Exam

- Next week
- Covers material from weeks 1-5
- Normal lecture time, different place
- **2-3pm Tuesday 30 August 42-115**
- Remember to bring a pencil and an eraser
- 2009 exam is on the website
  - No answers
  - Only covered weeks 1-4

Overview: aims

- have a feel for the limitations of logic for dealing with uncertainties (logic is the theme for chapters 7-10)
- understand the basics of probability theory and know concepts like "prior", and "conditional" ("posterior") probability
- understand some principles for inferring knowledge in the presence of uncertainties
  - through marginalization,
  - through Bayes' rule.
- understand what independence is and how it affects inference
- know what probabilistic inference can be used for

Overview: topics

- Logical reasoning
  - The Wumpus world
- Uncertainty
  - The Wumpus world
  - Diagnosis
- Probability theory
  - utility theory, decision theory, prior probability, conditional probability, axioms, inference
- Bayes' rule
- Examples
  - probabilities, Bayes' rule

Logical reasoning (1)

- Aristotle's (384-322 BC) observation on *syllogisms* (logical inference / inference rules):
  - All students are overworked
    Mary is a student
    So, Mary is overworked
- Can be analysed in terms of form (*symbols* and *inference*):
  - Can substitute "person" for "student", "mortal" for "overworked" and "Socrates" for "Mary"
  - Humans are mortal
    Socrates is human
    Thus, Socrates is mortal

Logical reasoning (2)

- What does it mean to "reason logically"?
  - Reason (verb): infer new facts from existing facts
  - Logic: rules for reasoning which produce valid outcomes from true premises
- Agents know facts about their world and use reasoning to select among available actions
- Agents accept new tasks via detailed descriptions of goals and adapt to changes by updating their knowledge

Logical reasoning (3)

- Agents need to know
  - (relevant parts of) the current world state,
  - how to infer implicit knowledge,
  - how the world changes,
  - what they want, and
  - what their actions produce
- By depending on "logical reasoning" one makes certain commitments in regard to representations and operations in the agent

Logical reasoning: Knowledge based agent (1)

- The agent must be able to:
  - Represent states, actions, etc.
  - Incorporate new percepts
  - Update internal representations of the world
  - Deduce hidden properties of the world
  - Deduce appropriate actions

Logical reasoning: Knowledge based agent (2)

- Knowledge (representation)
- Reasoning (about knowledge)

- Knowledge should be expressed in general terms, able to be recombined for different purposes

Logical reasoning: Knowledge based agent (3)

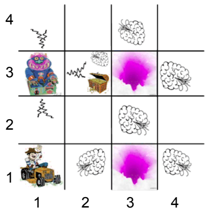
```
function KB-AGENT(percept) returns an action
   static: KB, a knowledge base
           t, a counter, initially 0, indicating time
   TELL(KB, MAKE-PERCEPT-SENTENCE(percept, t))
   action ← ASK(KB, MAKE-ACTION-QUERY(t))
   TELL(KB, MAKE-ACTION-SENTENCE(action, t))
   t ← t + 1
   return action
```

Logical reasoning: Knowledge based agent (4)

- General knowledge
- Current percepts
- Inferences about the world
- Selecting actions

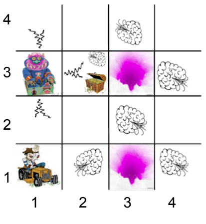
- Able to deal with partially observable environments

The Wumpus world (1)



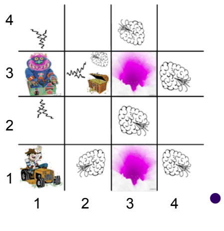
- World
  - 4 by 4 grid of squares
  - You can only perceive what is in your "square"
  - There's one monster, a Wumpus, somewhere in the world
  - Each square (other than [1,1]) can be a pit (probability 0.2)
  - You die if you enter a square with a Wumpus or a pit
  - You start at [x,y] = [1,1] which is a safe square

The Wumpus world (2)



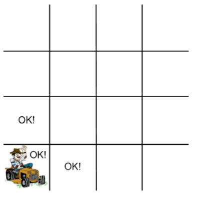
- Percepts
  - Next to the Wumpus, a stench is perceived
  - Next to a pit, a breeze is perceived
  - Gold visibly glitters if in your square
  - Hitting a wall means perceiving a bump
  - The Wumpus screams when it dies (heard everywhere)

The Wumpus world (3)



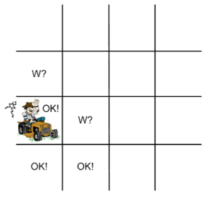
- Actions
  - Go forward
  - Turn left or right
  - Grab objects
  - Shoot one arrow
  - Climb to the next level
- Goal
  - Find the gold and survive

The Wumpus world (4)



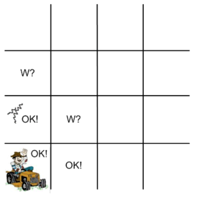
- Nothing is perceived, therefore it is safe to go north and east.
- Let's go north.

The Wumpus world (5)



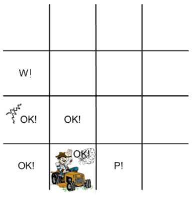
- A stench is perceived. Hence, a Wumpus is in either of the neighbouring squares.
- Let's go back.

The Wumpus world (6)



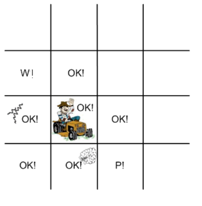
- And let's go east…

The Wumpus world (7)



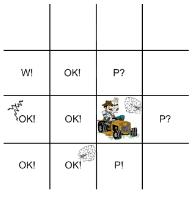
- A breeze is perceived. Therefore a pit is in one of the neighbouring squares. A breeze was not perceived when we were in [1,2], hence the pit is in [3,1]. A stench is not perceived. Therefore, there is no Wumpus in [2,2], hence it must be in [1,3].
- Let's go north (continue our search for gold).

The Wumpus world (8)



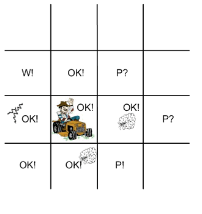
- Nothing is perceived.
- Let's go east.

The Wumpus world (9)



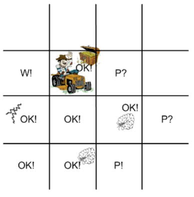
- A breeze is perceived. Therefore a pit is in either [4,2] or [3,3].
- Let's go back west.

The Wumpus world (10)



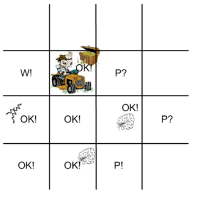
- Again, nothing is perceived.
- Let's go north

The Wumpus world (11)



- A glitter is perceived. Therefore there is gold here.
- Pick up the gold and go to the next level.

The Wumpus world (12)



There is a need for

- a language that can express: there's a Wumpus in either [1,3] or [2,2], and there can't be a pit in [2,2] (when we design the "agent")
- an inference mechanism that can combine knowledge gained at different times in different places, and
- a place to accumulate this knowledge

Propositions

- Assertions that such-and-such is the case
- Formal languages for stating propositions
  - Propositional logic (Chapter 7)
  - First-order logic (Chapter 8)

Propositional logic (1)

- Sentences constructed from proposition symbols and logical connectives
- Also called Boolean logic

Propositional logic (2)

- Proposition symbols
  - P, Q, R, stand for propositions that are TRUE or FALSE
- Proposition symbols in the Wumpus World
  - S[x,y] = a stench is perceived in square x,y
  - W[x,y] = a wumpus is in square x,y
  - …

Propositional logic (3)

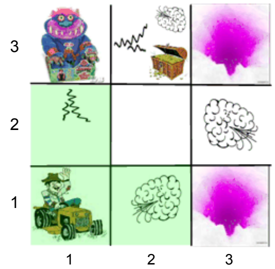
- Logical connectives
  - $\neg$ (not, negation)
  - $\wedge$ (and, conjunction)
  - $\vee$ (or, disjunction)
  - $\Rightarrow$ (implies, implication)
  - $\Leftrightarrow$ (if and only if, biconditional)
  - etc

Propositional logic: Logical connectives

- Truth table

P	Q	$\neg P$	$P \wedge Q$	$P \vee Q$	$P \Rightarrow Q$	$P \Leftrightarrow Q$
false	*false*	*true*	*false*	*false*	*true*	*true*
false	*true*	*true*	*false*	*true*	*true*	*false*
true	*false*	*false*	*false*	*true*	*false*	*false*
true	*true*	*false*	*true*	*true*	*true*	*true*

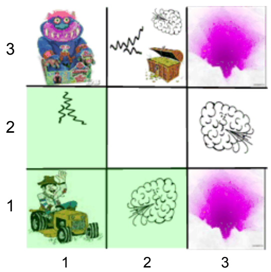
Propositional logic in the Wumpus world (1)



Let's assume that we have established all percepts in the shaded area:

- It does not stink in [1, 1], [2, 1].
- It stinks in [1, 2].
- There's not a breeze in [1, 1], [1, 2].
- There's a breeze in [2,1].

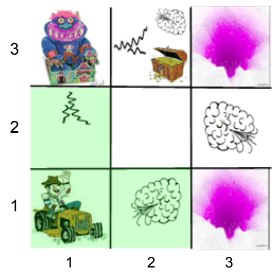
Propositional logic in the Wumpus world (2)



This conception of the world can be represented in propositional logic as:

```
¬S[1,1]
¬S[2,1]
S[1,2]
¬B[1,1]
¬B[1,2]
B[2,1]
```

Propositional logic in the Wumpus world (3)



We also have access to some basic facts of how things work in the world

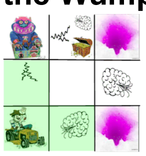
```
Rule 1: ¬S[1,1] ⇒ ¬W[1,1] ∧
   ¬W[1,2] ∧ ¬W[2,1]

Rule 2: ¬S[2,1] ⇒ ¬W[1,1] ∧
   ¬W[2,1] ∧ ¬W[2,2] ∧
  ¬W[3,1]

Rule 3: ¬S[1,2] ⇒ ¬W[1,1] ∧
   ¬W[1,2] ∧ ¬W[2,2] ∧
  ¬W[1,3]

Rule 4: S[1,2] ⇒ W[1,3] ∨
   W[1,2] ∨ W[2,2] ∨ W[1,1]
...
```

Propositional logic in the Wumpus world (4)



```
Rule 1: ¬S[1,1] ⇒ ¬W[1,1] ∧
 ¬W[1,2] ∧ ¬W[2,1]

Rule 2: ¬S[2,1] ⇒ ¬W[1,1] ∧
 ¬W[2,1] ∧ ¬W[2,2] ∧ ¬W[3,1]

Rule 3: ¬S[1,2] ⇒ ¬W[1,1] ∧
 ¬W[1,2] ∧ ¬W[2,2] ∧ ¬W[1,3]

Rule 4: S[1,2] ⇒ W[1,3] ∨
 W[1,2] ∨ W[2,2] ∨ W[1,1]
```

So what inferences can be made?

- `Rule 1 ∧ ¬S[1,1]` (modus ponens)
  `¬W[1,1] ∧ ¬W[1,2] ∧ ¬W[2,1]`
- `Rule 2 ∧ ¬S[2,1]` (modus ponens)
  `¬W[1,1] ∧ ¬W[2,1] ∧ ¬W[2,2] ∧¬W[3,1]`
- `Rule 4 ∧ S[1,2]` (modus ponens)
  `W[1,3] ∨ W[1,2] ∨ W[2,2] ∨ W[1,1]`
- `¬W[1,1]` ∧ previous (unit resolution)
  `W[1,3] ∨ W[1,2] ∨ W[2,2]`
- `¬W[2,2]` ∧ previous (unit resolution)
  `W[1,3] ∨ W[1,2]`
- `¬W[1,2]` ∧ previous (unit resolution)
  `W[1,3]` **(the Wumpus is in square [1,3])**

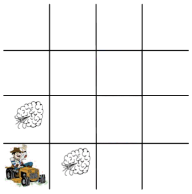
However, nothing is certain, except death and taxes…[1]

- The complexity of the world prohibits complete characterisation
- Experts often express their knowledge vaguely
  - Meningitis causes a stiff neck in about 50% of cases.
  - Birds fly (but ostriches don't)
- Data exhibits noise, and inconsistencies
  - 32% of respondents think this is a "seven", 68% think it is a "one"

- Some things are truly random (e.g. radioactive decay)

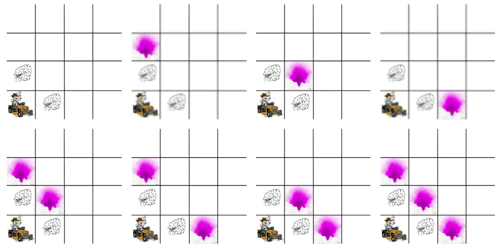
1: Daniel Defoe, 1726; Benjamin Franklin, 1789

Uncertainty in the Wumpus world (1)



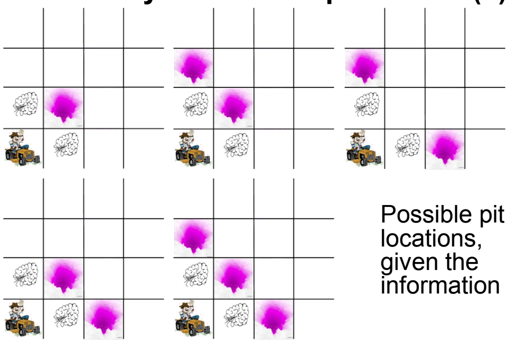
- Some situations cannot be solved by logic, due to the complexity of the world
- Where are the pits?
- In this case, have to take a chance and enter one of the squares to find out

Uncertainty in the Wumpus world (2)



Possible pit locations

Uncertainty in the Wumpus world (3)



Possible pit locations, given the information

Uncertainty in diagnosis

p: a patient

$$\forall p\ Symptom(p, Toothache) \Rightarrow Disease(p, Cavity).$$

Wrong ! Toothache can be associated with other diseases.

$$\forall p\ Symptom(p, Toothache) \Rightarrow Disease(p, Cavity) \vee Disease(p, GumDisease) \vee Disease(p, Abscess) \vee \ldots$$

Would have to be exhaustive : a huge list !

$$\forall p\ Disease(p, Cavity) \Rightarrow Symptom(p, Toothache).$$

Also wrong !
Might not have any pain (yet !)

Logical / crisp rules fail because:

- Too many qualifications to list
  - (Laziness)
- Knowledge is incomplete
  - (Theoretical ignorance)
- Not all tests can be done
  - (Practical ignorance)

Making rational decisions under uncertainty (1)

- A rational decision depends on the importance of various goals and the likelihood that they will be achieved
- Agents have preferences between possible outcomes, with regard to success and other features of the outcome
  - e.g. getting to the airport on time

Making rational decisions under uncertainty (2)

- Probability theory
  - Assigns to each sentence a numerical degree of belief between 0 and 1
  - Indicates how likely a successful outcome is
- Utility theory
  - every state has a degree of utility to an agent
  - the agent will prefer states with higher utility

Making rational decisions under uncertainty (3)

- Decision theory = Probability theory + Utility theory
- *an agent is rational if and only if it chooses the action that yields the highest expected utility, averaged over all the possible outcomes of the action*
- Principle of Maximum Expected Utility (MEU)

Making rational decisions under uncertainty (3)

- A decision-theoretic agent
  - an agent that uses decision theory to select actions

```
function DT-AGENT(percept) returns an action
    static:   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
```

Who wants to be a millionaire?

- In this game: Utility could be money; options are
  - keep the $16,000, or
  - answer the question
    - correctly, and win $32,000, with a “free” hit at $64,000;
    - incorrectly, and win $1,000

Question 10 for $32,000: What is the Italian word for a square or marketplace?

A. Presto	**10%**	**7%**
B. Pisa	**0%**	**0%**
C. Piazza	**50%**	**46%**
D. Plaza	**40%**	**47%**

Propositions: Random variables (1)

- Random variables are features of the world
- Part of the world whose status is initially unknown
- Each random variable has a domain (or support) of possible values
  - E.g. <true, false>

Propositions: Random variables (2)

- Boolean (true or false, e.g. Cavity)
  - *Cavity=true* is *cavity*
  - *Cavity=false* is *¬cavity*
- Discrete (a countable domain, e.g. *Weather*⟨sunny, rainy, overcast⟩)
  - *Weather=sunny* is *sunny*
- Continuous (real numbers, e.g. 3.14159265…)
  - *X*=4.02
  - *X*<4.02

Degree of belief and Probability Theory

- Probability provides a way of summarizing the uncertainty
- A sentence/proposition α is assigned a degree of belief between 0 and 1
  - we expect that α is true with a probability $P(\alpha)$
- P(true)=1, P(false)=0
- The degree of belief is based on some evidence and probabilistic reasoning

True and estimated probabilities:

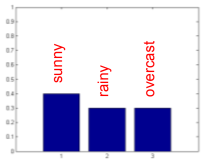
- Event = a set of possible outcomes for a random variable.
- P(Event) = probability that any of those outcomes occurred.
- Assume there is a true probability distribution, i.e. all possible events have true probabilities.
- We often don't know the true probabilities, but can estimate them via randomly sampling the population of interest. In this, we collect data and analyse it, e.g. look at relative frequencies of events of interest.

Probability

- Frequentist
  - Probabilities are determined from experiments
- Objectivist
  - Probabilities are real aspects of the universe
- Subjectivist
  - Probabilities are degrees of belief

Prior probability: Boolean / Discrete

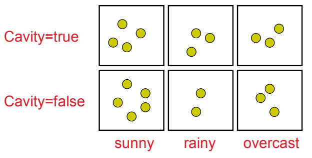
- The probability of a proposition α in the absence of any other information is $P(\alpha)$
- Unconditional or prior probability
  - P(Cavity=true) = 0.1
  - P(Engineering_student=true) = 0.15
  - **P**(Weather) = ⟨sunny=0.4, rainy=0.3, overcast=0.3⟩



Probability distribution
(adds up to 1.0)

Prior probability: Joint Probability Distribution

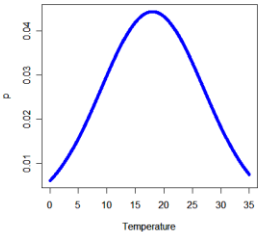
- **P**(Weather, Cavity)=
  ⟨ Cavity=true ⟨ sunny=0.20, rainy=0.15, overcast=0.15 ⟩
  Cavity=false ⟨ sunny=0.25, rainy=0.10, overcast=0.15 ⟩ ⟩



A full joint distribution covers the complete set of random variables.

Prior probability: Continuous

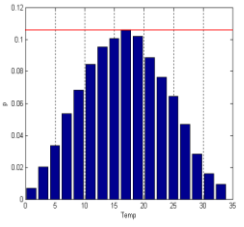
- The probability of a proposition α in the absence of any other information is $P(\alpha)$
- P(Temp=17) = 0 ! (since Temp is continuous)



Prior probability: Continuous

- The probability of a proposition α in the absence of any other information is $P(\alpha)$
- P(Temp=17) = 0 ! (since Temp is continuous)
- Use probability *density* functions p() for continuous variables:

e.g. p(17)=0.1059



Prior vs Conditional Probability

- With more evidence, we may be able to make a better estimate than the prior probability about the probability of α
  - Conditional or Posterior Probability

Conditional (posterior) probability (1)

- Using evidence

  $P(\alpha|\beta)$ = "probability of α given β".

  The probability of α given that all we know is β

  P(Engineering_student | ITEE_student) = 0.5

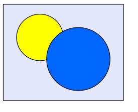
  P(Weather | Summer) =

  ⟨sunny=0.6, rainy=0.2, overcast=0.2⟩

Conditional (posterior) probability (2)

- Defined in terms of unconditional probabilities (priors)

$$P(\alpha \mid \beta) = \frac{P(\alpha \wedge \beta)}{P(\beta)}$$



Product rule:

$$P(\alpha \wedge \beta) = P(\alpha \mid \beta) P(\beta)$$

$$P(\alpha \wedge \beta) = P(\beta \mid \alpha) P(\alpha)$$

For distributions:

$$\mathbf{P}(X, Y) = \mathbf{P}(X \mid Y)\mathbf{P}(Y)$$

$$\mathbf{P}(X, Y) = \mathbf{P}(Y \mid X)\mathbf{P}(X)$$

Axioms of probability (1)

- All probabilities are between 0 and 1

$$0 \le P(\alpha) \le 1$$

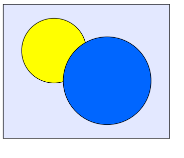
- True / Valid propositions have probability 1; False / Unsatisfiable propositions have probability 0

$$P(true) = 1$$
$$P(false) = 0$$

Axioms of probability (2)

- The probability of a disjunction is given by

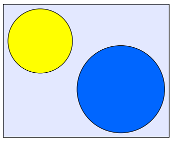
$$P(\alpha \vee \beta) = P(\alpha) + P(\beta) - P(\alpha \wedge \beta)$$



Axioms of probability (3)

- If α and β are *exclusive* $P(\alpha \wedge \beta) = 0$

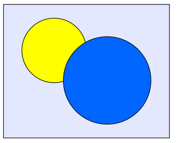
$$P(\alpha \vee \beta) = P(\alpha) + P(\beta)$$



Axioms of probability

$$0 \le P(\alpha) \le 1$$

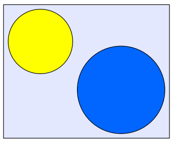
$$P(true) = 1$$
$$P(false) = 0$$



$$P(\alpha \vee \beta) = P(\alpha) + P(\beta) - P(\alpha \wedge \beta)$$

If α and β are *exclusive* $P(\alpha \wedge \beta) = 0$

$$P(\alpha \vee \beta) = P(\alpha) + P(\beta)$$



Probabilistic inference

- Computation from observed evidence of posterior probabilities for query propositions
- Full joint distribution is the ‘knowledge base’
- Answers to questions about the knowledge base can be derived through
  - Marginalisation / summing out:
  - Conditioning

Probabilistic inference

- Marginalisation
  - Summing out all other variables from any joint distribution containing A

$$\mathbf{P}(\mathrm{A}) = \sum_{\beta \in \mathrm{B}} \mathbf{P}(\mathrm{A}, \beta)$$

- Conditioning
  - Involves conditional probabilities instead of joint probabilities

$$\mathbf{P}(\mathrm{A}) = \sum_{\beta \in \mathrm{B}} \mathbf{P}(\mathrm{A} \mid \beta) P(\beta)$$

Collecting evidence for the “world” we want to understand…

Howard did a good job = true	*Voted Liberal = true*	*Male = true*
True	True	True
False	False	True
True	False	False
False	False	False
True	True	False
True	True	True
False	False	True
False	False	False
…	…	…
0.48	0.39	0.51

Full joint distribution:

Howard did a good job = true	*Voted Liberal = true*	*Male = true*	*P(event)=* **P(Good_job, Liberal, Male)**
True	True	True	0.19
True	True	False	0.17
True	False	True	0.05
True	False	False	0.07
False	True	True	0.01
False	True	False	0.02
False	False	True	0.26
False	False	False	0.23
			1.00

Conditional Probability

Howard did a good job = true	Voted Liberal = true	Male = true	P(event)
True	True	True	0.19
True	True	False	0.17
True	False	True	0.05
True	False	False	0.07
False	True	True	0.01
False	True	False	0.02
False	False	True	0.26
False	False	False	0.23
			1.00

If a person says Howard did a good job, with what probability did he/she vote Liberal?

$$P(Liberal \mid Good\_job) = \frac{P(Liberal \wedge Good\_job)}{P(Good\_job)} = \frac{0.19+0.17}{0.19+0.17+0.05+0.07} = 0.75$$

Marginalisation

- What is the probability of someone saying that Howard did a good job?

Howard did a good job	Voted Liberal	P(event)
True	True	0.36
True	False	0.12
False	True	0.03
False	False	0.49
		1.00

$$\mathbf{P}(Good\_job)$$
$$=< P(Good\_job = true, Liberal = true), P(Good\_job = false, Liberal = true) >$$
$$+ < P(Good\_job = true, Liberal = false), P(Good\_job = false, Liberal = false) >$$
$$=< 0.36, 0.03 > + < 0.12, 0.49 >$$
$$=< 0.48, 0.52 >$$

Conditioning

- What is the probability of someone saying that Howard did a good job?

Howard did a good job	Voted Liberal	P(event)
True	True	0.36
True	False	0.12
False	True	0.03
False	False	0.49
		1.00

$$\mathbf{P}(Good\_job)$$
$$=< P(Good\_job = true \mid Liberal = true)P(Liberal = true), P(Good\_job = false \mid Liberal = true)P(Liberal = true) >$$
$$+ < P(Good\_job = true \mid Liberal = false)P(Liberal = false), P(Good\_job = false, Liberal = false)P(Liberal = false) >$$
$$=< 0.36/0.39*0.39, 0.03/0.39*0.39 >$$
$$+ < 0.12/0.61*0.61, 0.49/.061*0.61 >$$
$$=< 0.48, 0.52 >$$

Independence (1)

- Also referred to as marginal independence and absolute independence
- If variables are independent, knowledge of one does not affect knowledge of another
- Independence written as:

Propositions	*Variables*
$P(a \mid b)=P(a)$	$\mathbf{P}(X \mid Y)=\mathbf{P}(X)$
$P(b \mid a)=P(b)$	$\mathbf{P}(Y \mid X)=\mathbf{P}(Y)$
$P(a \wedge b)=P(a)P(b)$	$\mathbf{P}(X,Y)=\mathbf{P}(X)\mathbf{P}(Y)$

- Weather is independent of dental problems
- Individual coin flips are independent

Independence (2)

- Add "blue eyes" to the material
  - Makes the joint distribution much larger
  - Does "blue eyes" influence how we think policy, or reversely, does policy influence eye colour? Hardly, so…

  P(Blue_eyes|Liberal)= P(Blue_eyes)
  P(Blue_eyes ∧ Liberal)= P(Blue_eyes)P(Liberal)
- Independence between variables simplifies rules and reduces information in joint distributions

Howard did a good job = true	Voted Liberal = true	Male = true	Blue eyes	P(event)
True	True	True	True	0.19/2
True	True	False	True	0.17/2
True	False	True	True	0.05/2
True	False	False	True	0.07/2
False	True	True	True	0.01/2
False	True	False	True	0.02/2
False	False	True	True	0.26/2
False	False	False	True	0.23/2
True	True	True	False	0.19/2
True	True	False	False	0.17/2
True	False	True	False	0.05/2
True	False	False	False	0.07/2
False	True	True	False	0.01/2
False	True	False	False	0.02/2
False	False	True	False	0.26/2
False	False	False	False	0.23/2
				1.00

Bayes' rule (1)

Product rule:

$$P(\alpha \wedge \beta) = P(\alpha \mid \beta)P(\beta)$$
$$P(\alpha \wedge \beta) = P(\beta \mid \alpha)P(\alpha)$$

$$P(\beta \mid \alpha)P(\alpha) = P(\alpha \mid \beta)P(\beta)$$

Bayes' rule: $$P(\beta \mid \alpha) = \frac{P(\alpha \mid \beta)P(\beta)}{P(\alpha)}$$

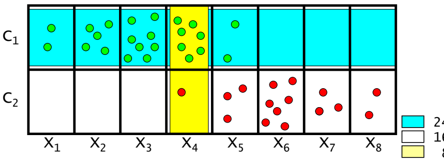
Bayes' rule (2)

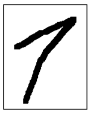
Bayes' rule: $P(\beta \mid \alpha) = \frac{P(\alpha \mid \beta) P(\beta)}{P(\alpha)}$

Multivalued variables: $\mathrm{P}(Y \mid X) = \frac{\mathrm{P}(X \mid Y)\mathrm{P}(Y)}{\mathrm{P}(X)}$

Given evidence, e: $\mathrm{P}(Y \mid X, e) = \frac{\mathrm{P}(X \mid Y, e)\mathrm{P}(Y \mid e)}{\mathrm{P}(X \mid e)}$

Bayes' rule for classification





24/40, $P(C_1)=.60$
16/40, $P(C_2)=.40$
8/40, $P(X_4)=.20$
Joint probability: 7/40, $P(C_1,X_4)\approx.18$
Posterior prob.: 7/24, $P(X_4|C_1)\approx.29$
Posterior prob.: 7/ 8, $P(C_1|X_4)\approx.88$

$P(C_k \mid X_n) = \frac{P(X_n \mid C_k) P(C_k)}{P(X_n)}$

e.g. $P(C_1 \mid X_4) = \frac{P(X_4 \mid C_1) P(C_1)}{P(X_4)} \approx \frac{.29 \cdot .60}{.20} \approx .88$

$P(C_2 \mid X_4) \approx \frac{.06 \cdot .40}{.20} \approx .12$

Diagnosis example (this time with probabilities)

- Meningitis causes stiff neck 50% of the time.
- One person in 50,000 has meningitis.
- One person in 20 has a stiff neck.
- If you have a stiff neck, what are the chances that you have meningitis?

$P(Stiff\_neck \mid Meningitis) = 0.50$

$P(Meningitis) = 1/50{,}000$

$P(Stiff\_neck) = 1/20$

$$P(Meningitis \mid Stiff\_neck) = \frac{P(Stiff\_neck \mid Meningitis) P(Meningitis)}{P(Stiff\_neck)} = \frac{0.50 \cdot 0.00002}{0.05} = 0.0002 = 1/5{,}000$$

Diagnosis example *continued*

$P(Meningitis \mid Stiff\_neck) + P(\neg Meningitis \mid Stiff\_neck) = 1$

$P(Meningitis \mid Stiff\_neck) = \frac{P(Stiff\_neck \mid Meningitis) P(Meningitis)}{P(Stiff\_neck)}$

$P(\neg Meningitis \mid Stiff\_neck) = \frac{P(Stiff\_neck \mid \neg Meningitis) P(\neg Meningitis)}{P(Stiff\_neck)}$

$$P(Stiff\_neck) = P(Stiff\_neck \mid Meningitis) P(Meningitis) + P(Stiff\_neck \mid \neg Meningitis) P(\neg Meningitis)$$

$$= 0.50 \cdot 0.00002 + P(Stiff\_neck \mid \neg Meningitis) \cdot 0.99998$$

Bayesian Reasoning Question from 2009 final exam (2)

- Random variables: G for sound –G for faulty; T1 for sound –T1 for faulty
- P(G) = 0.95 P(-G) = 0.05
- P(T1|G) = 0.82 P(-T1|G) = 0.18
- P(T1|-G) = 0.07 P(-T1|-G) = 0.93

- Want to find P(-G|T1)
- P(-G|T1) = (P(T1|-G)P(-G)) / P(T1)

$P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$ $\quad P(A) = P(A \mid B) P(B) + P(A \mid \neg B) P(\neg B)$

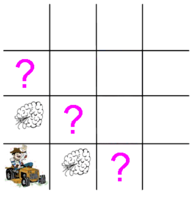
Bayesian Reasoning Question from 2009 final exam (3)

- P(T1) = P(T1|G)P(G) + P(T1|-G)P(-G)
- = 0.82 x 0.95 + 0.07 x 0.05
- = 0.779 + 0.0035
- = 0.7825

- P(-G|T1) = P(T1|-G)P(-G) / P(T1)
- = 0.07 x 0.05 / 0.7825
- = 0.0044728

$P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$ $\quad P(A) = P(A \mid B) P(B) + P(A \mid \neg B) P(\neg B)$

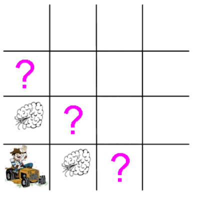
Uncertainty in the Wumpus world (4)



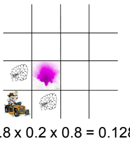
- Some situations cannot be solved by logic, due to the complexity of the world
- Where are the pits?
- In this case, have to take a chance and enter one of the squares to find out – which one is the best option?

Uncertainty in the Wumpus world (5)

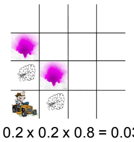
- Variables:
  - Breeze in square [i,j] (B_{ij})
  - Pit in square [i,j] (P_{ij})
- What is the probability that there is a pit in [1,3],[2,2], and [3,1] given the knowledge:
  - B_{11} = false
  - B_{12} = true
  - B_{21} = true
  - P_{11} = false
  - P_{12} = false
  - P_{21} = false



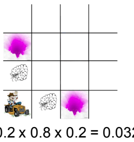
Uncertainty in the Wumpus world (6)



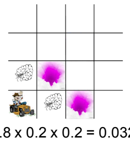
0.8 x 0.2 x 0.8 = 0.128



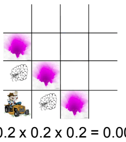
0.2 x 0.2 x 0.8 = 0.032



0.2 x 0.8 x 0.2 = 0.032



0.8 x 0.2 x 0.2 = 0.032

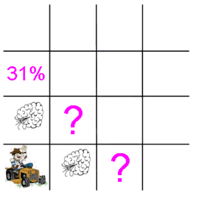


0.2 x 0.2 x 0.2 = 0.008

Relative probabilities of possible pit locations, given the information

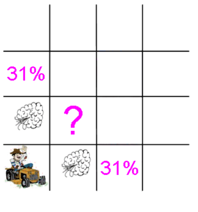
Uncertainty in the Wumpus world (7)

- P_{13} = true:
  - $\alpha(0.032 + 0.032 + 0.008) = \alpha(0.072)$
- P_{13} = false:
  - $\alpha(0.128 + 0.032) = \alpha(0.160)$
- $P_{13} = \alpha \langle 0.072,0.160 \rangle$
- $P_{13} = \langle 0.31,0.69 \rangle$



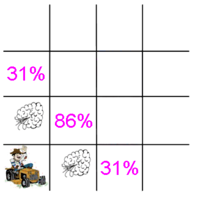
Uncertainty in the Wumpus world (8)

- P_{31} = true:
  - $\alpha(0.032 + 0.032 + 0.008) = \alpha(0.072)$
- P_{31} = false:
  - $\alpha(0.128 + 0.032) = \alpha(0.160)$
- $P_{31} = \alpha \langle 0.072,0.160 \rangle$
- $P_{31} = \langle 0.31,0.69 \rangle$



Uncertainty in the Wumpus world (9)

- P_{22} = true:
  - $\alpha(0.128 + 0.032 + 0.032 + 0.008) = \alpha(0.2)$
- P_{22} = false:
  - $\alpha(0.032)$
- $P_{22} = \alpha \langle 0.2,0.032 \rangle$
- $P_{22} = \langle 0.86,0.14 \rangle$



Prosecutor's fallacy (1)

- Court case
- Blood test matches a suspect
- Matching profile for blood test occurs in 1 out of 1000 people
- Prosecutor claims that this means that the probability of guilt is 0.999

But this doesn't take into account the prior probability of guilt!

Prosecutor's fallacy (2)

- Bayes' Theory:

$$P(\beta \mid \alpha) = \frac{P(\alpha \mid \beta)P(\beta)}{P(\alpha)}$$

$$P(\alpha) = P(\alpha \mid \beta)P(\beta) + P(\alpha \mid \neg\beta)P(\neg\beta)$$

$$P(\beta \mid \alpha) = \frac{P(\beta \mid \alpha)P(\beta)}{P(\alpha \mid \beta)P(\beta) + P(\alpha \mid \neg\beta)P(\neg\beta)}$$

$$P(I \mid E) = \frac{P(E \mid I)P(I)}{P(E)}$$

Prosecutor's fallacy (3)

$$P(I \mid E) = \frac{P(E \mid I)P(I)}{P(E \mid I)P(I) + P(E \mid \neg I)P(\neg I)}$$

- E = evidence
- I = innocence
- P(E) = probability of the evidence
- P(I) = probability of innocence

$$P(I \mid E) = \frac{P(E \mid I)P(I)}{P(E)}$$

Prosecutor's fallacy (4)

$$P(I \mid E) = \frac{P(E \mid I)P(I)}{P(E \mid I)P(I) + P(E \mid \neg I)P(\neg I)}$$

- Option 1: equal probability of guilt and innocence (in other words, a lot of other evidence)
  - P(I) = 0.5 (Probability of innocence)
  - P(¬ I) = 0.5 (Probability of guilt)
  - P(E| ¬I) = 1.0 (Probability of evidence given guilt)
  - P(E|I) = 0.001 (Probability of evidence given innocent)
  - P(I|E) = 0.001 * 0.5 / (0.001*0.5 + 1.0*0.5)
  - =0.0005 / (0.0005 + 0.5)
  - =0.000999
  - ≈0.001

$$P(I \mid E) = \frac{P(E \mid I)P(I)}{P(E)}$$

Prosecutor's fallacy (5)

$$P(I \mid E) = \frac{P(E \mid I)P(I)}{P(E \mid I)P(I) + P(E \mid \neg I)P(\neg I)}$$

- Option 2: no other evidence -> probability of guilt is 1 / number of people in the city (e.g. 10000)
  - P(I) = 9999/10000 (Probability of innocence)
  - P(¬ I) = 1/10000 (Probability of guilt)
  - P(E| ¬I) = 1.0 (Probability of evidence given guilt)
  - P(E|I) = 0.001 (Probability of evidence given innocent)
  - P(I|E) = 0.001 * 0.9999 / (0.001 * 0.9999 + 1.0 * 0.0001)
  - =0.09999 / (0.09999 + 0.0001)
  - =0.999

Further reading about probabilities (not covered in this course)

- Probabilistic reasoning (Chapter 14)
- Probabilistic reasoning over time (Chapter 15)
- Making simple decisions (Chapter 16)
- Making complex decisions (Chapter 17)

Applications of Probability

- Robots planning under uncertainty (deciding what to do next)
- Speech recognition
- Computational biology
- Probability-based processor
- …

ebiquity.umbc.edu/blogger/2010/08/18/probability-based-processor-might-speed-ai-applications/

www.aaai.org/aitopics/pmwiki/pmwiki.php/AITopics/Uncertainty

Summary

- Logic (Knowledge based agents)
- Propositional Logic
- Probability Theory (Decision-Theoretic agent)
- Random Variables and Domains
- Prior and Conditional Probabilities
- Independence
- Bayes' Rule