Introduction to Artificial Intelligence
COMP 3501 / COMP 4704-4
Lecture 11: Uncertainty

Prof. Nathan Sturtevant

Lecture Overview

- Return HW 1/Midterm
- Short HW 2 discussion
- Uncertainty / Probability

Uncertainty

- Previous approaches dealt with relatively certain worlds
- Couldn't make sensible moves in a card game with reasonably large stochasticity
- Cannot handle the uncertainty of the real world
  - What if we get hit by a meteor?
  - What if the sun goes supernova?
  - What if my car breaks down/explodes/get stolen?

Uncertainty

- Rational behavior must depend on quantifying uncertainty and acting accordingly
- Don't need to make contingency plans for a supernova
- Example:
  - Toothache ⇒ Cavity
  - Toothache ⇒ Cavity ∨ GumProblem ∨ Abscess …
- Casual rule:
  - Cavity ⇒ Toothache

Example

- A toothache doesn't mean a cavity, and a cavity doesn't mean a toothache
- Cannot strictly reason in this way
  - Need a degree of belief
    - If I have a toothache, how certain should I be that I have a cavity?
    - If I have toothaches, how certain should I be that I don't have a cavity?

Probability

- Probability summarizes the uncertainty we have about the world [eg from ignorance]
- Probability could come from measured sources or expert judgement
- Probability represents changes given current knowledge of the world
  - We either have a cavity or don't [ground truth]
  - What do we believe about it given just that our tooth hurts?

Probability and Rationality

- Do we always want a plan that maximizes the probability of success?
  - Each plan has a cost associated with it
  - The true cost may depend on the user
    - How much money & time do you have?
  - Utility represents preferences between costs and outcomes
- Choose plan with the *maximum expected utility*

Basic Probability

- In logic, we talked about models of the world
- In probability, we also have models
  - Each model has a probability
  - The sum of probabilities over all models is 1
  - $\sum_{w \in \Omega} P(w) = 1$
- Example: throw a die

Basic Probability

- A proposition or an event represents the set of possibilities/worlds over which we measure probability
- For a proposition $\phi, P(\phi) = \sum_{w \in \phi} P(w)$
- Example: roll two 6-sided dice
  - P(Sum=11) = P([5, 6]) + P([6, 5]) = 1/36 + 1/36 = 1/18
- Can perform calculation independent of other probabilities

Prior Probabilities

- The prior probability of an event is the probability in the absence of other information
  - What is the prior probability of rolling doubles?
  - What is the prior probability of rolling doubles sixes?

Example Problems

- What is the prior probability of any given 5-card hand?
- What is the probability of a royal flush?
- What is the probability of a straight flush?
- What is the probability of 4 of a kind?
- What is the probability of getting exactly one pair?

Conditional Probability

- The conditional probability is the probability of an event given other information about the world
  - P(*roll doubles* | *die 1 is a 6*)
  - P(*roll double 3* | *die 1 is a 6*)
- Note that the conditional information doesn't change the prior probability
  - A women has 10 daughters. What is the chance that her 11^{th} child is a daughter?
  - What is the probability of a women having all 11 children be daughters?

Conditional probabilities

- Conditional probabilities can be computed from joint distributions and prior probabilities
  - $P(a|b) = P(a \wedge b) / P(b)$ [as long as $P(b) > 0$]
  - P(*double* | *die 1 is a 5*) = P(*doubles* ∧ *die 1 is a 5*) / P(*die 1 is a 5*)
- Product rule:
  - $P(a \wedge b) = P(a|b)\, P(b)$

Examples

- Given that you have seen your first card, the A of Hearts
- What is the probability of a royal flush?
- What is the probability of a straight flush?
- What is the probability of 4 of a kind?
- What is the probability of getting exactly one pair?

Propositions and Probability

- Variables are called random variables [upper case]
- Can talk about all probabilities for a random variable
- **P**(Weather) = <0.6, 0.1, 0.29, 0.01>
  - P(Weather = sunny) = 0.6
  - P(Weather = rain) = 0.1
  - P(Weather = cloudy) = 0.29
  - P(Weather = snow) = 0.01
- **P**(Weather, Cavity) is a joint probability distribution

Foundations of probability

- Kolmogorov's axioms:

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

$$P(a \vee b) = P(a) + P(b) - P(a \wedge b)$$

- If an agent has beliefs that aren't consistent with these axioms, then the agent will not act rationally

Inference

- Given a full joint distribution, how can we answer queries?

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	0.108	0.012	0.072	0.008
¬cavity	0.016	0.064	0.144	0.576

- What is P(cavity ∨ toothache)?
- What is P(cavity)?

Marginalization

- Marginalization is the process of summing given possible values for other variables

$$P(Y) = \sum_{z \in Z} P(Y, z)$$

- Conditioning is similarly defined with conditional distributions

$$P(Y) = \sum_{z \in Z} P(Y|z)P(z)$$

Examples

- What is the probability of a cavity, given a toothache?
  - P(cavity | toothache) = P(cavity ∧ toothache) / P(toothache)
- What is the probability of no cavity, given a toothache?
  - P(¬cavity | toothache) = P(¬cavity ∧ toothache) / P(toothache)
- P(toothache) is just used for normalization
  - Can computed probabilities without it!

Independence

- Sometimes variables do not have a relationship
  - P(toothache, catch, cavity, cloudy) = P(cloudy | toothache, catch, cavity) · P(toothache, catch, cavity)
  - P(toothache, catch, cavity, cloudy) = P(cloudy) · P(toothache, catch, cavity)
- a and b are independent if P(a|b) = P(a); P(a∧b) = P(a)·P(b)

Bayes Rule

- $P(a \wedge b) = P(a|b)\,P(b) = P(b|a)\,P(a)$
- $P(b|a) = P(a|b)P(b) / P(a)$
- Useful for determining unknown probabilities
  - P(*cause*|*effect*) = P(*effect*|*cause*)P(*cause*) / P(*effect*)
- Meningitis causes a stiff neck. What is the probability a patient has meningitis given a stiff neck?
  - $P(s \mid m) = 0.7$; $P(m) = 1/50000$; $P(s) = 0.01$

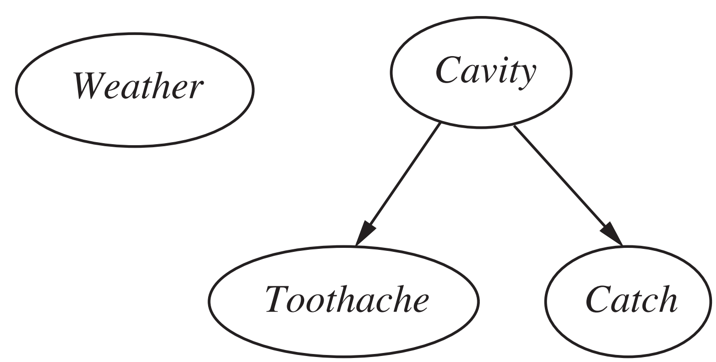
Bayesian Network

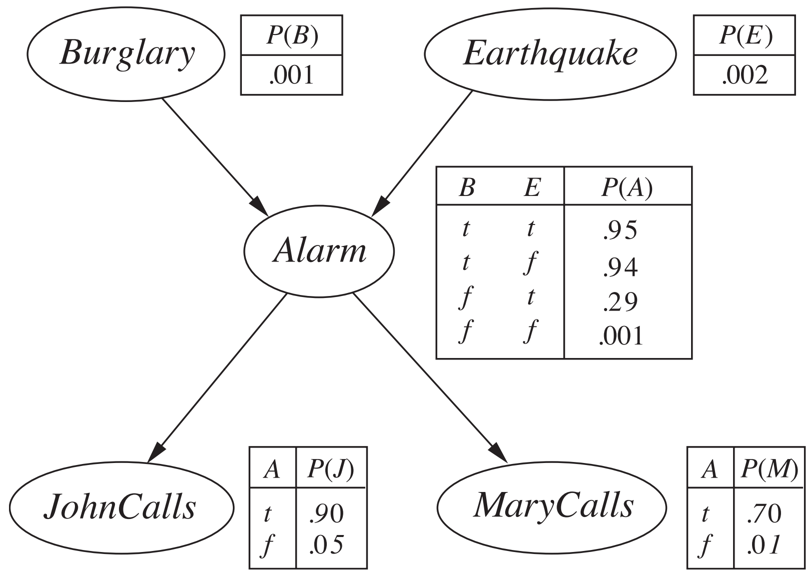
- A Bayesian network is a DAG:
  - Each node corresponds to a random variables
  - Directed edges link nodes
    - Edge goes from *parent* to child
  - Each node has a conditional probability dist.
    - $P(X_i \mid Parents(X_i))$
- Topology of network represents causality

Chain Rule

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

Example network





Building joint distribution

- Joint distribution is made up of 2^n lines in table
- Each line is defined by:

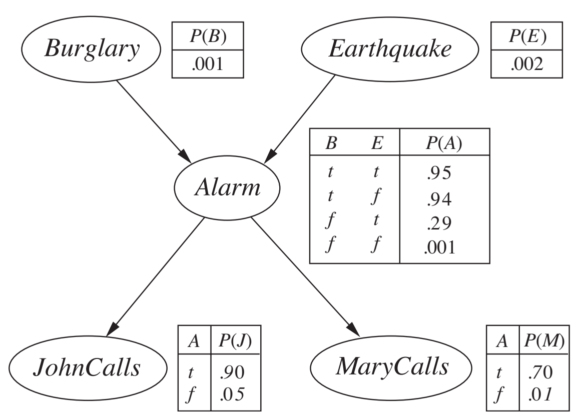
$$P(X_1 = x_1 \wedge \cdots \wedge X_n = x_n)$$

- Which can be computed as:

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i))$$

Example

- What is the probability of an alarm, but no burglary, or earthquake, and John and Mary both call?



Constructing Bayesian Networks

- Nodes in the network with no parents have no conditional distribution
- Construct table by ordering nodes in network
  - Incrementally add nodes to network
- Each time a node is added
  - Add links from any existing nodes
    - Iff the existing node has a causal relationship with new node
    - The variables are dependent