CS 1571 Introduction to AI
Lecture 20

Uncertainty

Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square

KB systems. Medical example.

We want to build a KB system for the **diagnosis of pneumonia**.

Problem description:

- **Disease:** pneumonia
- **Patient symptoms (findings, lab tests)**:
  - Fever, Cough, Paleness, WBC (white blood cells) count, Chest pain, etc.

Representation of a patient case:

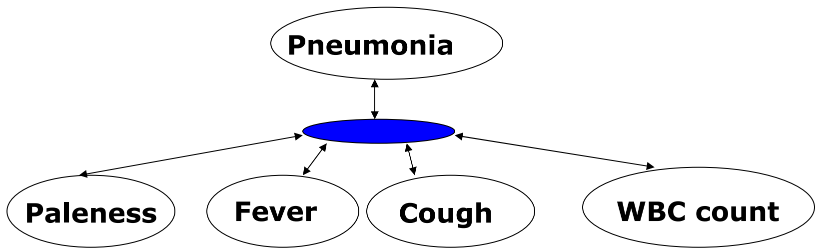
- Statements that hold (are true) for the patient.

  E.g: Fever =*True*
  Cough =*False*
  WBCcount=*High*

Diagnostic task: we want to decide whether the patient suffers from the pneumonia or not given the symptoms

Uncertainty

To make diagnostic inference possible we need to represent knowledge (axioms) that relate symptoms and diagnosis



Problem: disease/symptoms relations are not deterministic

– **They are uncertain (or stochastic) and** vary from patient to patient

Uncertainty

Two types of uncertainty:

- **Disease ⟶ Symptoms uncertainty**
  – A patient suffering from pneumonia may not have fever all the times, may or may not have a cough, white blood cell test can be in a normal range.
- **Symptoms ⟶ Disease uncertainty**
  – High fever is typical for many diseases (e.g. bacterial diseases) and does not point specifically to pneumonia
  – Fever, cough, paleness, high WBC count combined do not always point to pneumonia

Uncertainty

Why are relations uncertain?

- **Observability**
  - It is impossible to observe all relevant components of the world
  - Observable components behave stochastically even if the underlying world is deterministic
- **Efficiency, capacity limits**
  - It is often impossible to enumerate and model all components of the world and their relations
  - abstractions can make the relations stochastic

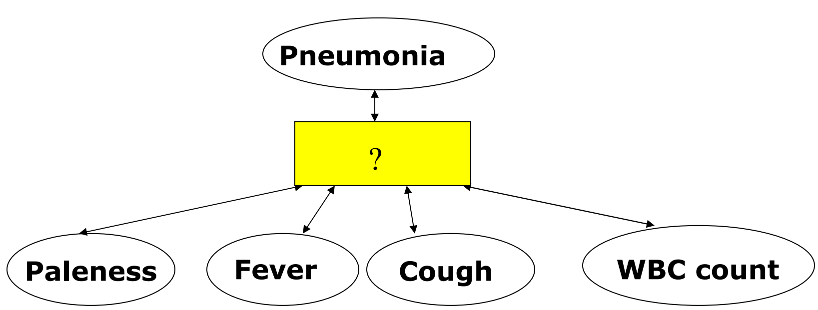
Humans can reason with uncertainty !!!

  - Can computer systems do the same?

Modeling the uncertainty.

Key challenges:

- How to represent the relations in the presence of uncertainty?
- How to manipulate such knowledge to make inferences?
  - **Humans can reason with uncertainty.**



Methods for representing uncertainty

Extensions of the propositional and first-order logic

– Use, uncertain, imprecise statements (relations)

Example: Propositional logic with certainty factors

Very popular in 70-80s in knowledge-based systems (MYCIN)

- **Facts (propositional statements)** are assigned a **certainty value** reflecting the belief in that the statement is satisfied:

$$CF(Pneumonia = True) = 0.7$$

- **Knowledge:** typically in terms of **modular rules**

If	1. The patient has cough, and 2. The patient has a high WBC count, and 3. The patient has fever
Then	**with certainty 0.7** the patient has pneumonia

Certainty factors

Problem 1:

- Chaining of multiple inference rules (propagation of uncertainty)

Solution:

- **Rules** incorporate tests on the **certainty values**

$$(A \text{ in } [0.5,1]) \wedge (B \text{ in } [0.7,1]) \rightarrow C \text{ with CF} = 0.8$$

Problem 2:

- Combinations of rules **with the same conclusion**

$$(A \text{ in } [0.5,1]) \wedge (B \text{ in } [0.7,1]) \rightarrow C \text{ with CF} = 0.8$$
$$(E \text{ in } [0.8,1]) \wedge (D \text{ in } [0.9,1]) \rightarrow C \text{ with CF} = 0.9$$

- What is the resulting *CF(C)* ?

Certainty factors

- **Combination of multiple rules**

$$(A \text{ in } [0.5,1]) \wedge (B \text{ in } [0.7,1]) \rightarrow C \text{ with } \mathrm{CF} = 0.8$$
$$(E \text{ in } [0.8,1]) \wedge (D \text{ in } [0.9,1]) \rightarrow C \text{ with } \mathrm{CF} = 0.9$$

- **Three possible solutions**

$$\left.\begin{aligned} CF(C) &= \max[0.9; 0.8] = 0.9 \\ CF(C) &= 0.9 * 0.8 = 0.72 \\ CF(C) &= 0.9 + 0.8 - 0.9 * 0.8 = 0.98 \end{aligned}\right\} \ ?$$

Problems:

- Which solution to choose?
- All three methods break down after a sequence of inference rules

Methods for representing uncertainty

Probability theory

- A well defined theory for modeling and reasoning in the presence of uncertainty
- A natural choice to replace certainty factors

Facts (propositional statements)

- Are represented via **random variables** with two or more values

  Example: *Pneumonia* is a random variable

  values: *True* and *False*
- Each value can be achieved **with some probability:**

$$P(Pneumonia = True) = 0.001$$

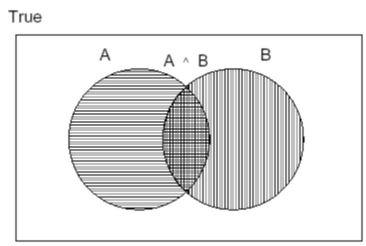
$$P(WBCcount = high) = 0.005$$

Probability theory

- Well-defined theory for representing and manipulating statements with uncertainty
- **Axioms of probability:**

  For any two propositions A, B.

  1. $0 \le P(A) \le 1$
  2. $P(True) = 1$ and $P(False) = 0$
  3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$



Modeling uncertainty with probabilities

Probabilistic extension of propositional logic.

- **Propositions:**
  - statements about the world
  - Represented by the assignment of values to **random variables**
- **Random variables:**

! – **Boolean** $Pneumonia$ is either $True, False$

Random variable **Values**

! – **Multi-valued** $Pain$ is one of $\{Nopain, Mild, Moderate, Severe\}$

Random variable **Values**

– **Continuous** $HeartRate$ is a value in $<0;250>$

Random variable **Values**

Probabilities

Unconditional probabilities (prior probabilities)

$P(Pneumonia) = 0.001$ or $P(Pneumonia = True) = 0.001$

$P(Pneumonia = False) = 0.999$

$P(WBCcount = high) = 0.005$

Probability distribution

- Defines probabilities **for all possible value assignments to a random variable**
- Values are mutually exclusive

$P(Pneumonia = True) = 0.001$
$P(Pneumonia = False) = 0.999$

$Pneumonia$	$\mathbf{P}(Pneumonia)$
$True$	0.001
$False$	0.999

Probability distribution

Defines probability for **all possible value assignments**

Example 1:

$P(Pneumonia = True) = 0.001$
$P(Pneumonia = False) = 0.999$

$Pneumonia$	$\mathbf{P}(Pneumonia)$
$True$	0.001
$False$	0.999

$$P(Pneumonia = True) + P(Pneumonia = False) = 1$$

Probabilities sum to 1 !!!

Example 2:

$P(WBCcount = high) = 0.005$
$P(WBCcount = normal) = 0.993$
$P(WBCcount = high) = 0.002$

$WBCcount$	$\mathbf{P}(WBCcount)$
$high$	0.005
$normal$	0.993
low	0.002

Joint probability distribution

Joint probability distribution (for a set variables)

- Defines probabilities for **all possible assignments of values to variables in the set**

Example: variables *Pneumonia* and *WBCcount*

$\mathbf{P}(pneumonia, WBCcount)$

Is represented by 2×3 matrix

Pneumonia / *WBCcount*

	high	*normal*	*low*
True	0.0008	0.0001	0.0001
False	0.0042	0.9929	0.0019

Joint probabilities

Marginalization

- reduces the dimension of the joint distribution
- Sums variables out

$\mathbf{P}(pneumonia, WBCcount)$ 2×3 matrix

Pneumonia / *WBCcount*

	high	*normal*	*low*	$\mathbf{P}(Pneumonia)$
True	0.0008	0.0001	0.0001	0.001
False	0.0042	0.9929	0.0019	0.999
$\mathbf{P}(WBCcount)$	0.005	0.993	0.002	

Marginalization (here summing of columns or rows)

Marginalization

Marginalization

- reduces the dimension of the joint distribution

$$P(X_1, X_2, \ldots X_{n-1}) = \sum_{\{X_n\}} P(X_1, X_2, \ldots X_{n-1}, X_n)$$

- We can continue doing this

$$P(X_2, \ldots X_{n-1}) = \sum_{\{X_1, X_n\}} P(X_1, X_2, \ldots X_{n-1}, X_n)$$

What is the maximal joint probability distribution?

- Full joint probability

Full joint distribution

- **the joint distribution for all variables in the problem**
  - It defines the complete probability model for the problem

Example: pneumonia diagnosis

Variables: *Pneumonia, Fever, Paleness, WBCcount, Cough*

Full joint defines the probability for all possible assignments of values to *Pneumonia, Fever, Paleness, WBCcount, Cough*

$$P(Pneumonia = T, WBCcount = High, Fever = T, Cough = T, Paleness = T)$$

$$P(Pneumonia = T, WBCcount = High, Fever = T, Cough = T, Paleness = F)$$

$$P(Pneumonia = T, WBCcount = High, Fever = T, Cough = F, Paleness = T)$$

$\ldots$ etc

Full joint distribution

- **Any joint probability for a subset of variables can be obtained via marginalization**

$$P(Pneumonia, WBCcount, Fever) =$$
$$\sum_{c,p=\{T,F\}} P(Pneumonia, WBCcount, Fever, Cough = c, Paleness = p)$$

- **Is it possible to recover full joint from the joint probabilities over a subset of variables?**