DRONACHARYA College of Engineering

INTELLIGENT SYSTEMS (CSE-303-F)

Section C

Uncertainity

Representing uncertain knowledge

- Points
- Symbolic and numerical uncertainty
- The Closed World Assumption
- Predicate completion
- Taxonomic hierarchies
- Abduction
- Truth maintenance
- Bayes’s rule
- Certainty factors
- Fuzzy sets

Symbolic and numerical uncertainty

- <u>Symbolic uncertainty</u>: defeasible reasoning.
  - Non-monotonic logic:
    make conclusions that cause no inconsistency.
  - Default logic.
  - Modal logic -- necessity and possibility:
    necessarily Φ ≡ ¬ possibly ¬ Φ
    possibly Φ ≡ ¬ necessarily ¬ Φ
- <u>Numeric uncertainty</u>: statistical reasoning, certainty factors.
  - Bayesian probability theory.
  - Fuzzy logic (theory of fuzzy sets).

Examples of approximate reasoning in real life

- Judging by general shape, not by details.
- Jumping to conclusions without sufficient evidence.
  (Consider a lawn sign, seen from afar.)
- Understanding language.
- ...
- ...

Why we need more than first-order logic

- Complete ↔ incomplete knowledge:
- the world cannot be represented completely,
there are exceptions and qualified statements.
- Generality ↔ specificity (and typicality):
- absolute statements ignore individual variety.
- Consistency ↔ inconsistency:
- conflicting views cannot be represented in first-order logic,
not within one theory.
- Monotonic ↔ defeasible reasoning:
- a change of mind cannot be represented.
- Absolute ↔ tentative statements:
- partial commitment cannot be represented,
statistical tendencies cannot be expressed.
- Finality ↔ openness of knowledge:
- learning should not only mean new theorems.

The Closed-World Assumption

- A complete theory in first-order logic must include either a fact or its negation.
- The Closed World Assumption (CWA) states that the only true facts are those that are explicitly listed as true (in a knowledge base, in a database) or are provably true.
- We may <u>extend</u> the theory by adding to it explicit negations of facts that cannot be proven.
- In Prolog, this principle is implemented by "finite failure", or "negation as failure".

Predicate completion

- The fact **`p(a)`** could be rewritten equivalently as
- $\forall x \; x = a \Rightarrow p(x)$
- A <u>completion</u> of this quantified formula is a formula in which we enumerate all objects with property **`p`**. In this tiny example:
- $\forall x \; p(x) \Rightarrow x = a$
- A larger example: the world of birds.
- $\forall x \; ostrich(x) \Rightarrow bird(x)$
- $\neg ostrich(Sam)$
- $bird(Tweety)$
- To complete the predicate **`bird`**, we say:
- $\forall x \; bird(x) \Rightarrow (ostrich(x) \vee x = Tweety)$
- This allows us to prove, for example, that
- $\neg bird(Sam)$

Predicate completion (2)

- How to achieve non-monotonic reasoning? Let us add a new formula:
- `∀x penguin(x) ⇒ bird(x)`
- `∀x ostrich(x) ⇒ bird(x)`
- `¬ostrich(Sam)`
- `bird(Tweety)`
- A new completion of the predicate `bird` would be:
- `∀x bird(x) ⇒`
- `(ostrich(x) ∨ penguin(x) ∨ x = Tweety)`
- We cannot prove `¬bird(Sam)` any more (why?).

Taxonomic hierarchies and defaults

- $\mathtt{thing(Tweety)}$
- $\mathtt{bird(x)} \Rightarrow \mathtt{thing(x)}$
- $\mathtt{ostrich(x)} \Rightarrow \mathtt{bird(x)}$
- $\mathtt{flying\text{-}ostrich(x)} \Rightarrow \mathtt{ostrich(x)}$
- The following set of formulae represents typicality and exceptions:
- $\mathtt{thing(x)} \wedge \neg\mathtt{bird(x)} \Rightarrow \neg\mathtt{flies(x)}$
- $\mathtt{bird(x)} \wedge \neg\mathtt{ostrich(x)} \Rightarrow \mathtt{flies(x)}$
- $\mathtt{ostrich(x)} \wedge \neg\mathtt{flying\text{-}ostrich(x)} \Rightarrow$
- $\neg\mathtt{flies(x)}$
- $\mathtt{flying\text{-}ostrich(x)} \Rightarrow \mathtt{flies(x)}$

Taxonomic hierarchies and defaults (2)

- This works, but is too specific. We need a way of showing exceptions explicitly. An exception is a departure from normality, and a way of blocking inheritance:
- $\mathbf{thing(x)} \wedge \neg \mathbf{abnormal_t(x)} \Rightarrow \neg \mathbf{flies(x)}$
- $\mathbf{bird(x)} \Rightarrow \mathbf{abnormal_t(x)}$
- $\mathbf{bird(x)} \wedge \neg \mathbf{abnormal_b(x)} \Rightarrow \mathbf{flies(x)}$
- $\mathbf{ostrich(x)} \Rightarrow \mathbf{abnormal_b(x)}$
- $\mathbf{ostrich(x)} \wedge \neg \mathbf{abnormal_o(x)} \Rightarrow \neg \mathbf{flies(x)}$
- $\mathbf{flying\text{-}ostrich(x)} \Rightarrow \mathbf{abnormal_o(x)}$
- $\mathbf{flying\text{-}ostrich(x)} \Rightarrow \mathbf{flies(x)}$

Taxonomic hierarchies and defaults (3)

- The taxonomy is now as follows:
- **flying-ostrich(x) ⇒ ostrich(x)**
- **flying-ostrich(x) ⇒ abnormal$_o$(x)**
- **ostrich(x) ⇒ bird(x)**
- **ostrich(x) ⇒ abnormal$_b$(x)**
- **bird(x) ⇒ thing(x)**
- **bird(x) ⇒ abnormal$_t$(x)**
- **thing(Tweety)**
- The properties of classes:
- **thing(x) ∧ ¬abnormal$_t$(x) ⇒ ¬flies(x)**
- **bird(x) ∧ ¬abnormal$_b$(x) ⇒ flies(x)**
- **ostrich(x) ∧ ¬abnormal$_o$(x) ⇒ ¬flies(x)**
- **flying-ostrich(x) ⇒ flies(x)**

Taxonomic hierarchies and defaults (4)

- This kind of formulae can be used to deduce properties of objects, if we can also supply the completion. For example, does Tweety fly?
- The completion:
- $\mathtt{thing(x)} \Rightarrow \mathtt{bird(x)} \vee \mathtt{x = Tweety}$
- $\mathtt{bird(x)} \Rightarrow \mathtt{ostrich(x)}$
- $\mathtt{ostrich(x)} \Rightarrow \mathtt{flying\text{-}ostrich(x)}$
- $\mathtt{abnormal_t(x)} \Rightarrow \mathtt{bird(x)}$
- $\mathtt{abnormal_b(x)} \Rightarrow \mathtt{ostrich(x)}$
- $\mathtt{abnormal_o(x)} \Rightarrow \mathtt{flying\text{-}ostrich(x)}$
- $\neg\mathtt{flying\text{-}ostrich(x)}$

Taxonomic hierarchies and defaults (5)

- The last formula is equivalent to
- **flying-ostrich(x) $\Rightarrow$ false**
- which reflects the fact that no taxonomical rule has **flying-ostrich** as a conclusion.
- We can now prove all of these:
- **¬flying-ostrich(Tweety)**
- **¬ostrich(Tweety)**
- **¬bird(Tweety)**
- **¬abnormal$_t$(Tweety)**
- so we can show that
- **¬flies(Tweety)**

Taxonomic hierarchies and defaults (6)

- This hierarchy again can be changed non-monotonically. Suppose that we add
- **bird(Tweety)**
- to the taxonomy.

- The new completion of the predicate **bird** will be:
- $\mathbf{bird(x)} \Rightarrow \mathbf{ostrich(x)} \vee \mathbf{x} = \mathbf{Tweety}$
- instead of
- $\mathbf{bird(x)} \Rightarrow \mathbf{ostrich(x)}$

- We will not be able to prove
- $\neg \mathbf{abnormal_t(Tweety)}$
- any more. We will, however, be able to prove
- $\neg \mathbf{abnormal_b(Tweety)}$

A quaker and a republican

- Quakers are pacifists. Republicans are not pacifists. Richard is a republican and a quaker. Is he a pacifist?
- These rules are ambiguous. Let us clarify:
- Only a typical quaker is a pacifist. Only a typical republican is not a pacifist.
- This can be expressed in terms of consistency:
- **∀x quaker(x) ∧ CONSISTENT(pacifist(x)) ⇒ pacifist(x)**
- **∀x republican(x) ∧ CONSISTENT(¬pacifist(x)) ⇒ ¬pacifist(x)**
- If we apply the first rule to Richard, we find he is a **pacifist** (nothing contradicts this conclusion), but then the second rule cannot be used—and vice versa. In effect, neither **pacifist(x)** nor **¬pacifist(x)** can be proven.

Abduction ... demonstrated on one example

- Abduction means systematic guessing: "infer" an assumption from a conclusion. For example, the following formula:
- **∀x rainedOn(x) ⇒ wet(x)**
- could be used "backwards" with a specific **x**:
- if **wet(Tree)** then **rainedOn(Tree)**
- This, however, would not be logically justified. We could say:
- **wet(Tree) ∧ CONSISTENT(rainedOn(Tree)) ⇒ rainedOn(Tree)**
- We could also attach probabilities, for example like this:
- **wet(Tree) ⇒ rainedOn(Tree) || 70%**
- **wet(Tree) ⇒ morningDewOn(Tree) || 20%**
- **wet(Tree) ⇒ sprinkled(Tree) || 10%**

Truth maintenance

- … demonstrated on one example

The first formula arrives: we build a partial network.

happy(a) $\Rightarrow$ happy(b)

happy(a)
truth value: UNKNOWN
justification: none

happy (a) $\rightarrow$ happy(b)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

happy(b)
truth value: UNKNOWN
justification: none

Truth maintenance (2)

A new fact: incorporate it into the network.
happy(a)

happy(a)
truth value: TRUE
justification: given

happy (a) → happy(b)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

happy(b)
truth value: TRUE
justification: deduced

Truth maintenance (3)

A new rule and a new fact:

happy(b) $\Rightarrow$ happy(c)

happy(d)

happy(a)
truth value: TRUE
justification: given

happy (a) $\rightarrow$ happy(b)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

happy(b)
truth value: TRUE
justification: deduced

happy (b) $\rightarrow$ happy(c)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

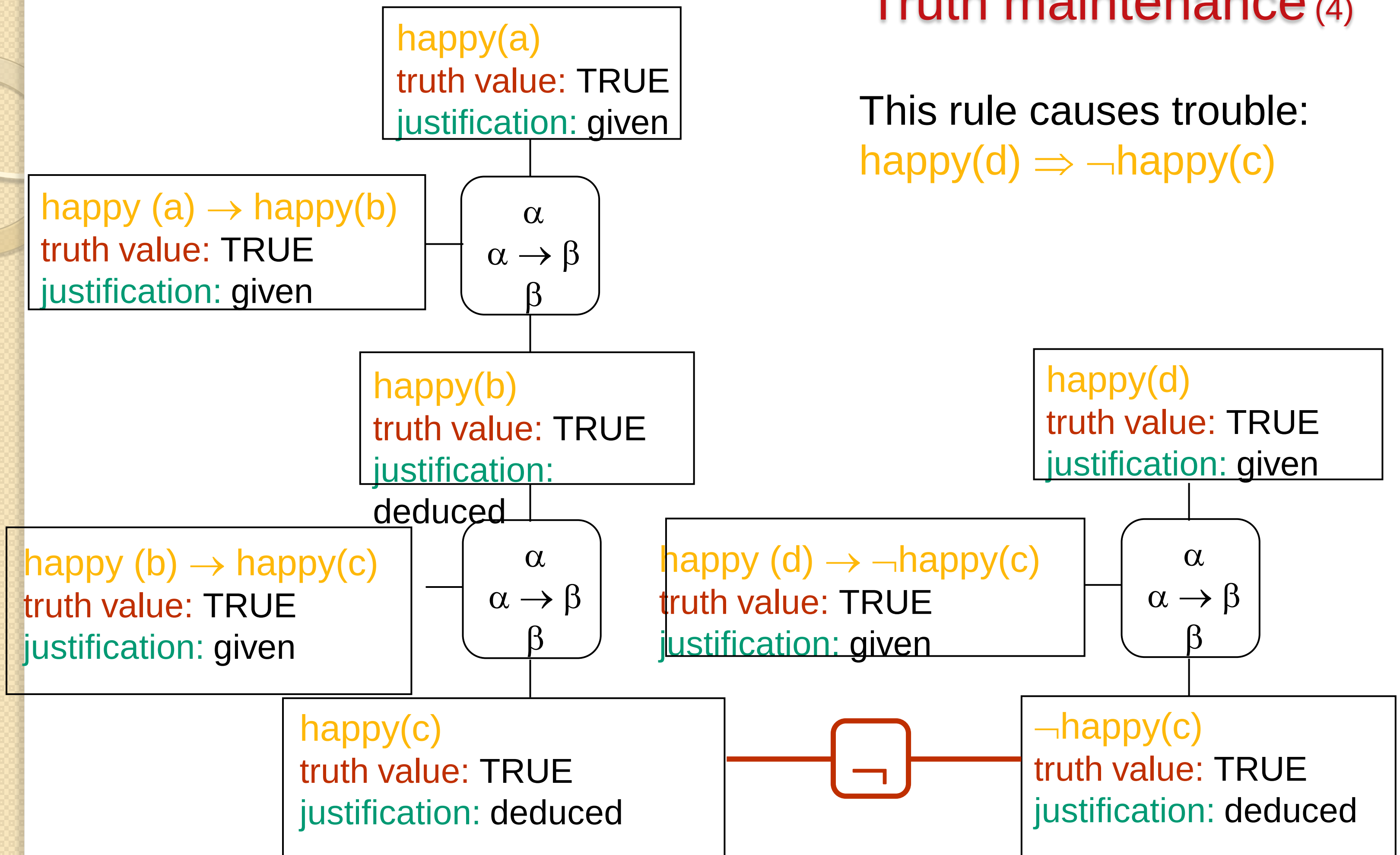
happy(c)
truth value: TRUE
justification: deduced

happy(d)
truth value: TRUE
justification: given

Truth maintenance (4)

This rule causes trouble:

happy(d) $\Rightarrow$ ¬happy(c)



Truth maintenance (5)

One possible fix: disable
happy(d) $\Rightarrow \neg$happy(c)

happy(a)
truth value: TRUE
justification: given

happy (a) $\rightarrow$ happy(b)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

happy(b)
truth value: TRUE
justification: deduced

happy (b) $\rightarrow$ happy(c)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

happy(c)
truth value: TRUE
justification: deduced

happy(d)
truth value: TRUE
justification: given

happy (d) $\rightarrow \neg$happy(c)
truth value: FALSE
justification: assumed

Truth maintenance (6)

Another fix: disable
happy(b) $\Rightarrow$ happy(c)

happy(a)
truth value: TRUE
justification: given

happy (a) $\rightarrow$ happy(b)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

happy(b)
truth value: TRUE
justification: deduced

happy (b) $\rightarrow$ happy(c)
truth value: FALSE
justification: assumed

happy(d)
truth value: TRUE
justification: given

happy (d) $\rightarrow$ $\neg$happy(c)
truth value: TRUE
justification: given

α
$\alpha \rightarrow \beta$
β

$\neg$happy(c)
truth value: TRUE
justification: deduced

Bayes's theorem

- Bayesian probability theory.
- Fuzzy logic (only signalled here).
- Dempster-Shafer theory (not discussed here).
- Bayes's theorem allows us to compute how probable it is that a hypothesis $\mathbf{H}_i$ follows from a piece of evidence **E** (for example, from a symptom or a measurement).
- The required data: the probability of $\mathbf{H}_j$ and the probability of **E** given $\mathbf{H}_j$ for <u>all</u> possible hypotheses.

Bayes's theorem (2)

- Medical diagnosis is a handy example. A patient may have a cold, a flu, pneumonia, rheumatism, and so on. The usual symptoms are high fever, short breath, runny nose, and so on.
- We need the probabilities (based on statistical data?) of all diseases, and the probabilities of high fever, short breath, runny nose in the case of a cold, a flu, pneumonia, rheumatism. [This is asking a lot!]
- We would also like to assume that all relationships between $\mathbf{H}_j$ and $\mathbf{E}$ are mutually independent. [This is asking even more!]

Bayes's theorem (3)

- The probability data:
- p($\mathbf{H}_i$ | $\mathbf{E}$)   the probability of $\mathbf{H}_i$ given $\mathbf{E}$.
- p($\mathbf{H}_i$)   the overall probability of $\mathbf{H}_i$.
- p($\mathbf{E}$ | $\mathbf{H}_i$)   the probability of observing $\mathbf{E}$ given $\mathbf{H}_i$.
- Bayes' theorem

$$p(\mathbf{H}_i \mid \mathbf{E}) = \frac{p(\mathbf{E} \mid \mathbf{H}_i) * p(\mathbf{H}_i)}{\sum_j p(\mathbf{E} \mid \mathbf{H}_j) * p(\mathbf{H}_j)}$$

Bayes's theorem (4)

- If we assume that all the conditional probabilities under summation are independent, we can simplify the formula:

$$p(\mathbf{H}_i \mid \mathbf{E}) = \frac{p(\mathbf{E} \mid \mathbf{H}_i) * p(\mathbf{H}_i)}{p(\mathbf{E})}$$

Bayes's theorem (5)

Example

- "A poker player closes one eye 9 times out of 10 before passing a hand. He passes 50% of the hands, and closes one eye during 60% of the hands. What is the probability that he will pass a hand given that he closes one eye?"
- $\mathbf{H}_j$: the player passes a hand.
- **E**: the player closes one eye.
- $p(\mathbf{E} \mid \mathbf{H}_j) = 0.9$
- $p(\mathbf{E}) = 0.6$
- $p(\mathbf{H}_j) = 0.5$
- $p(\mathbf{H}_j \mid \mathbf{E}) = 0.9 * 0.5 / 0.6 = 0.75$

Odds calculation

- Yet another version of Bayes's formula is based on the concepts of odds and likelihood.

$$p(\mathbf{H} \mid \mathbf{E}) = \frac{p(\mathbf{E} \mid \mathbf{H}) * p(\mathbf{H})}{p(\mathbf{E})}$$

$$p(\neg\mathbf{H} \mid \mathbf{E}) = \frac{p(\mathbf{E} \mid \neg\mathbf{H}) * p(\neg\mathbf{H})}{p(\mathbf{E})}$$

Odds calculation (2)

- These two formulae give this:

$$\frac{p(\,\mathbf{H} \mid \mathbf{E}\,)}{p(\,\neg\mathbf{H} \mid \mathbf{E}\,)} = \frac{p(\,\mathbf{E} \mid \mathbf{H}\,) * p(\,\mathbf{H}\,)}{p(\,\mathbf{E} \mid \neg\mathbf{H}\,) * p(\,\neg\mathbf{H}\,)}$$

The odds of event **e**:

$$O(\,\mathbf{e}\,) = \frac{p(\,\mathbf{e}\,)}{p(\,\neg\mathbf{e}\,)} = \frac{p(\,\mathbf{e}\,)}{1 - p(\,\mathbf{e}\,)}$$

We note that $p(\,\mathbf{H} \mid \mathbf{E}\,) + p(\,\neg\mathbf{H} \mid \mathbf{E}) = 1$.

Odds calculation (3)

$$O(\mathbf{H} \mid \mathbf{E}) = \frac{p(\mathbf{E} \mid \mathbf{H})}{p(\mathbf{E} \mid \neg\mathbf{H})} * O(\mathbf{H})$$

Define the fraction as the likelihood ratio $\lambda(\mathbf{E}, \mathbf{H})$ of a piece of evidence **E** with respect to hypothesis **H**:

$$O(\mathbf{H} \mid \mathbf{E}) = \lambda(\mathbf{E}, \mathbf{H}) * O(\mathbf{H})$$

An intuition: how to compute the new odds of **H** (given additional evidence **E**) from the previous odds of **H**.

$\lambda > 1$ <u>strengthens</u> our belief in **H**.

Odds calculation (4)

<u>Example</u>

25% of students in the AI course get an A.

80% of students who get an A do all homework.

60% of students who do not get an A do all homework.

75% of students who get an A are CS majors.

50% of students who do not get an A are CS majors.

Irene does all her homework is the AI course.

Mary is a CS major and does all her homework.

What are Irene's and Mary's odds of getting an A?

Let A = "gets an A".

C = "is a CS major".

W = "does all homework".

Odds calculation (5)

Example

$p(A) = 0.25$

$p(W \mid A) = 0.8$ $p(W \mid \neg A) = 0.6$

$p(C \mid A) = 0.75$ $p(C \mid \neg A) = 0.5$

$$O(A \mid W) = \frac{p(A \mid W)}{p(\neg A \mid W)} = \frac{p(W \mid A) * p(A)}{p(W \mid \neg A) * p(\neg A)}$$

$$= \frac{0.8 * 0.25}{0.6 * 0.75} = \frac{4}{9}$$

Odds calculation (6)

Example

$p(A) = 0.25$

$p(W \mid A) = 0.8 \qquad p(W \mid \neg A) = 0.6$

$p(C \mid A) = 0.75 \qquad p(C \mid \neg A) = 0.5$

$$O(A \mid C \wedge W) = \frac{p(A \mid C \wedge W)}{p(\neg A \mid C \wedge W)} = \frac{p(C \wedge W \mid A) * p(A)}{p(C \wedge W \mid \neg A) * p(\neg A)}$$

$$= \frac{p(C \mid A) * p(W \mid A) * p(A)}{p(C \mid \neg A) * p(W \mid \neg A) * p(\neg A)} = \frac{0.75 * 4}{0.5 * 9} = \frac{2}{3}$$

The Stanford certainty factor algebra

Textbook, section 9.2.1

- MB(H | E): the measure of belief in H given E.
- MD(H | E): the measure of disbelief in H given E.
- Each piece of evidence must be either <u>for</u> or <u>against</u> a hypothesis:
- either $0 < MB(H \mid E) < 1$ while $MD(H \mid E) = 0$,
- or $0 < MD(H \mid E) < 1$ while $MB(H \mid E) = 0$.
- The certainty factor is:
- $CF(H \mid E) = MB(H \mid E) - MD(H \mid E)$

The Stanford certainty factor algebra (2)

- Certainty factors are attached to premises of rules in production systems (it started with MYCIN). We need to calculate the CF for conjunctions and disjunctions:
- $CF(P_1 \wedge P_2) = \min(CF(P_1), CF(P_2))$
- $CF(P_1 \vee P_2) = \max(CF(P_1), CF(P_2))$
- We also need to compute the CF of a result supported by two rules with factors CF_1 and CF_2:
- $CF_1 + CF_2 - CF_1 * CF_2$ when $CF_1 > 0$, $CF_2 > 0$,
- $CF_1 + CF_2 + CF_1 * CF_2$ when $CF_1 < 0$, $CF_2 < 0$,

$$\frac{CF_1 + CF_2}{1 - \min(|CF_1|, |CF_2|)}$$

when signs differ.

Read: textbook, section 9.2.2

Fuzzy sets

A crisp set $C \subseteq S$ is defined by a characteristic function $\chi_C(s): S \to \{0, 1\}$.

$$\chi_C(s) = \begin{cases} 0 \text{ if } s \notin C \\ 1 \text{ if } s \in C \end{cases}$$

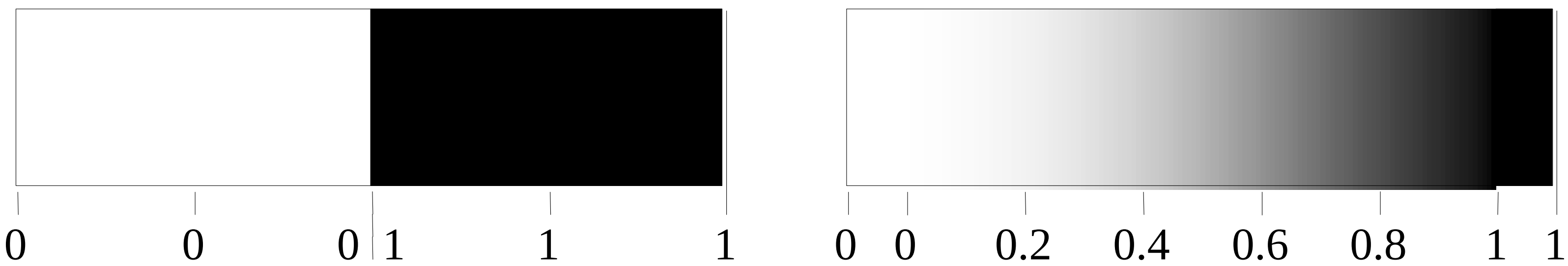
$$\mu_F(s) = \begin{cases} 0.0 \text{ if } s \text{ is not in } F \\ 0.0 < m < 1.0 \text{ if } s \text{ is partially in } F \\ 1.0 \text{ if } s \text{ is totally in } F \end{cases}$$

A fuzzy set $F \subseteq S$ is defined by a membership function $\mu_F(s): S \to [0.0, 1.0]$.

$\mu_F(s)$ describes to what *degree* s belongs to F: 1.0 means "definitely belongs", 0.0 means "definitely does not belong", other values indicate intermediate "degrees" of belonging.

Fuzzy sets (2)

Range of logical values in Boolean and fuzzy logic



(*a*) Boolean Logic. (*b*) Multi-valued Logic.

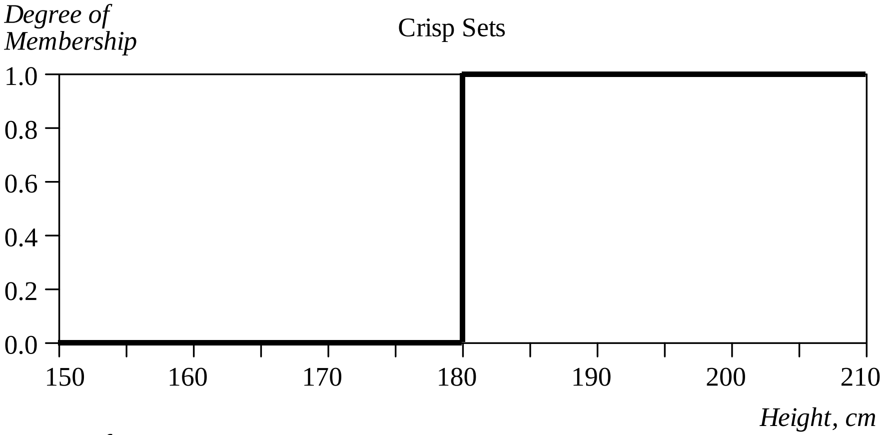
Fuzzy sets (3)

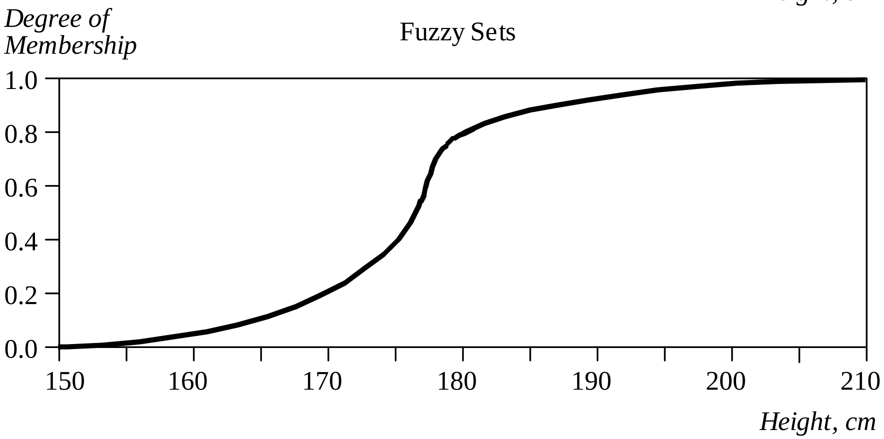
- Consider *N*, the set of positive integers.
- Let $F \subset N$ be the set of "small integers".
- Let μ_F be like this:
- $\mu_F(1) = 1.0$
- $\mu_F(2) = 1.0$
- $\mu_F(3) = 0.9$
- $\mu_F(4) = 0.8$
- ...
- $\mu_F(50) = 0.001$
- ...
- μ_F defines a probability distribution for statements such as "X is a small integer".

Fuzzy sets (4)

Tall men

Name	Height, cm	Degree of Membership	
		Crisp	*Fuzzy*
Chris	208	1	1.00
Mark	205	1	1.00
John	198	1	0.98
Tom	181	1	0.82
David	179	0	0.78
Mike	172	0	0.24
Bob	167	0	0.15
Steven	158	0	0.06
Bill	155	0	0.01
Peter	152	0	0.00

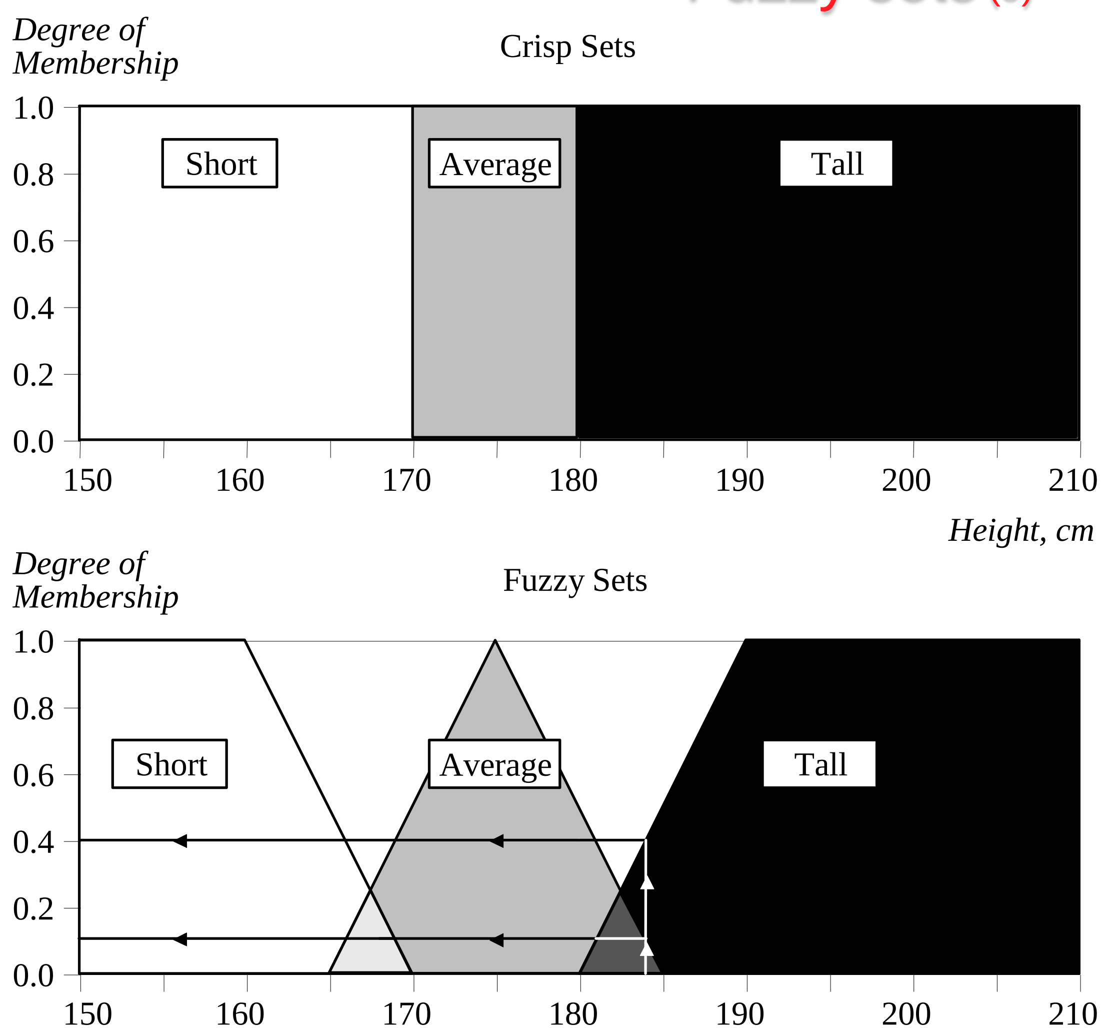




Fuzzy sets (5)

Sets of short, average and tall men

.. and a man 184 cm tall



Fuzzy sets (6)

Basic operations on fuzzy sets

$\mu_{\neg A}(x) = 1 - \mu_A(x)$

$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) = \mu_A(x) \cap \mu_B(x)$

$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)) = \mu_A(x) \cup \mu_B(x)$

This is the tip of a (fuzzy) iceberg. We have fuzzy “logic” and fuzzy rules, fuzzy inference, fuzzy expert systems, and so on.

Even fuzzy cubes...

http://ceeserver.cee.cornell.edu/asce/ConcreteCanoe/Icebreaker/pics/nonraceday/fuzzy_cubes.JPG