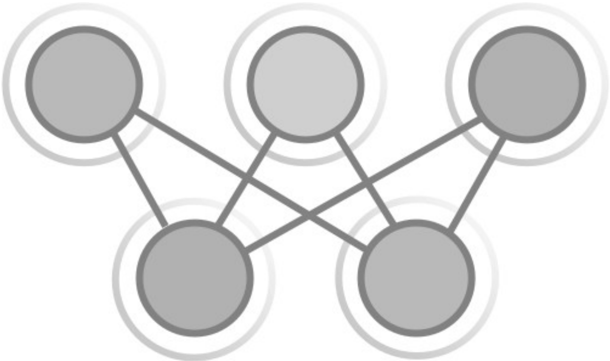
Prof.dr.sc. Bojana Dalbelo Bašić

Fakultet elektrotehnike i računarstva
Zavod za elekroniku, mikroelektroniku, računalne i inteligentne sustave

www.zemris.fer.hr/~bojana
bojana.dalbelo@fer.hr

Modeling uncertainty



V1.1

- Solving problems intelligently requires that we handle and compute with uncertain data

- Possible causes:
  - Data is unavailable or missing;
  - Data exists, but it's unclear or unreliable;
  - Data representation is imprecise;
  - Default values may have exceptions;

MODELING UNCERTAINTY IN KNOWLEDGE-BASED SYSTEMS

- Knowledge based systems that take uncertainty into account must cope with the following:

- How to represent vague data
- How to combine vague data
- How to infer conclusions from vague data

- There are four major numerically-oriented models for dealing with uncertainty:

1. The Bayes scheme

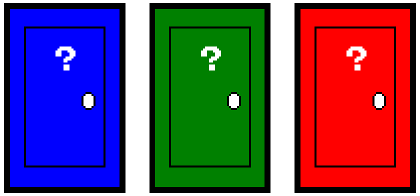
2. Fuzzy sets and fuzzy logic

3. Certainty factors

4. Dempster-Shafer theory

Monty Hall problem

- You're a participant in a game show presented by Monty
- You can choose among **three doors.** Behind one of them there's a car, while behid the other two there are goats. The car and the goats are randomly distributed behind te doors before the show starts.



- <u>The rules of the game</u>: after you chose one door, the door remains closed. The show presenter, Monty Hall, who knows what is hidden behind each of the doors, opens up one of the two remaining doors.

Monty Hall problem

- If behind one of the remaining two doors there's a car, Monty will open up the door behind which there's a goat.
- If goats are behind both doors, Monty will choose at random which door to open.
- After Monty opens up a door behind which there's a goat, he asks you whether you stick to your first choice, or you'd like to change your mind.

- For example, you've chosen door 1, Monty opens up door 3, behind which there's a goat

  Will you change to door 2, or stick to door 1?

.... Bayes scheme

BAYES SCHEME

- The oldest uncertainty modelling method
- Based on traditional probability theory

The basics

random experiment $\rightarrow$ random event x_i

elementary

complex

- $X = \{x_1, x_2, \ldots, x_n\}$ – **space of elementary events**
- An event is a subset of X
- P(X) – **space of all possible events**

BAYES SCHEME

- X – **certain event**
- ∅ – **impossible event**
- $\sim x_i$ – it is not the case that x_i (**opposite event**)
- $x_i \wedge x_j$ occurrence of both events x_i **and** x_j (**intersection**)
- $x_i \vee x_j$ occurence of event x_i **or** x_j (**union**)

- The classical laws of probability theory also apply to events (commutativity, associativity, DeMorgan's laws, etc.)

If $x_i \wedge x_j = \emptyset$ – **mutually exclusive events**

Definition

Probability p is a function, $p : P(X) \rightarrow [0, 1]$

- $0 \leq p(x_i) \leq 1$, for $\forall x_i \in X$, such that $p(X) = 1$
- If $x_1, x_2, \ldots, x_k$ are mutually exclusive, then $p(\cup_i x_i) = \Sigma_i\, p(x_i)$

- From this it follows $\Rightarrow p(x_i) + p(\sim x_i) = 1$ (1)

Definition

- Events x_1 and x_2 are **independent** if

$p\,(x_1 \wedge x_2) = p(x_1)\, p(x_2)$

Remark: If events x and y are independent, then

$$p(x|y) = p(x) \text{ and } p(y|x) = p(y)$$

Derivation of the Bayes rule

By definition, the probability of event x given y is

$$p(x|y) = \frac{p(y \wedge x)}{p(y)} \qquad (3)$$

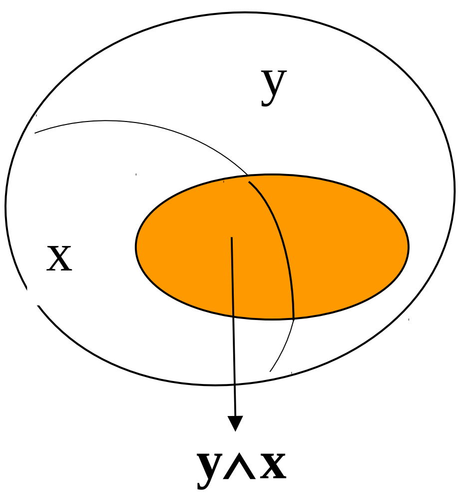
- From (3) $\Rightarrow p(y \wedge x) = p(x|y)p(y)$ (4)
- Due to commutativity: $p(x \wedge y) = p(x|y)p(y)$ (5)
- By substituting (5) in (2) $\Rightarrow$ $p(y|x) = \frac{p(x|y)p(y)}{p(x)}$ (6)
- (6) the simplest form of the Bayes rule

Conditional probability

- Let x and y be events, i.e., x, y $\subset$ X.
- Let us assume that **x** occurred
- *What is the probability of **y** if we know that **x** has occurred ?*
- We call this **conditional probability - p(y|x)**.

Definition

- **Condiitonal probability** is given by p(y|x) = $\frac{p(x \wedge y)}{p(x)}$. (2)



X – set of elementary events

BAYES SCHEME

$$p(y|x) = \frac{p(x|y)\,p(y)}{p(x)}$$

Rule denominator:

- $p(x) = p(x \wedge X) = p(x \wedge (y \vee \sim y)) = p((x \wedge y) \vee (x \wedge \sim y)) =$ according to (ii) from the definition, $(x \wedge y) \cap (x \wedge \sim y) = \emptyset \Rightarrow$ $p(x) = p(x \wedge y) + p(x \wedge \sim y)$

- Earch of the above terms can be rewritten as:

$$p(x \wedge y) = p(x|y)p(y),$$

$$p(x \wedge \sim y) = p(x|\sim y)p(\sim y)$$

Thus the denominator p(x) can be written as:

- $p(x) = p(x|y)p(y) + p(x|\sim y)p(\sim y)$, (7)

- This gives us the well-known Bayes rule:

$$p(y|x) = \frac{p(x|y)\,p(y)}{p(x|y)\,p(y) + p(x|\sim y)\,p(\sim y)} \qquad (8)$$

BAYES SCHEME IN EXPERT SYSTEMS

IF hypothesis H is true ,

THEN conclusion/evidence E is true, with probability p

Instead of p(y|x) **.... p(E|H) = *p***

The above IF-THEN rule can be interpreted as follows:

H in rule (x in formula) denotes a **hypothesis** $\rightarrow$ **H**,

E in rule (y in formula) denotes a **fact or evidence** $\rightarrow$ **E**

BAYES SCHEME IN EXPERT SYSTEMS

IF hypothesis H is true ,

THEN conclusion/evidence E is true, with probability p

Example:

IF *the patient has flu* (H)

THEN *the patient has rhinitis* (E) with probability 0.75

$$p(E|H) = 0.75$$

- VICE VERSA: If E is true (*patient has rhinitis*) , what can we infer about H (*patient has flu*)?

$$p(H|E) \quad ?$$

Which inference rule is this?
Is it a sound rule?

$$\begin{array}{l} E \\ \underline{H \to E} \\ H \end{array}$$

BAYES SCHEME IN EXPERT SYSTEMS

$$p(H|E) = \frac{p(E|H)\,p(H)}{p(E)} \quad (9)$$

$$p(H|E) = \frac{p(E|H)\,p(H)}{p(E|H)\,p(H) + p(E|\sim H)\,p(\sim H)} \quad (10)$$

BAYES SCHEME IN EXPERT SYSTEMS

Example:

- Does Ivan have flue (hypothesis), if he has rhinitis (fact) ?

$$p(\text{flue}|\text{rhinitis}) =$$

$$\frac{p(\text{rhinitis}|\text{flu})\,p(\text{flu})}{p(\text{rhinitis}|\text{flu})\,p(\text{flu}) + \mathrm{p}(\text{rhinitis}|\sim\text{flu})\,p(\sim\text{flu})}$$

Suppose we know that:

- p(H) = p(Ivan has flu) = 0.2 $\Rightarrow$ p(~H) = 0.8
- p(E|H) = p(Ivan has rhinitis | Ivan has flu) = 0.75
- p(E|~H) = p(Ivan has rhinitis | Ivan doesn't have flu) = 0.2

Then:

- p(E) = p(Ivan has rhinitis) = (0.75)(0.2) + (0.2)(0.8) = 0.31
- (10) Bayes rule $\Rightarrow$ p(H|E) = p(Ivan has flu | Ivan has rhinitis) = $\frac{0.75 \cdot 0.2}{0.31} = 0.48387$

- Using (10) we can also determine the probability of the hypothesis *Ivan has flue* given fact that *Ivan doesn't have rhinitis:*

- p(H|~E) = $\dfrac{p(\sim E|H)\,p(H)}{p(\sim E)} = \dfrac{(1-0.75)(0.2)}{(1-0.31)} = 0.07246$

By comparing p(H|E) and P(H) **we conclude that:**

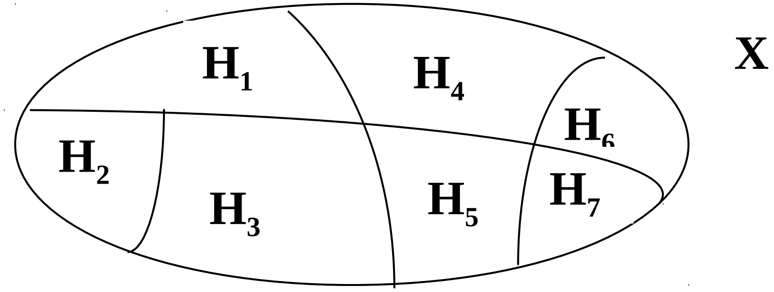
- The fact that Ivan has rhinitis increases the probability that he has flue by approximately 2.5 times

By comparing p(H|~E) and P(H) **we conclude that:**

- The fact that Ivan does not have rhinitis decreases the probability that he has flue by approx. 2.8 times

- Generalization of the Bayes formula to the case of *m* hypotheses, $H_1, H_2, \ldots, H_m$, where $H_1, H_2, \ldots, H_m$ are mutually exclusive,
  i.e. $H_i \cap H_j = \varnothing$ for $i \neq j$

  and union of $H_1, H_2, \ldots, H_m$ covers the whole space X
  i.e. $\cup H_i = X$



(H_i that satify this are called the **complete system of events**)

BAYES SCHEME IN EXPERT SYSTEMS

$$p(H_i|\, E) = \frac{p(E|H_i)\,p(H_i)}{p(E)} = \frac{p(E|H_i)\quad p(H_i)}{\sum_{k=1}^{m} p(E|H_k)\,p(H_k)}\,,\ i=1,\ldots m. \qquad (10a)$$

- Generalization of the Bayes formula to the case of *m* hypotheses $H_1, H_2, \ldots, H_m$ and *n* facts (evidences) $E_1, E_2, \ldots, E_n$.
- $p(H_i|\, E_1\, E_2 \ldots E_n) = \frac{p(E_1\, E_2 \ldots E_n|H_i)\,p(H_i)}{p(E_1 E_2 \ldots E_n)}$ = (cond. indep.)

$$= \frac{p(E_1|H_i)\ \mathrm{p}(E_2|H_i)\ldots\ \mathrm{p}(E_n|H_i)\ \mathrm{p}(H_i)}{\sum_{k=1}^{m} p(E_1|H_k)\,p(\mathrm{E2}|H_k)\ldots p(E_n|H_k)\,p(H_k)} \qquad (11)$$

Monty Hall problem

- Bayes analysis

Notation:

A_i, i=1, 2, 3 – the car is behind door *i*

M_{ij}, i, j=1, 2, 3 – presenter Monty has chosen door *j* after the participant has chosen door *i*

E.g.. A_1 denotes “the car is behind door 1”, while

M_{13} denotes “the presenter has opened up door 3 after the participand has chosen door 1"

Monty Hall problem

- The car can be behind any of the doors, thus each of the doors has the same *prior* probabilty to hide the car.
- In our case, *prior* means “before the show starts” or “before we see a goat”.
- Hence, *prior* probability of event A_i equals:

$$P(A_i) = 1 / 3, \qquad i = 1, 2, 3$$

Monty Hall problem

- If the presenter can choose between two doors, both doors have the same probability of being chosen

- The presenter will choose between the two doors that the participant has not chosen. Moreover, the presenter will always chose the door behid which there is no car.

- These rules determine the conditional probability of event M_{ij} given event A_i:

Monty Hall problem

- Conditional probability M_{ij} given A_i:

$$P(M_{ij}|A_k) = \begin{cases} 0 & \text{if } i = j \text{ (the presenter will not open the door chosen by the participant),} \\ 0 & \text{if } j = k \text{ (the presenter will not open the door that hide the car)} \\ 1/2 & \text{if } i = k \text{ (if the participant has chosen a door with the car, the remaning two are equally probable)} \\ 1 & \text{if } i \neq k \text{ and } j \neq k \text{ (only one door remaining)} \end{cases}$$

Monty Hall problem

- We can solve this problem by computing the *posterior probability of the win*, conditioned by the fact the presenter has opened up one door
  Without loss in generality, we can assume that the player has chosen door 1, while the presenter has opened door 3, behind which there is a goat. In other words, the presenter has accomplished event M_{13}.
- Aposterior winning probability, without changing the choice, under the condition M_{13}, is $P(A_1 \mid M_{13})$.
  Using Bayes theorem, we can express this probability as:

$$P(A_1|M_{13})=\frac{P(M_{13}|A_1)P(A_1)}{P(M_{13})}$$

Monty Hall problem

- Given the previous assumptions, we derive:

$$P(M_{13}|A_1)P(A_1) = \frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$$

- Normalization constant in the denominator can be expressed as:

$$P(M_{13}) = P(M_{13}, A_1) + P(M_{13}, A_2) + P(M_{13}, A_3)$$

$$\begin{aligned} P(M_{13}) = {} & P(M_{13}|A_1)P(A_1) + \\ & P(M_{13}|A_2)P(A_2) + \\ & P(M_{13}|A_3)P(A_3) \end{aligned}$$

$$P(M_{13}) = \frac{1}{2} \times \frac{1}{3} + 1 \times \frac{1}{3} + 0 \times \frac{1}{3} = \frac{1}{2}.$$

Monty Hall problem

- Hence:

$$P(A_1|M_{13}) = \frac{1}{6} / \frac{1}{2} = \frac{1}{3}$$

- Note that the *posterior* probability that the participant has chosen the correct door (as his/her first choice) is equal to the *prior* probability
- In short, what the presenter has done has not introduced any new information regarding this event. The actions of the presenter merely redistribute the probability between the remaining two doors.

Monty Hall problem

- The probability of the win, if the participant switches his choice to door 2, can be derived from the following condition (the car must be behind one of the three doors):

$$P\left(A_1 | M_{13}\right) + P\left(A_2 | M_{13}\right) + P\left(A_3 | M_{13}\right) = 1$$

- Behind door 3 there is no car (this is what the presenter has shown us), thus the corresponding posterior probability is 0. We can also derive this using Bayes:

$$P\left(A_3 | M_{13}\right) = \frac{P\left(M_{13} | A_3\right) P\left(A_3\right)}{P\left(M_{13}\right)} = \left(0 \times \frac{1}{3}\right) / \frac{1}{2} = 0$$

Monty Hall problem

- It therefore follows:

$$P\left(A_2 \middle| M_{13}\right) = 1 - \frac{1}{3} - 0 = \frac{2}{3}.$$

- This demonstrates that the winning strategy is always to switch the choice.

Important remark:

- Bayes formula (11) is derived under the assumption that the facts E_i are mutually independent given any hypothesis
- This condition can limit the applicability of the Bayes scheme

- Reminder: Two events are independent iff:

$$p(A \cap B)=p(A)\,p(B)$$

Example:

- Two symptoms A and B can both be indicative of an illness with probability p. However, if both symptoms are present simultaneously, they may reinforce each other (or oppose each other)

- *H_1 Ivan has a cold*
- *H_2 Ivan has allergy*
- *H_3 Ivan has light sensitivty*

Three mutually exclusive hypotheses

- *E_1 Ivan has rhinitis*
- *E_2 Ivan has cough*

evidences/ facts

BAYES SCHEME IN EXPERT SYSTEMS

	Prior and conditional probabilities		
	i = 1	*i=2*	*i=3*
	(cold)	*(allergy)*	*(light sensitivity)*
$p(H_i)$	*0.6*	*0.3*	*0.1*
$p(E_1 \| H_i)$	*0.3*	*0.8*	*0.3*
$p(E_2 \| H_i)$	*0.6*	*0.9*	*0.0*

- *If if is determined that the patient has rhinitis, we can compute the* ***posterior probabilities*** *of hypotheses H_i, i=1,3 using (10a)*

BAYES SCHEME IN EXPERT SYSTEMS

- $p(H_1| E_1) = \dfrac{0.3 \cdot 0.6}{0.3 \cdot 0.6 + 0.8 \cdot 0.3 + 0.3 \cdot 0.1} = 0.4$
- $p(H_2| E_1) = \dfrac{0.8 \cdot 0.3}{0.3 \cdot 0.6 + 0.8 \cdot 0.3 + 0.3 \cdot 0.1} = 0.53$
- $p(H_3| E_1) = \dfrac{0.3 \cdot 0.1}{0.3 \cdot 0.6 + 0.8 \cdot 0.3 + 0.3 \cdot 0.1} = 0.06$
- Probability has decreased H_1 (initially 0.6) and H_3 (initially 0.1)
- Probability has increased H_2 (initially 0.3)

in presence of evidence E_1

BAYES SCHEME IN EXPERT SYSTEMS

- If it now also turns out that the patient has a cough, then (n=2 and formula (11))
- $p(H_1| E_1 E_2) = \dfrac{0.3 \cdot 0.6 \cdot 0.6}{0.3 \cdot 0.6 \cdot 0.6 + 0.8 \cdot 0.9 \cdot 0.3 + 0.3 \cdot 0.0 \cdot 0.1} = 0.33$
- $p(H_2| E_1 E_2) = \dfrac{0.8 \cdot 0.9 \cdot 0.3}{0.3 \cdot 0.6 \cdot 0.6 + 0.8 \cdot 0.9 \cdot 0.3 + 0.3 \cdot 0.0 \cdot 0.1} = 0.67$
- $p(H_3| E_1 E_2) = \dfrac{0.3 \cdot 0.0 \cdot 0.1}{0.3 \cdot 0.6 \cdot 0.6 + 0.8 \cdot 0.9 \cdot 0.3 + 0.3 \cdot 0.0 \cdot 0.1} = 0.00$
- Hypothesis H_3 (light sensitivity) is excluded, while H_2 is much more probable than H_1

(DIS)ADVANTAGES OF BAYES SCHEME

Advantages

- Very good theoretical foundations
- Most elaborated among all models for dealing with uncertainty

Disadvantages

- A large amount of dana is needed to determine all the probabilites

- All probabilities must be defined!

(DIS)ADVANTAGES OF BAYES SCHEME

Example

- We have a system with 50 hypotheses and 300 evidences. How many probabilities must be defined?

  (assuming that the hypotheses are mutually exclusive and that the evidences are conditionally independent)

- We must define all the prior and conditional probabilites – how ?

  (statistics? , experts?)

(DIS)ADVANTAGES OF BAYES SCHEME

- It is difficult to build an inference justification system for a an expert system based on the Bayes scheme

Application example of the Bayes scheme

- PROSPECTOR expert system developed at Stanford University (Duda *et al.*, 1979) for the detection of mineral based on geographic characteristics

FUZZY LOGIC

FUZZY LOGIC and FUZZY SETS

- For fuzzy sets, two laws don't appy:
  - Law of the excluded middle
  - Law of contradiction

- Fuzzy logic = computing with words

- Limitations of two-valued logics

An example

- A car mechanic describes using his own words how he/she infers why the car won't start. This is a passage of this description pertaining to the accumulator battery age as a one potential cause of the problem.

FUZZY LOGIC and FUZZY SETS

- *"… <u>Occasionally</u>, it can happen that the accumulator battery is <u>too weak</u> to ignite the motor, but has just enaough power to keep the lights on for a certain amount of time. This happens <u>very rarely</u>, but the lights being on seems to be at odds with the accumulator battery being empty. What you must consider in situations like that one is the age of the accumulator battery. For an <u>old accumulator</u> the capacity drop off will be significant. A <u>new</u> accumulator will be much more resilient."*

FUZZY LOGIC and FUZZY SETS

Knowledge is expressed in linguistic terms whose meaning is often not clearly defined, such as: *occasionally, too weak, very rarely, old, new, ...*

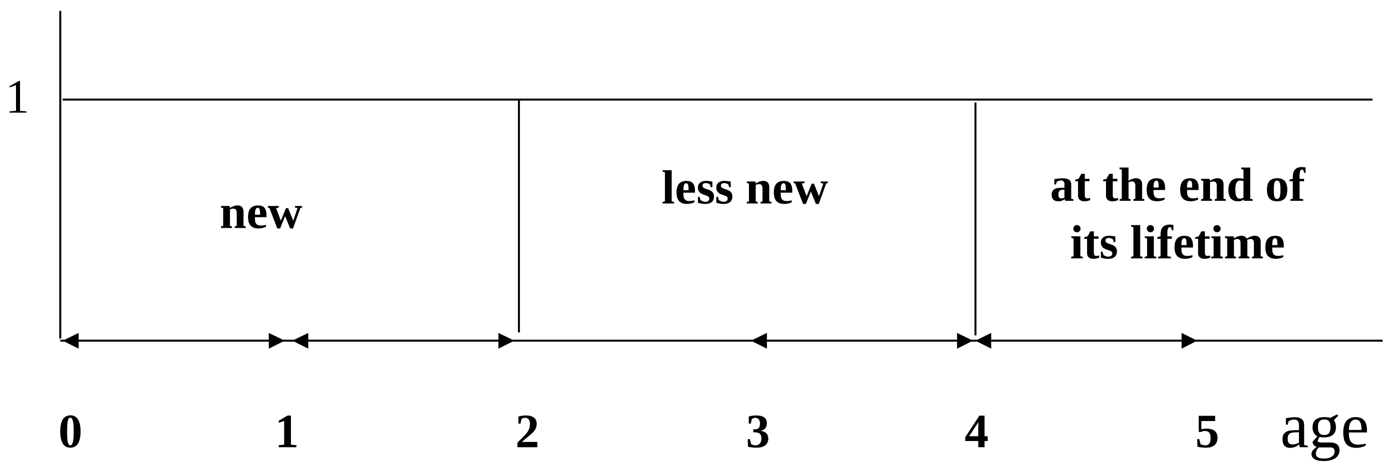
- To encode this knowledge in a technical system, we must know what is assumed by expressions such as *new* or *old*
- Here is how a mechanic tries to explain this:

"If an accumulator is <u>2 years old or less</u>, I would consider it ***new****. If it's between <u>two and four years</u>, I would say it's* ***less new****. If it is <u>older than four years</u>, I consider it to be* **at the end of its lifetime***…"*

FUZZY LOGIC and FUZZY SETS

- Auto mechanic has tried to explain what he/she understands when using vague linguistic expressions:

LINGUISTIC EXPRESSION	MEANING
“new”	less than 2 years
“less new”	between 2 and 4 yrs.
“at the end of its lifetime”	more than 4 years



FUZZY LOGIC and FUZZY SETS

Although it is possible to precisely define the meaning of linguistic expressions, in this case the expressions are **defined with clear-cut boundaries** on a time scale [0, ∞]

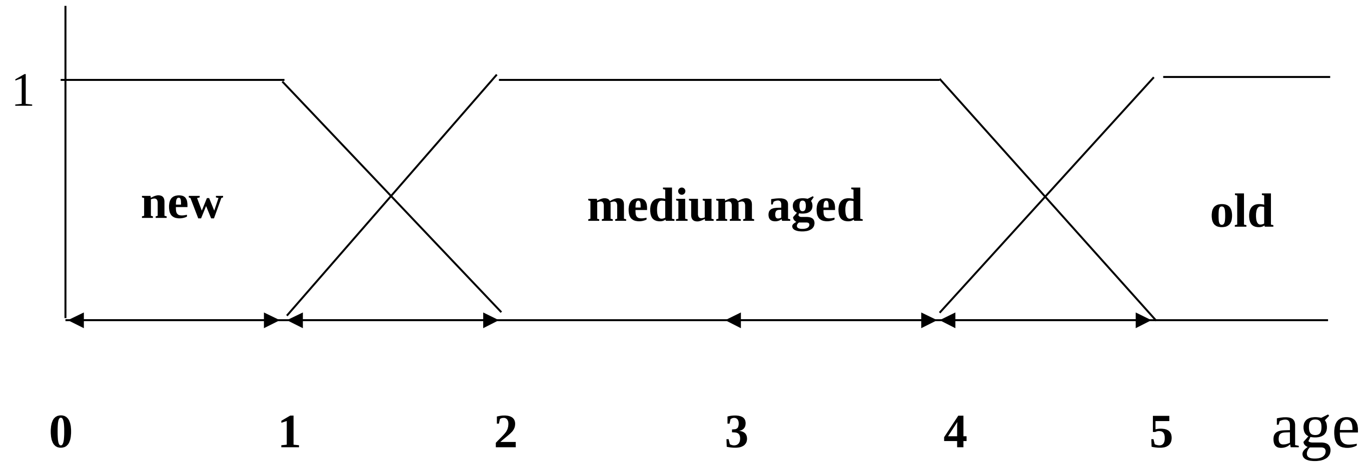
Is it natural to define the expressions such as "old" and "new" with such sharp boundaries?
Does it means that a 2-year accumulator is new today, but tomorrow it isn't new anymore?

Clear-cut boundaries give unrealistic models. Aging is a continuous process in which there are no sudden jumps.

FUZZY LOGIC and FUZZY SETS

Instead of using **clear-cut boundries** to define the vague linguistic expressions, it would be more natural to talk about **the extent** in which an accumulator is old or new

- If, instead of talking about whether an element is belong to a set or not, we talk about **the extent** in which it belongs to this set, we talk about a **fuzzy set**



FUZZY LOGIC and FUZZY SETS

1. **Knowledge is subjective**: another auto mechanic would perhpas define the boundaries differently, e.g. 3 and 5 years instead of 2 and 4 years. **Thus, the same linguistic expression can have a different meaning**

1. **On the other hand, different linguistic expressions can have the same meaning.** E.g., instead of “at the end of its lifetime”, we could have used “old”

1. ***Knowledge is contextually dependent****: old*

accumulator

human

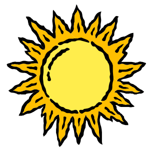
FUZZY LOGIC and FUZZY SETS

- Classical logic cannot adequatly model the knowledge expressed with words. This knowledge is vague and contextually dependent.
- Put differently: classical logic cannot determine to what extent an elements belongs to a set.
- These problems are solved by fuzzy logic and fuzzy sets, respectively

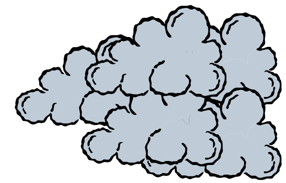
> Fuzzy logic is a mathematical model for the representation and inference with human knowledge, which is expressed in words, in a vague and imprecise manner

Why classical logic is not enough ?

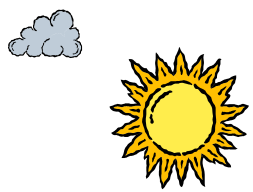
p = “Today is a sunny day”



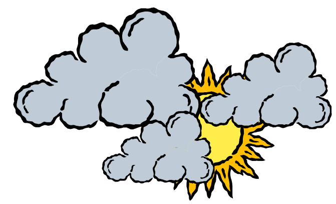
P is true



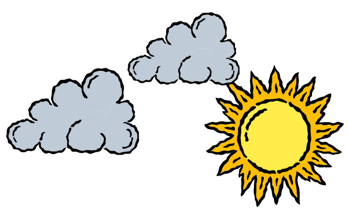
P is false



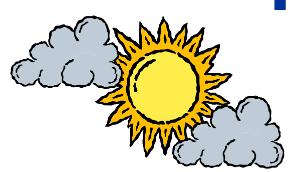
?



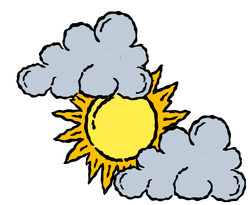
?



?



?



?

Two truth values are inadequate for modelling the inference based on human knowledge. This knowledge is often incomplete, vague, expressed in natural language, and a matter of degree.

Vagueness - Fuzziness

Around **1920.** Bertrand Russell introduced the term ***vagueness***

1937. Max Black – paper:

“Vagueness: An Exercise in Logical Analysis”, Philosophy of Science, Vol.4, 1937.

The paper did not draw much attention.

1965. **Lotfi A. Zadeh** – paper “**Fuzzy sets**”, *Information and Control*, Vol.8, No.4, pp. 338-353, June 1965.

Vagueness - Fuzziness

Around **1920.** Bertrand Russell introduced the term ***vagueness***

1937. Max Black – paper:

“Vagueness: An Exercise in Logical Analysis”, Philosophy of Science, Vol.4, 1937.

The paper did not draw much attention.

1965. Lotfi A. Zadeh – paper “**Fuzzy sets**”, *Information and Control*, Vol.8, No.4, pp. 338-353, June 1965.

Fuzzy set

A set with unclear, “soft” boundaries !

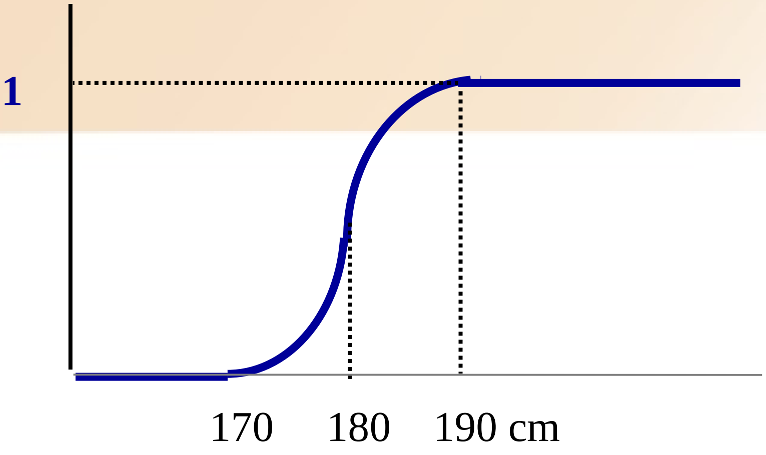
Classical set $\mu_A : X \rightarrow \{0, 1\}$

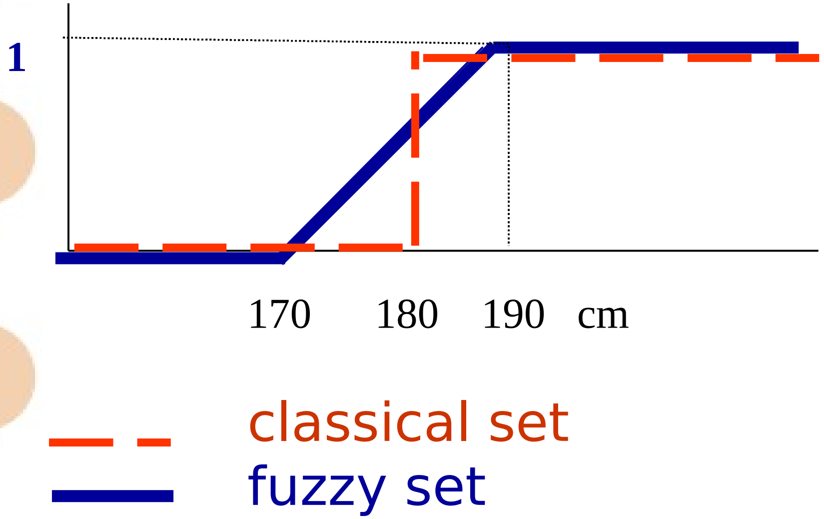
Fuzzy set $\mu_{\mathbf{A}}$ **:** $X \rightarrow$ **[0 , 1]**

$\mathbf{A} = \{(\mathbf{x}, \mu_{\mathbf{A}}(x)) \mid \mathbf{x} \in \mathbf{X}, \ \mu_{\mathbf{A}}(x) \in [0,1]\}.$

Example

Set of tall people





Defining the membership function of a fuzzy set is:

- **subjective**
- **contextually dependent**

LINGUISTIC VARIABLE

- Informally, a **linguistic variable** is a variable that, instead of taking on numeric values, takes on values in the form of words or phrases

- A linguistic variable links together the **natural language expressions** of and their **quantifications as fuzzy sets**

LINGUISTIC VARIABLE

- *Definition* (linguistic variable) [Zadeh, 1975]

  A tuple (x, Tx, U, G, Mx), where:

  x – the name of the linguistic variable;

  Tx – the set of linguistic terms;

U – universe of discourse (the set of numbers from the real physical domain; continuous or discrete);

  G – the formal grammar, i.e., the set of syntax rules that generate T from a set of basic terms;

  Mx – the semantic function that gives quantative meaning (interpretation) to linguistic terms. Mx assigns to each term $\forall$ x $\in$ T a fuzzy set defined on U

LINGUISTIC VARIABLE

Example

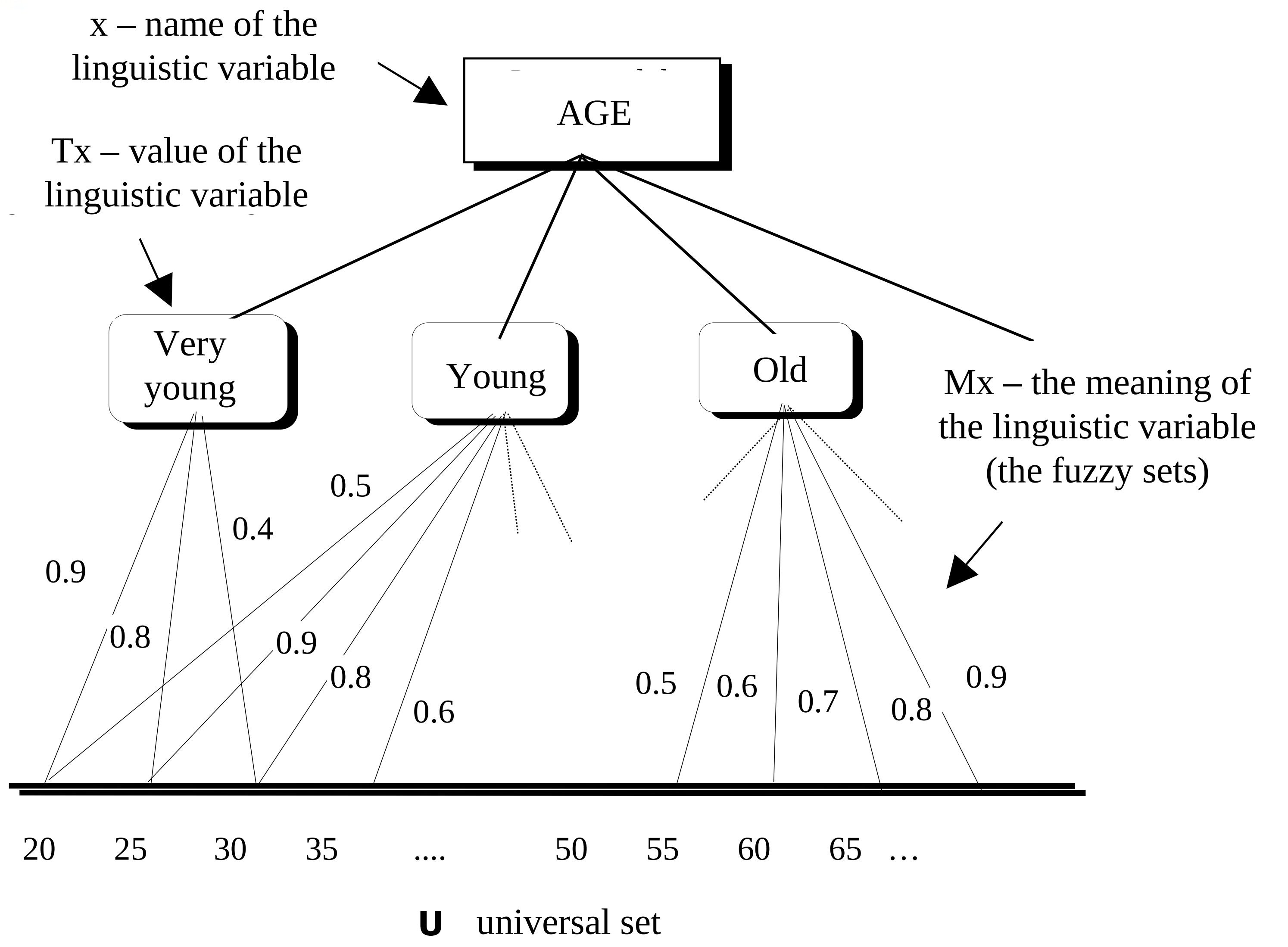
- Linguistic variable: AGE

- x= AGE

- T(AGE) = young + not young + very young + not very young + very very young + … + middle-aged + old + not old + very old + very very old + … + not old and not middle-aged + …

- U = {0, 1, 2, 3, 4, …, 99, 100}

LINGUISTIC VARIABLE



LINGUISTIC VARIABLE

- The meaning of basic terms (e.g., *young* and *old*) are subjective and context-dependent. These meaning must be defined in advance.

- Using only the basic terms, we can use the grammar to derive all the other values of terms from the set T. E.g., the terms *very young, not very young, very very young,* etc.

- *Examples of other linguistic variables: temperature, truth, height, weight,* etc.

LINGUISTIC MODIFIERS

Fuzzy logic = computing with words
(L. Zadeh)

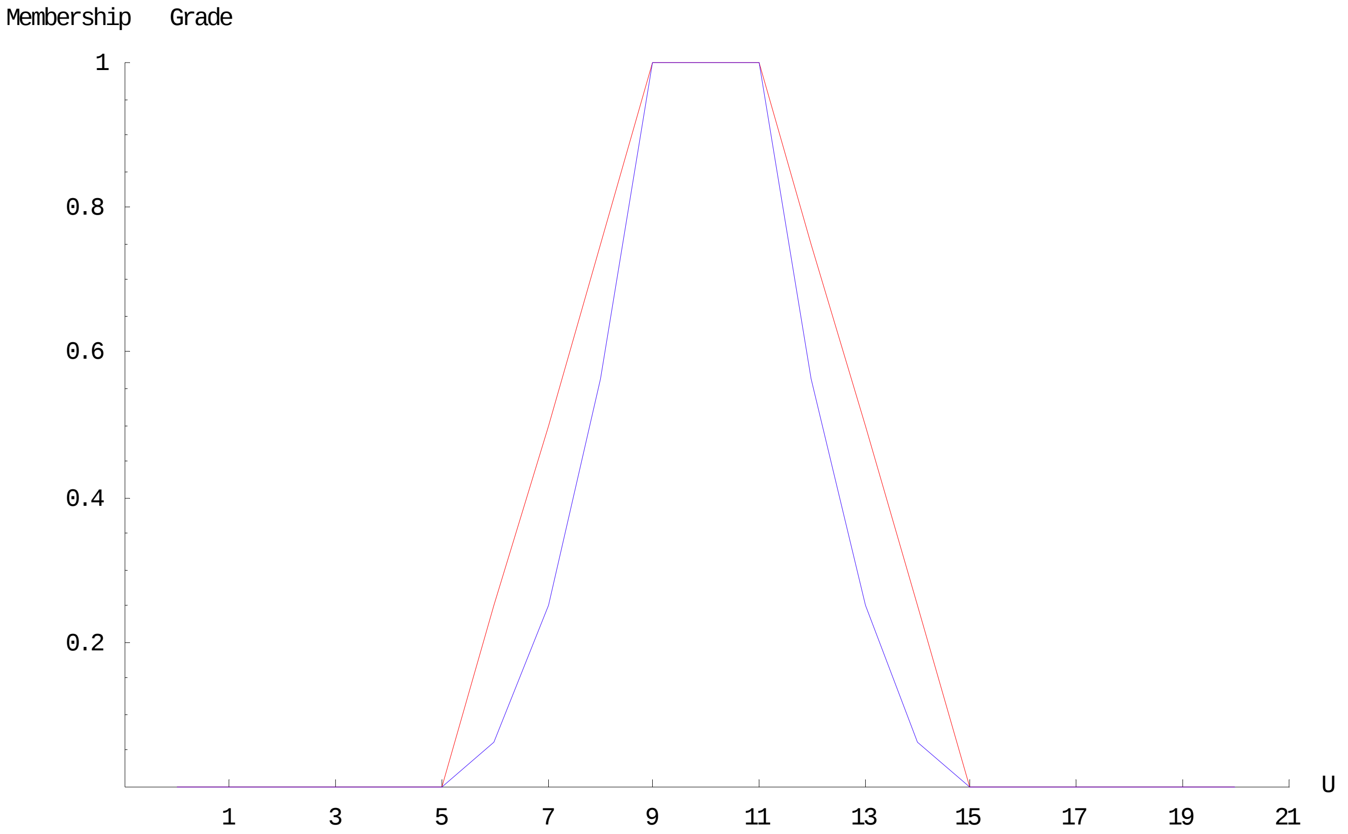
- **Concentration** of a fuzzy set is obtained by squaring the membership function
- It correponds to the linguistic term "*VERY*“

- Let **A** be a fuzzy set. Then the concentration of A is defined as:

$$\mathrm{Con}(\mathbf{A}) = \mu_{\mathbf{A}}(x)^2$$

LINGUISTIC MODIFIERS

- A = “Cold”
- Con(A) = “Very cold”



LINGUISTIC MODIFIERS

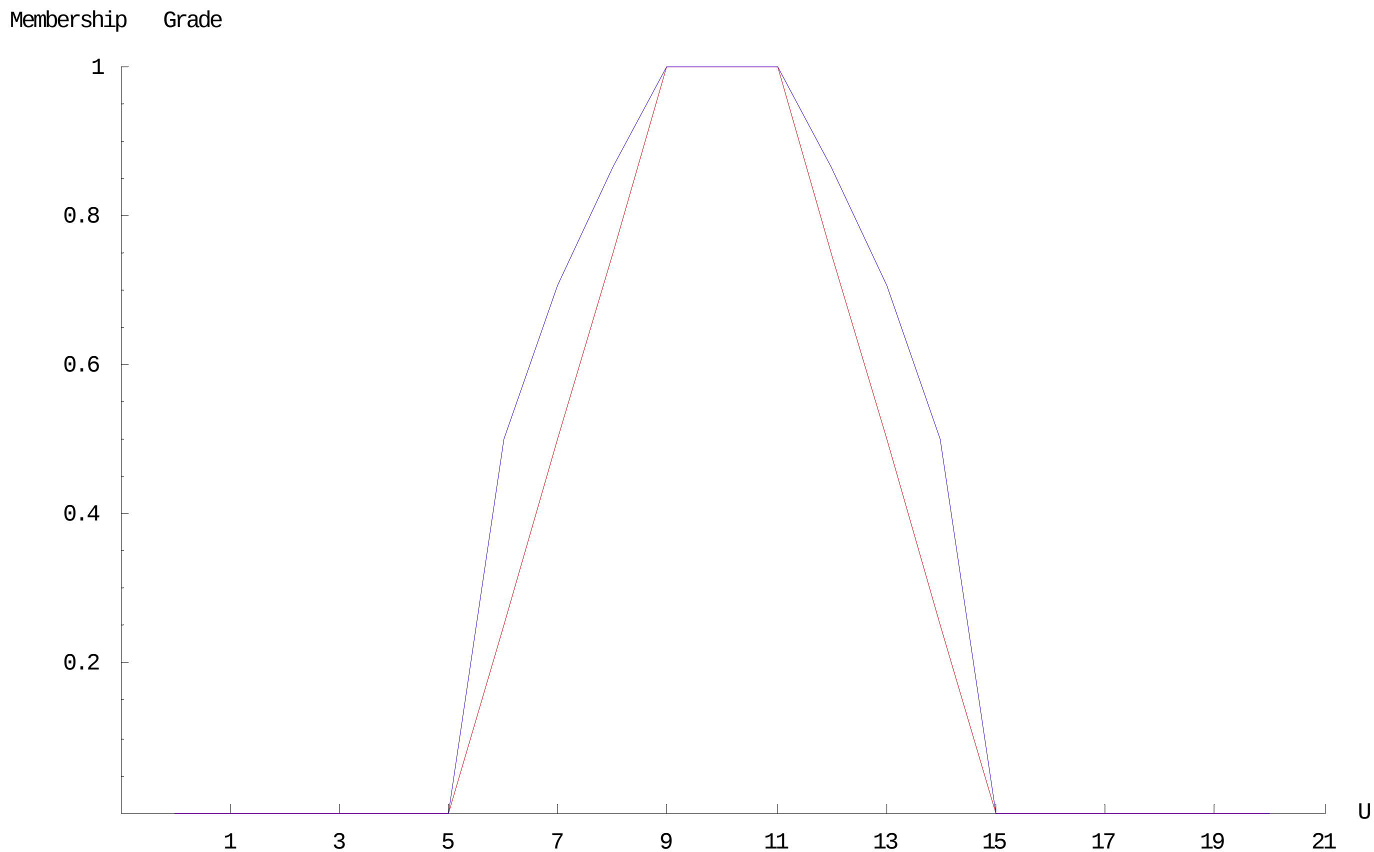
- **Dilation** of a fuzzy set is obtained by taking the square root of the membership function
- It corresponds to the linguistic term “MORE OR LESS”

- Let A be a fuzzy set. Then dilation of A is defined as:

$$\mathrm{Dil}(\mathbf{A}) = \mu_{\mathbf{A}}(x)^{1/2}$$

Example

- Dil(A) = “More or less cold”



The relation between FUZZY SETS and FUZZY LOGIC

- A fuzzy set **membership function** and the **truth value** of a proposition are related to each other as follows:

- The truth value of "*Element x belongs to set A*" is equivalent to the degree of membership with which *x* belong to the set fuzzy set A, i.e,. $\mu_A(x)$

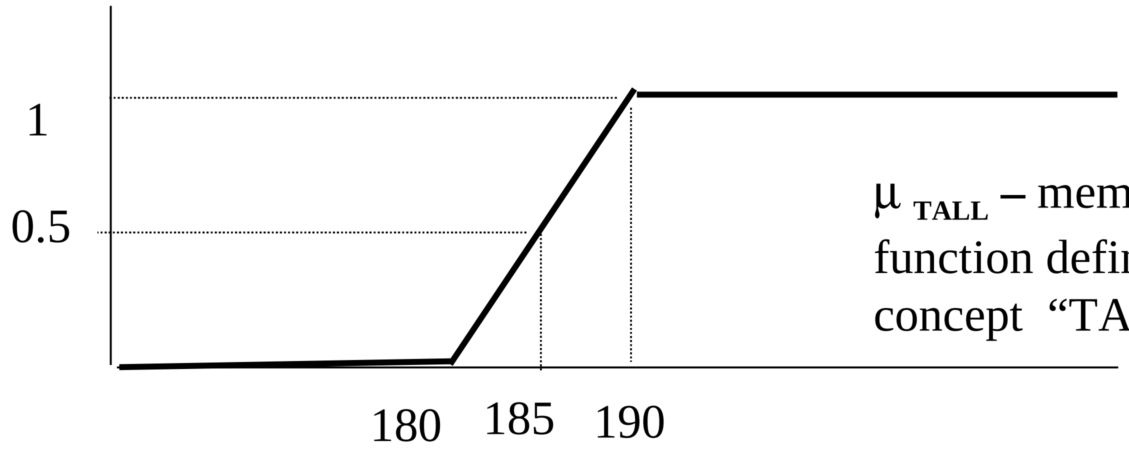
And vice versa:

- The degree of membership of x in fuzzy set A is equivalent to the truth evalue of proposition "*Element x belongs to set A*"

The relation between FUZZY SETS and FUZZY LOGIC

Example

- Assum John's height is 185 cm
- Someone has stated “John is tall”
- What is the truth value of this proposition?
- Assume the set of tall people is defined using $\mu_{TALL}(x)$



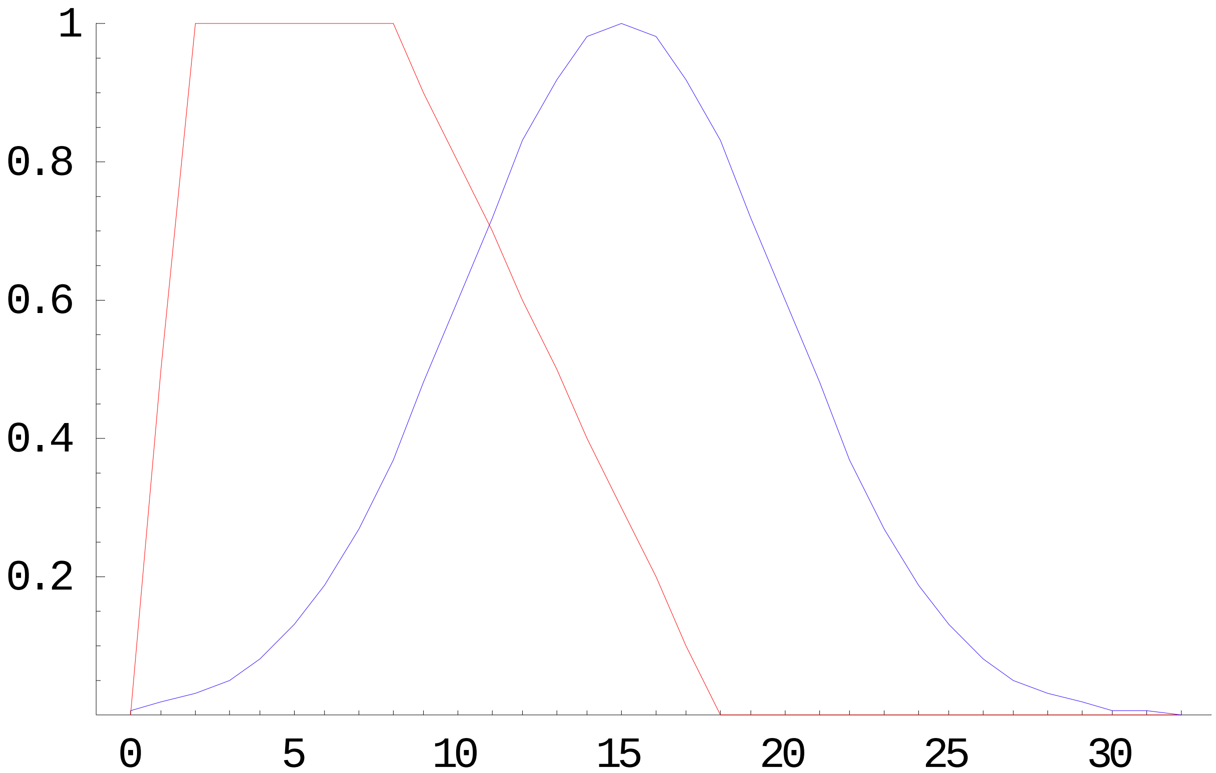
μ_{TALL} – membership function defining the concept “TALL PEOPLE”

The relation between FUZZY SETS and FUZZY LOGIC

- Then:
- The truth value of proposition "John is tall" is equivalent to the degree of membership with which the element 185 belongs to the fuzzy set of tall people
- $\mu_{TALL}(185) = 0.5$

OPERATIONS on fuzzy sets

- Let **A** and **B** be fuzzy sets defined over the same domain

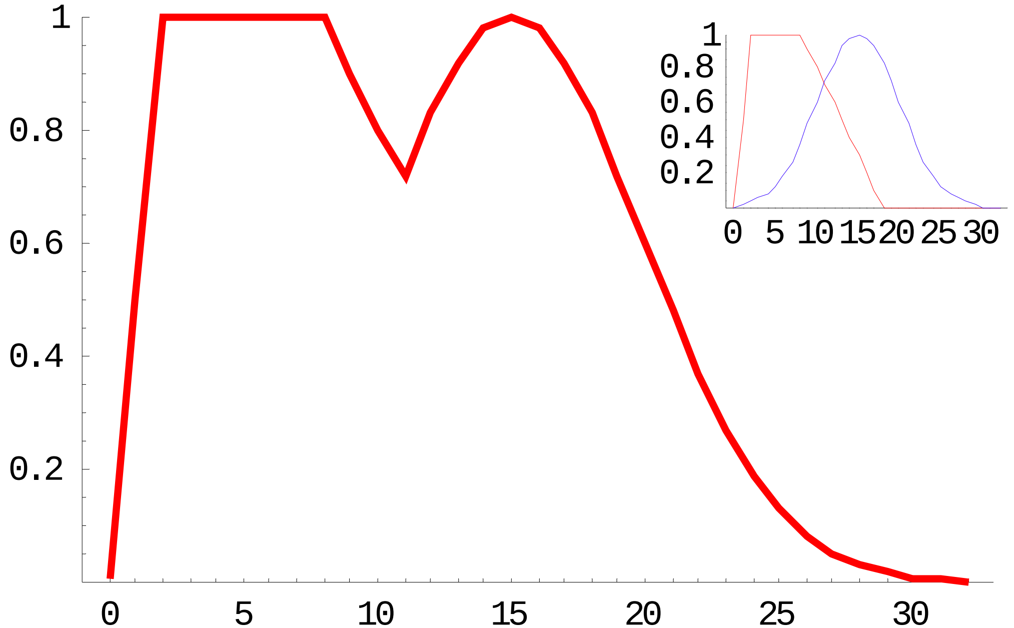


OPERATIONS on fuzzy sets

- The union **A** and **B** is the fuzzy set **A** ∪ **B**,

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

- The union corresponds to the linguistic term ***OR***

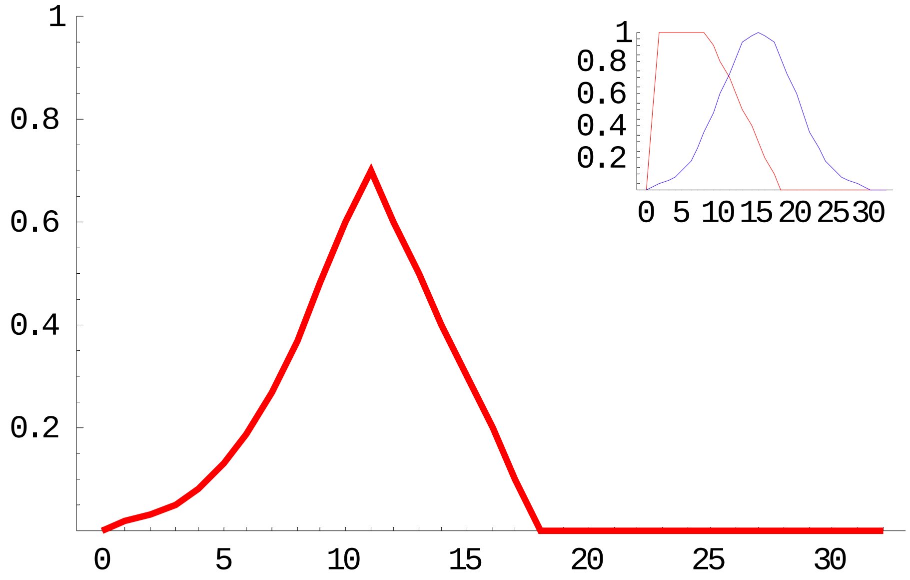


OPERATIONS on fuzzy sets

- The intersection of **A** and **B** is the fuzzy set **A** $\cap$ **B**

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

- The intersection corresponds to the linguistic term ***AND***



OPERATIONS on fuzzy sets

- The negation of **A** is defined as:

$$\forall x \in X,\ \mu_{A^c}(x) = 1 - \mu_A(x)$$

Which of the classical set theory laws **don't apply** in fuzzy logic?

- De Morganov's laws?
- Commutativity?
- Distributivity?
- Law of excluded middle?
- Law of contradiction?
- ...

REASONING IN FUZZY LOGIC

- In clasical logic we can use sound inference rules (such as modus ponens) to derive new formulae (new statements)
- In fuzzy logic: **GENERALIZED MODUS PONENS**

Let A, A1, B, B1 be fuzzy sets

Premise	*x* is **A1**
Implication	If *x* is **A** then *y* is **B**
Conclusion	*y* is **B1**

REASONING IN FUZZY LOGIC

- This is very much different from the traditional modus ponens:

We use imprecise expressions that are represented as fuzzy sets (**A, A1, B, B1**)

Fuzzy sets **A** and **A1** as well as **B** and **B1** (i.e., the corresponding expressions) need NOT be IDENTICAL!

Note: A and A1 as well as B are B1 defined on the same universal set

REASONING IN FUZZY LOGIC

Example

Premise	Ivan is a **tall** person.
Implication	If a person is **tall**, then the person has a **large weight**.
Conclusion	Ivan has a **more or less large weight**.

Premise	Banana is **very yellow**.
Implication	If the banana is **yellow**, then it is **ripe**.
Conclusion	Banana is **very ripe.**

REASONING IN FUZZY LOGIC

- In order to understand the generalized modus ponens, we introduce the concept of a **fuzzy relation**

- **Fuzzy set A –** defined by the membership function $\mu_A(x) : X \rightarrow [0,1]$, where X is the universal set, while $\mu_A(x)$ is a number between 0 and 1, which determines to what extent element *x* belongs to fuzzy set **A**

- **Fuzzy relation R –** defined by the membership function $\mu_R: X \times Y \rightarrow [0, 1]$, where $\mu_R(x, y)$ determines to what extent the elements x (from set X) and y (from set Y) are related

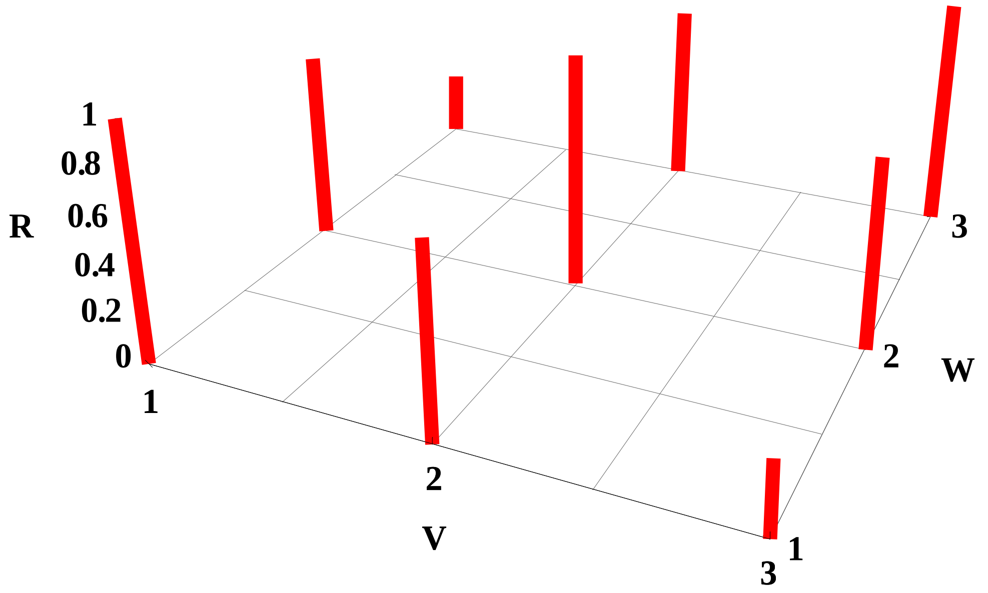
Example

- Fuzzy relation **"approximately equal"** defined on the universal set V, W ={1,2,3}
- R={ ((1, 1), 1), ((1, 2), .8), ((1, 3), .3), ((2, 1), .8), ((2, 2), 1), ((2, 3), .8), ((3, 1), .3), ((3, 2), .8), ((3, 3), 1) }
- In matrix form:

	1	2	3
1	**1**	**0.8**	**0.3**
2	**0.8**	**1**	**0.8**
3	**0.3**	**0.8**	**1**

REASONING IN FUZZY LOGIC

- Interpretation: 1 is “approximately equal” to 3 with value 0.3, while 2 is “approximately equal” to 2 with value 1



Remark

- If this were a classical (non-fuzzy, crisp) relation, it would have ones as diagonal and zeros as off-diagonal elements

REASONING IN FUZZY LOGIC

- If **A** is a fuzzy set on X, and **B** is a fuzzy set on Y, then **A** x **B is a fuzzy relation** on the universal set X x Y defined by $\mu_{\mathbf{A} \times \mathbf{B}}: X \times Y \rightarrow [0, 1]$,

$$\mu_{\mathbf{A x B}}(x, y) = \min(\mu_{\mathbf{R}}(x), \mu_{\mathbf{R}}(y))$$

Every fuzzy relation respresents an implication.

Implication *"If x is **A** then y is **B**"* is defined by a **fuzzy relation A x B** on X x Y.

REASONING IN FUZZY LOGIC

Example

- Implication “**If** Ivan is tall **then** Ivan has a large weight" is defined by a fuzzy relation “HEIGHT x WEIGHT” as follows:
- Given a set of “tall” people and a set of “heavy-weight” people. Then, according to the above definition, the fuzzy relation “tall and heavy-weighted " people is given by:

REASONING IN FUZZY LOGIC

Example

- Implication “**If** Ivan is tall **then** Ivan has a large weight" is defined by a fuzzy relation “HEIGHT x WEIGHT” as follows:
- Given a set of “tall” people and a set of “heavy-weight” people. Then, according to the above definition, the fuzzy relation “tall and heavy-weighted " people is given by:

REASONING IN FUZZY LOGIC

“heavy people” (in kg)

“tall people” (in cm)

		0	0.1	0.2	0.3	0.5	0.7	0.9	1	1
		60	65	70	75	80	85	90	95	100
0	170	0	0	0	0	0	0	0	0	0
0.3	175	0	0.1	0.2	0.3	0.3	0.3	0.3	0.3	0.3
0.5	180	0	0.1	0.2	0.3	0.5	0.5	0.5	0.5	0.5
0.8	185	0	0.1	0.2	0.3	0.5	0.7	0.8	0.8	0.8
1	190	0	0.1	0.2	0.3	0.5	0.7	0.9	1	1
1	195	0	0.1	0.2	0.3	0.5	0.7	0.9	1	1
1	200	0	0.1	0.2	0.3	0.5	0.7	0.9	1	1

For example: $\mu_{\text{HEIGHT x WEIGHT}}(175, 90) = \min\{0.3, 0.9\} = 0.3$

REASONING IN FUZZY LOGIC

- An example of reasoning using **modus ponens:**

Premise	v is a **small** number
Implication	v and w are **approx. equal**
Conclusion	w is **more or less small**

How can one derive this?

- Universal sets: V, W ={1, 2, 3}
- We define a fuzzy set **A** = ”**small number**” from the premise as follows:

$$\mu_{\textbf{small number}} = 1/1 + 0.5/2 + 0.1/3$$

REASONING IN FUZZY LOGIC

- The implication is represented by a fuzzy relation "**approx. equal**":

$$
R_{\text{"approx. equal"}} = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & 1 & 0.8 & 0.3 \\ 2 & 0.8 & 1 & 0.8 \\ 3 & 0.3 & 0.8 & 1 \end{array}
$$

- To apply modus ponens, we use the **Zadeh's rule of min-max composition** and appy it between the premise and the implication relation
- Composition **A o R** yields a new fuzzy set **B.** This fuzzy set is the conclusion. It is defined as follows:
- $\mu_B(w) = \max(\min(\mu_A(v), \mu_R(v,w)))$

REASONING IN FUZZY LOGIC

A "small number"

1	2	3
1	0.5	0.1

R ="approx. equal numbers"

	1	2	3		
	1	0.8	0.3	1	
0	0.8	1	0.8	2	=B
	0.3	0.8	1	3	

$\boldsymbol{\mu_B(1)} = \max(\ \min(\mu_A(v), \mu_R(v,1))\ )=$

$\max\{\ \min(1,1), \min(0.5,0.8), \min(0.1,0.3)\} =$

$\max\{1, 0.5, 0.1\} = \mathbf{1}$

$\boldsymbol{\mu_B(2)} = \max(\ \min(\mu_A(v), \mu_R(v,2))\ )=$

$\max\{\ \min(1,0.8), \min(0.5, 1), \min(0.1,0.8)\} =$

$\max\{0.8, 0.5, 0.1\} = \mathbf{0.8}$

$\mu_B(\mathbf{3}) = \max(\min(\mu_A(v), \mu_R(v,32)))=$

$\max\{\min(1,0.3), \min(0.5, 0.8), \min(0.1,1\} =$

$\max\{0.3, 0.5, 0.1\} = \mathbf{0.5}$

- The result of the composition is the fuzzy set:

  B = A o R = 1/1 + 0.8/2 + 0.5/3

- We can now assign a linguistic term to such a fuzzy set. For example: "**more or less small**". Thus,

Conclusion w is **more or less small.**

- The process of assigning linguistic terms to fuzzy sets is called **linguistic approximation**

REASONING IN FUZZY LOGIC

- In summary: