Uncertainty

CHAPTER 13

Outline

- Uncertainty
- Probability
- Syntax and Semantics
- Inference
- Independence and Bayes' Rule

Questions

Write your answers please (you can give average, you can define range, you can define limits)

1. What is the speed for a fast car?
2. What is the weight of a fat person?
3. What is the age of an old person?
4. What is the height of a tall building?

5. What about, spicy, dangerous, sleepy, funny?

Classical SETs vs Fuzzy Sets

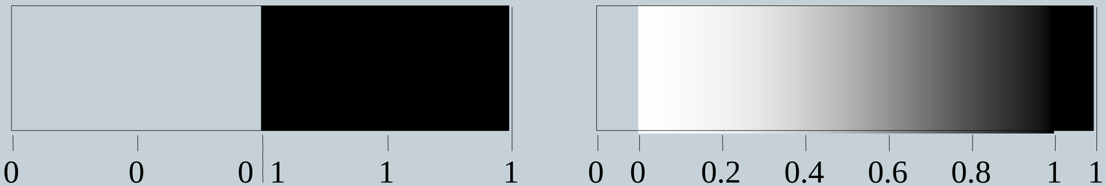
- List of members {1,2,3}, {s,m,i,t,h}, {s,a,d,i}
- Rule of Membership {prime numbers less than 100}, {names of students in the class}
- Function of Membership

A= {Temperatures at which water will boil}

boil(t) = {1,0} , where 1 means it is boiling, 0 means not boiling

Here Comes the Fuzzy Sets: for boil(x), how much does x belongs to the function?

Now we have interval [0,1]

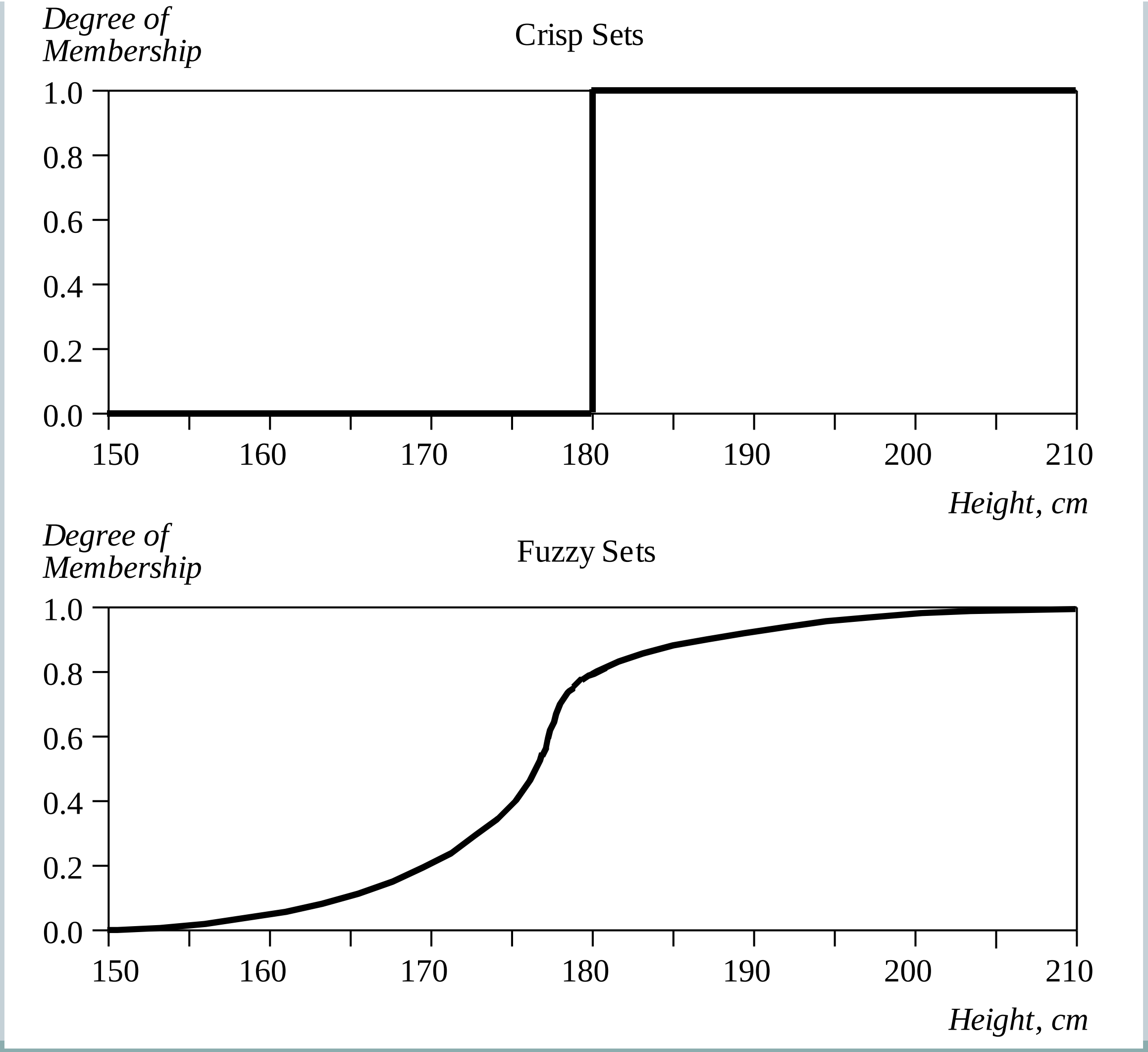


(*a*) Boolean Logic. (*b*) Multi-valued Logic.

Tall?

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

Crisp vs Fuzzy



Crisp vs Fuzzy Functions

- Let X be the universe of discourse and its elements be denoted as x. In the classical set theory, **crisp set A of X is defined as function $f_A(x)$ called the characteristic function of A**:

$f_A(x) : X \rightarrow \{0, 1\}$, where $$f_A(x) = \begin{cases} 1, \text{if } x \in A \\ 0, \text{if } x \notin A \end{cases}$$

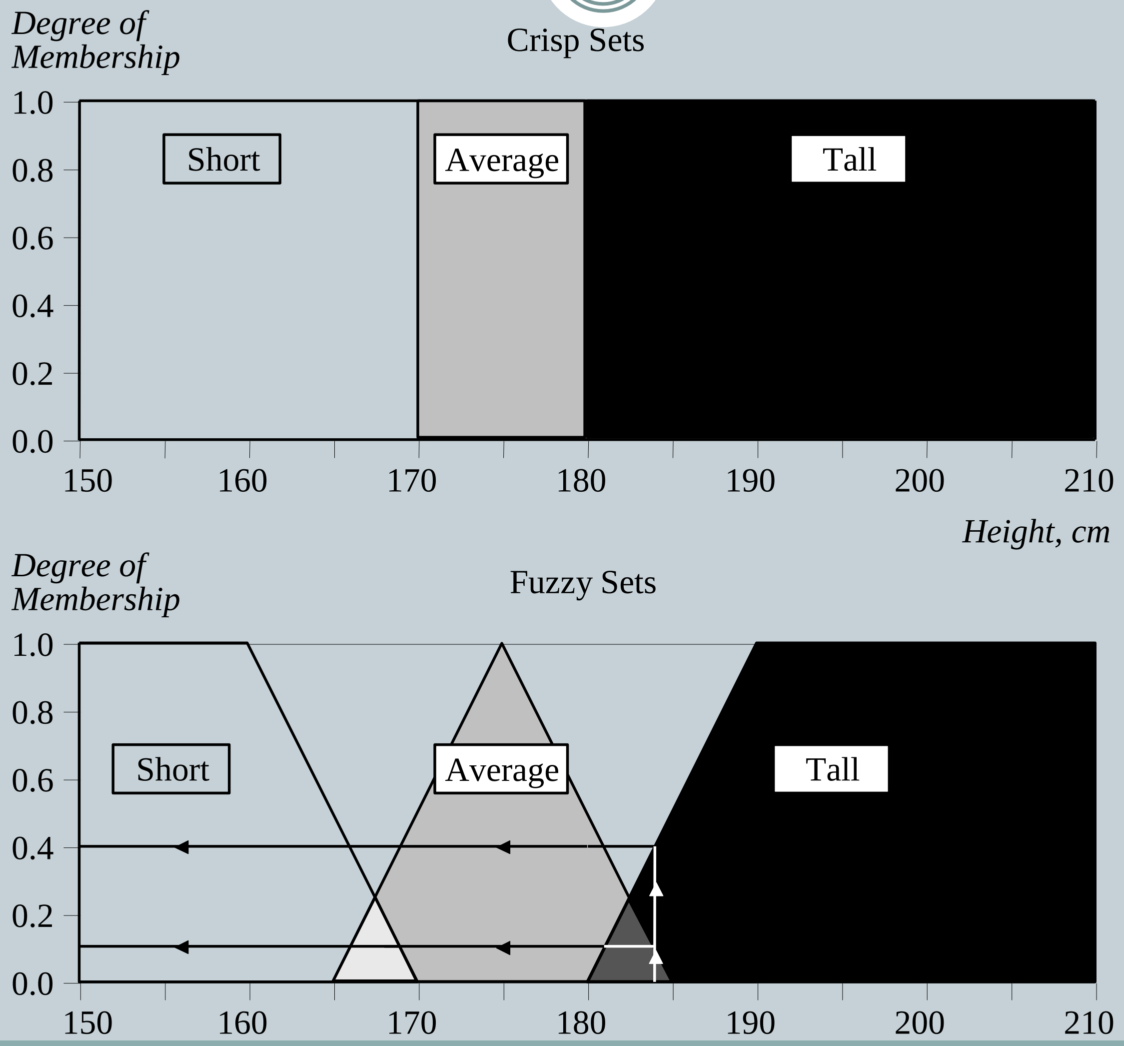
- In the fuzzy theory, fuzzy set A of universe X is defined by function $\mu_A(x)$ called the membership function of set A

$\mu_A(x) : X \rightarrow \{0, 1\}$, where

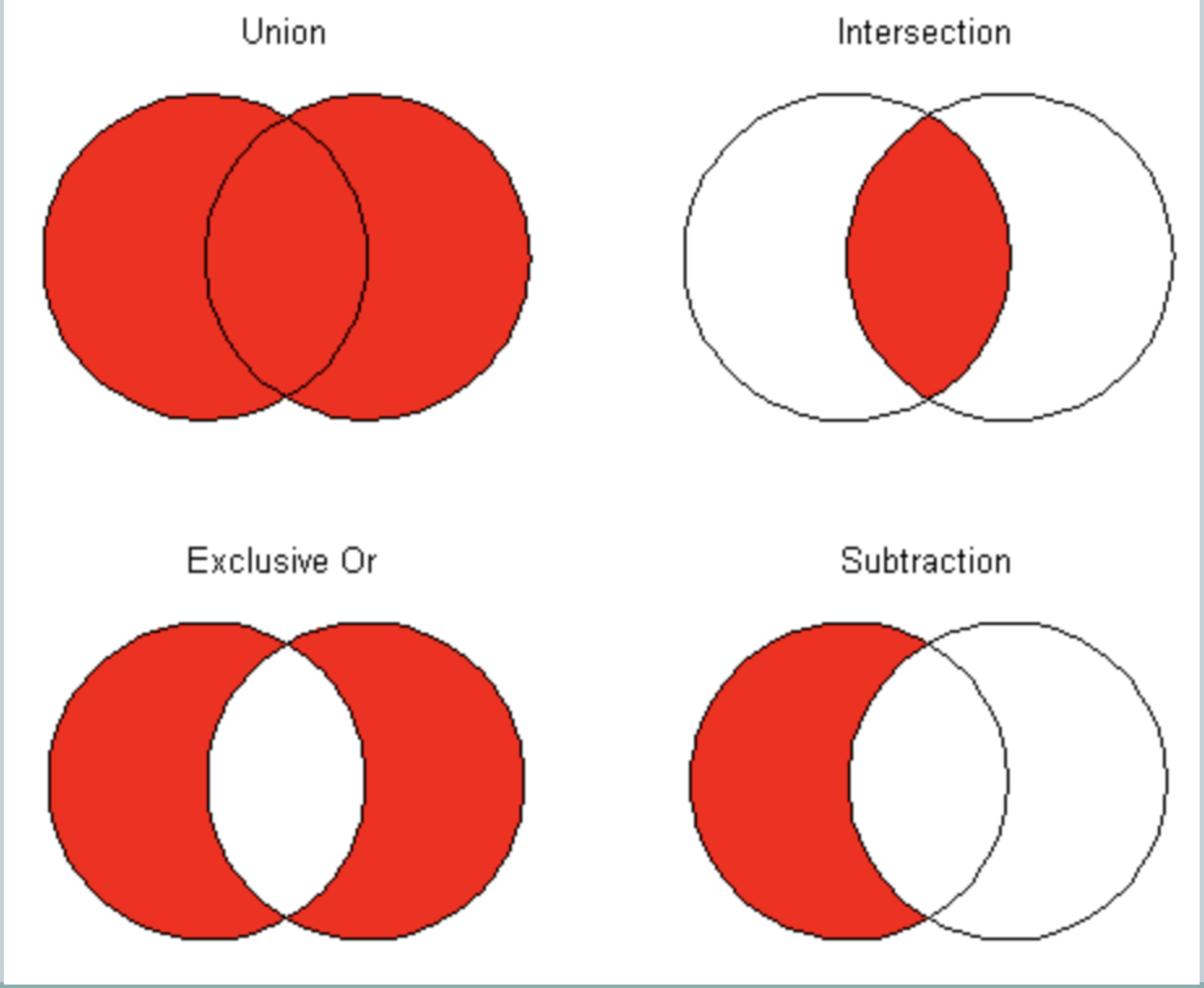
$\mu_A(x) = 1$ if x is totally in A;
$\mu_A(x) = 0$ if x is not in A;
$0 < \mu_A(x) < 1$ if x is partly in A.

degree of membership, also called **membership value**, of element x in set A.

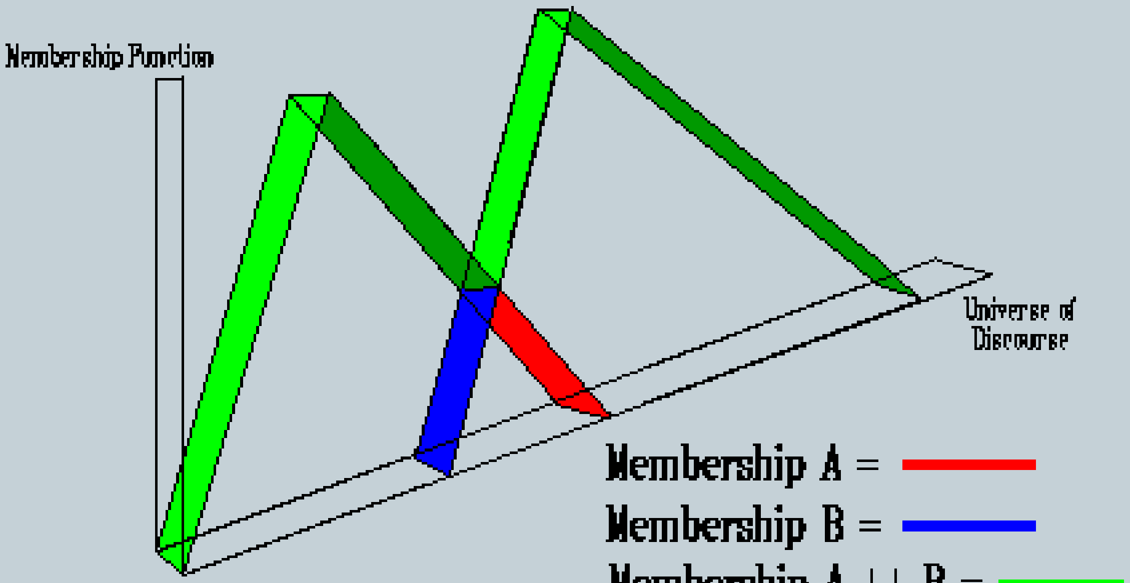
Fuzzy Set Representation



Set Operations

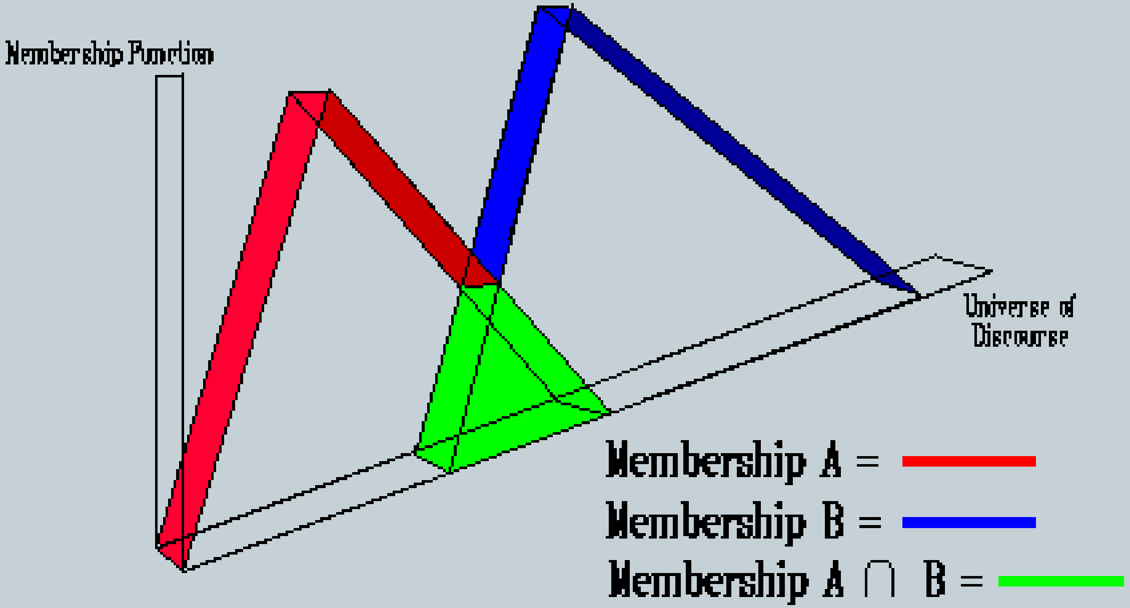


Fuzzy Sets Union



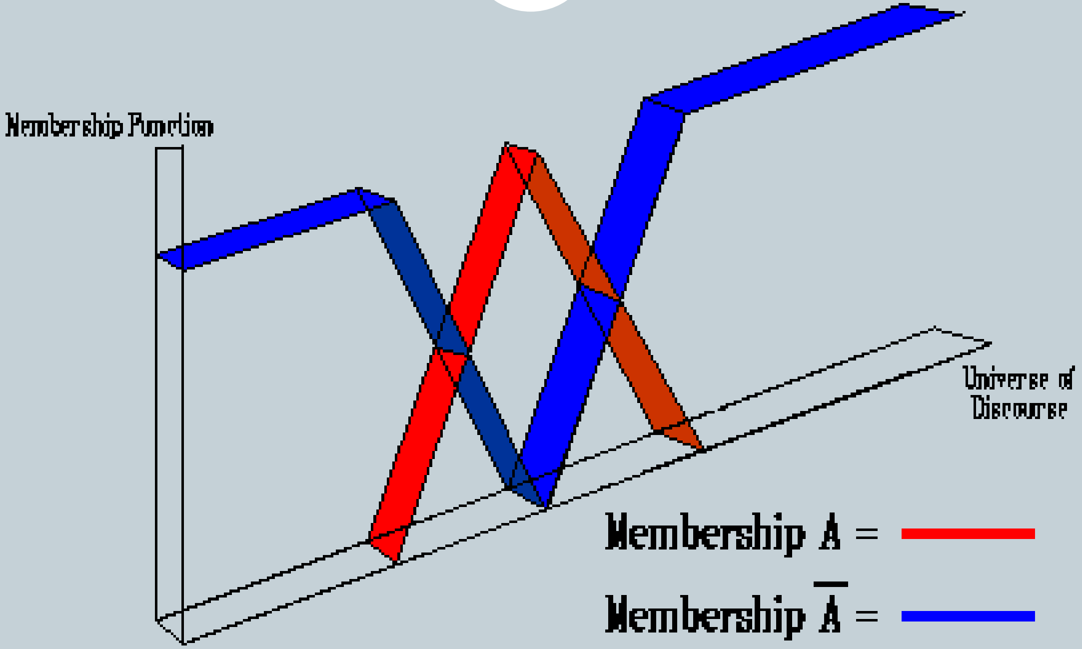
$$\mu_{A \cup B} = \max(\mu_A, \mu_B)$$

Fuzzy Sets Intersection



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

Complement in Fuzzy Sets



$$\mu_{\overline{A}} = 1 - \mu_A$$

Uncertainty

Let action A_t = leave for airport $_t$ minutes before flight
Will A_t get me there on time?

Problems:

1. partial observability (road state, other drivers' plans, etc.)
2.
2. noisy sensors (traffic reports)
3.
3. uncertainty in action outcomes (flat tire, etc.)
4.
4. immense complexity of modeling and predicting traffic
5.

Hence a purely logical approach either

1. risks falsehood: "A_{25} will get me there on time", or
2. leads to conclusions that are too weak for decision making:

"A_{25} will get me there on time if there's no accident on the bridge and it doesn't rain and my tires remain intact etc etc."

(A_{1440} might reasonably be said to get me there on time but I'd have to stay overnight in the airport …)

Methods for handling uncertainty

- Default or nonmonotonic logic:
  - Assume my car does not have a flat tire
  - Assume A_{25} works unless contradicted by evidence
- Issues: What assumptions are reasonable? How to handle contradiction?

- Rules with fudge factors:
  - $A_{25} \mid\rightarrow_{0.3}$ get there on time
  - $Sprinkler \mid\rightarrow_{0.99} WetGrass$
  - $WetGrass \mid\rightarrow_{0.7} Rain$
- Issues: Problems with combination, e.g., *Sprinkler* causes *Rain*??

- Probability
  - Model agent's degree of belief
  - Given the available evidence,
  - A_{25} will get me there on time with probability 0.04

Probability

Probabilistic assertions summarize effects of

- laziness: failure to enumerate exceptions, qualifications, etc.
-
- ignorance: lack of relevant facts, initial conditions, etc.
-

Subjective probability:

- Probabilities relate propositions to agent's own state of knowledge

  e.g., $P(A_{25} \mid \text{no reported accidents}) = 0.06$

These are not assertions about the world

Probabilities of propositions change with new evidence:

e.g., $P(A_{25} \mid \text{no reported accidents, 5 a.m.}) = 0.15$

Making decisions under uncertainty

Suppose I believe the following:

$P(A_{25}$ gets me there on time | …) = 0.04
$P(A_{90}$ gets me there on time | …) = 0.70
$P(A_{120}$ gets me there on time | …) = 0.95
$P(A_{1440}$ gets me there on time | …) = 0.9999

- Which action to choose?
- Depends on my preferences for missing flight vs. time spent waiting, etc.
    - Utility theory is used to represent and infer preferences
    - Decision theory = probability theory + utility theory

Syntax

- Basic element: random variable
- Similar to propositional logic: possible worlds defined by assignment of values to random variables.
- Boolean random variables
- e.g., *Cavity* (do I have a cavity?)
- Discrete random variables
- e.g., *Weather* is one of <*sunny,rainy,cloudy,snow*>
- Domain values must be exhaustive and mutually exclusive
- Elementary proposition constructed by assignment of a value to a
- random variable: e.g., *Weather* = *sunny*, *Cavity* = *false*
- (abbreviated as $\neg$*cavity*)
- Complex propositions formed from elementary propositions and standard logical connectives e.g., *Weather* = *sunny* $\vee$ *Cavity* = *false*

Syntax

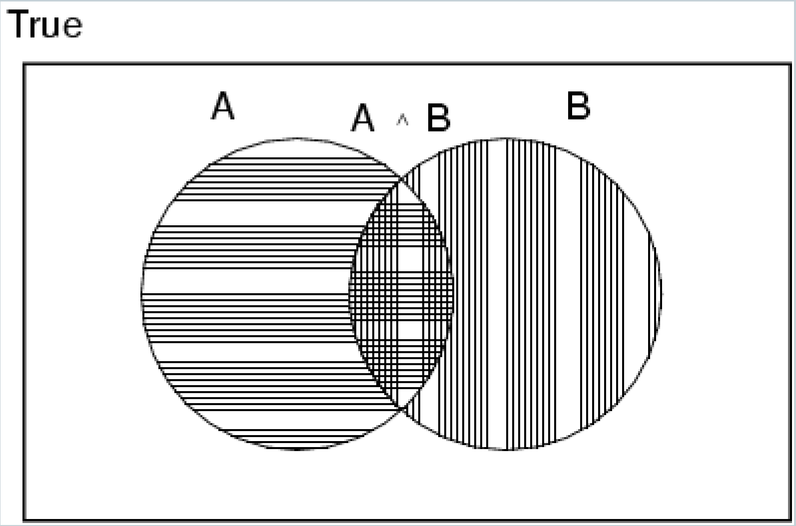
- Atomic event: A complete specification of the state of the world about which the agent is uncertain
-

  E.g., if the world consists of only two Boolean variables *Cavity* and *Toothache*, then there are 4 distinct atomic events:

$$Cavity = false \wedge Toothache = false$$
$$Cavity = false \wedge Toothache = true$$
$$Cavity = true \wedge Toothache = false$$
$$Cavity = true \wedge Toothache = true$$

- Atomic events are mutually exclusive and exhaustive
-

Axioms of probability

- For any propositions *A*, *B*
-
  - $0 \le P(A) \le 1$
  - $P(true) = 1$ and $P(false) = 0$
  - $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$
  - 

Prior probability

- Prior or unconditional probabilities of propositions
- e.g., P(*Cavity* = true) = 0.1 and P(*Weather* = sunny) = 0.72 correspond to belief prior to arrival of any (new) evidence

- Probability distribution gives values for all possible assignments:
- **P**(*Weather*) = <0.72,0.1,0.08,0.1> (normalized, i.e., sums to 1)
- Joint probability distribution for a set of random variables gives the probability of every atomic event on those random variables
- **P**(*Weather*,*Cavity*) = a 4 × 2 matrix of values:

Weather =	sunny	rainy	cloudy	snow
Cavity = true	0.144	0.02	0.016	0.02
Cavity = false	0.576	0.08	0.064	0.08

- Every question about a domain can be answered by the joint distribution

Conditional probability

- Conditional or posterior probabilities
- e.g., $P(cavity \mid toothache) = 0.8$

  i.e., given that *toothache* is all I know

- (Notation for conditional distributions:
- $\mathbf{P}(Cavity \mid Toothache)$ = 2-element vector of 2-element vectors)

- If we know more, e.g., *cavity* is also given, then we have
- $P(cavity \mid toothache, cavity) = 1$

- New evidence may be irrelevant, allowing simplification, e.g.,
- $P(cavity \mid toothache, sunny) = P(cavity \mid toothache) = 0.8$
- This kind of inference, sanctioned by domain knowledge, is crucial