Class note (11/26/2001) by Rui Yamada rui@cc.gatech.edu

Probabilistic theory

	Ontology(reality)	Epistemology(knowledge)
Propositional logic	Facts	True, False (Unknown)
Predicate calculus	Fact, Relation	True, False(Unknown)
Probability	**Facts**	**Degree of Belief**
Fussy logic	Degree of Truth	Degree of Belief

Utility

Desirability of state

Happiness in a state

Rationality

(from viewpoint of agent, rationality→ goal directed)

Decision-theoretic view

What does decision theoretic agent do?

function: DT (Decision Theoretic) agent percept, returns action.

knowledge: a set of probabilistic beliefs about the world

1. Calculate updated probability for current state based on available evidence including current percept and previous actions
2. Calculate outcome probability for actions
3. Select an action with the highest expected utility based on outcome probability and utility information
4. Return the action.

Question:

If each action is primitive (compared with goal), it is difficult to return one action.

Answer:

Use propagating probability (and return sequence of actions)

Who wants to be a millionaire?

Choice1: no toss, you can get $1 million.

Choice2: coin toss, Head→$0, Tail→$3 million

Assuming fair coin and toss

Expected value is…

1/2 x (0) + 1/2 x (3milion) = $1.5 million (Choice 2)

or

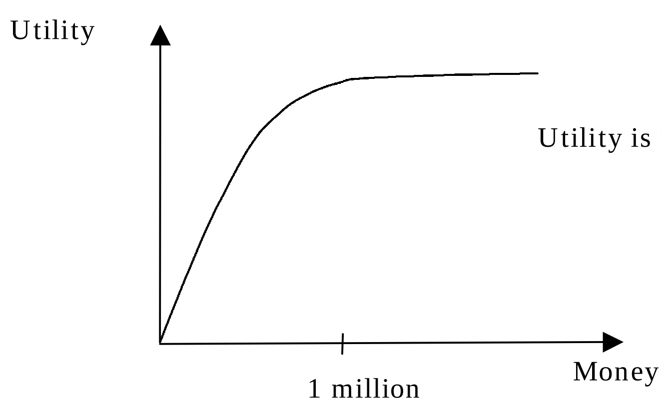
$1 million (Choice 1)

Goal: maximizing amount of money (Choice2>Choice1)

→ This doesn't work, why?

Multimillionaire says

"First million dollar is very cool, but next million dollars doesn't matter."



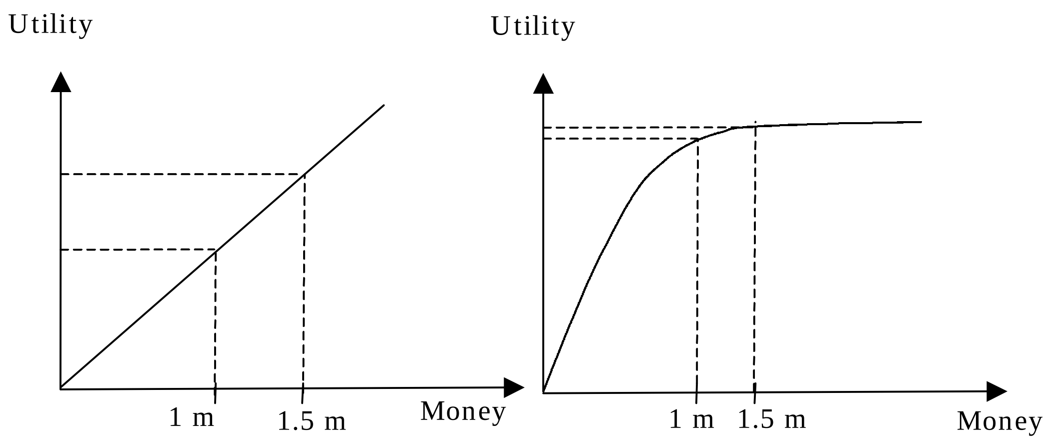
Expected utility (no toss) = u (k + $1 million)

Expected utility (toss) = 1/2 u (k + 0) + 1/2 u (k + $3 million)

Value k is different for each person.

For rich people, 1 million and 1.5 million don't make big difference.

For poor people, 1 million and 1.5 million make a difference.



Unconditional Probability

0 <= P(a) <= 1

Conditional Probability

P (cavity | toothache) = 0.002

hypothesis given

Product rule

$$P(A \cap B) = P(B|A)P(A)$$

$$P(A \cap B) = P(A|B)P(B)$$

Bayes' Theorem

$$P(B|A)P(A) = P(A|B)P(B)$$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Abduction

Probabilistic theory is used by vision, speech recognition or medical diagnosis etc.

Example: Medical diagnosis

effect → cause

stiff neck →meningitis

P (meningitis | stiff neck) = ?

$$P(meningitis | stiffneck) = \frac{P(stiffneck | meningitis)P(meningitis)}{P(stiffneck)}$$

P(stiffneck | meningitis) = 1/2

P(meningitis) = 1/5000

P(stiff neck) = 1/20

Then

P(meningitis | stiff neck) = 0.0002

Anthrax→ not enough sample → not suitable for this theorem because of unstable population

Should be… stable population, probability is known and probabilities are stable.

This is why speech recognition system is tuned for specific accent.

Conditional independence assumption

$$P(cavity | toothache \cap caught\_hook) = P(cavity | toothache)$$

Cavity doesn't depend on caught_hook.

Bayes' network

This is based on

Bayes' theorem

and conditional independence assumption.

Bayes' network is product of AI.

Suppose multiple variables

How do we apply Bayes' theorem?

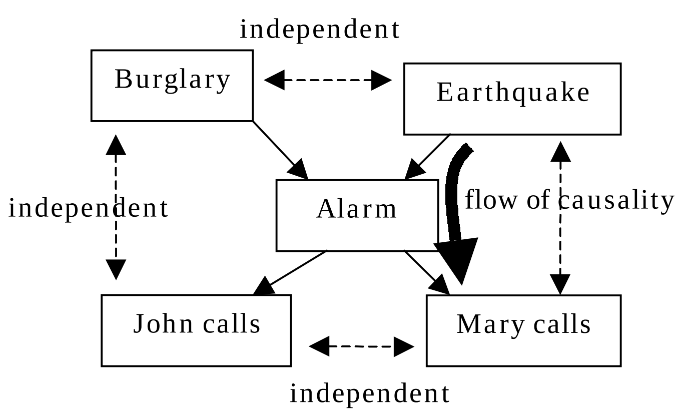
example:

You have alarm system in your house.

You have two neighbors John and Mary.

John and Mary call you if alarm is on.

Alarm detects either burglary or earthquake.



Question is

Given "John calls", what is the chance that earthquake occurred?

Joint Probability Distribution

system characterized by n variables   atomic event x1,x2, …, xn

B, not E, A, J, not M

n = 2 cavity and toothache

C, T

not C, not T

C, not T

not C, T

	T	not T
C	0.04	0.06
not C	0.01	0.89

Joint probability → very powerful

This will be explained in next class.