Class 1 - May 24

Exam Pattern:
- Major 50%
- Quiz 2 (15%, 15%)
- Assignment 20%

Faculties:
- Dr Vignesh Murlidharan
- Dr Dipanjan Roy

Random Variable: In Math it is a function which maps state space to sample space
X: Ω -> E where Ω (state space) maps to E (sample space) variable
- For a coin toss state space Ω -> Head/Tail and Variable E -> 1/0

for Probability P(X); probability distribution of coin toss

Probability $P(X)$

Probability distribution in case of coin
$H \rightarrow 1/2 \qquad T \rightarrow 1/2$

=> See the extra reference on next page

⇒ Conditional Probability: Possibility of happening an event, based on the existence of a previous event
Probability of a stock is going high if the index of the sector is going high

$$\boxed{P(X \mid Y) = \frac{P(X, Y)}{P(Y)}}$$

{ $P(X, Y) \rightarrow$ Joint probability

⇒ Product Rule (Extension of Conditional Probability)

$$P(X \mid Y) = \frac{P(X, Y)}{P(Y)}$$

$$\Rightarrow P(X \mid Y) \cdot P(Y) = P(X, Y)$$

$$\Rightarrow \boxed{P(X, Y) = P(X \mid Y) \cdot P(Y)}$$

Extra Reference:

In the context of a coin toss, the sample space represents all possible outcomes of the experiment, which are {Heads, Tails}. The random variable could be defined as the outcome of a coin toss, with each outcome (Heads or Tails) being assigned a numerical value. The state space is not typically used in a simple coin toss scenario as it's a term more often associated with dynamic systems evolving over time, .

Elaboration:

- Sample Space: The sample space for a single coin toss is simply the set of all possible outcomes. In this case, it's {H, T}, where H represents Heads and T represents Tails.

- Random Variable: A random variable is a function that assigns a numerical value to each outcome in the sample space. For example, we could define a random variable X where X = 1 if the outcome is Heads and X = 0 if the outcome is Tails.

- State Space:
  State space, in contrast, is used to describe the set of all possible states of a system over time. In a simple coin toss scenario, there's no evolution over time, so the concept of state space isn't applicable

⇒ Independence between two random variable

$$\boxed{P(X,Y) = P(X)P(Y)}$$

Explanation: In two coins toss the probability of coming head on each coin is: .5. The probability of coming head on both or any combination(HH,HT,TT, TH) is .5*.5 =0.25.25(1/4)

Mutual Exclusive	vs	Independent Events
- Can't happen together		Can happen together, but are Independent & doesn't influence other instruction

⇒ Expected value of a Random variable X

$$E(X) = \sum_{x \in X} x \cdot P(X = x)$$

X can be considered as weight
Check next page for example

⇒ Baye's Inference

$$\underbrace{P(H \mid E)}_{\text{Posterior}} = \frac{\overbrace{P(E \mid H)}^{\text{Likelihood}} \cdot \overbrace{P(H)}^{\text{prior}}}{P(E)}$$

/ divide / Normalize by probability of Evidence

H - Hypothesis / Event
E - Evidence

$P(H \mid E)$ → Probability of a Hypothesis or events given there is some evedence for the event

$P(E \mid H)$ → You will get Probability of evidence given the events

This is How the probability of hypothesis change as additional evidence $\{P(E \mid H)\}$ appears

	Frequentiest	vs	Bayesian
Example	Searching phone based on vibration sound coming from a place		Searching Phone based on some prior (places where the phone lost earlier) and you give more weightage

The expected value of a discrete random variable represents the average value of its possible outcomes, weighted by their probabilities. The formula for calculating expected value is E(X) = Σ (x * P(X = x)), where x represents a possible value and P(X = x) is the probability of that value occurring.

Example:

Consider a fair six-sided die. The possible outcomes are 1, 2, 3, 4, 5, and 6. Each outcome has a probability of 1/6.

- Step 1:

Identify the possible values of the random variable (die roll): 1, 2, 3, 4, 5, 6.

- Step 2:

Determine the probability of each value (1/6 in this case).

- Step 3:

Multiply each value by its probability:

  - 1 * (1/6) = 1/6
  - 2 * (1/6) = 2/6
  - 3 * (1/6) = 3/6
  - 4 * (1/6) = 4/6
  - 5 * (1/6) = 5/6
  - 6 * (1/6) = 6/6

- Step 4:

Add the products: (1/6) + (2/6) + (3/6) + (4/6) + (5/6) + (6/6) = 21/6 = 3.5

Therefore, the expected value of a single roll of a six-sided die is 3.5.

⇒ Utility Theory & function :

Utility Theory + Probability Theory = Decision - Theoritic Agent

Some agents are goal oriented & doesn't consider Uncertainity.

Decision Theoritic Agents consider that Uncertainity

Decision Theory : Choosing action based on desirability of Immediate outcomes

$P(\text{Result}(a) = s' \mid a, e)$ can be written as $\rightarrow s' \mid a, s)$

a — action

s' — state

e — Prior Evidence

action taking state s to s'

Baise Nets ?

Goal oriented Agent

In utility theory, agent has to take Decision and can't Just Do the timepass

Agents Preference : captured by a utility function (Averaged by expected utility

$$EU(a \mid e) = \sum_{s'} P(\text{Result}(a) = s' \mid a, e)\, V(s')$$

- => See Extra Ref on next page

The essence of MEU - Maximum Expected Utility

- The Agent acts to maximize a utility function by achieving the highest performance score

Extra Reference:

$$EU(a \mid e) = \sum_{s'} P(\mathrm{RESULT}(a) = s' \mid a, e) \cdot U(s')$$

In Simple Words:

This formula tells an AI system **how to choose the best action** by calculating its **Expected Utility**, based on:

What **action** it can take,

What it **knows** about the world (its current **evidence**),

What possible **future states** might result from the action,

And how **good** (or bad) each future state is.

Breakdown of the Formula

Symbol	Meaning	In Simple Terms
a	An **action**	Something the AI might do (e.g., "go left", "buy stock", "send message")
e	**Evidence** or current information	What the AI currently knows (like sensor readings or observations)
s'	A **possible future state**	What the world might look like after taking action a
$P(\mathrm{RESULT}(a) = s' \mid a, e)$	Probability that taking action a in current state e leads to future state s'	How likely this outcome is
$U(s')$	Utility of future state s'	How good or desirable that future state is
$\sum_{s'}$	Sum over all possible future states	Add up the weighted utilities of all possible outcomes

Basis of utility theory — 6 axioms

① – Orderability - Prefer A over B or B over A $A \succ B$ $B \succ A$

or Be indifferent between them, $\underbrace{A \sim B}_{\text{fine with Any.}}$

② - Transitivity

$A \succ B$, $B \succ C$, $A \succ C$

A will prefer A over C

③ **Continuity**

$A \succ B \succ C$, then $\exists p$ such that $[p, A ; 1-p, C] \sim B$

If B is in between A & C then there exist some probability p such that

Agent is indifferent between getting B for sure and getting a lottery between A & C

④ **Substitutability**

If $A \sim B \Rightarrow [p, A; 1-p, C] \sim [p, B; 1-p, C]$

One (A) can be substituted for other (B) without changing the agents preference

Example: 70% Shimla 30% Kerala
70% Manali 30% Kerala

⑤ **Monotonicity**

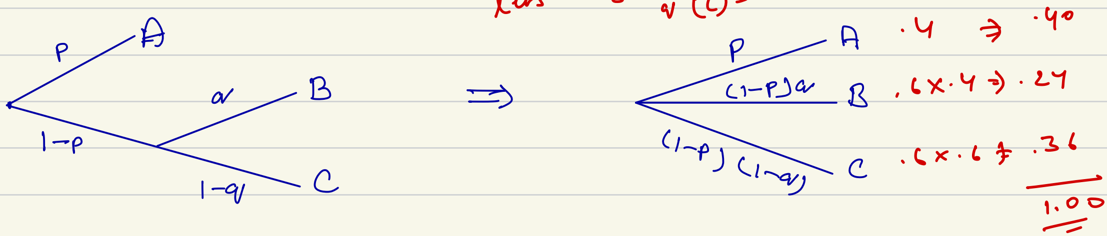
If agent prefers A over B, it will choose the lottery where probability of A winning is higher.

If $A \succ B$

$[p, A; 1-p, B] \succ [q, A; 1-q, B]$ if and only if $p > q$

⑥ Decomposibility

$P(A) = .4$
$P(\neg A) = .6 \Rightarrow q(B) + q[C]$
lets say $q(B) = .4$
$q(C) = .6$



A compound lottery can be reduced to an equivalent simple lottery by compounding the total probabilities

$$[p, A; 1-p, [q, B; 1-q, C]] \sim [p, A; (1-p)q, B; (1-p)(1-q), C]$$

- **Orderability**: Given any two lotteries, a rational agent must either prefer one or else rate them as equally preferable. That is, the agent cannot avoid deciding. As noted on page 412, refusing to bet is like refusing to allow time to pass.

  Exactly one of $(A \succ B)$, $(B \succ A)$, or $(A \sim B)$ holds.

- **Transitivity**: Given any three lotteries, if an agent prefers A to B and prefers B to C, then the agent must prefer A to C.

  $$(A \succ B) \land (B \succ C) \Rightarrow (A \succ C).$$

- **Continuity**: If some lottery B is between A and C in preference, then there is some probability p for which the rational agent will be indifferent between getting B for sure and the lottery that yields A with probability p and C with probability $1-p$.

  $$A \succ B \succ C \Rightarrow \exists p \; [p,A;\; 1-p,C] \sim B.$$

- **Substitutability**: If an agent is indifferent between two lotteries A and B, then the agent is indifferent between two more complex lotteries that are the same except that B is substituted for A in one of them. This holds regardless of the probabilities and the other outcome(s) in the lotteries.

  $$A \sim B \Rightarrow [p,A;\; 1-p,C] \sim [p,B; 1-p,C].$$

  This also holds if we substitute $\succ$ for $\sim$ in this axiom.

- **Monotonicity**: Suppose two lotteries have the same two possible outcomes, A and B. If an agent prefers A to B, then the agent must prefer the lottery that has a higher probability for A (and vice versa).

  $$A \succ B \Rightarrow (p > q \Leftrightarrow [p,A;\; 1-p,B] \succ [q,A;\; 1-q,B]).$$

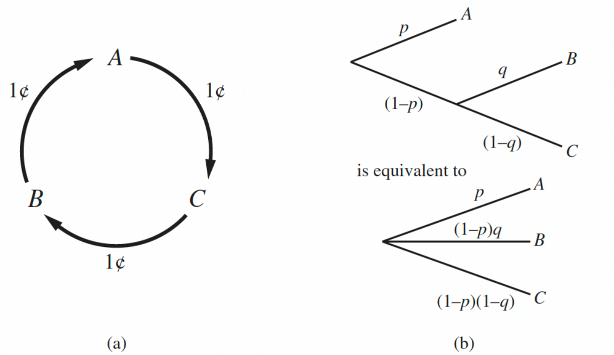


Figure 15.1 (a) Nontransitive preferences $A \succ B \succ C \succ A$ can result in irrational behavior: a cycle of exchanges each costing one cent. (b) The decomposability axiom.

- **Decomposability**: Compound lotteries can be reduced to simpler ones using the laws of probability. This has been called the "no fun in gambling" rule: as Figure 15.1(b) shows, it compresses two consecutive lotteries into a single equivalent lottery.[2]

  $$[p,A;\; 1-p,[q,B;\; 1-q,C]] \sim [p,A;\; (1-p)q,B;\; (1-p)(1-q),C].$$

Violations

There are some violations as well. for example violations in transitivity can b

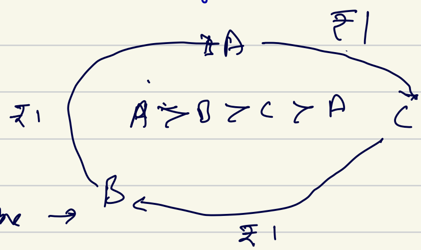
$A \succ B \succ C \succ A$

$\Downarrow$

This is violations because as per transitivity

$A \succ C$

in circular behavior will extract all the value of agent $\rightarrow$



Utilities are function over preferences $\Rightarrow$ $\boxed{\text{Preferences lead to Utility}}$

$\Rightarrow$ Existence of utility function

$$U(A) > U(B) \Leftrightarrow A \succ B$$

$$U(A) = U(B) \Leftrightarrow A \sim B$$

$\Rightarrow$ Expected Utility: Sum probability of each outcome multiply by utility of that outcome

$$U([p_1, S_1;\ p_2 S_2 \ \ldots\ ;p_n S_n]) = \sum_i p_i U(S_i)$$

$\Rightarrow$ Theory says Utility exist but is not unique

$$U'(S) = aU(S) + b$$

Utility function changes, but agents preferences remains same for that problem

Utility of Money: Money is a substitute of utility?

- monotonic preference towards money.
- But money doesn't behave like a utility function

Gamble on coin flip for more money?

\$100000 to \$2500000 on coin flip

Expected monitory value

$EMV(A) = 1.75$ Million $\Rightarrow \frac{1}{2}(1M) + \frac{1}{2}(2.5M)$

$EMV(B) = 1$ Million $= .5M + 1.25M$

$= 1.75M$

or

$1.(1M) + \frac{1}{2}(1.5M)$

$= 1 + .75$

$= 1.75M$

person has S_k in Bank account

$$EU(A) = \frac{1}{2} U(S_k) + \frac{1}{2} U(S_k + 2.5M)$$

$$EV(B) = U(S_k + 1M)$$

✅ **Correct formula:**

$$\frac{1}{2}U(S_k) + \frac{1}{2}U(S_k + 2.5M)$$

- This says:
  - With 50% probability, you end up with **just** your original money S_k → utility is $U(S_k)$.
  - With 50% probability, you end up with $S_k + 2.5M$ → utility is $U(S_k + 2.5M)$.

You're taking the expected **utility** of the two total **final wealth** outcomes.

❌ **Your version:**

$$U(S_k) + \frac{1}{2}U(2.5M)$$

This has two big problems:

1. **You're not weighting the $U(S_k)$ term by probability.**
   - The original $U(S_k)$ only happens **50% of the time** (when the gamble is lost).
   - So it should also be multiplied by $\frac{1}{2}$.
   - If you write just $U(S_k)$, you're incorrectly assuming you *always* get S_k, regardless of the gamble outcome.
2. **You're treating the $2.5M as a standalone value instead of a change to your current wealth.**
   - In reality, you don't just get $2.5M out of nowhere — it's **added to** your existing S_k.
   - So the final amount should be $S_k + 2.5M$, and the correct utility is $U(S_k + 2.5M)$, not just $U(2.5M)$.

Setting some utility values

Random / Assume

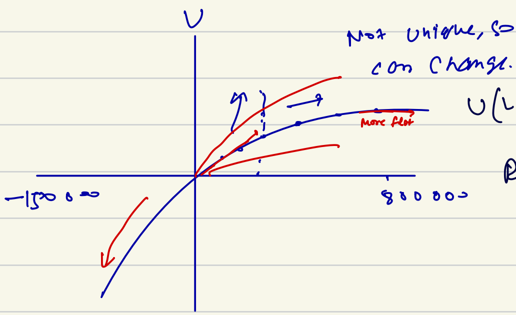
$$EU(A) = \frac{1}{2} \underbrace{U(S_k)}_{5} + \frac{1}{2} \underbrace{U(S_k + 2.5M)}_{9}$$

$$EV(B) = U(S_k + 1M) \rightarrow 8$$

$$EV(A) = \frac{1}{2} \times 5 + \frac{1}{2} \times 9 = 2.5 + 4.5 = \boxed{7}$$

$$EV(B) = \boxed{8} ✓$$

$$MEV \Rightarrow EV(B) > EV(A)$$



$U(L) < (Exp\,V(L))$

→ Risk-averse

Another utility function

Class - 2 May 11

$\boxed{\text{Rational Deisions \& Post-decision disappointment}}$

- The Rational way to choose the best action

$$a^* = \text{argmax}_a\, EU(a|e^s)$$

- MEU :- Choose the action that gives the highest expected utility, given the evidence

$$EU = \sum_s \underbrace{p(s|a,e)}_{\text{Model of the world — Uncertain, Oversimplifies}} \cdot U(s)$$

Extra Ref in Next page

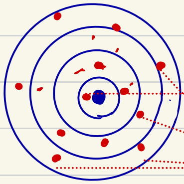
- EU is calculated by combining
  - Probability of each outcome from model
  - Utility (value) of each outcome : True utilities are interactable so we used unbiased estimates $\widehat{EU}$

$$E(\widehat{EU} - EU) = 0 \;\{E\text{-error},\ \widehat{EU}\text{: estimated},\ EU\text{: True}\}$$

"Intractable" means it's **too difficult (or impossible) to compute**

- Some times $\widehat{EU}$ might be more and some times EU would be more so we would consider average E = +ve or −ve

- Things will still go wrong.
- Real outcome is significantly worse than what estimated
- This is where optimizers curse enters



- True skill is worse than their best
- when picking best, overestimating the true value because selecting among the noisy

Use Baye's Rule
- Start with a prior estimate of EU, P(EU)
- Use the expected utility as evidence
- Estimate the posture distribution

$$P(EU|\widehat{EU}) \underset{\text{Propotonal}}{\propto} P(\widehat{EU}|EU) \cdot P(EU)$$

Extra ref:

n decision theory and AI, when choosing an action, the agent doesn't know exactly what the result of that action will be. Each action can lead to **multiple possible states**, each with a different probability and utility.

The summation is necessary because:

Uncertainty over outcomes: Since we can't be sure which state will result from an action, we need to account for **all possible outcomes** weighted by their likelihood.

Expected value concept: This is similar to expected value in probability. You're computing the **average utility** you'd get from action a

a

a, considering all possible results and their probabilities.

Making rational decisions: Summing gives a single expected utility value for each action. The agent then chooses the action a

a

a that **maximizes** this expected utility.

Human Judgement and Irrationality

Decision Theory: How a rational agent should act (Normative)

Descriptive Theory: How a actual agent (like humans) really do act

- The applicions of both theories will enhanced, if coincides
- Evidence is to contrary - Humans are predictably irrational

20% chance of \$4000

25% chance of \$3000

C > D (Poll Result)

Normative analysis

$.2 \times 4000 = 800 \qquad .25 \times 3000 = 750$

$EU(\$4000) > EU(\$3000)$

$0.2 \times U(\$4000) > 0.25 \times U(\$3000)$ → ①

Multiply 1 to 4 times to make it .8

to comare with other Scenario

$.8\,U(\$4000) > 1 \times U(\$3000)$

A 80% chance of winning \$4000

B 100% " " \$3000

B > ~~A~~ (Poll result)

Normative analysis.

$.8\,U(\$4000) < U(\$3000)$

Contradictory

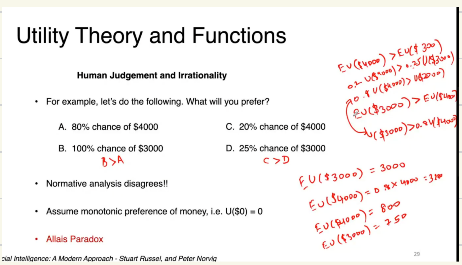
Allais Paradox

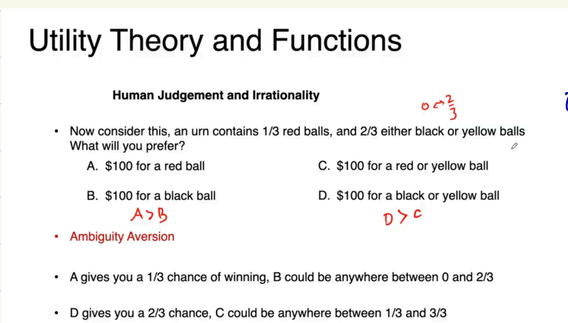
Why?

① certainity effect

② Distrust (is com fair?)

③ Dont want to Regret loosing





Ellesbord Paradox

Uncertainity impact judgement

Framing effect: How the fact is presented impacts agents choice

10% death chance vs 90% surviving chance

Anchoring effect: Easy to make relative decision

Utility Theory and Functions

Thinking Fast and Slow
Daniel Kahneman

Other interesting Effects

- Framing Effect: Exact wording of a decision problem can have big impact on agent's choices
- Treatment has "90% chance of survival" rather than "10% chance of death"
- Anchoring Effect: People are more comfortable making relative utility judgements than absolute ones

$U(s') = aU(s) + b$

- Serving a $200 wine bottle – then serving a $55 one which seems like a bargain.

Decision Theory & Network

Decision Theory → Acting in uncertain world

MEU
Rational agent
Maximize agents Expected Utility

Utility Theory ↓ what does the agent want	+	Probability Theory What should the agent believe based on evidence	=	Decision – Theoritic Agent How should the agent act in uncertain world

Decision Theory and Networks

Important to reason in an uncertain world

- However, more importantly it is necessary to make decisions in an uncertain world.
- Decision Theory → Acting in an uncertain world

MEU

Rational Agent maximizes agent's Expected Utility

Utilities	+	Bayesian Networks	+	Actions	=	Decision-theoretic Agent
↓ What does the agent want?		↓ What should the agent believe based on evidence?				↓ How should the agent act in an uncertain world?

Decision Theory and Networks

Constructing a Decision Network

- A package pick-up robot – as a running example

- A robot must choose its route to pick up packages. There are two routes: a short one and long one. The short route is a bit treacherous – the robot can encounter an accident. However, the robot can wear protective gear which does not change the probability of an accident but will reduce the damage in case an accident occurs. Unfortunately, the protective gear will add more weight to the robot and slow it down. The aim is to pick up the package as soon as possible while minimizing the damage caused by an accident.

- **What should the robot do?**

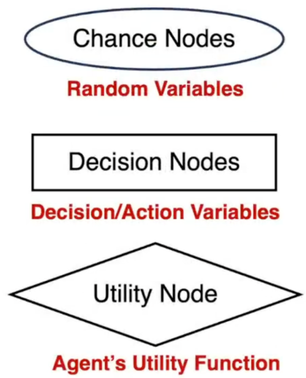
Decision Theory and Networks

Constructing a Decision Network

- Let's see what are the variables

- What are the random variables here?
  - A: whether an accident occurs or not

- What are the decision variables here (actions)?
  - S: whether the robot chooses the short route
  - P: whether the robot uses the protective gear



Class - 3

May 17

Decision Theory & network

Packag pickup problem from last class - continue

Random variable →

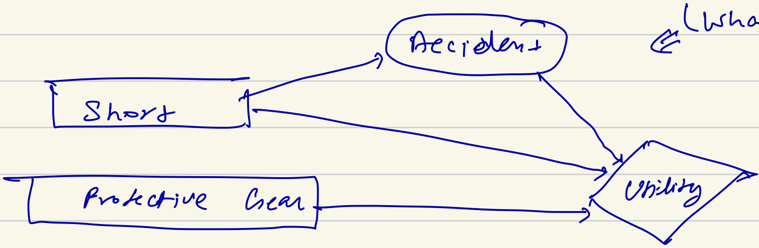
- whether an accident occurs or not : Probability

Decision variable

- whether the robot chooses short route
- " " " uses the protective gear

Robots Decision Network (Graphical Network)

(what will affect what)



Long Route → Acciden won't occur

Short Route → Accident may occur - Probability q

$P(A|S) = q \longrightarrow 0 \leq q \leq 1$

$P(\neg A|S) = 1-q$ (No accident for shorter route)

$P(A|\neg S) = 0$ (for longer route)

A Critical Question

- When an accident does NOT happen, which of the following statements is true?

1. The robot prefer not to wear the protective gear than to wear it. $S, \neg P$
2. The robot prefers the long route over the short route
3. Both (A) and (B) are true
4. Both (A) and (B) are false

An example Utility function for the robot

Variables	State	$U(w_i)$
¬P, ¬S, ¬A	w_0 slow, no weight	6
¬P, ¬S, A	w_1 impossible	
¬P, S, ¬A	w_2 quick, no weight	10
¬P, S, A	w_3 severe damage	0
P, ¬S, ¬A	w_4 slow, extra weight	4
P, ¬S, A	w_5 impossible	
P, S, ¬A	w_6 quick, extra weight	8
P, S, A	w_7 moderate damage	2

How does the robot's Utility/Happiness depend on the random variables and decision variables?

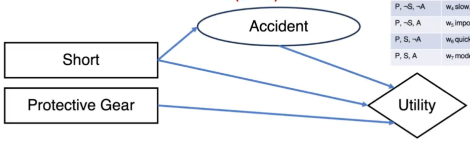
- When an accident does NOT happen, does the robot prefer not to wear the protective gear or to wear it?
- When an accident does NOT happen, does the robot prefer the short route or the long route?

- When an accident does happen, does the robot prefer not to wear the protective gear or to wear it?
- When an accident does happen, does the robot prefer the short route or the long route?

Decision Theory and Networks

The complete Robot Decision Network

- $P(A \mid S) = q, 0 \le q \le 1$
- $P(A \mid \neg S) = 0$



Variables	State	$U(w_i)$
¬P, ¬S, ¬A	w_0 slow, no weight	6
¬P, ¬S, A	w_1 impossible	
¬P, S, ¬A	w_2 quick, no weight	10
¬P, S, A	w_3 severe damage	0
P, ¬S, ¬A	w_4 slow, extra weight	4
P, ¬S, A	w_5 impossible	
P, S, ¬A	w_6 quick, extra weight	8
P, S, A	w_7 moderate damage	2

This all will help to identify that what the robot should do

How do we choose an action

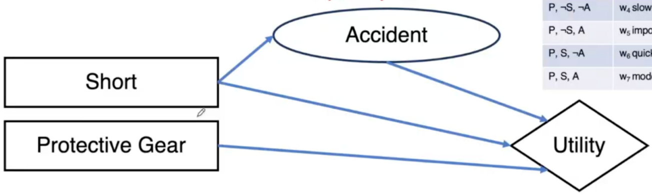
Evaluating the Robot Decision Network

- How do we choose an action?
  - Set evidence variables for the current state
  - For each possible value of the decision node
    - Set decision node to that value
    - Calculate the posterior probabilities for the parent nodes of the utility node
    - Calculate expected utility for the action
  - Return action with the highest expected utility

Decision Theory and Networks

The complete Robot Decision Network

- $P(A \mid S) = q, 0 \le q \le 1$
- $P(A \mid \neg S) = 0$



Variables	State	$U(w_i)$
$\neg P, \neg S, \neg A$	w_0 slow, no weight	6
$\neg P, \neg S, A$	w_1 impossible	
$\neg P, S, \neg A$	w_2 quick, no weight	10
$\neg P, S, A$	w_3 severe damage	0
$P, \neg S, \neg A$	w_4 slow, extra weight	4
$P, \neg S, A$	w_5 impossible	
$P, S, \neg A$	w_6 quick, extra weight	8
P, S, A	w_7 moderate damage	2

We can make 4 decision here. S+P, S Not P, Not S + P, not S Not P

1) $\neg P$ $\neg S$, Agents utility ?

$EU = \sum_{s'} P(s' \mid a, e)\, U(s')$

$EU(\neg P, \neg S) = P(W_0 \mid \neg P, \neg S)\, U(W_0) + P(W_1 \mid \neg P, \neg S) . U(W_1)$

$= P(\neg P, \neg S, \neg A \mid \neg P, \neg S)\, U(W_0) + P(\neg P, \neg S, A \mid \neg P, \neg S) . U(W_1)$ – Putting w_0, w_1 from table

$= P(\neg A \mid \underline{\neg P, \neg S})\, U(W_0) + P(\underline{A \mid \neg P, \neg S}) . U(W_1)$ – since $\neg P, \neg S$ is already we need not to find the probability of that

$= P(\neg A \mid \neg S) . U(W_0) + P(A \mid \neg S) . U(W_1)$ – since long route doesn't have accidents, so wearing protection gear would have no impact

$= 1 \cdot U(w_0) + 0 \cdot U(w_1)$

$= 1 \cdot 6 + 0 \cdot (\ ) = 6$

Q2 $EU(\neg P, S)$

$= EU = \sum_{s'} P(s' \mid a, e) \cdot U(s')$

$EU(\neg P, S) = P(w_2 \mid \neg P, S) \cdot U(w_2) + P(w_3 \mid \neg P, S) \cdot U(w_3)$

$= P(\neg P, S, \neg A \mid \neg P, S) \cdot U(w_2) + P(\neg P, S, A \mid \neg P, S) \cdot U(w_3)$

$= p(\neg A \mid \neg P, S) \cdot U(w_2) + p(A \mid \neg P, S) \cdot U(w_3)$

$= p(\neg A \mid S) \cdot U(w_2) + P(A \mid S) \cdot U(w_3)$

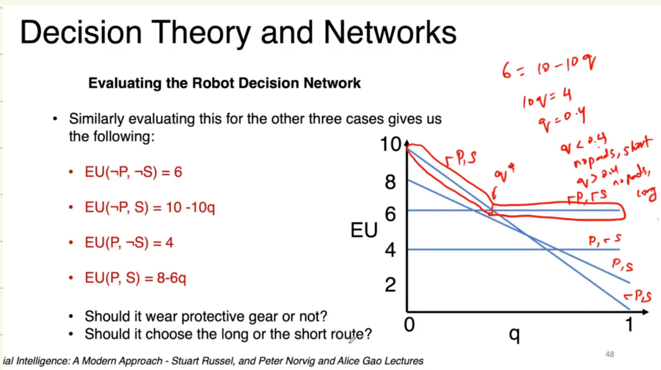
$= (1-q) \cdot U(w_2) + p(q) \cdot U(w_3)$

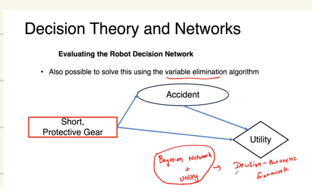
$= (1-q)\,10 + p(q) \cdot 0$

$= (1-q)10 \Rightarrow 10 - 10q$

Bayesian Networ
variable elimination
(VE) algo

- EU(¬P, ¬S) = 6
- EU(¬P, S) = 10 -10q
- EU(P, ¬S) = 4
- EU(P, S) = 8-6q

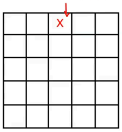




Decision Theory and Networks

Value of Information

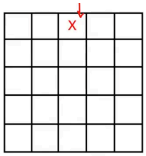
- Most important part of decision-making is knowing what question to ask.
- Information value theory enables agent to choose what information to acquire.
- Take this example for instance:
  - Seismologist gives clue
  - Company buys it for C/n dollars
  - In case block has oil:
    - Profit: C – C/n = (n-1)C/n



If found in first block

Value of Information

- Take this example for instance:
  - n blocks of land
  - 1 of them contains oil worth C dollars
  - Each block cost – C/n
  - Company buys it for C/n dollars
  - In case block has oil:
    - Profit: C – C/n = (n-1)C/n
  - In case block has no oil:
    - Profit: C/(n-1) – C/n = C/n(n-1)



$$Expected\ Profit = \frac{1}{n}.\frac{(n-1)C}{n} + \frac{n-1}{n}.\frac{C}{n(n-1)}$$

$$= \frac{C}{n}$$

$$\frac{1}{20}\left(\frac{20-1}{20}\right) \times 1000$$

$$\frac{19}{20 \times 20} \times 1000$$

5

2.5

38+9.5

$= \boxed{47.5}$

1000 – 50 = 950

– 50

Decision Theory and Networks

Value of Perfect Information (VPI)

- Let agent's current evidence be e.

$$EU(\alpha|\mathbf{e}) = \max_a \sum_{s'} P(\text{RESULT}(a) = s' \mid a, \mathbf{e})\, U(s')$$

- After new evidence e_j is obtained,

Ej -> RV all possible ejs from e

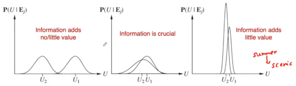
$$EU(\alpha_{e_j}|\mathbf{e}, e_j) = \max_a \sum_{s'} P(\text{RESULT}(a) = s' \mid a, \mathbf{e}, e_j)\, U(s')$$

- Ej is a random variable so the value of obtaining that information is

$$VPI_{\mathbf{e}}(E_j) = \left(\sum_k P(E_j = e_{jk}|\mathbf{e})\ EU(\alpha_{e_{jk}}|\mathbf{e}, E_j = e_{jk})\right) - EU(\alpha|\mathbf{e})$$

Value of Perfect Information (VPI)

- More intuitively it means the following, consider only 2 actions with utilities U_1 and U_2



Question 1
Not yet answered
Marked out of 1.00
Flag question

Which of the following expressions represents the correct version of the Bellman optimality form for the action-value function? **(1 Mark)**

- A. $Q_\pi(s,a) = \max_a P(s',r|s,a)(r + \gamma V_\pi(s'))$
- B. $Q_\pi(s,a) = \max_a P(s',r|s,a)(r + \gamma \sum_{a'} \pi(a|s) Q_\pi(s',a'))$
- C. $Q_\pi(s,a) = \sum_{s',r} P(s',r|s,a)(r + \gamma \max_{a'} Q_\pi(s',a'))$
- D. $Q_\pi(s,a) = \pi(a|s) \sum_{s',r} P(s',r|s,a)\left(r + \gamma \max_{a'} Q_\pi(s',a')\right)$

Question 2
Not yet answered
Marked out of 1.00
Flag question

In the Temporal difference algorithm for Value iteration, if alpha is the learning rate and gamma is the discount factor, which among the following properties do they exhibit? **(1 Mark)**

- A. Gamma and alpha both account for the past experiences
- B. Gamma and alpha both account for future experiences
- C. Gamma accounts for the past and alpha for the future
- D. Gamma accounts for the future and alpha for the past

Time left 0:56:56

Question 3
Not yet answered
Marked out of 1.00
Flag question

A rational agent is presented with outcomes A and B. Which of the following must be true? **(1 Mark)**

- A. The agent prefers both A and B equally.
- B. The agent can prefer A to B, B to A, or be indifferent
- C. The agent can avoid choosing between A and B.
- D. The agent should always prefer the option with the highest monetary value.

Question 4
Not yet answered
Marked out of 1.00
Flag question

What property here is not true about Passive Reinforcement learning? **(1 Mark)**

- A. It is a model-free approach of performing RL
- B. It can lead to policies that are better than the one initialized
- C. State transitions are available as samples
- D. One does not have to wait till the terminal state to update the value of a state.

Time left 0:55:48

Question 5
Not yet answered
Marked out of 2.00
Flag question

I am in a state S1 and make an action A1 and transition to state S2 to get a reward of +2. From S2, I make an action A2 and transition to state S3 and get a reward of -5. Finally, from S3 with action A3, I make a transition to state T, which is terminal state where I get a reward of +7. Compute the return at the state S1, S2 and S3 with a discount factor of 0.7. **(2 Marks)**

G(S1) = ______ **(Upto 2 Demcimal Places)**

G(S2) = ______ **(Upto 2 Demcimal Places)**

G(S3) = ______ **(Upto 2 Demcimal Places)**

Question 7
Not yet answered
Marked out of 1.00
Flag question

You're deciding whether to run a diagnostic test before choosing a treatment. The test costs ₹1000 and perfectly reveals the patient's condition. Without the test, you have: • 50% chance of correct treatment: +₹10,000 utility • 50% chance of incorrect treatment: −₹10,000 utility With the test, treatment is always correct. Should you buy the test? **(1 Mark)**

- A. Yes, the test removes uncertainty and increases expected utility
- B. No, the test reduces your net utility
- C. Only if test cost is refunded later
- D. It doesn't matter — utilities are symmetric

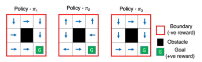
Question 8
Not yet answered
Marked out of 1.00
Flag question

What is the information that we know or can estimate from episodes of state-action-reward-next state, in case of a model-based RL approach? **(1 Mark)**

- A. a. Transition model of the environment
- B. Policy of the agent
- C. Probability distribution of the rewards
- D. None of the above

Question 9
Not yet answered
Marked out of 1.00
Flag question

9. Consider the following grid world and three different policies to navigate through it as shown in the image. The arrow represents the action at that state which has the highest value. Which scenario seems to reflect the most optimal policy for this grid world? (**1 Mark**)



- A. Policies p1 and p2
- B. Policy p2 only
- C. Policy p3 only
- D. Policies p2 and p3

Information
Flag question

A lottery ticket costs ₹2. There are two possible prizes: a ₹20 prize with probability 1/40 and a ₹5,00,000 prize with probability 1/10,00,000. All other outcomes yield ₹0. Assume your current wealth is k and your utility function satisfies: U(Sk) = 0; U(Sk+20) = 10 x U(Sk+2) No assumptions are made about U(Sk+500,000).

Question 10
Not yet answered
Marked out of 1.00
Flag question

a. What is the expected monetary value (EMV) of the ticket? **(1 Mark)**

- A. ₹0.75
- B. −₹1.00
- C. −₹1.50
- D. −₹2.00

Question 11
Not yet answered
Marked out of 1.00
Flag question

b.As a follow-up, suppose U(Sk+2) = 0.02. What is the expected utility of buying the ticket (ignoring the ₹5,00,000 prize)? **(1 Mark)**

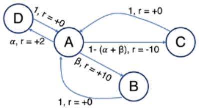
- A. 0.005
- B. −0.015
- C. 0.2
- D. Cannot be determined without knowing U(Sk+500,000)

Quiz-1

Back

Information
Flag question

11. A robot has been deployed in Mars for autonomous navigation. The Martian base has four zones, A, B, C and D. The robot aim to collect energy resources without getting harmed by radiation. The action space is search or stay. The following is the transition graph of the robot MDP and the robot starts at A.



Answer the following questions based on the graph:

Question 12
Not yet answered
Marked out of 1.00
Flag question

What are the terminal states in this MDP? (1 Mark)

- A. No terminal states
- B. States A and D
- C. Only State A
- D. States B, C and D

Question 13
Not yet answered
Marked out of 1.00
Flag question

What are the states where the robot is likely to find the energy resources? (1 Mark)

- A. States B, C and D
- B. States A and D
- C. States B and C
- D. States B and D

Question 14
Not yet answered
Marked out of 1.00
Flag question

Which of the following transitions in this MDP are deterministic? (1 Mark)

- A. D ® A
- B. A ® D
- C. A ® B
- D. None of the transitions are deterministic

Question 15
Not yet answered
Marked out of 1.00
Flag question

On evaluating this MDP with a = 0, which states do you expect the robot to prefer based on average expected return? **(1 Mark)**

- A. State C
- B. State B
- C. State D
- D. Depends on b

Question 16
Not yet answered
Marked out of 1.00
Flag question

A policy π for the robot is: (1 Mark)

- A. A table of state-action-reward mappings
- B. A function that maps states to actions
- C. A list of all terminal states
- D. A scalar representing the total reward

Question 17
Not yet answered
Marked out of 1.00
Flag question

If the robot occasionally chooses a random action instead of the one with the highest estimated value, it is: (1 Mark)

- A. Following a greedy policy
- B. Exploiting the environment
- C. Exploring to improve policy estimation
- D. Ignoring the MDP model

Previous page

Question 18

Not yet answered

Marked out of 1.00

Flag question

. Which of the following scenarios best illustrates the optimizer's curse? (1 Mark)

- A. A pharmaceutical company selects one experimental drug from 5,000 candidates based on estimated trial outcomes. The selected drug had the most promising results in early tests, but fails in large-scale clinical trials.
- B. A weather forecasting model predicts snow with 90% probability, and snow falls as predicted.
- C. A doctor recommends a treatment with a 95% success rate, and the patient recovers as expected.
- D. A student uses the most popular online study guide and performs just as the guide predicted.

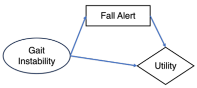
Time left 0:27:25

Information

Flag question

13. A decision-theoretic agent is designed to predict falls in elderly population. The agent takes into account in its decision-making the sensor data which informs gait instability (if it is present or not). Based on this factor it issues an alert to intimate the person regarding the fall. The figure below shows the decision network and the utility table for the fall-prediction agent.

The Fall-prediction Decision Network



Variables	State	$U(w_i)$
¬GI, ¬A	w_0 no instability, no fall alert	+10
¬GI, A	w_1 no instability, fall alert	-10
GI, ¬A	w_2 instability, no fall alert	-500
GI, A	w_3 instability, fall alert	+100

Answer the following questions based on the above:

Question 19

Not yet answered

Marked out of 2.00

Flag question

a. If the initial probability of gait instability is 0.6, what is the expected utility of raising and not raising an fall alert?

EU(Raise Alert) =

EU(Don't Raise Alert) =

Question 20

Not yet answered

Marked out of 1.00

Flag question

If the probability of no gait instability is 0.7, should the rational fall-prediction agent raise an alert?

- A. Yes, raise alert
- B. No, don't raise alert
- C. Depends on the utility
- D. None of the above