RDS2 G3 - Monday

Tutor: Dr. Lim Yee Mei (ymlim@tarc.edu.my), Dr. Christopher Lazarus (christopherl@tarc.edu.my)

Tutorial 14 (Fuzzy Logic & Modelling and Simulation)

Part I: Fuzzy Logic

1. Following shows a sample crisp rule.

IF temperature > 37.5 THEN fever = TRUE

a. Explain the difference between crisp rule and fuzzy rule.

Lim Ming Jun

In crisp rules, the result can only be TRUE or FALSE, and we can also use 1 and 0 to represent the answer. There is no other possible answer to this logic. However, the result of the fuzzy rule can be partially true or partially false, instead of entirely true or entirely false. This means that the result can be very close to 1 (e.g. 0.94) or very close to 0 (e.g. 0.04) or any values between 0 and 1 (e.g. 0.5, 0.6, 0.4)

b. Suggest how do you convert the crisp rule above into a fuzzy rule.

Lim Ming Jun

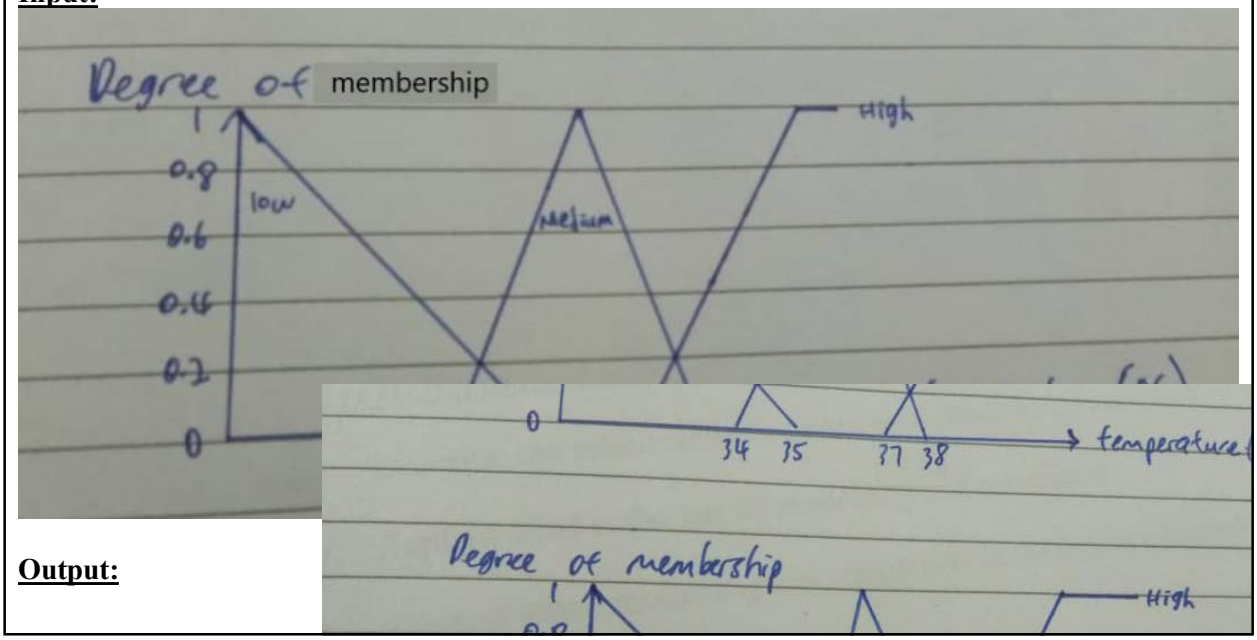
IF temperature is HIGH THEN fever is SERIOUS

c. Design the appropriate fuzzy sets to represent the input and output variables respectively.

Lai Pei Xuan

fever(x) =
{ 0,
  f(temperature(x),
  1,
  If temperature(x) <= 37.5,
  If temperature(x) >37.5}
IF temperature is high
THEN fever is high

Input:

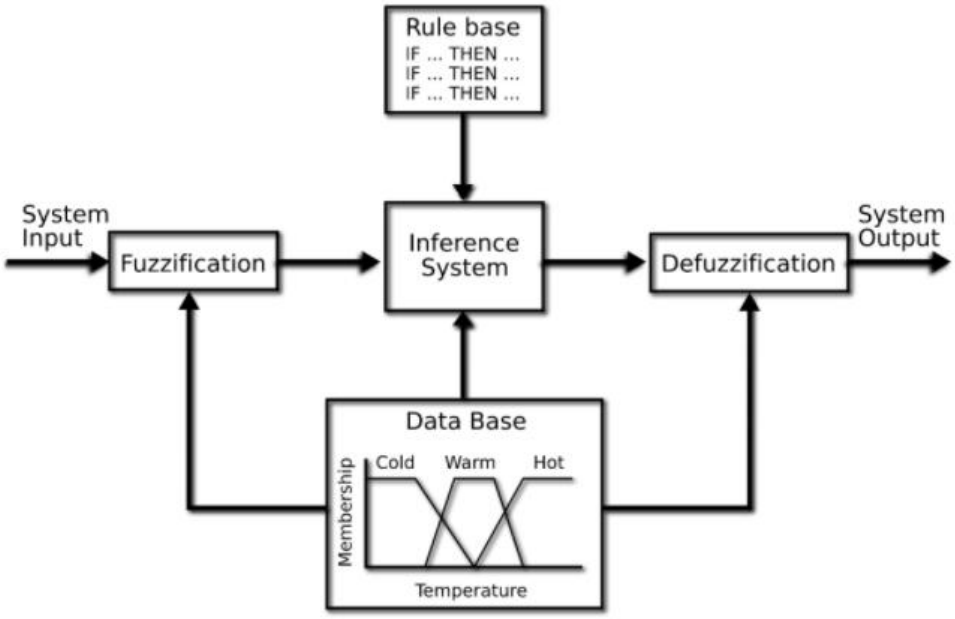


Output:

2. Illustrate the general structure of the Fuzzy Inference System and explain each of the components.

Lai Pei Xuan

Principal Component Analysis (PCA) can be used to identify two significant components



1. Fuzzification
   - Fuzzification is the process of transforming a crisp set to a fuzzy set or a fuzzy set to fuzzier set.
2. Rule- Base
   - It contains fuzzy IF-THEN rules.
3. Data Base
   - It defines the membership functions of fuzzy sets used in fuzzy rules.
4. Inference System (Decision-Making Unit)
   - It performs operation on rules
5. Defuzzification
   - Defuzzification is the process of reducing a fuzzy set into a crisp set or converting a fuzzy member into a crisp member.

3. Assume that you are designing an automated air conditioning system using Fuzzy Logic.

a. Suggest an uncertainty in the above scenario with explanation.

Lim Chia Chung

One uncertainty would be different levels of coldness. For example, what would we consider as cold? 18 degrees, 21 degrees?

Principal Component Analysis (PCA) can be used to identify two significant components

b. Generate the TWO (2) input and TWO (2) output fuzzy membership functions that are able to model the air conditioning system.

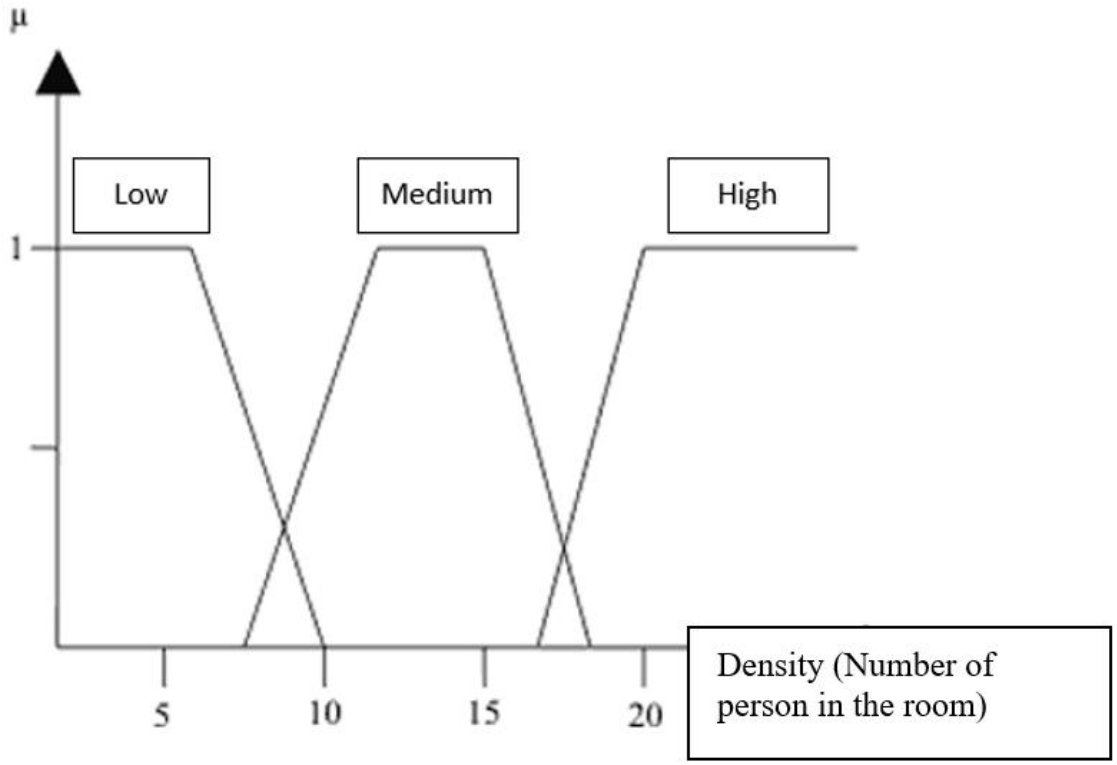
Lim Chia Chung

Student can provide any appropriate membership function, example:

Input: Temperature of the room, humidity, density, etc..

Output: Volume of the fans, cooling temperature, etc..

Example of membership function:



c. Suggest TWO (2) fuzzy rules that associate to the Fuzzy membership functions generated in Q3.b.

Lee Jun Xian

If the Temperature is COLD and the Density is LOW, Then the cooling temperature is HIGH

Principal Component Analysis (PCA) can be used to identify two significant components

If the Temperature is HOT and the Density is HIGH, Then the cooling temperature is LOW

Part II: Modelling and Simulation

1. Describe the relationship between modelling and simulation.

Lee Jun Xian

Modeling is the act of building a model. A simulation is the process of using a model to study the behavior and performance of an actual or theoretical system. In a simulation, models can be used to study existing or proposed characteristics of a system.

2. Explain what is the use of modelling and simulation.

Chin Jun Wai

The use of modelling and simulation is to build a model (e.g., physical, mathematical, or logical representation of a system, entity, phenomenon, or process) as a basis for simulations to develop data utilized for managerial or technical decision making.

3. Suggest when modelling and simulation is needed in a manufacturing company.

Chin Jun Wai

When a manufacturing company is trying to choose the optimal level of an operation for machines to produce the highest number of products with the criteria of reducing the manufacturing time and cost and then they will optimize the machines accordingly based on the results provided right after using modeling and simulation on the machines.

Tutorial 13

Part I: Probability Theory

Principal Component Analysis (PCA) can be used to identify two significant components

1. Given the problems below, determine the posterior probability (P(H|E)) based on the conditional probability and prior probability given in every of the questions. Provide the conclusions based on your finding.

a. A witness sees a crime involving a taxi in Chow Kit Street. The witness says that the taxi is yellow. It is known from previous research that witnesses are correct 80% of the time when making such statements. The police also know that 85% of the taxis in Chow Kit are yellow, the other 15% being red. What is the probability that a yellow taxi was involved in the crime?

Tan Kai Yuan

$$P(\text{yellow} \mid \text{witness}) = \frac{P(\text{witness} \mid \text{yellow}) \times P(\text{yellow})}{P(\text{witness} \mid \text{yellow}) \cdot P(\text{yellow}) + P(\text{witness} \mid \text{red}) \cdot P(\text{red})}$$

$$= \frac{0.8 \times 0.85}{0.8 \times 0.85 + 0.2 \times 0.15}$$

$$= \frac{0.68}{0.68 + 0.03}$$

$$= \frac{0.68}{0.71}$$

$$= 0.9577$$

$$= 95.77\%$$

b. A rare genetic disease is discovered. Although only one in a million people carry it, you consider getting screened. You are told that the genetic test is extremely good; it is 99.99% specific (it gives a false positive result only 0.01% of the time). Having recently learned Bayes' theorem, you decided not to take the test. Why?

(From Durbin et.al. 1998."Biological Sequence Analysis", Cambridge University Press)

Tan Kai Yuan

Principal Component Analysis (PCA) can be used to identify two significant components

$P(p|\neg d) = 0.01/100$

$P(d) = 1/1{,}000{,}000$

$P(\neg d) = 1 - 1/1{,}000{,}000$

$P(p|d) = 1 - 0.01$
$= 0.99$

$$P(\text{disease}|\text{positive}) = \frac{P(\text{positive}|\text{disease}) \cdot P(\text{disease})}{P(p|d) \cdot P(d) + P(p|\neg d) \cdot P(\neg d)}$$

$\approx$ near to 0 %

Even though the test is positive, the probability of having the disease is still very low.

2. Four of the rules in a rule-based system are given as follows:

R1:	If there was drizzle last night, then there is 60% chance that the grass in the backyard is wet the next morning.
R2:	If there was rain last night, then there is 80% chance that the grass in the backyard is wet the next morning.
R3:	Drizzle occurs in 160 days in a year.
R4:	Rain occurs in 120 days in a year.

Assume that the drizzle and rain occur exclusively, and there are 365 days in a year.

(i) Specify the above rules in probability assertion.

Lim Yih Feng

P(wet|drizzle)R1 = 0.6
P(wet|rain)R2 = 0.8
P(drizzle)R3 = 160/365 = 0.4383
P(rain)R4 = 120/365 = 0.3287

(ii) If you see the grass in the backyard is wet in the morning, by using Bayes Theorem what is the event that most probably occurred last night?

Lim Yih Feng

Principal Component Analysis (PCA) can be used to identify two significant components

P(drizzle | wet) = P(wet | drizzle) * P (drizzle)/P(wet | drizzle) * P(drizzle) + P(wet | rain) * P (rain)
= 0.6 * 0.4383 / 0.6 * 0.4383 + 0.8 * 0.3287
=0.263 / 0.263 + 0.263
= 0.263 / 0.526
= 0.5

P(rain | wet) = P(wet | rain) * P (rain) /P(wet | drizzle) * P(drizzle) + P(wet | rain) * P (rain)
=0.8 * 0.3287 / 0.6 * 0.4383 + 0.8 * 0.3287
=0.263 / 0.263 + 0.263
= 0.263 / 0.526
= 0.5

Conclusion : There is an even chance that one of the events occurs, it does not have a significance difference that shows which event occurs.

(iii) After a conclusion is drawn from Q2. (ii), your grandfather asserts that sometimes your neighbour would spray water in the backyard at night.

- Do you think this assertion would affect the conditional probability?
- Do you think this assertion would affect your initial conclusion as well?

  Explain your answers.

Kow Yee Hui

- Do you think this assertion would affect the conditional probability?

  Yes. It is because the probability of these two events are the same. It shows that when your neighbour would spray water in the backyard at night then you will see the grass in the backyard is wet in the morning.

- Do you think this assertion would affect your initial conclusion as well?
  Yes. Because I think the backyard will be dry in the morning if spraying water at night.

3. Part of the rules in a liver disease diagnosis system is shown below.

Principal Component Analysis (PCA) can be used to identify two significant components

R1:	IF	patient has loss of appetite
	THEN	patient has liver disease
R2:	IF	patient is a alcoholic
	AND	patient is a smoker
	THEN	patient has liver disease

(i) Considering only the following information is given to you.

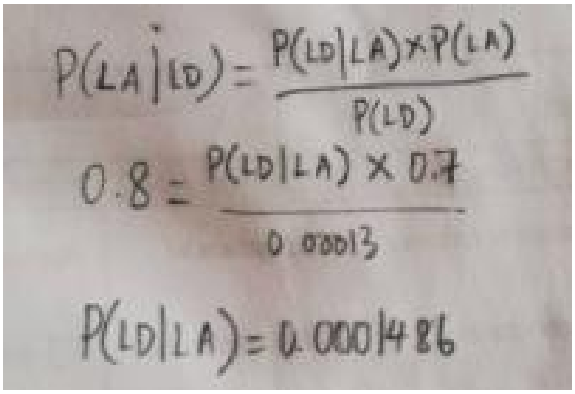
P(patient has loss appetite | patient has liver disease) = 0.8

P(patient has liver disease) = 0.00013

Assume that a patient rated that he has loss appetite 70% of the time, demonstrate how do you predict P(patient has liver disease | patient has loss appetite) using Naïve Bayes Algorithm.

Lim Jun Rong

Let P(LA|LD) = P(patient has loss appetite | patient has liver disease)
P(LD|LA) = P(patient has liver disease | patient has loss appetite)
P(LA) = P(patient has loss appetite)
P(LD) = P(patient has liver disease)



4. A police officer sees you run a red light late one Saturday night. Before pulling you over, the officer tries to guess what caused you to run the red light and comes up with the following possibilities:

a. You are drunk

Principal Component Analysis (PCA) can be used to identify two significant components

b. You are in a hurry

c. You are not wearing your glasses

d. You felt you could do it and get away with it

The policeman assigns the following evidential probabilities to these hypotheses given that the incident happened late on a Saturday night

P(run light late Sat. night | drunk driver) = 0.45

P(run light late Sat. night | driver in a hurry) = 0.60

P(run light late Sat. night | driver didn't see light) = 0.15

P(run light late Sat. night | didn't see cop/thinks can get away with it) = 0.05

He further assigns the following prior probabilities:

P(drunk driver) = 0.10

P(hurried driver) = 0.33

P(driver not wearing glasses) = 0.10

P(driver feels he can get away with things) = 0.25

Compute the conditional probabilities for the four possibilities. What decision is the police officer most likely to make?

Lai Xin Yi

Joint conditional probability

$\sum_i^4 \quad P(H_i) * P(H_i) = 0.45 * 0.10 + 0.60 * 0.33 + 0.15 * 0.10 + 0.05 * 0.25 = 0.27$

Where H_1= drunk driver

H_2= hurried driver
H_3= driver not wearing glasses
H_4= driver feels he can get away with things
E = run light late Sat. Night

Principal Component Analysis (PCA) can be used to identify two significant components

H_1:

$$P(E) = \frac{P(H_1) * P(H_1)}{\sum_i^4 \quad P(H_i) * P(H_i)} = \frac{0.45 * 0.10}{0.27} = 0.166$$

H_2:

$$P(E) = \frac{P(H_2) * P(H_2)}{\sum_i^4 \quad P(H_i) * P(H_i)} = \frac{0.6 * 0.33}{0.27} = 0.733$$

H_3:

$$P(E) = \frac{P(H_3) * P(H_3)}{\sum_i^4 \quad P(H_i) * P(H_i)} = \frac{0.15 * 0.10}{0.27} = 0.056$$

H_4:

$$P(E) = \frac{P(H_4) * P(H_4)}{\sum_i^4 \quad P(H_i) * P(H_i)} = \frac{0.05 * 0.25}{0.27} = 0.046$$

The decision most likely made by the police officer is the driver is in a hurry. Because the joint conditional probability is the highest, which is 0.733

Part II: Certainty Factor

5. You receive a phone call from your classmate, James, informing you that your name is in the pass list. Preparing to tell your parents the good news, you recall that James is a liar who loves to play jokes on others. Thus, you decide to call another classmate, Sean. Although Sean is known for being occasionally careless, you believe that Sean is far more reliable.

If James' call is true, then there is a 0.5 certainty that your name is in pass list.
If Sean's call is true and Sean is not careless is true, then there is a 0.9 certainty that your name is in the pass list.
If your name is in the pass list, then there is a 0.99 certainty that you pass the examination.
The certainty of James' call is 1.0.
The certainty of Sean's call is 1.0.
The certainty of Sean is not careless is 0.5.

Based on the information provided, would you pass the examination?

Leong Yit Wee

A = James call = 1
B = name in pass list
C = Sean call = 1
D = Sean is not careless = 0.5

If A then B = CF(r1) = CF(B) = 1*0.5 = 0.5
If C and D then B = CF(r2) = CF(B) = min(1,0.5) * 0.9 = 0.45

CF(r1,r2) = 0.5+0.45 - (0.5*0.45) = 0.725
0.99*0.725=0.71775

I believe that I have 71.775% to pass the exam.

6. Assume that a simple inference system is built using certainty factors. Demonstrate how to compute the certainty factor that a patient has liver disease based on the following rules.

R1: IF patient has loss of appetite [CF 0.7]
THEN patient has liver disease [CF 0.5]

R2: IF patient is a alcoholic [CF 0.8]
AND patient is a smoker [CF 1.0]
THEN patient has liver disease [CF 0.9]

Combining Certainty Factors:

$$CFc(CF_1, CF_2) = \begin{cases} CF_1 + CF_2(1 - CF_1) & \text{if both} > 0 \\ CF_1 + CF_2(1 + CF_1) & \text{if both} < 0 \\ CF_1 + CF_2 / (1 - \min(|CF_1|, |CF_2|)) & \text{if one} < 0 \end{cases}$$

Lee Kah Wei

Principal Component Analysis (PCA) can be used to identify two significant components

7. The Ebola outbreak has raised concerns of international spread in 2014, similar to the SARS epidemic in more than 10 years ago. Assume that you are assisting a medical centre to develop an online expert system that consults people on recognizing the symptoms of these diseases. Part of the rules of the system is as follows:

R1:	IF	body temperature exceeds 37 °C
	AND	headache presents AND muscle aches presents
	THEN	fever presents (CF 0.8)
R2:	IF	body temperature exceeds 37 °C
	THEN	fever presents (CF 0.2)
R3:	IF	body temperature exceeds 38 °C
	THEN	fever presents (CF 1.0)
R4:	IF	cough presents AND fever presents
	THEN	symptom of SARS (CF 0.5)
R5:	IF	cough presents AND cough starts 2-3 days after other symptoms
	AND	fever presents
	THEN	symptom of SARS (CF 0.6)
R6:	IF	(cough presents OR nausea presents OR stomach pain presents)
	AND	fever presents
	THEN	symptom of Ebola (CF 0.5)
R7:	IF	cough presents AND nausea presents AND stomach pain presents
	AND	fever presents
	THEN	symptom of Ebola (CF 0.8)

John has recently returned from oversea and he started to show the following symptoms. Based on R1 to R7, demonstrate the inference steps and conclude what disease that John could be infected.

Principal Component Analysis (PCA) can be used to identify two significant components

Body temperature exceeds 37 °C	CF = 1.0
Headache presents	CF = 1.0
Muscle aches presents	CF = 1.0
Body temperature exceeds 38 °C	CF = 1.0
Cough presents	CF = 1.0
Cough starts 2-3 days after other symptoms	CF = 0.8
Nausea presents	CF = 0
Stomach pain presents	CF = 0.9

Remark: To calculate a combined certainty factor, use the following equation:

$$cf(cf_1, cf_2) = \begin{cases} cf_1 + cf_2 \times (1 - cf_1) & \text{if } cf_1 > 0 \text{ and } cf_2 > 0 \\ \dfrac{cf_1 + cf_2}{1 - min\left[|cf_1|, |cf_2|\right]} & \text{if } cf_1 < 0 \text{ or } cf_2 < 0 \\ cf_1 + cf_2 \times (1 + cf_1) & \text{if } cf_1 < 0 \text{ and } cf_2 < 0 \end{cases}$$

Ong T'nsam

CF(R1) = 0.8
CF(R2) = 0.2
CF(R3) = 1.0
CF(R4) = min(1,1) *0.5 = 0.5
CF(R5) =min(1,0.8,1) * 0.6 = 0.48
CF(R6) = max(1,0,0.9,1) * 0.5 = 0.5
CF(R7) =min(1,0,0.9,1) * 0.8 = 0

Fever :
CF(R1,R2) = 0.8+0.2-(0.8*0.2) = 0.84
CF(R1,R2,R3) = 0.84 + 1.0 - (0.84*1) = 1

SARS:
CF(R4,R5) = 0.5+0.48-(0.5*0.48) = 0.74

Ebola:
CF(R6,R7)= 0.5+0-(0.5*0)=0.5

Conclusion : the disease that John could be infected is fever since CF(fever) = 1

8. Given the report of an observation at 7am as below, demonstrate the inference and conclude the weather of tomorrow based on certainty factors.

Principal Component Analysis (PCA) can be used to identify two significant components

Place of observation: Miami, Florida (tropical climate) (CF=1)
Observation data:

- Cloud height > 19,000 feet (CF = 0.8)
- Clouds are small, rounded puffs, appeared in long rows. (CF=0.8)
- Cloud patterns are sheet-like. (CF = 0.2)
- Color of cloud is white. (CF=0.7)
- Sun can shine through. (CF=0.9)
- Humidity: warm and sticky (CF=0.7)

Remark: The certain factor (CF) for other unknown data is 0.

```
IF      Cloud_Height > 18,000 feet [P1.1]
THEN    Cloud_Group is Cirrus [H1] (CF= 0.8)

IF      Cloud_Height is between 6,500 feet to 18,000 feet [P2.1]
THEN    Cloud_Group is Alto [H2] (CF=0.8)
```

```
IF      Cloud_Group is Cirrus [P3.1]
AND     Cloud_Pattern is sheet-like [P3.2]
AND     Cloud_Color is white [P3.3]
AND     Moon or Sun can shine through [P3.4]
THEN    Cloud_Type is Cirrostratus [H3] (CF=0.8)

IF      Cloud_Group is Cirrus [P4.1]
AND     Cloud_Pattern is small, rounded puffs with long rows [P4.2]
AND     Cloud_Color is white or gray [P4.3]
THEN    Cloud_Type is Cirrocumulus [H4] (CF=0.8)

IF      Cloud_Group is Alto [P5.1]
AND     Cloud_Color is gray or blue gray [P5.2]
AND     Moon or Sun can shine through [P5.3]
THEN    Cloud_Type is Altostratus [H5] (CF=0.7)

IF      Cloud_Group is Alto [P6.1]
AND     Cloud_Color is gray [P6.2]
THEN    Cloud_Type is Altocumulus [H6]  (CF=0.8)
```

Principal Component Analysis (PCA) can be used to identify two significant components

```
IF      Cloud_Type is Cirrostratus [P7.1]
THEN    Rain or snow storm occurs within 12-24 hours [H7] (CF=0.7)

IF      Cloud_Type is Cirrocumulus [P8.1]
AND     Climate_Zone is NOT of Tropical [P8.2]
THEN    Weather is fair but cold [H8] (CF=0.7)

IF      Cloud_Type is Cirrocumulus [P9.1]
AND     Climate_Zone is of Tropical [P9.2]
THEN    Hurricane is approaching [H9] (CF=0.6)

IF      Cloud_Type is Altostratus [P10.1]
THEN    Storms with continuous rain or snow is reaching [H10] (CF=0.7)

IF      Cloud_Type is Altocumulus [P11.1]
AND     Morning is warm and sticky [P11.2]
THEN    thunderstorms by late afternoon [H11] (CF=0.7)
```

Figure 3.1 Part of the rules in the expert system

(Reference: boatsafe.com)

Kong Mun Jun

CF(P1): 0.8
CF(P2): 0.8

CF(P3): 0.2x0.8 = 0.16
CF(P4): 0.8x0.8 = 0.64

CF(P7): 0.7
CF(P9): 0.8x0.6 = 0.48

Conclude : Rain or Snow Storm occurs within 12-24 hours
Hurricane is approaching

(Dont know how to do)

H1: 0.8 * 0.8 = 0.64 (CF(Cirrus) = 0.64)
H2: CF = 0 (CF(Alto) = 0)
H3 = min(0.64, 0.2, 0.7, 0.9) * 0.8 = 0.16 (CF(cirrostratus) = 0.16)
H4: min(0.64, 0.8, 0.7) * 0.8 = 0.512 (CF(cirrocumulus) = 0.512)
H5: min(0, 0, 0.9) * 0.7 = 0 (CF(altostratus) = 0)
H6: min(0, 0) * 0.8 = 0
H7: 0.16*0.7 = 0.112 (CF(rain or snow storm) = 0.112)
H8: min(0.512, -1) * 0.7 = -0.7 (CF(weather is fair) = -1)
H9: min(0.512, 1) * 0.6 = 0.3072 (CF(hurricane) = 0.3072)
H10: min(0, 0.7) * 0.7 = 0 (CF(thunderstorm) = 0)

Principal Component Analysis (PCA) can be used to identify two significant components

Conclusion: hurricane may be approaching (CF 0.3072)

Tutorial 12 (Expert System and Recommender System)

Part I: Expert Systems

1. Given a customer has a problem deciding whether or not to purchase a car. He consulted an expert for advice and the consultant suggested some rules as follows.

Rule 1:
IF The condition of car is poor
OR The price of the car is high
THEN Don't buy the car

Rule 2:
IF Mileage on the car exceeds 100,000
AND The car is city driven
AND The body of the car is bad
THEN The condition of the car is poor

Rule 3:
IF The car has dents
THEN The body of the car is bad

Rule 4:
IF The car has rust
THEN The body of the car is bad

a) Convert the given rules (Rule 1 to Rule 4) into first-order predicate logic representation.

Chin Jun Wai

Rule1: condition(car, poor) ∨ price(car, high) → ¬buy(car)

Rule2: condition(car, > 100,000 mileage) Λ condition(car, city driven) Λ condition(body, bad) → car(bad)

Rule2: mileage(car, X) Λ X>10,000 Λ drive(car, city) Λ body(car, bad) → condition(car, poor)

Rule3: condition(car, dents) → body(bad)

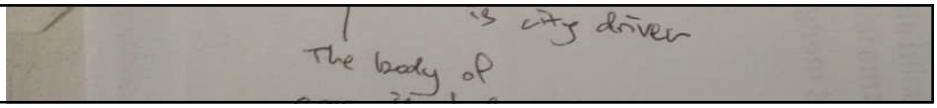
Rule3: has(car, dents) → body(car, bad)

Rule4: condition(car, rust) → body(bad)

Rule4: has(car, rust) → body(car, bad)

b) Assume the hypothesis is “Don’t buy the car”, draw an inference network (and-or graph) that supports the problem above.

Kong Mun Jun

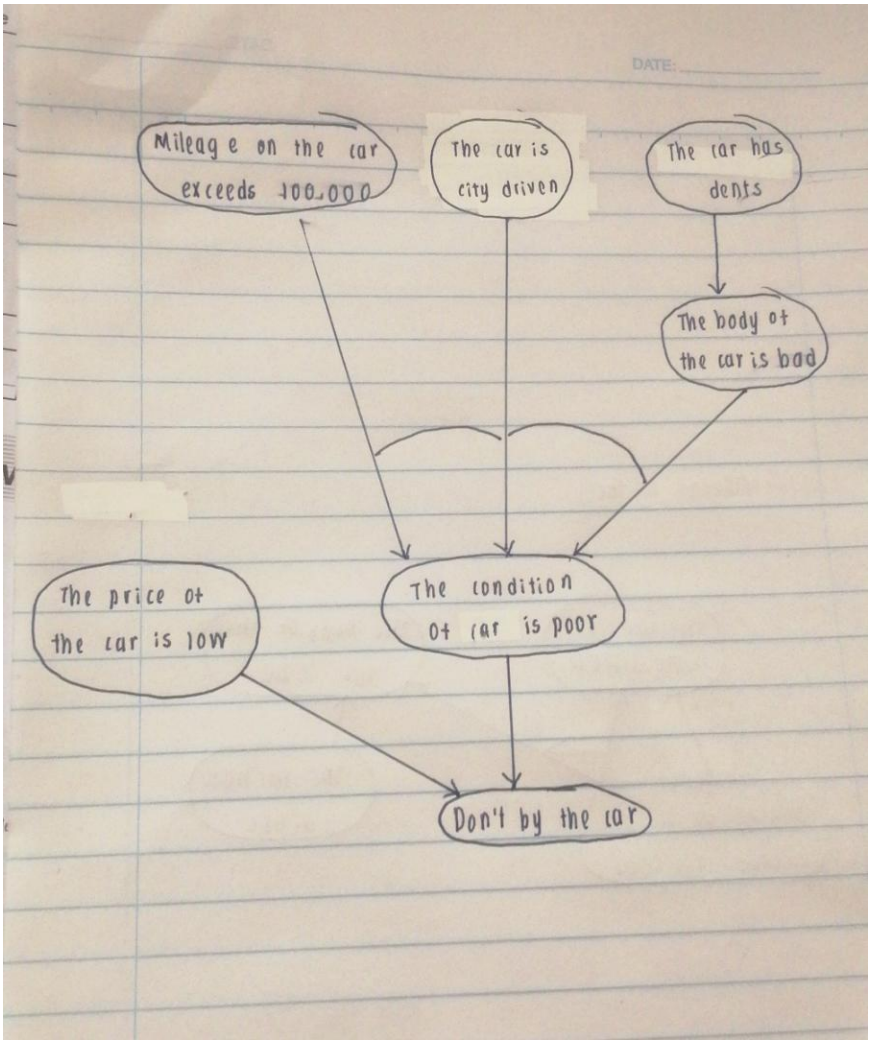


c) Given the observation by the expert as follows, demonstrate the inference process by using forward-chaining.

Mileage: 160,000 miles
Location of car: KL city
Dents on car: right rear door and back door
Rust on car: not noticed
Price of car: $1000 (remark: considered low)

Kow Yee Hui

Principal Component Analysis (PCA) can be used to identify two significant components



If the user chose Forward Chaining, then the processes involved are:
Rule 3 is fired :
E6 – The car has dents is true Then E5 – the car body is bad is true
Rule 2 is fired:
E3 – mileage > 100k is true,
E4 – car is driven in city is true,
E5 – the car body is bad is true then E1 – the car condition is poor is true
Rule 1 is fired:
E1 – Car condition is poor is true
E2 – Car price is high is false (but is OR so this rule is acceptable) then H1 – Don't buy is true

d) If you are going to develop the above expert system, which inference mechanism are you going to implement? Forward or backward chaining? Justify your answer.

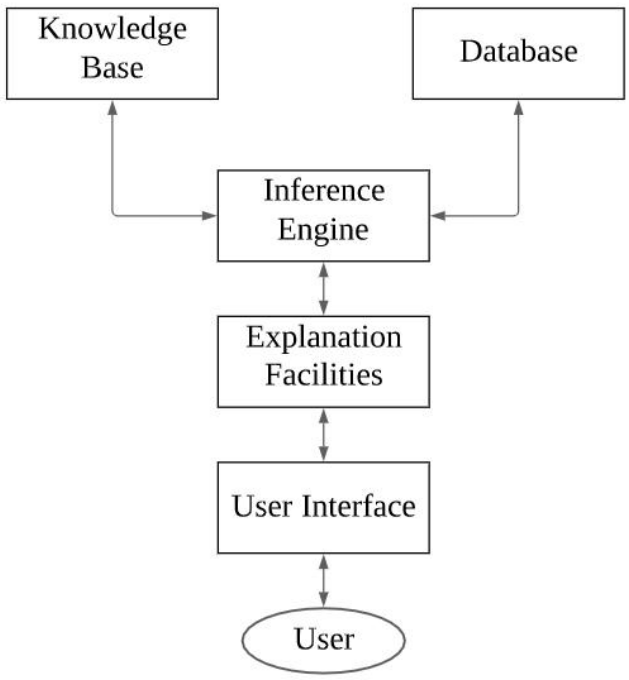
Lai Pei Xuan

Forward chaining
- It is a strategy of an expert system to answer the question, "What can happen next?"
- The inference mechanism follows the chain of conditions and derivations and finally deduces the outcome.
- It considers all the facts and rules, and sorts them before concluding to a solution.
- This strategy is followed for working on conclusion, result, or effect.
- For example given is the prediction of the car values as an effect of changes in conditions of the car.

2. A basic structure of a rule-based expert system involves the components such as knowledge base, database, inference engine, explanation facilities, user interface, and user.

a) Draw a diagram to illustrate the relations between the components in an expert system as given above, and describe the functions of each component.

Lai Xin Yi



Knowledge Base
- It holds the expert's problem-solving knowledge. It stores the knowledge elicited from the expert such as rules, facts and descriptions of objects

Database
- The set of facts will match with the rules in the knowledge base

Inference Engine
- It tries to read the input from the user and it will map with the rule that we set (knowledge base).
- It acquires and manipulates the knowledge from the knowledge base to arrive at a particular solution.
- It may use forward-chaining and backward-chaining to help in deducting the problem to find the solution. This component is also helpful for formulating conclusions.

	Explanation Facilities - It is used to explain its reasoning and justify its advice, analysis or conclusion User Interface - It provides interaction between the user and the Expert system. - This component takes the user's query in a readable form and passes it to the inference engine. User - The user presents the problem and has the conclusions presented by the user interface .

b) Define the fundamental characteristics of an expert system.

Lee Jun Xian

1. Narrow but specialized domain
2. High-quality performance
3. Speed
4. Apply heuristics to guide reasoning and reduce search space
5. Capable to explains
6. Employ symbolic reasoning when solving a problem

c) Differentiate between a conventional program and a rule-based expert system. Therefore describe the merits of developing an expert system when compared with the conventional programs.

Lee Kah Wei

d) What is the reason that an expert-system separates its knowledge base from the inference mechanism in the system?

Leong Yit Wee

Easier to maintain :
Because the Expert System enhances the quality of problem solving by adding new rules or adjusting old ones in the knowledge base. When new knowledge is acquired, changes are easy to accomplish.

Easy to understand :
Because the Expert System provides a clear separation of knowledge from its processing.

3. Given the well-form formulas (wffs) as follows:

R1:	breathes(X, gills) → lives(X, sea)
R2:	has(X, backbone) ∧ has(X, fin) ∧ breathes(X, gills) → fish(X)
R3:	lives(X, sea), has(X, backbone), has(X, fin) → Seahorse(X)
R4:	fish(X) ∧ has(X, fin) ∧ two_eyes(X, two_sides) → shark(X)
R5:	fish(X) ∧ two_eyes(X, one_side) → flounder(X)

A subject with ID "X1001" was observed and some facts were gathered as listed below:

- X1001 has backbone
- X1001 has fins
- X1001 breathes through gills
- X1001's two eyes on two sides of its head

a) From the observation above, conflicts may occur when you infer the rules above. Analyze what is the potential conflict that might occur.

Lim Chia Chung

Normally the conflict will occur when more than one rule is matched on the facts asserted. Conflict occurs since the fact R2 and R3 have a slightly similar fact. Means that "X1001" can be fish and seahorse.

b) Suggest and discuss THREE (3) approaches of conflict resolution.

Lim Jun Rong

1. First in first serve

- First in first serve is one of the simplest conflict resolution strategies in use. It involves the firing of the first rule that matches the content of the working memory or the facts asserted. It is more suitable to be used in simple applications where it will not cause any logical problems. First in first serve is a system which will always give precedence to the first classification applied to a rule.

2. Last in first serve
   - Last in first serve works on the similar principles to first in first serve. It will take the last conclusion it comes to as the final conclusion or goal. It is highly dependent on the ordering of rules and the order facts that are asserted.

3. Prioritisation
   - Prioritisation is a conflict resolution strategy that emphasises the rules that are more important. It is said that this strategy directs the inference to reason with information in the order of importance.

c) Derive a conclusion for the problem above for each of the conflict resolution approaches that you have suggested in Q3(b) by using forward chaining.

Lim Ming Jun

First in first serve
?
Last in first serve
?
Prioritisation
?

Part II: Recommender Systems

4. There are two common types of recommendation engine algorithms, i.e. collaborative filtering models and content-based models. Briefly explain the differences between them. Provide examples of applications to aid your answers.

Lim Yih Feng

collaborative filtering models	content-based models
- User's past behavior - Similar decisions made by other users	- a series of discrete - pre-tagged characteristics of an item

Principal Component Analysis (PCA) can be used to identify two significant components

- Predict items that user may interest in - Eg: Reddit, YouTube, and Last.fm are typical examples of collaborative filtering based media.	- Recommend items with similar properties - Eg :Shopee, Lazada

5. Discuss the limitations of collaborative filtering models.

Ong T'nsam

-Collaborative filtering model requires a large amount of information about a user to make accurate recommendations(cold start problem).
-Cold start : For a new user or item, there is not enough data to make accurate recommendations.

6. Discuss the limitations of content-based models.

Ong T'nsam

-Content-based filtering model is limited in scope, for example, it can only make recommendations that are similar to the original seed.
-For users with thousands of purchases, however it is impractical to base a query on all the items.
-The algorithm must use a subset or summary of the data, reducing quality.

7. There are three types of collaborative filtering models as follows. Differentiate these three types with the aid of suitable examples.

a. User-user

b. item-item

c. user-item

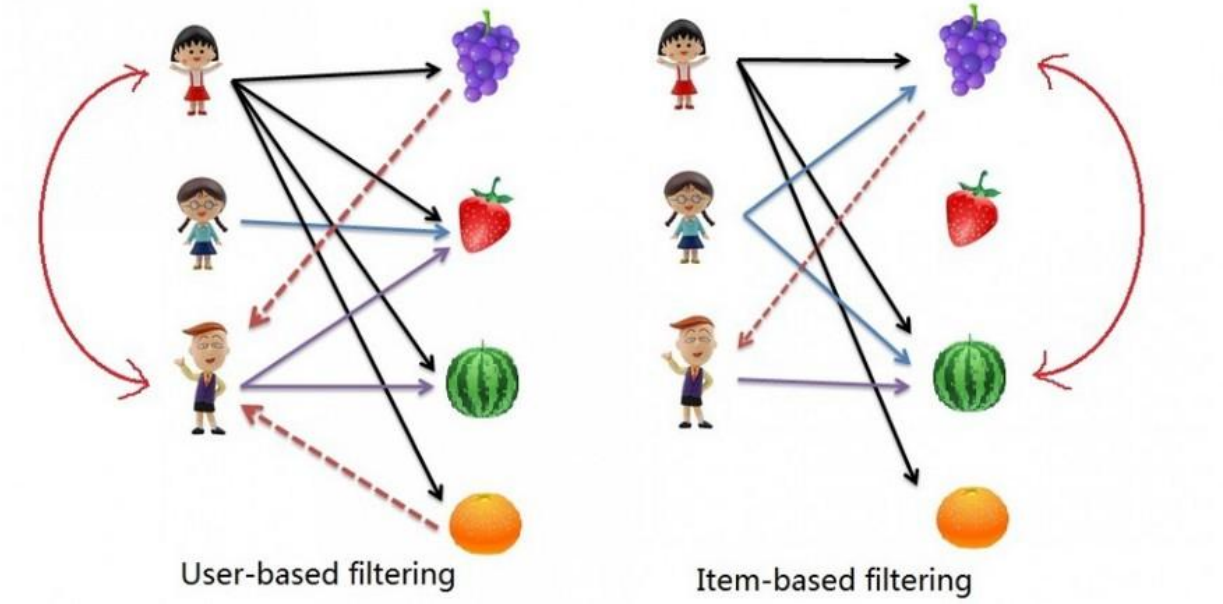
Tan Kai Yuan

User-user collaborative filtering models focus on finding a similar set of customers whose purchased and rated items overlap with the user's purchased and rated items, it is harder to scale because of the

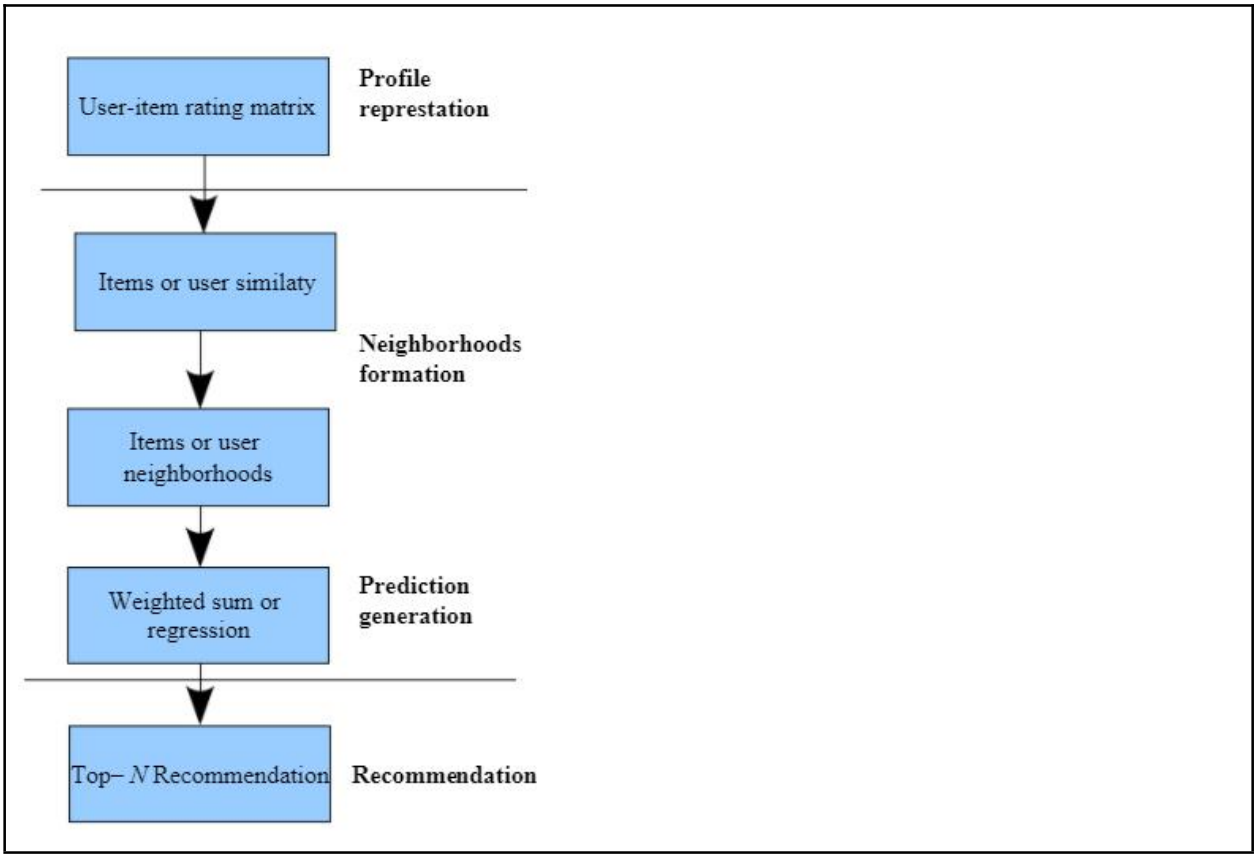
dynamic nature of users.

Item-item collaborative filtering models focus on finding similar items, not similar customers. It can be computed offline and served without constantly retraining.

User-item based CF uses a missing value algorithm to calculate the rating prediction and is not like other approaches, user-item based eliminates the neighborhood formation.



Principal Component Analysis (PCA) can be used to identify two significant components



Tutorial 11 (Image Processing)

1. Noise reduction is very important in image processing.

a. Discuss the above statement.

Chin Jun Wai

Image noise can be defined as the unwanted signal in the image.

The importance of noise reduction:

1. To recover from the Image noise that might be caused by different intrinsic (i.e., sensor) and extrinsic (i.e., environment) conditions which are often not possible to avoid in practical situations.
2. To eliminate unintended information during feature extraction.
3. To ensure the smoothness and the best performance of later preprocessing steps.

b. Identify TWO (2) types of noises in an image.

Kong Mun Jun

Poisson Noise:

The appearance of this noise is seen due to the statistical nature of electromagnetic waves such as x-rays, visible lights and gamma rays. The x-ray and gamma ray sources emitted number of photons per unit time. These rays are injected into the patient's body from its source, in medical x-rays and gamma rays imaging systems. These sources are having random fluctuations of photons. The Resulting gathered image has spatial and temporal randomness. This noise is also called as quantum (photon) noise or shot noise.

Types of Impulse Noise:

There are three types of impulse noises. Salt Noise, Pepper Noise, Salt and Pepper Noise.

Salt Noise: Salt noise is added to an image by addition of random bright (with 255 pixel value) all over the image.

Pepper Noise: Salt noise is added to an image by addition of random dark (with 0 pixel value) all over the image.

Salt and Pepper Noise: Salt and Pepper noise is added to an image by addition of both random bright (with 255 pixel value) and random dark (with 0 pixel value) all over the image.This model is also known as data drop noise because statistically it drop the original data values [5]. Source: Malfunctioning of camera's sensor cell.

https://medium.com/image-vision/noise-in-digital-image-processing-55357c9fab71#:~:text=2.1%20Types%20of%20Impulse%20Noise%3A&text=Salt%20Noise%2C%20Pepper%20Noise%2C%20Salt,value)%20all%20over%20the%20image.

c. Name a filter to perform noise reduction.

Kow Yee Hui

Gaussian Filter
A Gaussian filter is a linear filter. It's usually used to blur the image or to reduce noise. If you use two of them and subtract, you can use them for "unsharp masking" (edge detection). The

Principal Component Analysis (PCA) can be used to identify two significant components

Gaussian filter alone will blur edges and reduce contrast.
https://www.mathworks.com/matlabcentral/answers/294211-why-is-gaussian-filter-used-in-image-filtering-what-are-its-advantages-compared-to-other-filters-li

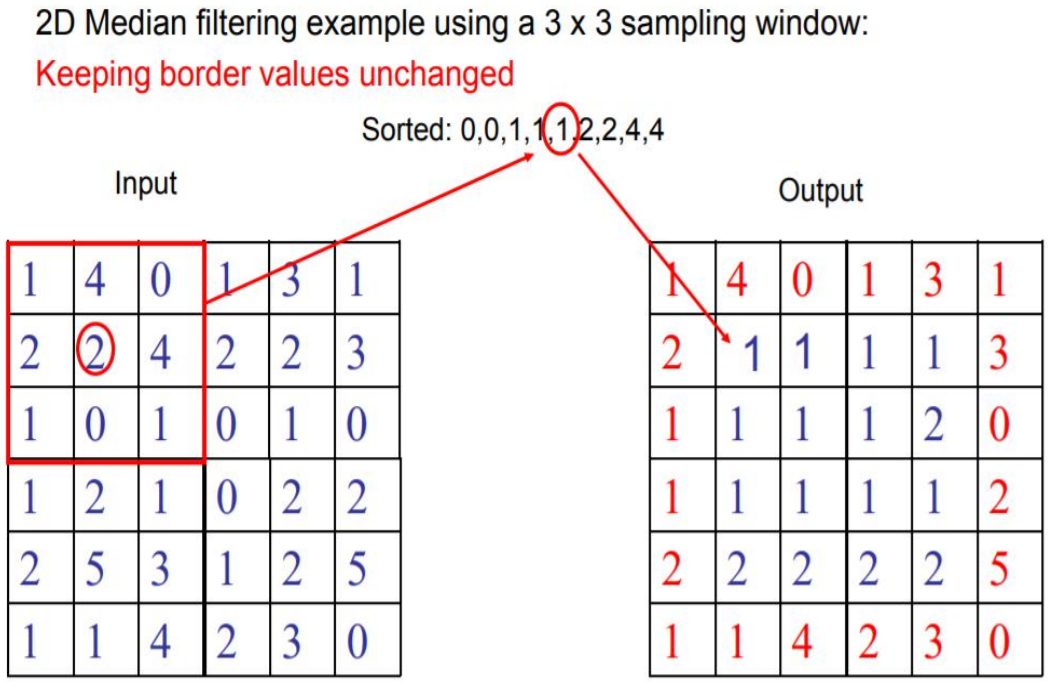
d. Demonstrate how image filtering works in reducing the noise by using median filter.

Lai Pei Xuan

- Median filtering is a nonlinear method used to remove noise from images.
- It is widely used as it is very effective at removing noise while preserving edges.
- It is particularly effective at removing 'salt and pepper' type noise.

How its works:
- The median filter works by moving through the image pixel by pixel, replacing each value with the median value of neighbouring pixels.
- The pattern of neighbours is called the "window", which slides, pixel by pixel over the entire image 2 pixel, over the entire image.
- The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.

Example:



Step1: Sorted the number in the red boxes
Step2: Find the median value from the sorted number
Step3: Replace the median value into the center of the red box value

Principal Component Analysis (PCA) can be used to identify two significant components

2. Morphological operation is important in image processing.

a. Explain FOUR (4) morphological operations.

Lai Xin Yi

1. Erosion
- Elimination of boundary pixels from objects in binary images
- Making the objects smaller, also called shrinking

2. Dilation
- Each background pixel that has a neighbor in the object is relabeled as an object pixel
- Making the objects bigger, also called growing

3. Opening
- A single erosion followed by a single dilation by the same operator

4. Closing
- A single dilation followed by a single erosion by the same operator

b. Suggest a morphological operation that could be used to recover the image as shown in figure below (the objective is to close the gap between the pixels of the number). Explain your Answer.

Lee Jun Xian

Morphological closing. By using this operation, single dilation is performed followed by single erosion by the same operator. The operation is used for filling small holes and gaps from an image while preserving the shape and size or the objects in the image. In this case, the gap in the image of the number will be filled while the size and shape is preserved.

3. Explain ONE (1) importance of morphological image analysis in image processing.

Lee Kah Wei

Principal Component Analysis (PCA) can be used to identify two significant components

Useful for removal of noise, improvement on image texture/image enhancement etc.

4. In morphological image analysis, morphological operators often probe an image with a **structuring** element. A 6x6 image is given in Figure 1 below.

	0	1	2	3	4	5
0	0	0	0	0	0	1
1	0	1	1	1	1	1
2	0	1	1	1	1	0
3	0	1	1	1	1	0
4	0	1	1	1	0	0
5	1	1	1	0	0	0

Figure 1. 6x6 binary image

(i) Based on the 3x3 structuring element with the origin in the middle as shown in Figure 2, perform an **erosion** effect on the 6x6 binary image as shown in Figure 1. Provide the resulting binary image.

1	1	1
1	1	1
1	1	1

Figure 2. 3x3 square structuring element

Leong Yit Wee

Erosion effect

0	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**
0	**0**	**1**	**1**	**0**	**0**
0	**0**	**1**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**
0	**0**	**0**	**0**	**0**	**0**

Principal Component Analysis (PCA) can be used to identify two significant components

(ii) Based on the 2x2 structuring element with the origin in the middle as shown in Figure 3, perform a dilation effect on the 6x6 binary image as shown in Figure 1. Provide the resulting binary image.

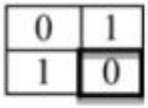


Figure 3. 2x2 square structuring element

Lim Chia Chung

0	0	0	0	0	0
0	0	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	1
0	1	1	1	1	0
01	1	1	1	0	0

5. You intend to create an intrusion detection system (the system will fire alarm if an unidentified person cross the boundary) by using computer vision technology. Describe the steps that could be useful in the implementation.

Lim Jun Rong

1. Video acquisition
   - The system will first get the video of the surroundings to make sure that there is anyone within the specified region.

2. Video slicing
   - It will then slice the video into a large number of images and make analysis on each of

the images.

3. Pre-processing
   - The system will then perform pre-processing process on the images such as background subtraction or image enhancement. The purpose of doing this is to remove the unwanted noises and to only focus on a human. The images of the human will also be enhanced so that it looks clearer and the system is able to recognise the face.

4. Feature extraction
   - After that, the system will then detect if there is any face and extract the important features such as forehead, eyes, nose, mouth and cheek from the face obtained. This is the most important step in computer vision as the system is not able to detect or recognise anyone if the feature extraction is done properly.

5. Perform machine learning
   - It will then undergo a series of training so that the system knows which part of the face is eye, nose, mouth and so on.

6. Perform classification, recognition and interpretation
   - Lastly, when a face is detected, the system will check the features that have been extracted previously and match it with the one in the database of the system. If none of the features are matched, it means that the individual does not belong to the organisation and he/she is an intruder. The system will then fire the alarm so that the people, especially the security guards are aware of the intrusion.

6. Preprocessing is one of the important steps in digital image analysis.

(i) Explain the importance of image preprocessing.

Lim Ming Jun

1. To improve the image visualization
2. Images can be given more sharpness and better visual appearance
3. To eliminate unintended information during feature extraction
4. To ensure the smoothness and the best performance of the later processing steps
5. To ensure the quality of the outputted image after feature extractions.

(ii) Identify the suitable preprocessing operations based on the images in Figure 4.

Provide brief description for each of the suggested operations.

Principal Component Analysis (PCA) can be used to identify two significant components

Image A (objects are less distinctive from its background)	**Image B** (salt and pepper in the image)	**Image C** (broken alphabet)
		

Figure 4. Sample images.

Lim Yih Feng (Image A)

Enhancing Contrast
Because the contrast color of the object is nearly the same as the background because the object color is too shallow, with the image contrast enhancement, the object contrast can be darker in order to show the object from the background .

Ong T'nsam (Image B)

Median Filtering
This is because salt and pepper noise contains random occurrences of black and white pixels, so median filtering can be used as it operates over a window by selecting the median intensity in the window.

Tan Kai Yuan (Image C)

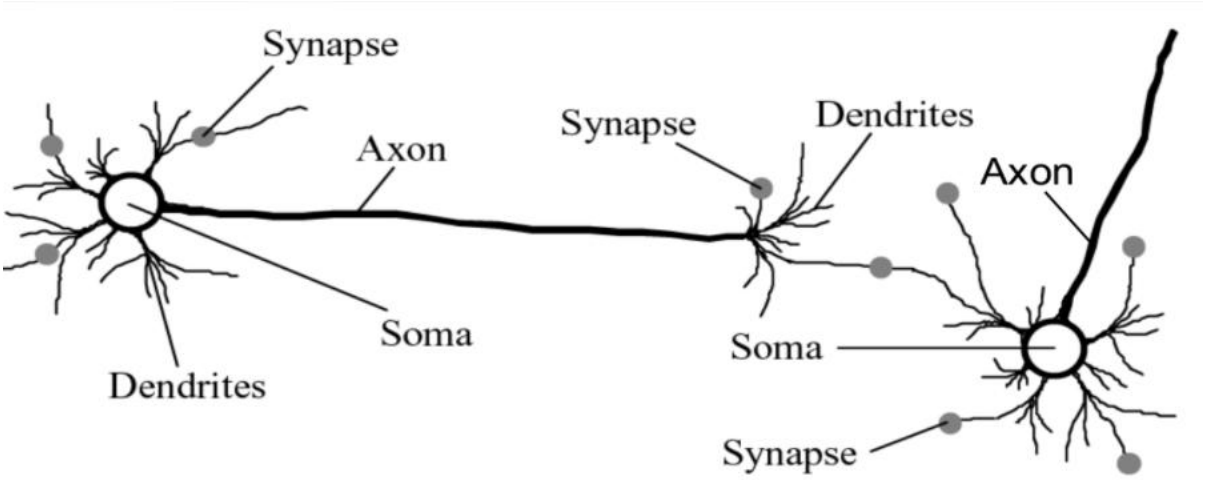
Morphological operation
Because it fills in small hole and eliminating small protrusions from their boundaries, dilation is good in remove the hole and bridging the gaps while erosion is good in removing irrelevant details

Tutorial 10 (Artificial Neural Network)

Q1. (a) Research in Artificial Neural Networks (ANNs) started with a view of mimicking the functioning of Biological Neural Networks (BNNs). Illustrate with diagrams, how the BNNs and ANNs are similar to each other.

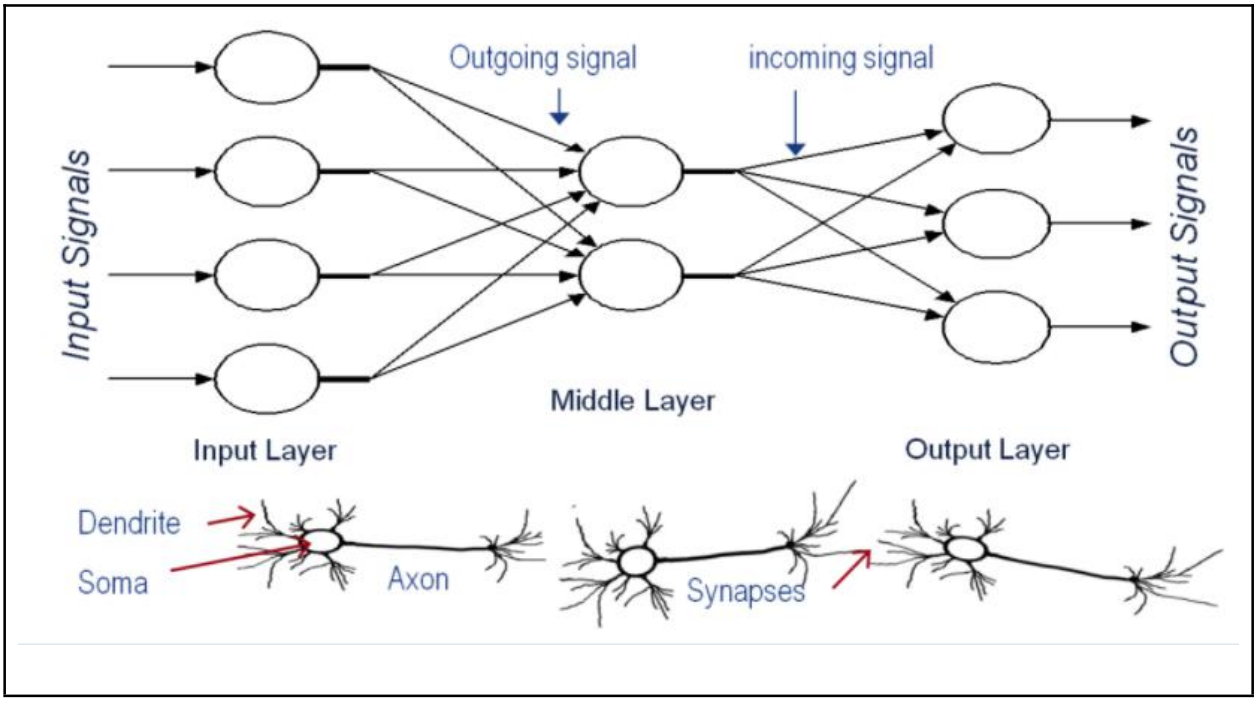
Chin Jun Wai (Diagram)

Biological Neural Networks (BNNs)



Artificial Neural Networks (ANNs)

Principal Component Analysis (PCA) can be used to identify two significant components



Kong Mun Jun (Explanation)

Biological Neuron	Artificial Neuron
Dendrites	Input
Cell Nucleus(Soma)	Node
Axon	Output
Synapse	Interconnections

https://towardsdatascience.com/what-is-the-relation-between-artificial-and-biological-neuron-18b05831036 Node aka Neuron, Interconnections aka Weights

Dendrite Dendrites are responsible for getting incoming signals from outside	**Input values or One input layer** We pass input values to a neuron using this layer. It might be something as simple as a collection of array values. It is similar to a dendrite in biological neurons.
Soma Soma is the cell body responsible for the processing of input signals and deciding whether a neuron should fire an output signal	**Activation Function(node)** Activation Function decides whether or not a neuron is fired. It decides which of the two output values should be generated by the neuron
Axon Axon is responsible for getting processed signals from neuron to relevant cells	**Output Layer** Output layer gives the final output of a neuron which can then be passed to other neurons in the network or taken as the final output value.
Synapse Synapse is the connection between an axon and other neuron dendrites	**Weights and Bias (interconnections)** Weights are a collection of array values which are multiplied to the respective input values. We then take a sum of all these multiplied values which is called a weighted sum. Next, we add a bias value to the weighted sum to get a final value for prediction by our neuron

(b) Figure 1 below shows the simplest neural network model that is known as perceptron, which can be used for classification tasks. Briefly explain the functions of each of the layers A, B and C as shown in the diagram.

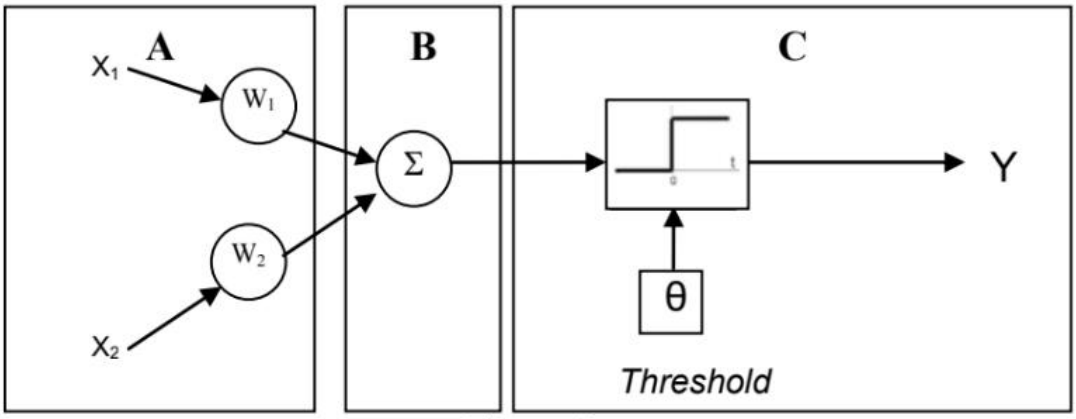


Figure 1

Principal Component Analysis (PCA) can be used to identify two significant components

Kow Yee Hui

- Layer A : Input Layer
  The Input Layer in Perceptron will take the inputs from the outside world to the network, multiply them by weight and also computes why they are important.
  https://missinglink.ai/guides/neural-network-concepts/perceptrons-and-multi-layer-perceptrons-the-artificial-neuron-at-the-core-of-deep-learning/
  Received data (X_1 and X_2)
  W_1 and W_2 used to express strength or importance of neuron input

Lai Pei Xuan

- Layer B : Weighted Sum *aka Linear Combiner* $\Sigma x,w$
  The input values are multiplied by the weights and summed up, to create one aggregate value which is fed into activation function.

Lai Xin Yi

- Layer C: Activation function/Output layer
- The result of the neuron's calculation is turned into an output signal, in this example two classes based on the threshold θ. The perceptron output is a classification decision. or to pass the output to other neuron For example, the iris is classified as "setosa" or "Versicolor"

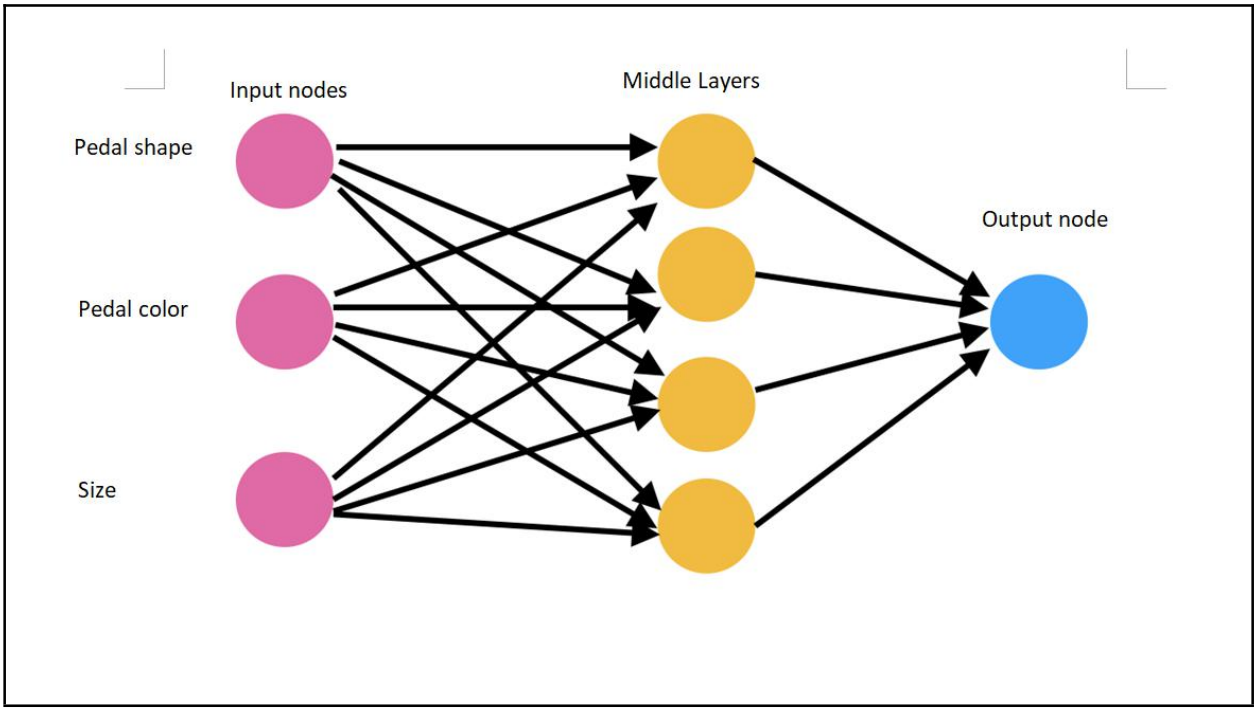
(c) Back propagation Multilayer Artificial Neural Network can be applied into a system that is able to recognize and interpret two different types of flowers. Discuss the prerequisite for the Neural Network system and design ONE (1) Neural Network structure for the purpose. State the advantage of back propagation network.

Lee Chun Xian (Explain)

The prerequisite for the Neural Network system refers to systems of neurons either organic or artificial in nature. It can adapt to change input so that the network generates the best possible result without the needs of redesign output criteria. Prepare training data for the flowers with labels for training the neural network. What preparation?
-Petals Colour, Flower Appearance and Petal Size.

Lee Jun Xian (Neural Network Structure)

Principal Component Analysis (PCA) can be used to identify two significant components



Petal not Pedal. Petal shape, color, size as input. Middle layer as hidden layer. Single output as binary output (two different types of flowers). Explore further the parameters/architecture of back propagation. How are the weights adjusted? What happens to the errors in the output?

Lee Kah Wei (Advantage of back propagation network)

Back propagation neural network is fast, simple and easy to program.fast? For how many data instances? Simple and easy to program are not advantages related to back propagation in relation to other ML algorithms. How about when we compare between feed forward and back propagation ANN?
It has no parameters to tune apart from the number of input. How about the architecture itself? How many hidden layers?
It is a flexible method as it does not require prior knowledge about the network.
It is a standard method that generally works well

Source: https://www.guru99.com/backpropogation-neural-network.html

Q2. Assume that the Sea Fisheries Department has approached you to analyze abalone from physical measurements. A set of data that contains 9 attributes and 4177 rows of records is given. Part of the data are shown in Table 1 below:

Principal Component Analysis (PCA) can be used to identify two significant components

Table 1: Partial Abalone Data

Sex	Length (mm)	Diameter (mm)	Height (mm)	Whole weight (g)	Shell weight (g)	No. of rings
F	0.59	0.455	0.145	1.063	0.25	8
F	0.595	0.47	0.155	1.121	0.155	11
F	0.595	0.45	0.15	1.114	0.25	11

Demonstrate how a multi-layer feedforward back-propagation neural network can be constructed to predict the number of rings of abalone. Illustrate the architecture of the neural network with appropriate labels of each layer in the diagram. In your diagram, also include all the attributes and output as shown in Table 1.

Leong Yit Wee

First, a training input pattern is presented to the network input layer. The network propagates the input pattern from layer to layer until the output pattern is generated by the output layer.

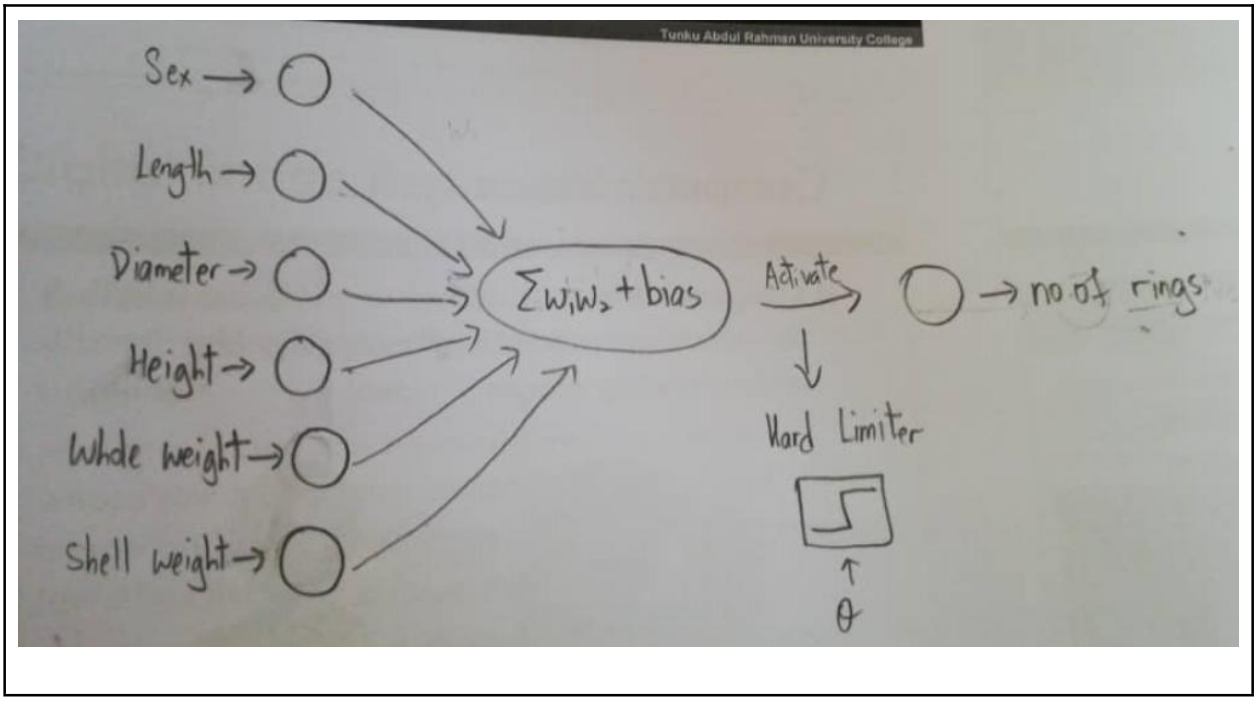
If this pattern is different from the desired output, an error is calculated and then propagated backwards through the network from the output layer to the input layer. The weights are modified as the error is propagated.

Lim Chia Chung (Diagram)

Draw an ANN, label the input and output.

Lim Jun Rong (Diagram)

Principal Component Analysis (PCA) can be used to identify two significant components



Input and output Fix the architecture in the middle of the diagram, refer to ANN diagram, should at least have another layer before output.

Q3. A combinational logic circuit are made up from basic logic AND, OR, NAND, NOR or NOT gates that are "combined" or connected together to produce more complicated switching circuit. The output of the gate will be determined by a combination of 1's and 0's which are present at the input terminals. The truth table of 3-input combinational logic circuit is given by Table 2 and Figure 2 where A, B and C are the inputs and Q is the output.

A	B	C	Q
0	0	0	1
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	0
1	1	1	1

Table 2

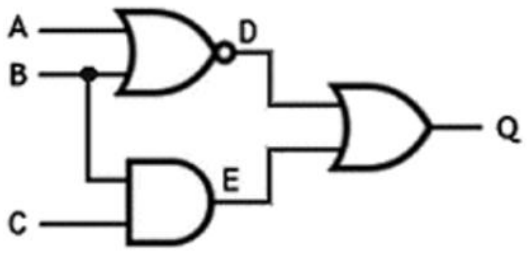


Figure 2

Principal Component Analysis (PCA) can be used to identify two significant components

Train the perceptron to perform 3-input combinational logic circuit operation. The formula of output Q is given by

$$Q = w_1A + w_2B + w_3C$$

where w_1, w_2 and w_3 are weight for A, B and C respectively. Assume that the threshold $\theta = 0.25$, learning rate $\alpha = 0.1$ and **step function** is used for activation function. **Calculate the final weights of w_1, w_2 and w_3 for epoch 1 only**. (Assume the initial weights $w_1 = 0.3$, $w_2 = 0.2$ and $w_3 = 0.1$)

Lim Ming Jun, Lim Yih Feng

	A	B	C	Q_D	w_1	w_2	w_3	Q_P	e	w_1	w_2	w_3
1	0	0	0	1	0.3	0.2	0.1	0	1	0.3	0.2	0.1
	0	0	1	1	0.3	0.2	0.1	0	1	0.3	0.2	0.2
	0	1	0	0	0.3	0.2	0.2	0	0	0.3	0.2	0.2
	0	1	1	1	0.3	0.2	0.2	1	0	0.3	0.2	0.2
	1	0	0	0	0.3	0.2	0.2	1	-1	0.2	0.2	0.2
	1	0	1	0	0.2	0.2	0.2	1	-1	0.1	0.2	0.1
	1	1	0	0	0.1	0.2	0.1	1	-1	0	0.1	0.1
	1	1	1	1	0	0.1	0.1	0	1	0.1	0.2	0.2

Fill up the rest of the slots.

Q4. In a supervised learning session, a 2-input neuron and the respective truth table are given in Figure 3 and Table , where x1 and x2 are the inputs and y is the output.

Principal Component Analysis (PCA) can be used to identify two significant components

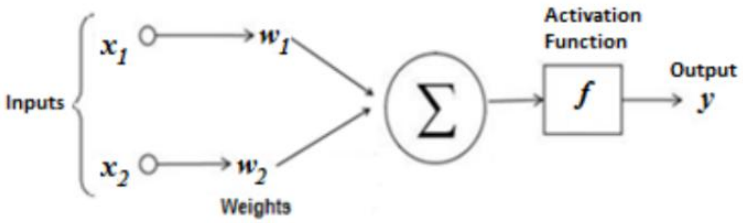


Figure 3: The 2-input neuron

x_1	x_2	*Desired Ouput,* y_d
0	0	-1
0	1	1
1	0	1
1	1	1

Table 3

The formula of output y is given by

$$y = sign(n) = \begin{cases} 1 \; if \; n \geq \theta \\ -1 \; if \; n < \theta \end{cases}$$

where $\boldsymbol{n = w_1x_1 + w_2x_2}$, and w_1 and w_2 are the weights for x_1 *and* x_2 respectively. The step function is used for activation function. The weights, $w_1 = 0.1$ and $w_2 = 0.2$, the threshold, $\theta = 0.2$ and the learning rate, $\alpha = 0.1$ are set initially for the training. Train the neuron and demonstrate the final weights of w_1 and w_2 **for epoch 1 only**.

Ong T'nsam , Tan Kai Yuan

epoch	inputs		Desired outputs Yd	Initial weights		Predicted output Y	error	Final weights	
	x1	x2		w1	w2			w1	w2
1	0	0	-1	0.1	0.2	-1	0	0.1	0.2
	0	1	1	0.1	0.2	1	0	0.1	0.2
	1	0	1	0.1	0.2	-1	2	0.3	0.2
	1	1	1	0.3	0.2	1	0	0.3	0.2

Principal Component Analysis (PCA) can be used to identify two significant components

Tutorial 9 (Machine Learning – Unsupervised Learning)

1. Draw a diagram to depict unsupervised learning model.

Chin Jun Wai

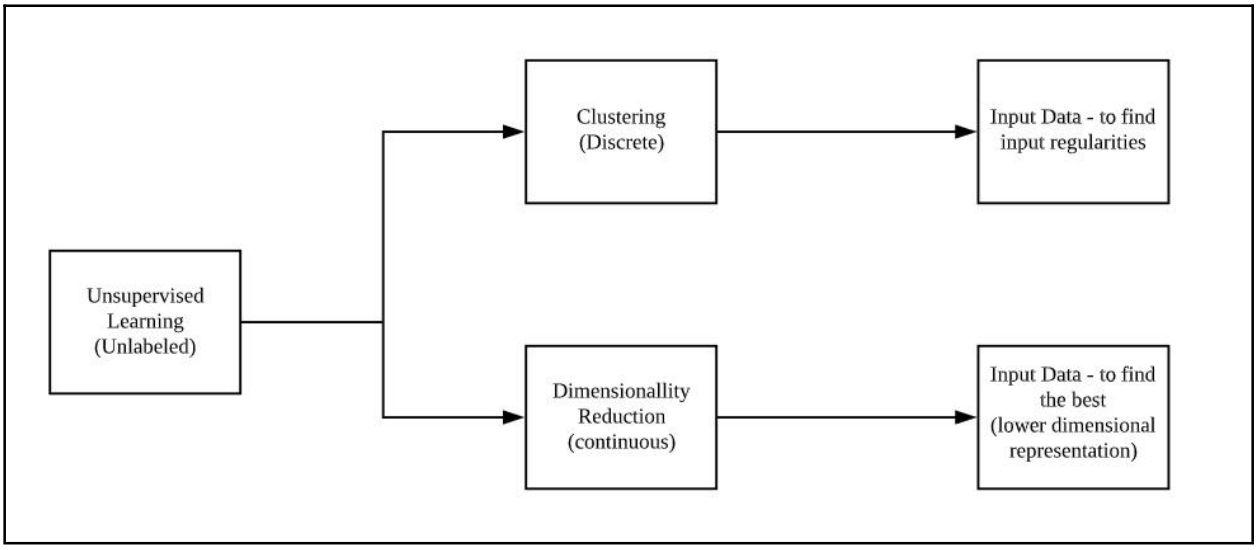


diagram from slide 3 of notes.

2. Considering a partial dataset as shown in Table 1:

Age	Education	Gender	Income
31	Doctoral	Male	150k
45	Bachelor's	Female	78k

26	Master's	Male	76k
…	…	…	…

a. Explain how a regression model is different from clustering.

Kong Mun Jun (ABSENT), Kow Yee Hui

The regression analysis is the statistical model which is used to predict the numeric data instead of labels. It also can identify the distribution trends by using the available data or historic data.
For example, predict a person's income from their age and education.

Clustering is the statistical model which used to partition the dataset into groups. It used to determine the grouping among unlabeled data.

b. Explain how a regression model and clustering can be applied on the dataset above.

Regression:
1. Use numeric data such as age to predict the income of a person.

Clustering:
1. Use unlabeled data such as gender and education level to determine the grouping of income.

3.Clustering is the grouping of unlabeled data, it can be used for knowledge discovery. Discuss about this and provide TWO (2) real world applications of clustering.

Lai Pei Xuan

Clustering is an unsupervised machine learning task that automatically divides the data into **clusters,** or groups of similar items. It does this without having been told how the groups should look ahead of time. As we may not even know what we're looking for, clustering is used for knowledge discovery rather than prediction. It provides an insight into the natural groupings found within data.

Clustering is guided by the principle that items inside a cluster should be very similar to each other, but

very different from those outside. The definition of similarity might vary across applications, but the basic idea is always the same—group the data so that the related elements are placed together.

Example: Business and Marketing (K-means Clustering)

- Segmenting customers into groups with similar demographics or buying patterns for targeted marketing campaigns

Link: https://hub.packtpub.com/introduction-clustering-and-unsupervised-learning/

Lai Xin Yi

Clustering is the search for distinct groups in the feature space. It is expected that these groups have different structures and that can be clearly differentiated. The clustering task separates the data into a number of partitions, which are volumes in the n-dimensional feature space. Clustering is a classification technique. (Given a vector of N measurements describing each pixel or group of pixels (i.e., region) in an image, a similarity of the measurement vectors and therefore their clustering in the N-dimensional measurement space implies similarity of the corresponding pixels or pixel groups. Therefore, clustering in measurement space may be an indicator of similarity of image regions, and may be used for segmentation purposes. Image segmentation example.)

Real world application of clustering

Image Segmentation (K- means clustering)

- Image segmentation is the process of partitioning a digital image into multiple distinct regions containing each pixel(sets of pixels, also known as superpixels) with similar attributes.
- By dividing the image into a few small segments, the system will process the image with the important segment instead of processing the entire image.
- A common technique is to look for discontinuities in pixel values, which typically indicate edges that define a region.
- Another common technique is to detect the similarities in the region of an image.
- The goal of Image segmentation is to change the representation of an image into something that is more meaningful and easier to analyze.

Explanation of K- means clustering

K-means clustering algorithm is an unsupervised algorithm and it is used to segment the interest area from the background. It clusters the given data into k-clusters or parts based on the k-centroids.The algorithm is used when you have unlabeled data(i.e. data without defined categories or groups). The goal is to find certain groups based on some kind of similarity in the data with the number of groups represented by K.

https://towardsdatascience.com/introduction-to-image-segmentation-with-k-means-clustering-83fd0a9e2fc3

Principal Component Analysis (PCA) can be used to identify two significant components

4.K-means is used to cluster the data in Table. 1. Randomly we choose two cluster centers (k=2) for two clusters; C1 = (1.0, 1.0) and C2 = (5.0, 7.0).

Table 1. Sample of data

Individual	Variable 1	Variable 2
1	1.0	1.0
2	1.5	2.0
3	3.0	4.0
4	5.0	7.0
5	3.5	5.0
6	4.5	5.0
7	3.5	4.5

a. List the clusters that each data will be assigned to.

Lee Chun Xian

Individual	Center 1	Center 2
1	0	7.21
2	1.118	6.103
3	3.606	3.606
4	7.211	0.000
5	4.717	2.500
6	5.315	2.062
7	4.301	2.915

C1=(√(1-2)^2+(1-2)^2)=1.4142
C2=(√(1-4)^2+(5-6)^2)=3.1628

Cluster 1 = {1,2,3}
Cluster 2 = {4,5,6,7}

slide 22

b. Identify the new centers for the first iteration.

Lee Jun Xian

C1 =(1/3(1.0 + 1.5+3.0), 1/3(1.0+2.0+4.0)) = (1.83,2.33)
C2 (1/4(3.0+5.0+3.5+4.5+3.5),⅛ (4.0+7.0+5.0+5.0+4.5)) = (3.9, 5.1)

slide 22

c. Suggest TWO (2) stopping conditions for K-means.

Lee Kah Wei

Centroids of newly formed clusters do not change

Leong Yit Wee

Maximum number of iterations are reached

d. Discuss on the limitation of K-means.

Lim Chia Chung

- Need to specify the number of clusters, but we do not know how many there really are.
- Results can change depending on the location of the initial centroids.
- K-means is not recommended if you have a lot of categorical variables.
- Assumes that clusters are spherical, distinct and approximately equal in sizes.

5.Explain the E-Step (Expectation) and M-Step (Maximization) in Gaussian Mixture Model and state the differences of E & M Step between GMM and K-Means.

Lim Jun Rong

- E-Step is performed to estimate the expected value for each latent variable. The probability distribution of the latent or hidden variables are often computed before it can be used for expectations.

- M-Step is used to optimise the parameters of the distribution using maximum likelihood. The probability distribution from the E-Step will be tweaked to include the new data.

6.Mean Shift is one of the popular clustering algorithms:

a. Demonstrate the algorithm for Mean Shift

Lim Ming Jun

Step 1: First, start with the data points assigned to a cluster of their own.

Step 2: Next, this algorithm will compute the centroids.

Step 3: In this step, location of new centroids will be updated.

Step 4: Now, the process will be iterated and moved to the higher density region.

Step 5: At last, it will be stopped once the centroids reach a position from where it cannot move further.
https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_clustering_algorithms_mean_shift.htm#:~:text=Working%20of%20Mean%2DShift%20Algorithm&text=Step%201%20%E2%88%92%20First%2C%20start%20with,to%20the%20higher%20density%20region.

b. Discuss the advantages and disadvantages of Mean Shift.

Lim Yih Feng

Advantages
- Does not assume number of clusters
- Just a single parameter(Window size)
- Finds variable number of modes
- Robust to outliers

Principal Component Analysis (PCA) can be used to identify two significant components

DIsadvantages
- Output depends on window size
- Computationally expensive

7.Explain why cluster evaluation is a hard problem. Identify TWO (2) methods based on internal information that we can use to evaluate a clustering algorithm.

Ong T'nsam

-Cluster evaluation is a hard problem because we do not know the correct clusters.
-One of the methods is inter-cluster separation(isolation), separation means that different cluster centroids should be far away from one another.

Tan Kai Yuan

-Clustering is difficult because we are just given dataset and there isn't necessarily a "correct" or ground truth solution that we can refer to
-Intra-cluster cohesion (compactness), Cohesion measures how near the data points in a cluster are to the cluster centroid. Sum of squared error (SSE) is a commonly used measure.

Tutorial 8

1. Differentiate between artificial intelligence (AI) and machine learning (ML).

CHIN JUN WAI

Artificial intelligence is a technology using which we can create intelligent systems that can simulate human intelligence. Artificial intelligence systems do not require to be pre-programmed and they use algorithms such as Reinforcement learning algorithms and deep learning neural networks which can work with their own intelligence.

Machine learning is about extracting knowledge from the data. Furthermore, it enables a computer system to make predictions or take some decisions using historical data without being explicitly programmed.

Machine Learning can be divided to three types:
- Supervised Learning

Principal Component Analysis (PCA) can be used to identify two significant components

- Unsupervised Learning
- Reinforcement Learning

2. Discuss the differences between parametric and non-parametric algorithms. Provide examples for each one.

CHONG JIA LOONG

Algorithms	Parametric	Non-parametric
Number of parameters	Fixed	Flexible
Computation	Faster, but makes stronger assumptions about the data. Mayonly work well if the assumptions turn out to be correct	Slower, but makes fewer assumptions about the data
Example	Linear regression	K-nearest neighbour

3. There are two main types of Machine Learning which are Supervised and Unsupervised Learning.
   - Discuss the differences between them.
   - Provide TWO (2) algorithms for each type Machine Learning.

KONG MUN JUN

Supervised learning, you train the machine using data which is well **"labeled**." It means some data is already tagged with the correct answer. It can be compared to learning which takes place in the presence of a supervisor or a teacher.

- Linear regression for regression problems.
- Random forest for classification and regression problems.
- Support vector machines for classification problems.

Unsupervised learning is a machine learning technique, where you do not need to supervise the model. Instead, you need to allow the model to work on its own to discover information. It mainly deals with the unlabelled data.

- k-means for clustering problems.
- Apriori algorithm for association rule learning problems.

https://www.guru99.com/supervised-vs-unsupervised-learning.html#2

https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/

Principal Component Analysis (PCA) can be used to identify two significant components

4. Assume that the Sea Fisheries Department has approached you to analyze abalone from physical measurements. A set of data that contains 9 attributes and 4177 rows of records is given. Part of the data are shown in Table 1 below:

Table 1: Partial Abalone Data

Sex	Length (mm)	Diameter (mm)	Height (mm)	Whole weight (g)	Shell weight (g)	No. of rings
F	0.59	0.455	0.145	1.063	0.25	8
F	0.595	0.47	0.155	1.121	0.155	11
F	0.595	0.45	0.15	1.114	0.25	11
F	0.595	0.46	0.14	1.0045	0.2515	9
F	0.605	0.49	0.15	1.1345	0.295	9
M	0.595	0.475	0.165	1.213	0.274	9
M	0.595	0.455	0.15	1.044	0.27	9
M	0.605	0.475	0.155	1.161	0.275	9
M	0.605	0.47	0.165	1.2315	0.2925	11
M	0.61	0.47	0.15	1.1625	0.3085	11

Data Source: http://mlr.cs.umass.edu/ml/datasets/Abalone

Reference: Asuncion, A. & Newman, D.J. (2007). UCI Machine Learning Repository [http://www.ics.uci.edu/~mlearn/MLRepository.html]. Irvine, CA: University of California, School of Information and Computer Science.

Between supervised learning or unsupervised learning, which type of machine learning is suitable to perform the following data analytics? Justify your answer. Then suggest **ONE (1)** suitable machine learning algorithm for each case.

a) To determine the association between sex and diameter that may affect whole weight.

KOW YEE HUI

Unsupervised Learning - Gaussian mixture model

- Gaussian mixture model is a probabilistic model for representing normal distributed subpopulations within an overall population.

b) To determine 3 different grades of abalone based on diameter and whole weight.

LAI PEI XUAN

Unsupervised Learning - K-means

- K-means method is given preference for its fast, efficient, and accurate features. The method is used to establish the grading standards for the abalone

c) To predict the age of an abalone.
(To determine the age of an abalone, we can check the number of rings, which is between 1 to 29)

LAI XIN YI

Supervised learning

Principal Component Analysis (PCA) can be used to identify two significant components

- K- Nearest Neighbour (KNN)
- because it can be classify the output in category / nominal (discrete label)
- it can use the range of the age to classify

d) To identify two principal factors (most significant attributes) that can be used so that you can produce a visualisation that predicts age.

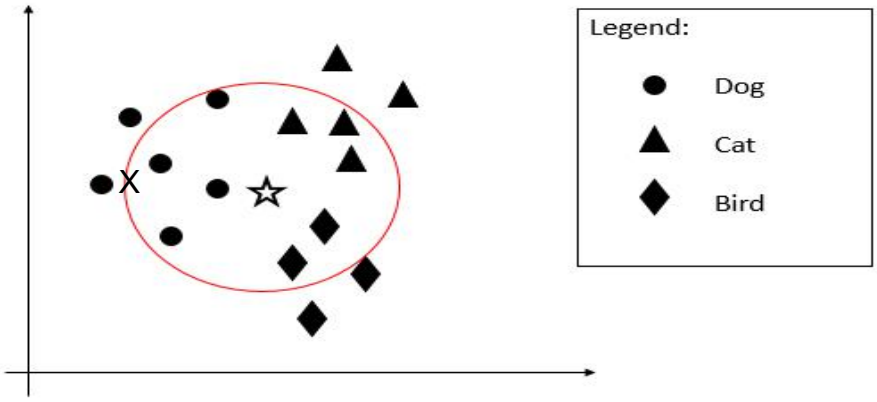
LEE CHUN XIAN

Diameter and number of rings of abalone.

1. Combine the information given
2. Reduce the unnecessary data is not related to predict age.

Principal Component Analysis (PCA) can be used to identify two significant components

5. K-Nearest Neighbors is used to classify data X into its respective class (Dog, Cat or Bird). Remark: K = 9.



a) Compute P(Dog | X), P(Cat | X), and P(Bird | X).
b) Identify which class that X belongs to with explanation.

LEE JUN XIAN

P(Dog| X) = 4/9 = 0.44
P(Cat | X) = 3/9 = 0.33
P(Bird | X) = 2/9 = 0.22

X belongs to class Dog because most of its nearest neighbour is class Dog.

Principal Component Analysis (PCA) can be used to identify two significant components

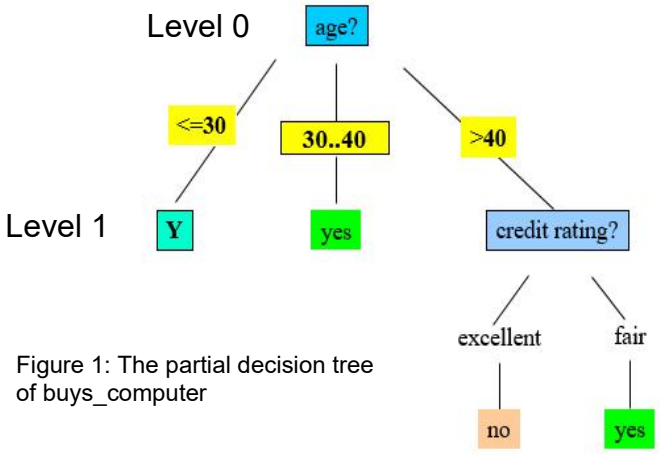


Figure 1: The partial decision tree of buys_computer

Table 2: The decision table of buys_computer

age	income	student	credit_rating	buys_computer
<=30	high	no	fair	no
<=30	high	no	excellent	no
31…40	high	no	fair	yes
>40	medium	no	fair	yes
>40	low	yes	fair	yes
>40	low	yes	excellent	no
31…40	low	yes	excellent	yes
<=30	medium	no	fair	no
<=30	low	yes	fair	yes
>40	medium	yes	fair	yes
<=30	medium	yes	excellent	yes
31…40	medium	no	excellent	yes
31…40	high	yes	fair	yes
>40	medium	no	excellent	no

6. Given the table and the corresponding decision tree as above.
   a) Determine the appropriate attribute Y (income, student or credit_rating) at level 1 by using Information Gain. Justify your answer.

LEE KAH WEI (income)

Income	P	N	I(P,N)
High	0	2	0
Medium	1	1	1
Low	1	0	0

E(Income_Level) = (⅕) * 0 + (⅖) * 1 + (⅖) * 0 = 0.4
Gain(Income_Level) =0.9710 - 0.4 = 0.5710

LEONG YIT WEE (student)

student	**P**	**N**	**I(P,N)**
yes	2	0	~~0.5917~~
no	0	3	~~0.9852~~

E(Student) = (⅗) * 0 + (⅖) * 0 = 0
Gain(Student) =0.9710 - 0 = 0.9710

LIM CHIA CHUNG (credit_rating)

Credit Rating	**P**	**N**	**Total**	**I(P,N)**

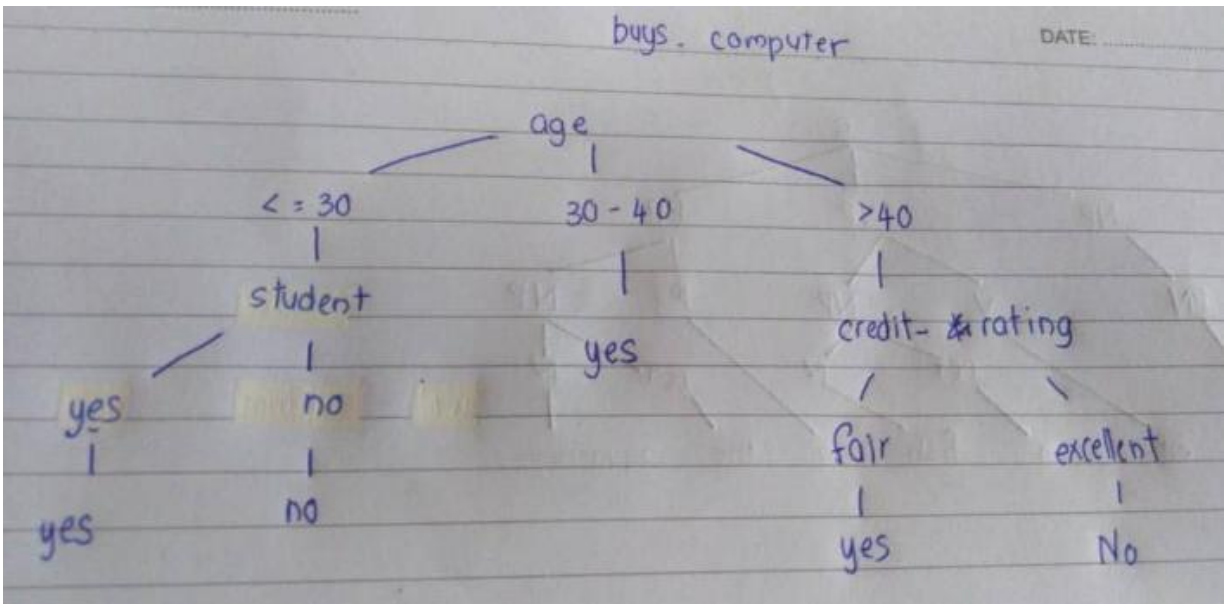
Principal Component Analysis (PCA) can be used to identify two significant components

Fair	1	2	8	0.81130.9183
Excellent	1	1	6	1

E(credit_rating) = (⅜) * 0.9813 + (⅝) * 1 = 0.9510
G(Credit_rating) = 0.9710 - 0.9510 = 0.02

b) Produce the complete decision tree based on the answer from Q6(a).

LIM JUN RONG



c) Given a new sample, X = < <=30, low, no, excellent >, predict the class it belongs to.

LIM MING JUN

no

d) Refer to Table 2, given a new sample, X = < <=30, low, no, excellent >, predict the class it belongs to, P(C|X), using **Naïve Bayes** classification.

LIM YIH FENG (P(Yes|X)

P(Yes) = 9/14
P(<=30) = 2/9
P(low) = 3/9 = 1/3
P(no) = 3/9 = 1/3
P(excellent) = 3/9 = 1/3
P(Yes | X) = 9/14 * 2/9 * 1/3 * 1/3 * 1/3
=0.00529

ONG T'NSAM (P(No|X))

Principal Component Analysis (PCA) can be used to identify two significant components

P(N)= 5/14
P(<=30) = 3/5
P(low) = 1/5
P(no) = 4/5
P(excellent) = 3/5
P(No|X) = P(<=30) x P(low) x P(no) x P(excellent) x P(N)
= 3/5 x 1/5 x 4/5 x 3/5 x 5/14
= 0.0206

To calculate the information gain:

$$I(p,n) = -\frac{p}{p+n}\log_2\frac{p}{p+n} - \frac{n}{p+n}\log_2\frac{n}{p+n}$$

$$E(A) = \sum_{i=1}^{v}\frac{p_i+n_i}{p+n}I(p_i,n_i)$$

$$Gain(A) = I(p,n) - E(A)$$

		Predicted		
		Cat	Dog	Rabbit
Actual class	Cat	5	3	0
	Dog	2	3	1
	Rabbit	0	2	11

Figure 2. Confusion Matrix

7. Assess the following performance of the classifier as shown in Figure 2:
   a. Overall accuracy
   b. Precision for cat classification
   c. Recall for dog classification

TAN KAI YUAN

A.Overall accuracy = Total Correctly Classified / Total Actual = (5+3+11) / 27
B.Precision for cat classification = Correctly Predicted Cat / Predicted Cat= 5 / 7
C. Recall for dog classification = Correctly Classified Dog / Actual Dog = 3 / 6

Principal Component Analysis (PCA) can be used to identify two significant components

Tutorial 7 (Week 7)

1. Discuss the THREE (3) major obstacles involved in Natural language understanding (NLU).

CHIN JUN WAI (obstacle 1)

Syntactic Problem
It is when a sentence or a word can carry out two or more possible meanings and leads to misunderstanding in Natural language understanding (NLU).

Example : She slaps the boy with the newspaper.

From the example above, we are unable to determine whether “She slaps with the newspaper” or she slaps “the boy with the newspaper”.

ONG T'NSAM (obstacle 2)

Semantic problem
-Same word which has different meaning.
-example : She steals a pair of shoes
She steals my heart.

From the example given, we cannot merely chain together the dictionary meanings of “steal” to both of the sentences as “steal” in both sentences have different meanings:
She steals a pair of shoes = She takes my shoes without my permission
She steals my heart = I fall in love with her

TAN KAI YUAN (obstacle 3)

Pragmatic problem
It is missing the intended meaning behind an expression and opt instead for the semantic (literal) meaning of it.
For example, consider a mom says to a son ‘is that your coat on the floor?’.
The son may answer it as yes (that is he may consider it as a question and this is why he said yes).
Consequently, this is a pragmatic problem because he got only the literal meaning of the sentence whereas the mom meant it as a request (go and pick it up)

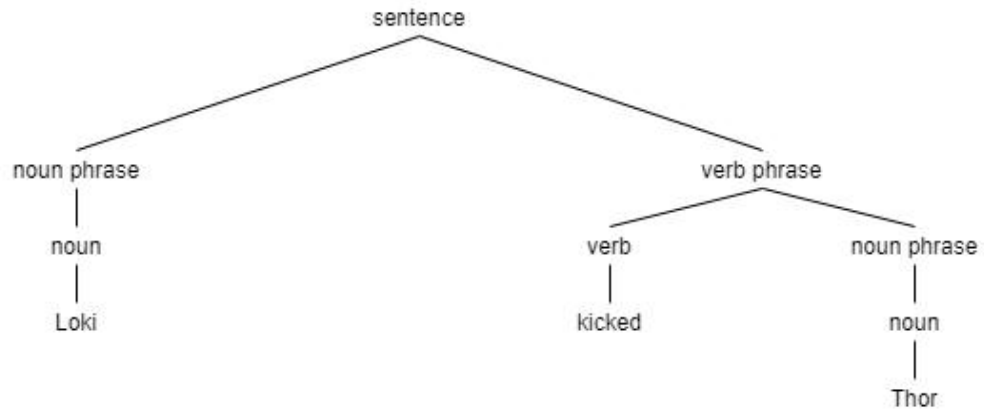
2. What are the three typical stages involved in a natural language processing (NLP) application development? Elaborate these three stages with the aid of appropriate diagrams.

CHONG JIA LOONG

1.Parsing

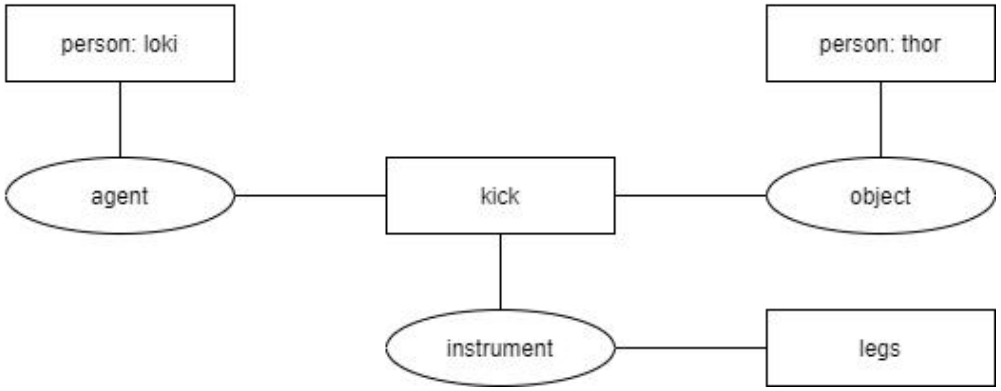
Principal Component Analysis (PCA) can be used to identify two significant components

parse tree:



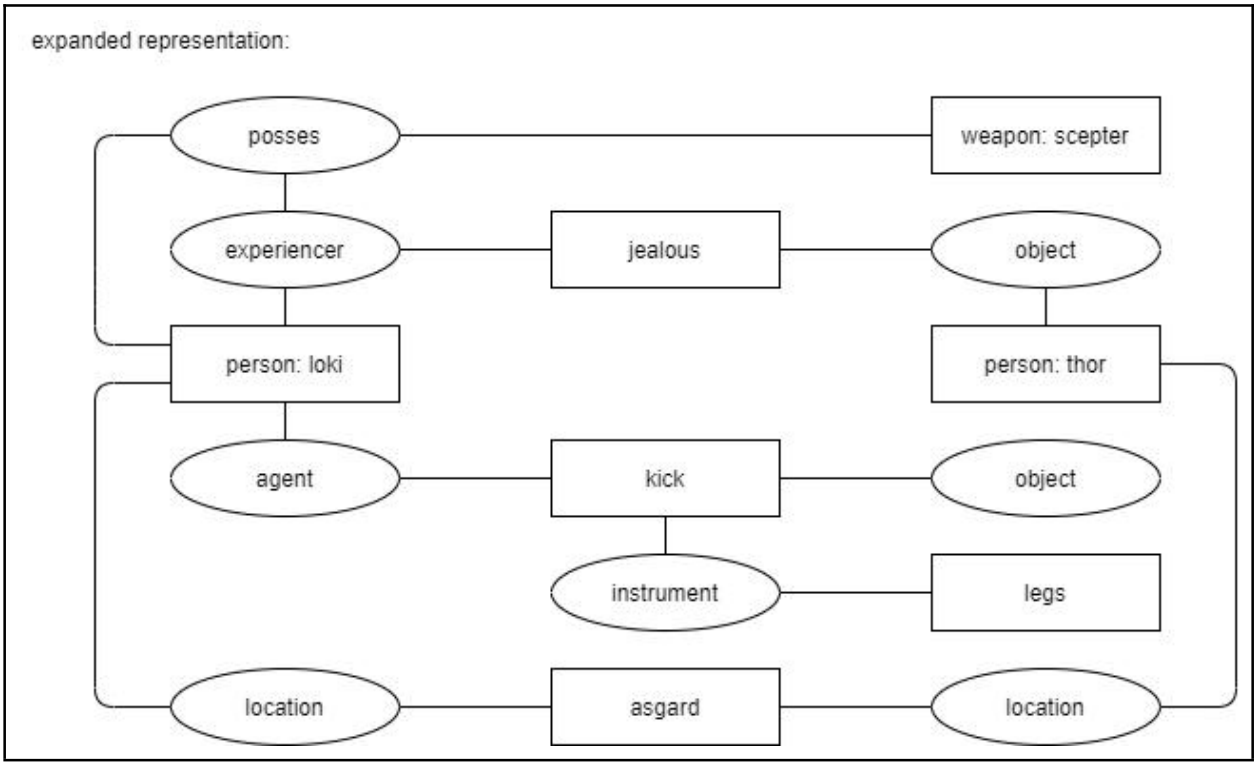
2.Semantic Interpretation

internal representation:



3.World Knowledge Representation

Principal Component Analysis (PCA) can be used to identify two significant components



(need to provide brief explanation)

3. Besides the three main phases involved in NLP as discussed in Q2, describe **three other significant analyses** that can be done to improve NLP applications.

KONG MUN JUN

Semantic Analysis

- Determines the possible meanings of a sentence by focusing on the interactions among word-level meanings in the sentence.
- Some people may think it's the level which determines the meaning, but actually all the levels do.
- The semantic analyser disregards sentence such as "hot ice-cream"

KOW YEE HUI

Discourse Analysis

- The meaning of any single sentence will depend upon those sentences.
- It will convey the meaning of a sentence by making a connection between component sentences.
- For example, the word "that" in the sentence "He wanted that" **depends upon the prior discourse context.**

https://datascience.foundation/sciencewhitepaper/natural-language-processing-nlp-simplified-a-step-by-step-guide

LAI PEI XUAN

Principal Component Analysis (PCA) can be used to identify two significant components

Lexical Analysis

- It involves identifying and analyzing the structure of **words**. Lexicon of a language means the collection of words and phrases in a language.Lexical analysis is dividing the whole chunk of text into paragraphs, sentences, and words.

https://www.tutorialspoint.com/artificial_intelligence/artificial_intelligence_natural_language_processing.htm

4. NLP application requires the use of knowledge about human languages. Suggest an example of NLP application and describe the difficulties that a researcher would face during the development of the application that you have suggested.

LAI XIN YI

Language Translation

1. Multiple meanings

English and Japanese are probably two of the most complex languages that exist on the planet, and although they are from very different ends of the spectrum in terms of the alphabet, spoken language and geography, the two suffer from one very similar problem that can create a real headache for any translator anywhere in the world. Words in each of these languages can often have a double meaning. For example, in English, scales can either mean a part of a fish or a kitchen utensil, where in Japan, simply writing the kanji is not enough, because things can be written one way and said differently. Both of these are common issues in their languages, and each needs an expert at hand at all times to find effective and accurate workarounds.

2.Grammatical problem

Grammatical issues include, for example, temporality, aspect (appearance from the perspective of verb development (rather than time itself), indicating how the verb expresses the process or the state of the verb expression), pronouns and whether there are questions to make the subject pronoun clear

https://www.languageconnect.net/blog/language-connect/the-top-5-problems-every-translator-has-had-to-face-head-on/

5. In NLP, representation is important as it can solve issues like canonical form of sentences and syntactic problem of a sentence.

a) Explain the meaning of canonical form of sentences. Provide examples to elaborate your answer.

LEE CHUN XIAN

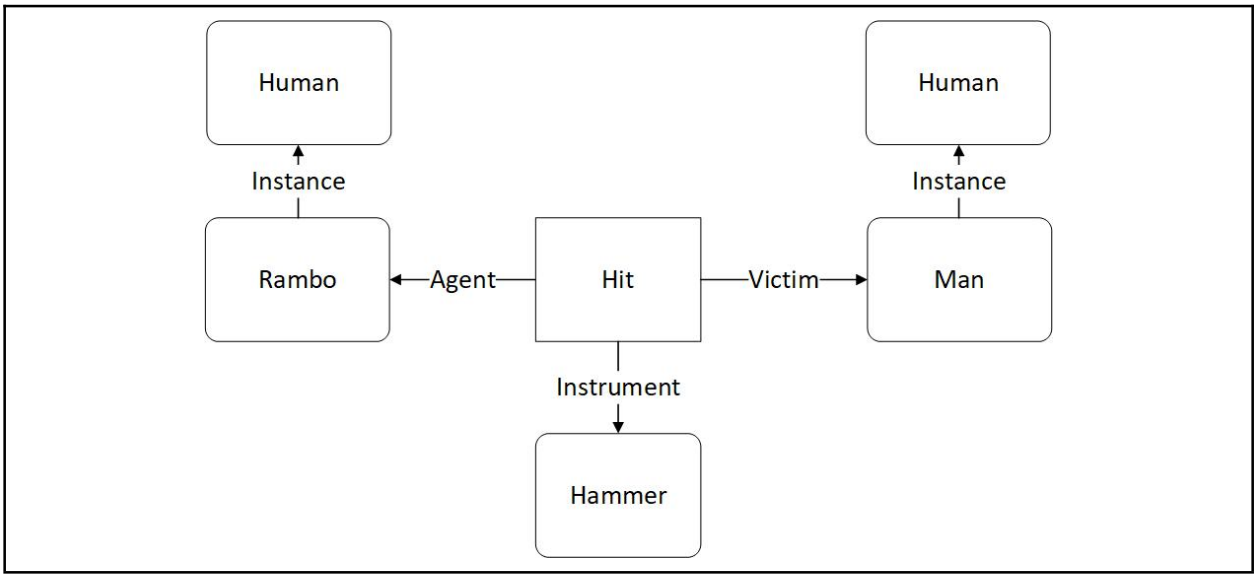
Two different sentences with the same meaning.

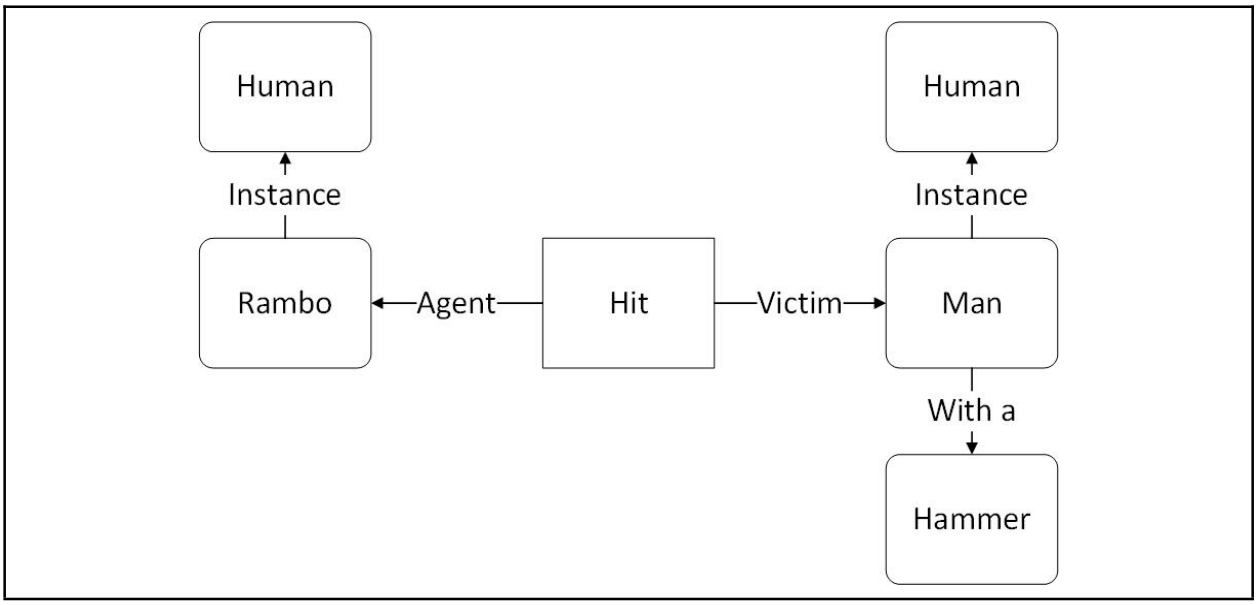
- Square bought a red colour bag.
- A red colour bag is purchased by Square.

b) By referring to the below statement, there is a syntactic problem. Identify the problem and then draw **TWO (2) different representations** (by selecting either the semantic network or the conceptual graph) to solve the syntactic problem.

"Rambo hit the man with a hammer"

LEE JUN XIAN





6. Parse tree is a popular tool used in one of the phases of natural language processing (NLP) called parsing.

Principal Component Analysis (PCA) can be used to identify two significant components

a) Explain the importance of parsing in NLP.

LEE KAH WEI

There are many NLP tasks where Parsing helps. For example, in sentiment analysis, parsing is crucial for understanding a few very similar sentences but with different words. And for the speech recognition, Parsing scores the strings with either a pass/fail or a likelihood score, to give a powerful language model for speech recognition. Similarly, parsing could help in spell checking, optical character recognition, text prediction, handwriting detection etc.

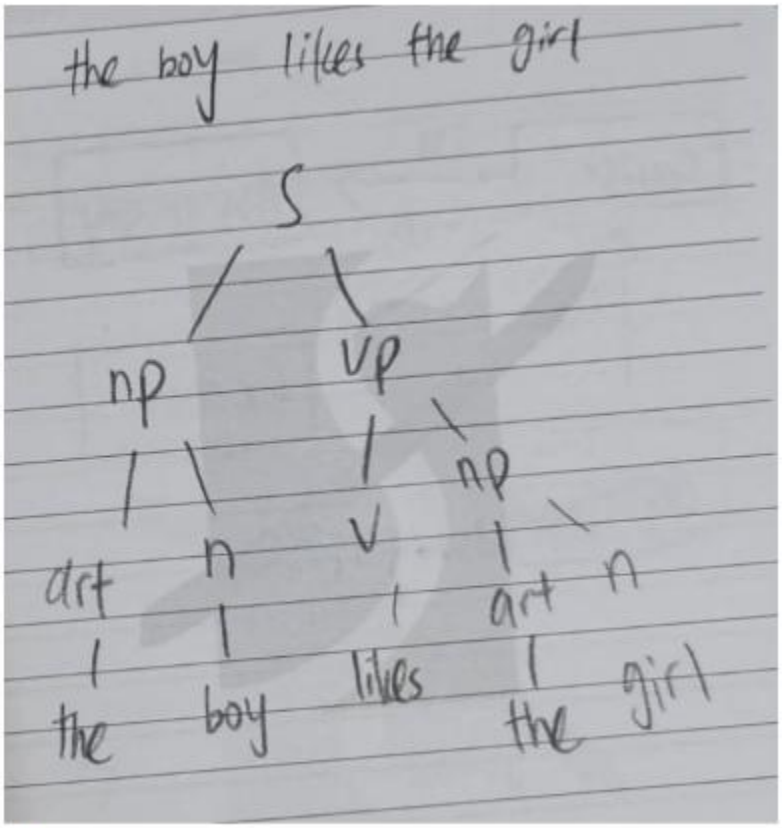
https://www.quora.com/How-does-parsing-help-in-Natural-Language-Processing#:~:text=Parsing%20scores%20the%20strings%20with,T9)%2C%20handwriting%20detection%20etc.

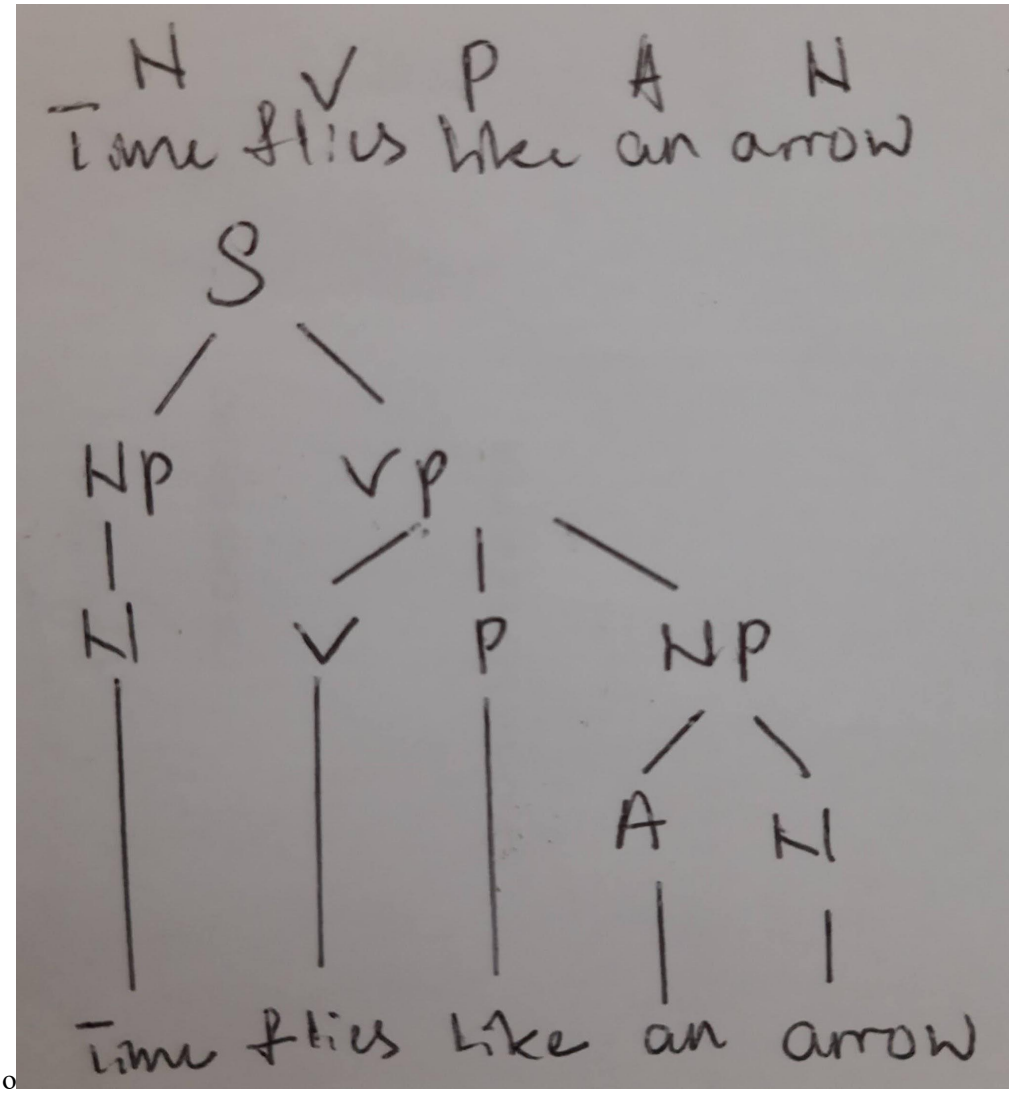
b) With the aid from the simple English grammar for simple transitive sentence as shown next page, draw the respective **parse trees** to verify the sentences "***the boy likes the girl***" and "***time flies like an arrow***".

```
sentence        ↔ noun-phrase verb_phrase
noun_phrase     ↔ noun
noun_phrase     ↔ article noun
verb_phrase     ↔ verb noun_phrase
verb_phrase     ↔ verb preposition noun_phrase
preposition     ↔ [like]
article         ↔ [a, an, the]
noun            ↔ [flies, time, arrow, boy, girl]
verb            ↔ [like, flies]
```

LEONG YIT WEE (***the boy likes the girl)***oLIM CHIA CHUNG (***time flies like an arrow***)

Principal Component Analysis (PCA) can be used to identify two significant components





o

7. Considering the sentence *S* = **"She beats George with one hand at the bank".**

   a) The sentence *S* consists of semantic ambiguity and syntactic ambiguity. Identify both of the ambiguities found from the sentence above.

LIM JUN RONG

Semantic ambiguity: The word "bank" in the sentence can be referred to the financial institution or the river bank. Thus, it may lead the reader to confusion.

Syntactic ambiguity: The sentence above carries two meanings. First is she (the girl) beats George with one of her hands. While the second meaning is she (the girl) beats George who only has one hand.

Principal Component Analysis (PCA) can be used to identify two significant components

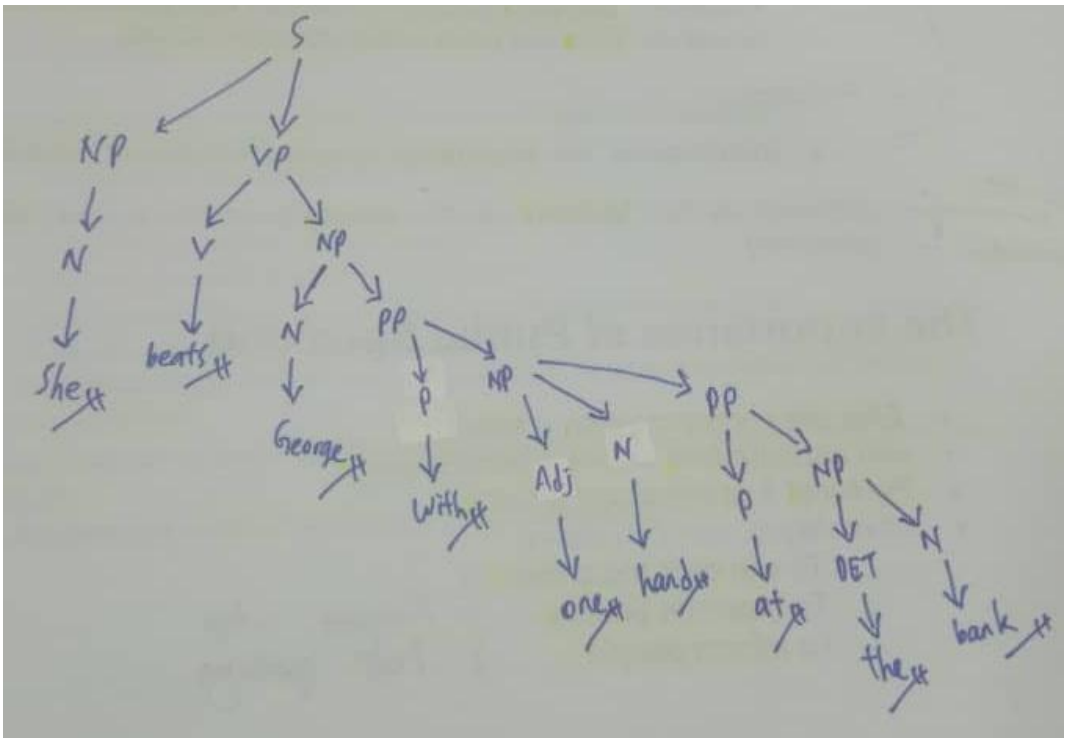
b) Given the grammar below, construct **ONE (1)** parse tree for the sentence *S*.

S = **"She beats George with one hand at the bank".**

Remark: grammar in the parentheses () means it is optional.

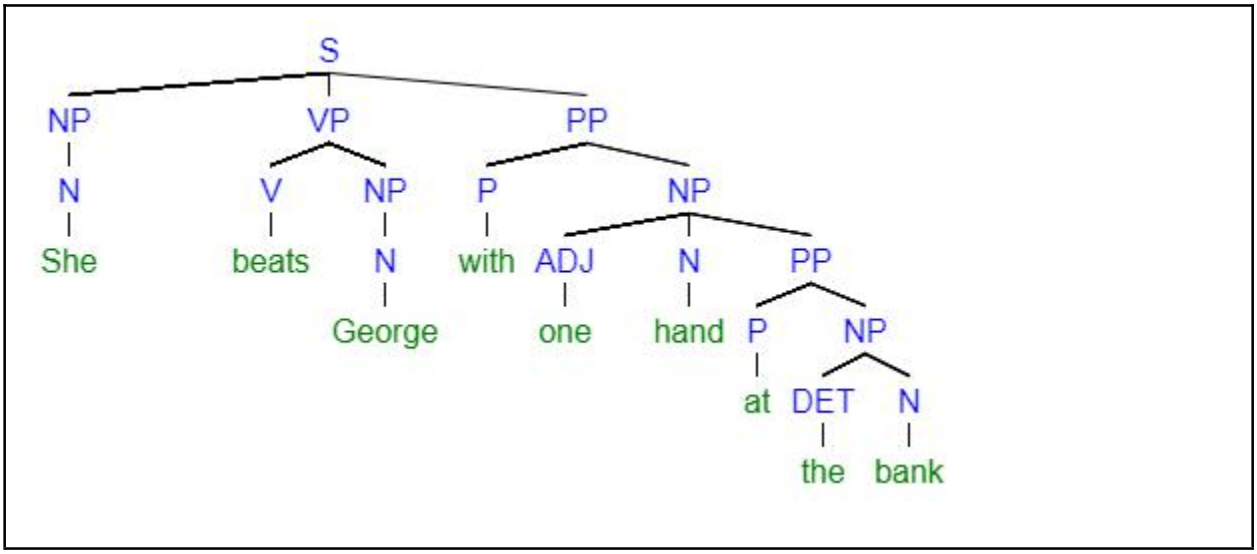
```
S -> NP VP (PP)
NP -> (DET) (ADJ) N (PP)
VP -> V NP (PP)
PP -> P NP
ADJ -> [one]
DET -> [a, the]
N -> [She, George, hand, bank]
V -> [beats]
P -> [with, at]
```

LIM MING JUN (parse tree 1)



LIM YIH FENG (parse tree 2)

Principal Component Analysis (PCA) can be used to identify two significant components



Tutorial 6 (Week 6)

Q1. Good Food Restaurant sells local food such as nasi lemak, roti canai, laksa, etc. **Each customer can make one or more food orders, and each food order may contain one or more food item.** Figure 1.1 shows an example of receipt to be produced for one food order.

Good Food Restaurant

Order No: Date:
Customer No:
Customer Name:

Order list

No.	Item Code	Description	Quantity	Unit Price	Subtotal Price
1					
2					
3					
4					
5					
6					
				Total	

Figure 1.1 Sample receipt

(i) Between semantic network and frames, justify which knowledge representation tool you would use to represent the information above.

CHIN JUN WAI

I would use frames as my representation tool because the basic characteristic of frames is that it represents the **related knowledge about an object** that has much default knowledge such as the item code, description and unit price. On the other hand,a semantic network is more useful in representing concepts or relationships between objects. Therefore, frames is more suitable to be my knowledge representation tool since for each food ordered by the customers is an object.

(ii) Illustrate the **representation of the given statements above** using the knowledge representation tool that you selected in Question 4(i) above. You diagram shall clearly demonstrate the relationships between customer, food order and food item.

CHONG JIA LOONG

KONG MUN JUN

Q2. Assume that L&L Food Bar wanted to develop an online membership system, which offers two different types of **Member**, i.e. **RegularMember** and **PremiumMember**. The **RegularMember** has the attributes of *memberID*, *name* and *discount.* Every **RegularMember** enjoys *discount* rate of 5%. Similarly, the **PremiumMember** contains *memberID*, *name* and *discount.* Every **PremiumMember** gets *discount* rate of 10%. In addition, **PremiumMember** also obtains extra10% *discount* on *birthdate*.

Principal Component Analysis (PCA) can be used to identify two significant components

(i) Between semantic network and frames, justify which knowledge representation tool you would use to represent the information above.

KOW YEE HUI

Frame. It is because :

1. Expressive power .
   - More easier to understand and use
   - Can represent stereotyped object
   - More detail than Semantic network.
2. Flexible
   - Can set up more than one slot for new attributes and relations.
   - Simply to include default data
   - Discover the missing value.
   - Easy to create specialized procedures.
3. Show inheritance
   - Easy to control and show inheritance.
4. Show constraints
   - Allow constraints to be set for value/ facets.

https://pdfs.semanticscholar.org/3050/f186dfd77fce3ab3d094abebd78411f5a0c1.pdf

LAI PEI XUAN

Semantic Network

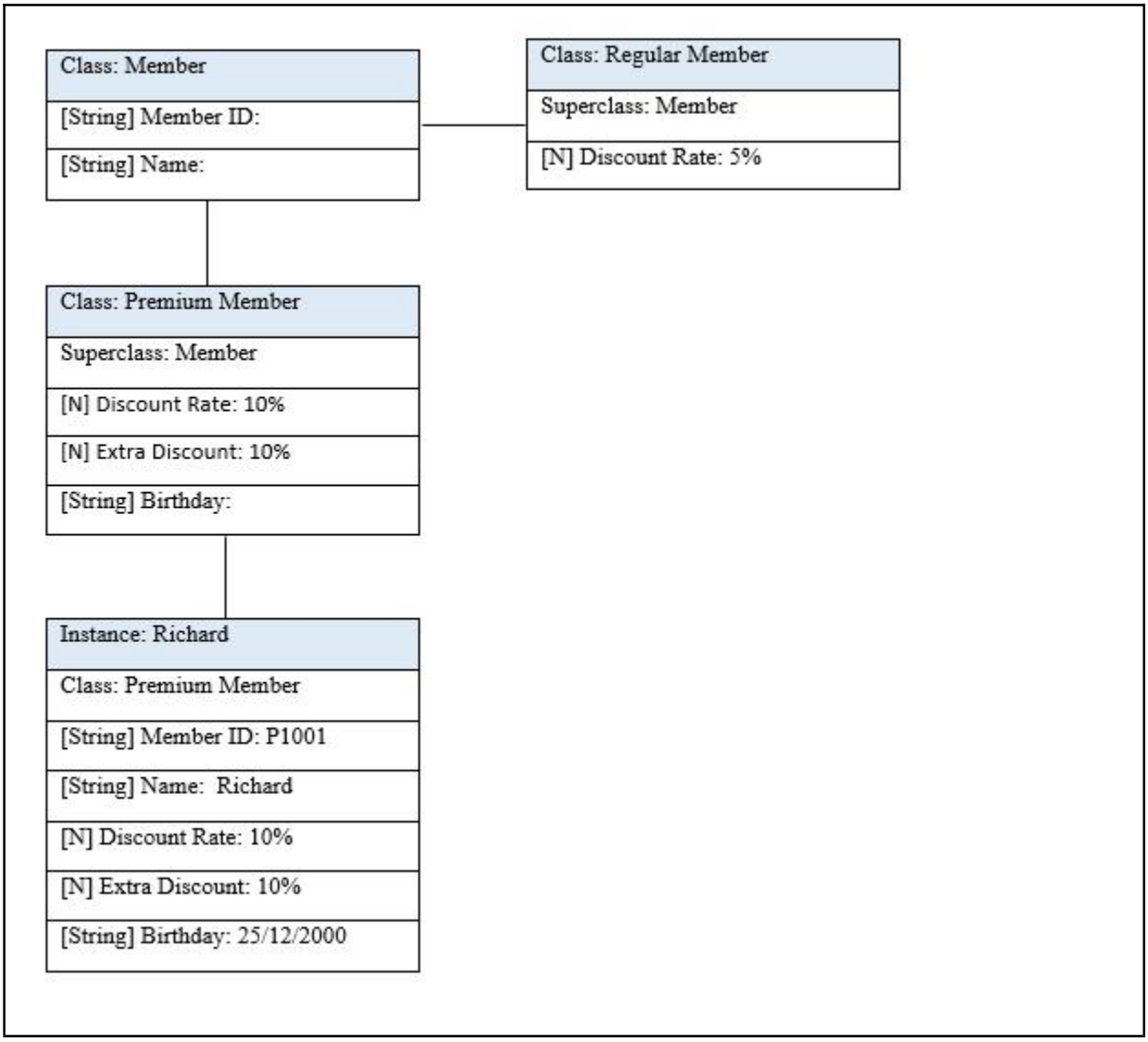
- The semantic network is a natural representation of knowledge
- The semantic network permits using of effective inference algorithm (graphical algorithm)
- The semantic network is simple and can be easily implemented and understood
- The semantic network convey meaning in a transparent manner

(ii) Illustrate the representation of the given statements above using the knowledge representation tool that you selected in Question 5(i) above. Your diagram shall clearly demonstrate the relationships between **Member**, **RegularMember** and **PremiumMember**. Demonstrate **ONE (1)** instance of **PremiumMember** (provide your own sample data).

LAI XIN YI

Frame

Principal Component Analysis (PCA) can be used to identify two significant components



LEE CHUN XIAN

Instance: David
Class: Premium Member
[String] Member ID: P1002
[String] Name: David
[N] Discount Rate: 10%
[N] Extra Discount: 10%
[String] Birthday: 21/09/1998

Q3. Every course consists of a course code, course title, number of credits and department code. Every course may contain a few assignments and exams. Each assignment has an assignment ID, assignment description, assignment mark and assignment due date. Each exam contains exam ID, exam mark and exam date. Every department has a department code, and department head.

(i) Compare semantic network and frames, propose which tool is more suitable to represent the information above. Justify.

LEE JUN XIAN

Frame is the better tool for the information above. There is more detailed information and there are no events or actions that occur. Both semantic networks and frames can show inheritance but frame shows the details of each objects and since there is no action or event occurs, which frames is not suitable to describe events, therefore frame are the better tool.

LEE KAH WEI

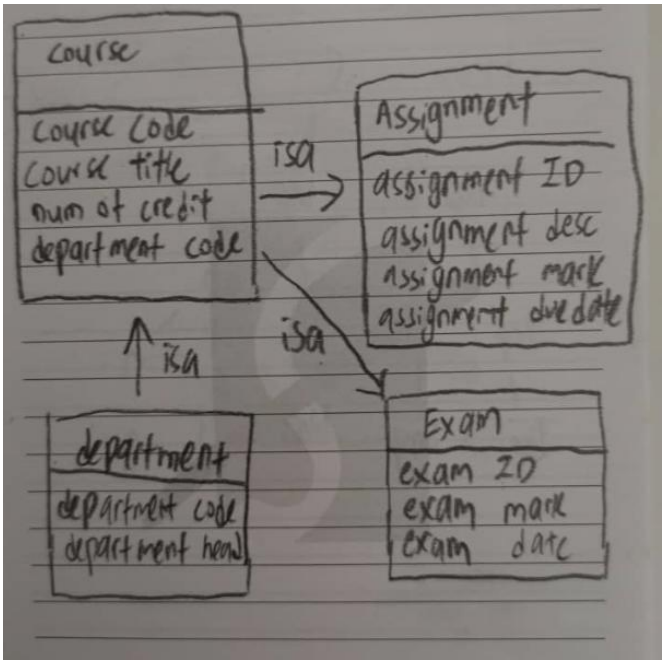
Semantic Networks. It is because a semantic network can be used as a typical connection between two different fields. For this case, the characteristics of assignment can be connected to course using semantic network but not with semantic frames.

(ii) Based on the selection in Q3(i), draw the representation of the information above.

LEONG YIT WEE

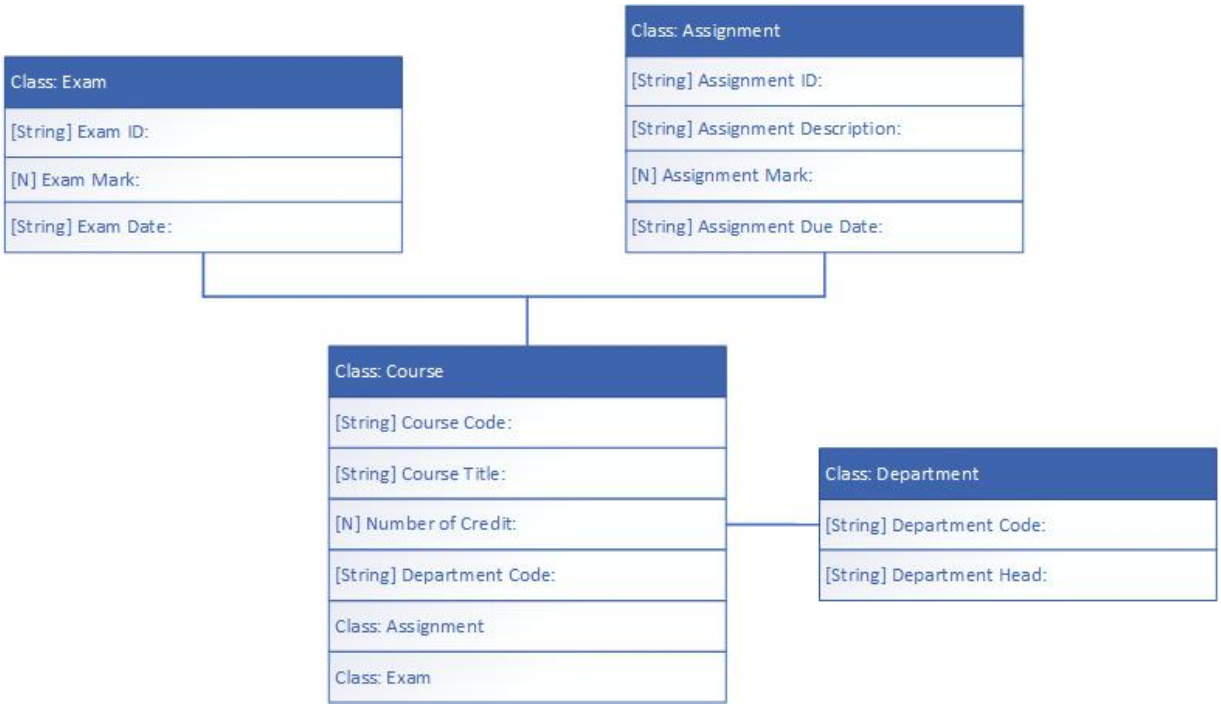
Semantic network

Principal Component Analysis (PCA) can be used to identify two significant components



LIM CHIA CHUNG

Frame



Q4. Refer to the following scenario and design a new *Student Class* frame that demonstrates the appropriate demons.

"When the student administrator hits the add button of the system, the system will require inputs for student's name, student's ID, and level. When a subject code is added, the mark of the subject is requested and the mark will be added to the student instance frame. When the mark is entered by the user, the system will make sure the mark is between 0 and 100. The subject code supports multi-values."

LIM JUN RONG

CLASS: Student
[String] Student Name: if_new: ask
[String] Student ID: if_new: ask
[String] Level: if_new: ask
[String] Subject Code: if added: ask mark multi-valued: True
[N] Mark: Range: 0-100

"When the student administrator hits the add button of the system, the system will require inputs for student's name, student's ID, and level. When a subject code is added, the mark of the subject is requested and the mark will be added to the student instance frame. When the mark is entered by the user, the system will make sure the mark is between 0 and 100. The subject code supports multi-values."

Principal Component Analysis (PCA) can be used to identify two significant components

LIM MING JUN

??

Q5. Describe the role of frame-based representation in the development of World Wide Web.

LIM YIH FENG

Frame-based representation is more preferable for the world wide web because its structure is much more organized and could easily represent a stereotyped object or concept. Frame-based representation can describe various attributes and characteristics of an object even in detail just by reading the frames values and attributes.

Advantages of frame-based representation is the reason why it is widely used in the world wide web.Frame-based representation has strong expressive power is because its structure is so simple and details to understand.Flexibility is also one of the key point that frame-based representation is used that it can easily set up slots and value and detect missing value. Showing inheritance and constraints value making user understands the structure better and developer would not be confused when developing.

Not so relevant. Good try

ONG T'NSAM

Frame-based representation allows for exceptions in particular instances. This gives frames an amount of flexibility that allow representations of real world phenomena to be reflected more accurately. A frame represents an object or a concept. Attached to the frame is a collection of attributes (slots), potentially having types (or value restrictions) and potentially filled initially with values. When a frame is being used the values of slots can be altered to make the frame correspond to the particular situation at hand.

A frame-based knowledge representation system is built on a relational database that is completely transparent to the user. A user at a client machine sends standard knowledge base queries across a distributed computer system and the system translates the queries into a language suitable for querying the database, such as Structured Query Language (SQL). The system stores a hierarchical data model that includes classes, particular instances of the classes, and relations among the classes and instances. Primitive objects, such as classes

and instances, are organized with their associated attributes into frames

References :
https://en.wikipedia.org/wiki/Frame_(artificial_intelligence)
https://www.ida.liu.se/ext/epa/ej/etai/2001/018/01018-etaibody.pdf
https://patents.google.com/patent/US6442566B1/en

Good example of application but not so relevant to the question. Good try

TAN KAI YUAN

The term “Semantic Web” encompasses efforts to build a new world wide web (WWW) architecture. The SemanticWeb will allow users to use more automated functions (reasoning, information and service discovery,autonomous agents, etc.) on the Web, easing the workload of humans. The Semantic Web will also pave the way for true "device independence" and customization of information content for individual users, since the information on the Web would now be contained in a "raw form" and any (context-dependent) presentation could be rendered on demand.

The fundamental idea of a frame system is rather simple: A frame represents an object or a concept.Attached to the frame is a collection of attributes (slots), potentially having types (or value restrictions) and potentially filled initially with values. When a frame is being used the values of slots can be altered to make the frame correspond to the particular situation at hand.

Ref:
https://www.researchgate.net/publication/229101030_The_role_of_frame-based_representation_on_the_Semantic_Web

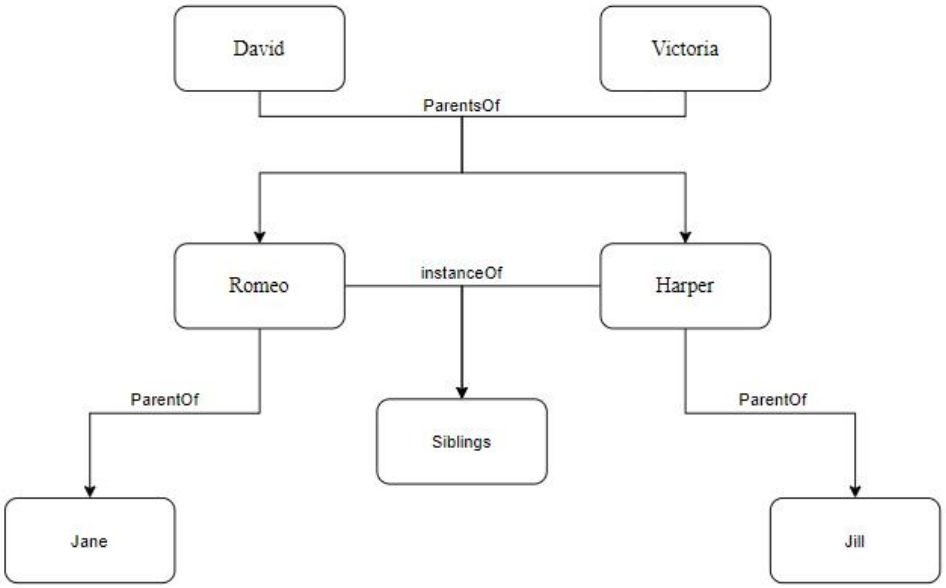
Tutorial 5

1. Semantic network is a knowledge base that represents *semantics* relations between concepts in a *network*. Draw a semantic network for the following facts and relations.

   “David and Victoria are the parents of Romeo and Harper. Romeo and Harper are brother and sister. Jane is the child of Romeo; Jill is the child of Harper.”

Principal Component Analysis (PCA) can be used to identify two significant components

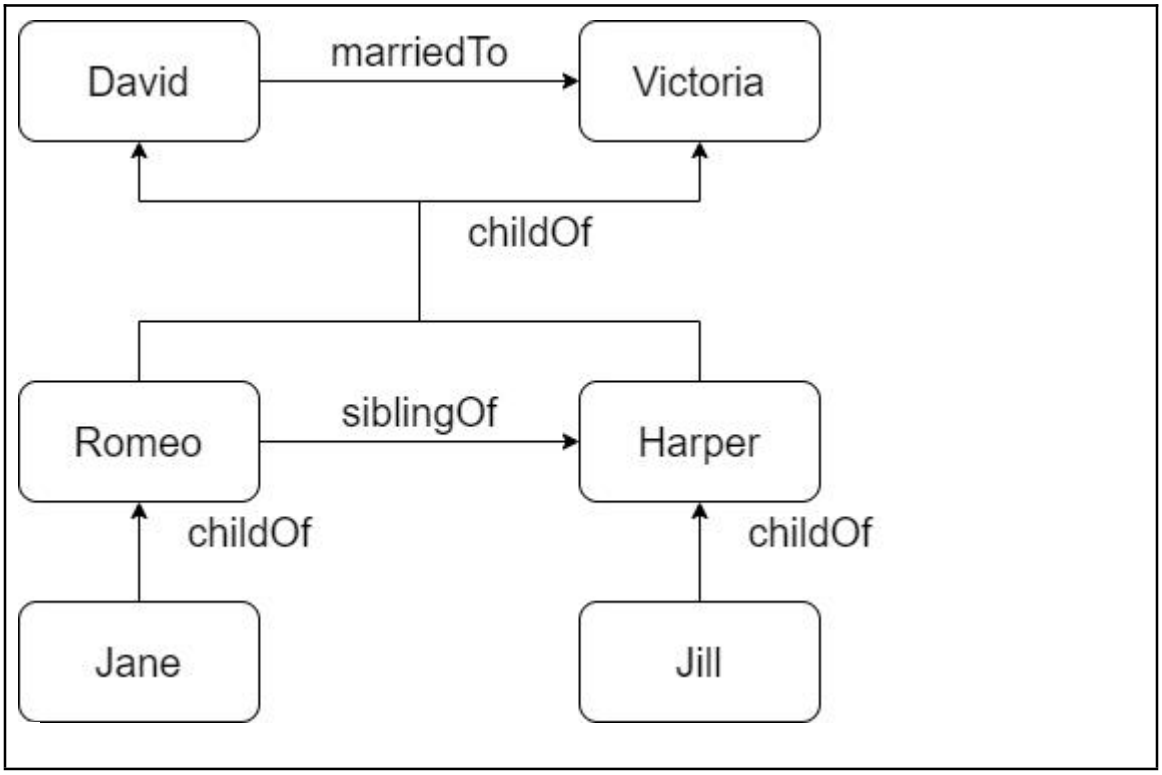
CHIN JUN WAI



CHONG JIA LOONG

(Corrected)

Principal Component Analysis (PCA) can be used to identify two significant components



2. How are semantic networks different from Conceptual Graph? Explain your answer with the help of **ONE(1)** diagram for **each** of the methods based on the following statement:

"A cat is on a mat."

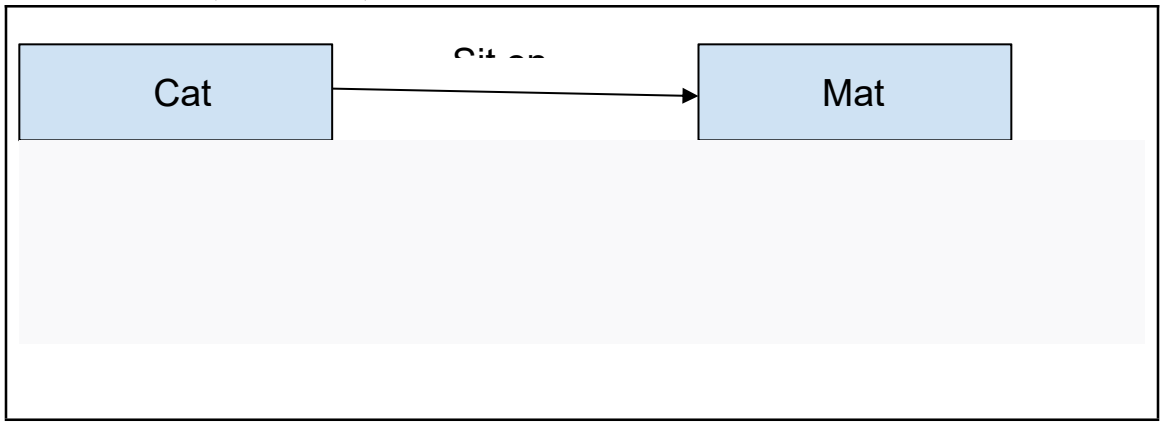
KONG MUN JUN (difference)

Semantic networks are used to represent inheritance relationships. **is-a** is used to represent a relationship between a subclass and a superclass.**has-a/has** represent relationships between objects.
^^semantic network?

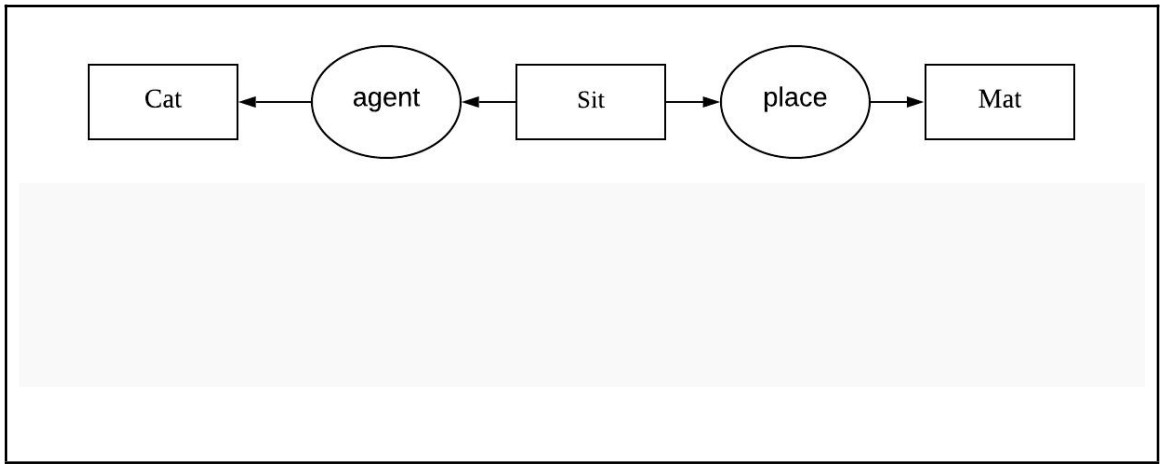
A conceptual graph consists of concepts, represented by boxes and conceptual relations are represented by ellipses.
(how about semantic network)?

Principal Component Analysis (PCA) can be used to identify two significant components

KOW YEE HUI (semantic net)



LAI PEI XUAN (conceptual graph)



3. The following statements are given.
   - ★ Albert is a human.
   - ★ Pepper is a robot.
   - ★ The construction of humans is biological, while robots are mechanical.
   - ★ Both humans and robots are autonomous systems.
   - ★ The behaviours of autonomous systems are mobile and adaptive.

   a. Between semantic network and frames, justify which knowledge representation tool you would use to represent the information above.

   LAI XIN YI

   Frames
   - ➔ Frame is used to represent stereotyped object
   - ➔ Frame representation is easy to understand and visualize
   - ➔ It is very easy to add slots for new attributes and relations

Principal Component Analysis (PCA) can be used to identify two significant components

➔ It is easy to include default data and to search for missing values ✓

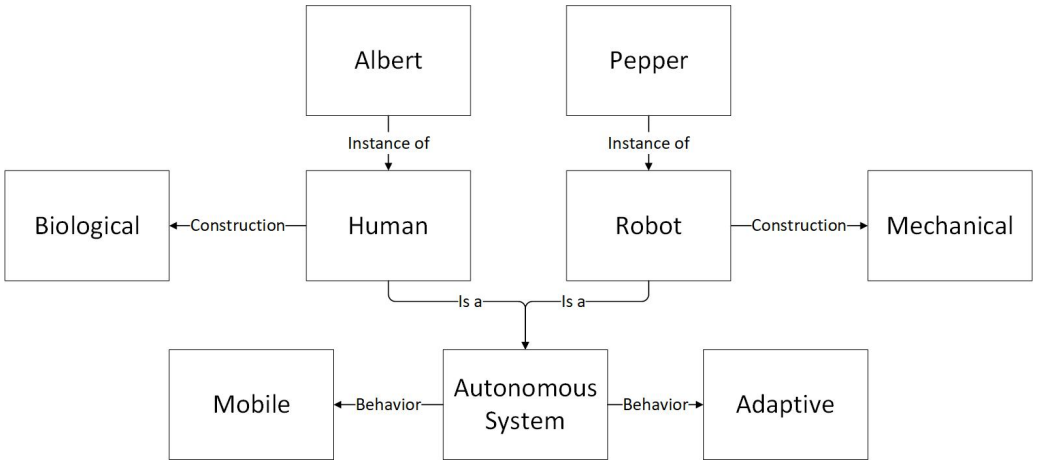
b. Discuss **ONE (1)** limitation of the knowledge representation tool that is selected in Question 3(a) above.

LEE CHUN XIAN

Difficult to program, especially making inference or hypothesis. ✓

c. Illustrate the representation of the given statements above using the knowledge representation tool that you selected in Question 3(a) above.

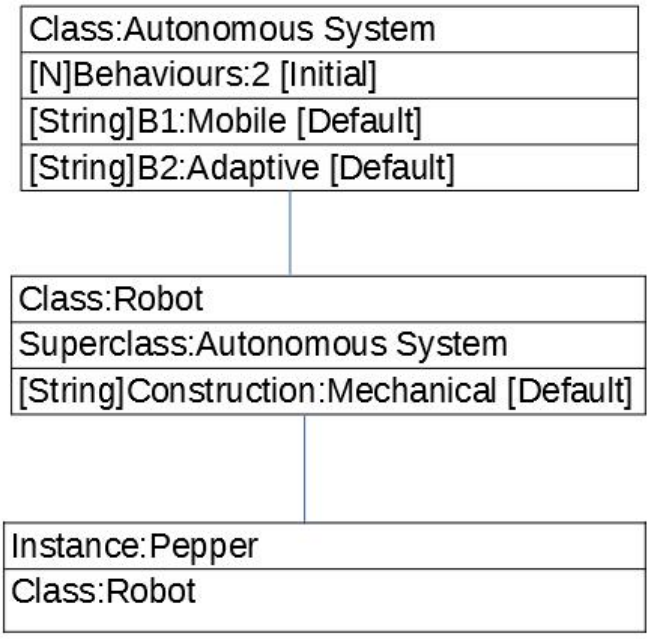
LEE JUN XIAN (Amended)



★ Albert is a human.
★ Pepper is a robot.
★ The construction of humans is biological, while robots are mechanical.
★ Both humans and robots are autonomous systems.
★ The behaviours of autonomous systems are mobile and adaptive.

LIM YIH FENG

- ★ Albert is a human.
- ★ Pepper is a robot.
- ★ The construction of humans is biological, while robots are mechanical.
- ★ Both humans and robots are autonomous systems.
- ★ The behaviours of autonomous systems are mobile and adaptive.



4. The following statements are given.

Bird is a living being that builds nests.
Insect is an invertebrate living being.
Ant is a kind of insect that builds nests.
Both dragonfly and fly are insects that have wings.

(i) Between semantic network and frames, justify which knowledge representation tool you would use to represent the information above.

LEE KAH WEI

Semantic Frames. As semantic frames is easier to understand as compared to semantic network.

Principal Component Analysis (PCA) can be used to identify two significant components

(ii) Discuss **ONE (1)** limitation of the knowledge representation tool that is selected in Q4(i) above.

LEONG YIT WEE

Frame
Limitation of frame is difficult to program especially making inference
Semantic network
Limitation of semantic network is there are no standard definition of link names and nets are not intelligent , dependent on creator.

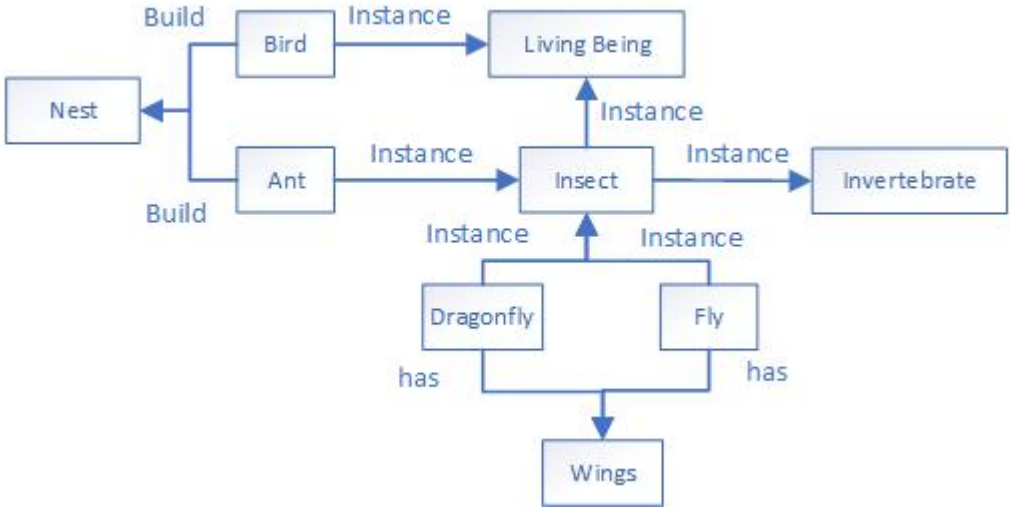
reference:https://www.slideshare.net/JennyGalino/semantic-networks-67206497#:~:text=Disadvantages%20%E2%80%A2%20There%20is%20no,asserts%20relationships%20and%20structural%20links.&text=Undistinguished%20nodes%20that%20represent%20classes%20and%20that%20represent%20individual%20objects.
✓

Principal Component Analysis (PCA) can be used to identify two significant components

(iv) Illustrate the representation of the given statements above using the knowledge representation tool that you selected in Q4(i) above.

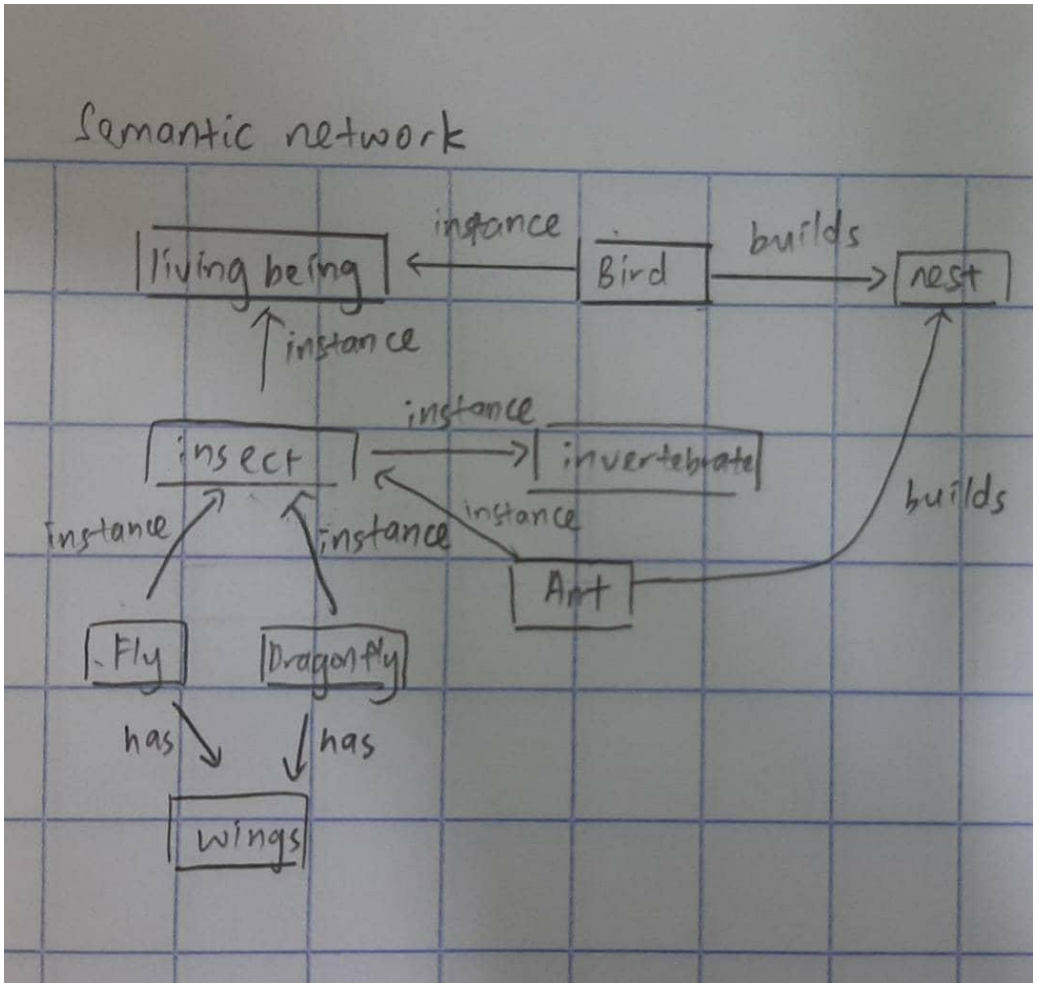
LIM CHIA CHUNG

Bird is a living being that builds nests.
Insect is an invertebrate living being.
Ant is a kind of insect that builds nests.
Both dragonfly and fly are insects that have wings.



ONG T'NSAM

Semantic network



5. Suggest and describe a real-life application where a semantic network is applied.

LIM JUN RONG

Search Engine (E.g Google)

- The application of semantic network solutions in search engines helps to improve the search result. The network allows the computer to describe the information in a formal way and to introduce interdependencies between different pieces of information. Such knowledge also allows searching not only for the key words entered by the user but also for related concepts, and shows how the relation is made. For instance, when the user enters “animal”, the search engine will not only display the random examples of animals, but instead, the images or information are displayed according to the category such

as bird, reptile, mammal and many more. Therefore, the information search is relatively broader and the user is able to retrieve more accurate and deeper information.

Reference: https://www.intechopen.com/books/engineering-the-computer-science-and-it/application-of-semantic-networks-in-natural-language-issues

LIM MING JUN

Recognition System (E.g: Apple Siri)

Siri, Apple's voice recognition assistant, is the best example of the Semantic Web. Siri, Apple's new voice recognition technology, is transformative — calling it "voice recognition" is vastly underselling its importance. It is a technology that has the potential to revolutionize computing, entertainment, home life, and work over the coming decade. And it is perhaps the biggest stake in the ground ever for the vaunted but seemingly elusive semantic Web. The semantic Web has been written off as many times, but Siri is a powerful realization of the semantic Web. Siri has to figure out what you mean, not just take your literal words and run them against an index of files. Siri has to comprehend that you mean "six" instead of "sex." and that's powerful.

Reference:
https://scholarlykitchen.sspnet.org/2011/11/03/siri-and-the-resurrection-of-the-semantic-web/

TAN KAI YUAN

Marketing Applications (e.g Knowledge System)
Marketing activities in an organisation consist of an entire set of actions and their purpose to sell products or services to the customers. The application of knowledge system is very suitable to support marketing activities in the environment of internet-based services to make more effective expenditure with respect to the results achieved across internet-based services. Every knowledge object relates to something that generates positive and negative interaction within the area outlined by the portal (keyword). The branching of keywords and knowledge objects linked to the portal as a result of application of semantic network for knowledge sharing in the field of

marketing.

Ref:
https://www.researchgate.net/publication/263619928_Application_of_semantic_network_for_knowledge_sharing_in_the_field_of_marketing

Tutorial 1

1. Who is the father of AI? Describe the reason why he was recognized so.

CHIN JUN WAI

John McCarthy is the father of AI. The reason is that he coined the term AI and furthermore he proposed an idea of making a machine that can work like a human. ✓

Link: https://www.independent.co.uk/news/obituaries/john-mccarthy-computer-scientist-known-as-the-father-of-ai-6255307.html ✓

John McCarthy, an American computer scientist pioneer and inventor, was known as the father of Artificial Intelligence (AI) after playing a seminal role in defining the field devoted to the development of intelligent machines

Link :https://blog.frase.io/who-are-the-godfathers-of-ai/

2. Identify one key event or major achievement of AI development in the year of
   - 1960-1969

CHONG JIA LOONG

Unimation's industrial robot Unimate worked on a General Motors automobile assembly line. GM first

used the machine for die casting handling and spot welding of car bodies. The first Unimate robot was installed at GM's Inland Fisher Guide Plant in Ewing Township, New Jersey in 1961 to lift hot pieces of metal from a die-casting machine and stack them.

✓

Link: https://en.wikipedia.org/wiki/Timeline_of_artificial_intelligence#1960s
Link: https://en.wikipedia.org/wiki/Unimation

- 1970-1979

KO ZHI XIN

Jaime Carbonell (Sr.) developed SCHOLAR, an interactive program for computer assisted instruction based on semantic nets as the representation of knowledge.
Link : https://en.wikipedia.org/wiki/Timeline_of_artificial_intelligence#1970s

- 1980-1989

KONG MUN JUN

AI program called the "Expert System" became mainstream in AI research. Completed by CMU for Digital Equipment Corporation and called XCON

✓

https://en.wikipedia.org/wiki/History_of_artificial_intelligence

- 1990-1999

KOW YEE HUI

Polly Behavior - based Robotics by lan Horswill in 1993. The first robot can use vision and operate at the speed like an animal.

✓

https://en.wikipedia.org/wiki/Timeline_of_artificial_intelligence#1990s

- 2000-2009

LAI PEI XUAN

Recommendation Technology based on tracking web activity or media usage brings AI to marketing ✓
https://en.wikipedia.org/wiki/Timeline_of_artificial_intelligence#2000s

LAI XIN YI

1. iRobot's Roomba autonomously vacuums the floor while navigating and avoiding obstacles in 2002
2. The Semantic Web aims at converting the current web,dominated by unstructured and semi-structured documents into a "web of data“ (Tim Berners-Lee, 2001)

✓

- 2010-2019

LEE CHUN XIAN

Artificial Intelligence successfully diagnosed Lung Cancer - Google AI (2019) ✓
https://www.google.com.my/amp/s/thinkml.ai/top-5-ai-achievements-of-2019/amp

LEE JUN XIAN

Machine Learning and Deep Learning become popular topics ✓

Principal Component Analysis (PCA) can be used to identify two significant components

3. Name an artificial intelligence application that is created by a Malaysian company. Briefly describe its AI functions.

LEE KAH WEI

Simultaneous Artificial Response and Action (SARA) which is created by Saratix and Custlr which they have partnered to launch it. Its function is to provide fashion businesses with services, including body profiling using only an A4 paper, fashion analytics and retail reports.

✓

http://www.conventuslaw.com/report/overview-of-artificial-intelligence-in-malaysia/
https://markets.businessinsider.com/news/stocks/saratix-the-new-fashion-articial-intelligence-1028336527

4. Differentiate between the following:
   - Systems that think like humans.

LEONG YIT WEE

System that thinks like humans is like a system able to think like humans, so the system can catch humans' minds and guess or give the correct recommendation to humans that they might like it . Example : Recommender

References :
https://people.eecs.berkeley.edu/~russell/intro.html#:~:text=A%20system%20is%20rational%20if%20it%20does%20the%20right%20thing.&text=This%20gives%20us%20four%20possible,Systems%20that%20think%20rationally.

Correct examples, but please use a more specific definition for this system. Cognitive computation

   - Systems that act like humans.

LIM CHIA CHUNG

The system that acts like humans won't really be based on a universal definition of what is to be human. As such it can only be based on a wrong definition of what is to be human. So acting like humans will only be superficial and any works done by it won't be for humanity, but for itself.

References:
https://www.quora.com/What-is-the-difference-between-a-system-that-thinks-like-human-and-system-that-acts-like-human

Good attempt, this sounds more like a critic. Please find a more specific definition for this system

- Systems that think rationally.

LIM JUN RONG

- Systems that think rationally often refers to systems that are able to think logically or in a correct manner. **The system should follow a set of rules in order to act accordingly. This law of thought has led to a new field known as "logic".**

Example: Expert system

Reference:
https://www.youtube.com/watch?v=XuE0GqYHPqQ&list=WL&index=38&t=0s
Good. Please provide one example

- Systems that act rationally.

LIM MING JUN

Acting rationally means acting to achieve one's goals, given one's beliefs or understanding about the world. An agent is a system that receives precepts from the environment and acts within that environment by maximizing the expected value of their performance measure given what they currently know. An intelligent agent is one that acts rationally with respect to its goals.

Principal Component Analysis (PCA) can be used to identify two significant components

Example: Adaptive System, AlphaGO
Good. Please provide one example
https://cs.lmu.edu/~ray/notes/introai/
http://www.sci.brooklyn.cuny.edu/~sklar/teaching/s10/cis20.2/notes/lecIV.1-notes.pdf

5. An agent is created for an intelligent air-conditioning system to automatically release the amount of chilled gas based on the surrounding temperature.
   - Define agent

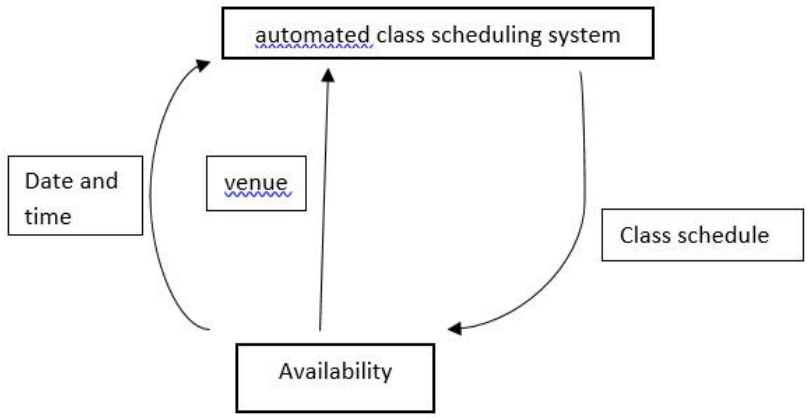
LIM YIH FENG

Agent is something that acts autonomously, sensitive(sense) to its environment, adapts to change, and creates/pursues goals. ✓

- Design a simple architecture of the following agents. You must clearly illustrate how the agent should react to the necessary input and produce output based on the problem above.

  i. Agent of an automated class scheduling system.

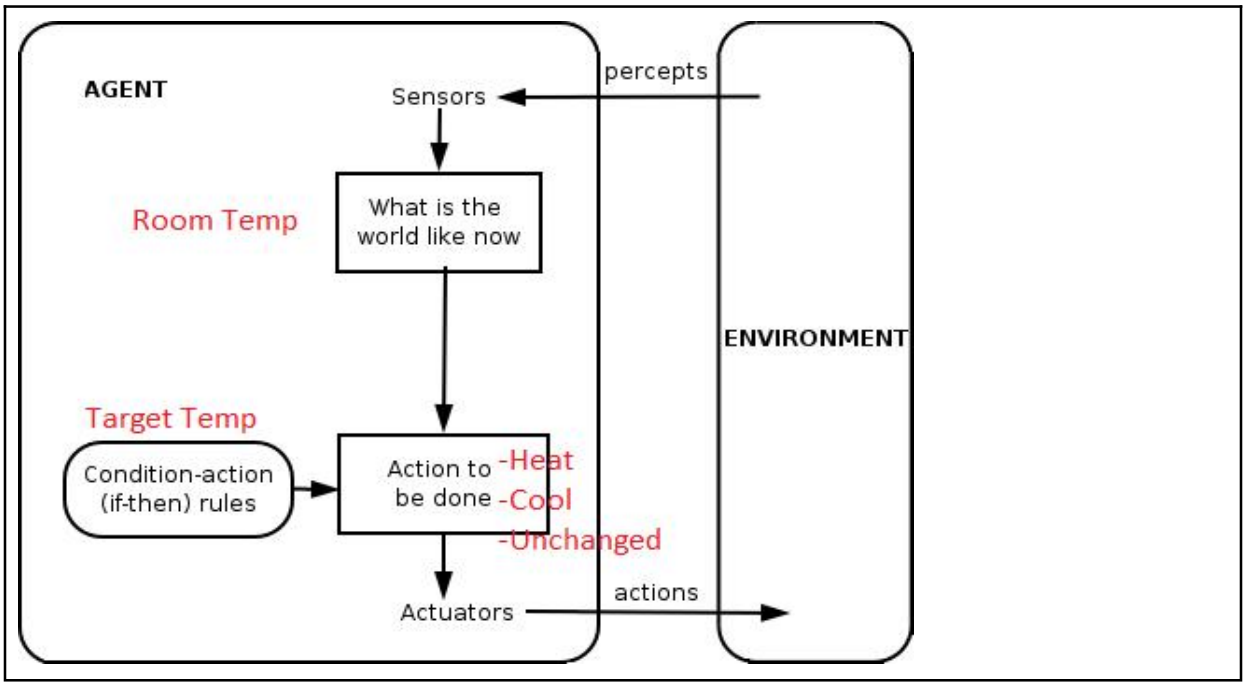
ONG T'NSAM



The system will receive input such as date and time, venue so that it can check the time and venue availability to avoid the class schedule from clashing with another class. After that, it will produce an updated class schedule.

ii. Agent of an intelligent air-conditioning system.

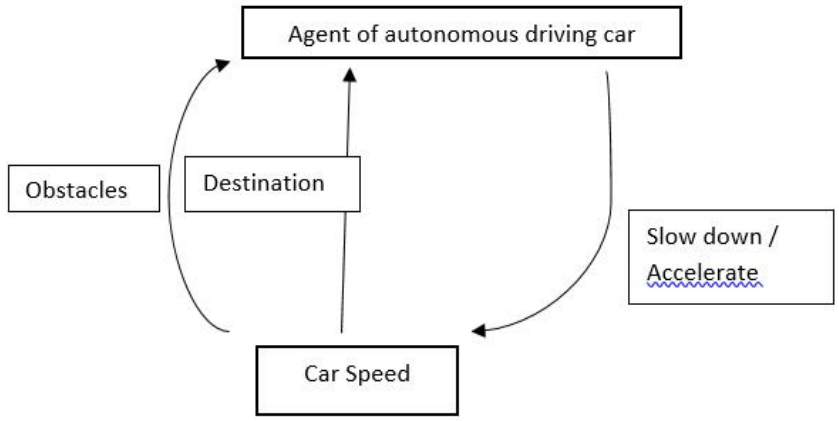
TAN KAI YUAN



Principal Component Analysis (PCA) can be used to identify two significant components

The system will detect the temperature in the environment. If the temperature does not fulfill the target, the system will function to change the temperature in the environment to fulfill the condition. Otherwise, the system will remain unchanged.
✓

iii. Agent of autonomous driving car.



The system will check the surroundings to make sure that there are no obstacles nearby. If there is, it will slow the car down to prevent it from hitting obstacles such as a divider or another vehicle. Otherwise, the agent will maintain the car speed or even accelerate. ✓

Tutorial 2

1. Explain the Turing Test.

TEH CHIN GUAN

A Turing Test is a test that is conducted among the human interrogator, human participants and a computer. It is named after an English computer scientist, Alan Turing. The purpose of this test is to study the behaviour of the computer and determine whether it has the ability to think like a human or not. According to Alan Turing, the computer is said to be intelligent if it passes the test. In order for the computer to pass the test , it must be able to deceive the human interrogator successfully or make itself indistinguishable from the human.

For example, the **human interrogator** will first have a conversation with a male and a female participant. Both the human participants must convince the human interrogator that they are female. The same idea applies to a **human participant and a computer**. Both the human participant and computer must be able to make the human **interrogator think that they are human**. If the **interrogator is not able to determine which is which, then the computer is said to pass the test and is intelligent.**

2. **Criticize** Turing's criteria for judging a computer's intelligence.

TAN KAI YUAN

Software could pass the Turing test simply by manipulating symbols of which they had no understanding. Without understanding, they could not be described as "thinking" in the same sense people are.
The first program with some claim to success was called ELIZA with only a fairly short and simple script, it managed to mislead many people by mimicking a psychologist, encouraging them to talk more and reflecting their own questions back at them.
Another early script PARRY took the opposite approach by imitating a paranoid schizophrenic who kept steering the conversation back to his own preprogrammed obsessions.
Their success in fooling people highlighted one weakness of the test

Ref: https://en.wikipedia.org/wiki/Turing_test#ELIZA_and_PARRY

5 minutes test
Judge
30% chance to deceive the human interrogator

3. Suggest how could this test be used (or modified) to assess other kinds of artificial intelligence besides a chatbot. Provide an example to elaborate your answer.

ONG T'NSAM

Turing test can be used to prevent automated systems from being used to abuse the site. If any software is able to read the distorted image accurately, so any system able to do so is likely to be a human.

Example : Captcha

Can you think of another example based on your creativity?

4. The Loebner Prize is an annual competition in artificial intelligence that awards those computer programs considered by the judges to be the most human-like, using format of a standard Turing Test.

The conversation scope between the programs and the judges has been unrestricted since 1995, and the duration of the conversation has been increased from 5 minutes to 25 minutes since 2010 (http://www.loebner.net/).

(i) Discuss **TWO (2)** reasons why Turing Test is considered **not effective enough** in assessing machine intelligence.

LIM YIH FENG

Some human behaviour is unintelligent

The Turing test requires that the machine be able to execute all human behaviours. It tests for behaviours that may not be considered intelligent at all, such as the susceptibility to insults,the temptation to lie or, simply, a high frequency of typing mistakes. If a machine cannot imitate these unintelligent behaviours in detail it fails the test.

LIM MING JUN

Some intelligent behaviour is inhuman

The Turing test does not test for highly intelligent behaviours, such as the ability to solve difficult problems or come up with original insights. In fact, it specifically requires deception on the part of the machine. If the machine is more intelligent than a human being it must deliberately avoid appearing too intelligent. If it were to solve a computational problem that is practically impossible for a human to solve, then the interrogator would know the program is not human, and the machine would fail the test. Since, it cannot measure intelligence that is beyond the ability of humans, the test cannot be used to build or evaluate systems that are more intelligent than humans.

References:
https://en.wikipedia.org/wiki/Turing_test#cite_note-66

(ii) Discuss **TWO (2)** challenges to build a computer program that can win the Grand Loebner Prize, in which judges totally cannot distinguish it from a real human.

LIM JUN RONG

Data Inaccuracy

Principal Component Analysis (PCA) can be used to identify two significant components

- The Artificial Intelligence (AI) learns from the data integrated by the programmer, unlike human learning. It does not have or have low ability to study the data and determine whether it is correct or not before processing. For example, if the computer programmer inserts incorrect data into the program, the program will just accept, process and store it without second thoughts. It will then generate the information when being asked to do so. As a result, the judge can easily distinguish between the AI and human.

Reference: https://www.scmp.com/business/china-business/article/2131903/biggest-limitation-artificial-intelligence-its-only-smar
t

LIM CHIA CHUNG

First of all, the inherent irrationality of some of the programs is easily misinterpreted in the setting of the test as "whimsical conversation". Moreover, most damaging to the integrity of the test, is the realization that the Turing Test "relies solely on the ability to fool people", and thus is a "sorely inadequate test of intelligence." In this case, certain programs were built to give nonsense responses regardless of the input, but under the guise of "whimsical conversation" its speech patterns can be quite convincing.

5. The Chinese room argument by John Searle is one of the best known and widely credited criticism of the Turing Test. Briefly explain John Searle's Chinese room concept.

LEONG YIT WEE

The conclusion of John Searle’s Chinese room concept is that programming a digital computer may make it appear to understand the language but could not produce a real understanding for the language.

For example , John Searle imagines himself alone in a room and understands nothing of Chinese but by following the computer program he is able to send appropriate strings of Chinese characters back out under the door and people outside the room are mistaken that have a Chinese Speaker inside the

room.

Reference : https://plato.stanford.edu/entries/chinese-room/

6. Try to chat with the following chatbots within a few minutes. Then discuss what are the characteristics/behaviors of a chatbot should have in order to mitsuku/deceive any human.

Mitsuku, the 5-time Loebner Prize winner - https://www.pandorabots.com/

Eliza, the first chatbot - https://web.njit.edu/~ronkowit/eliza.html (not the original Eliza website)

CHIN JUN WAI , CHONG JIA LOONG , KO ZHI XIN , KONG MUN JUN , KOW YEE HUI , LAI PEI XUAN , LAI XIN YI , LEE CHUN XIAN , LEE JUN XIAN , LEE KAH WEI

1. The chatbot should have more knowledge than the human so that it can act like the professional human to suggest or give advice to the human. (LAI PEI XUAN)

2. The chatbot should have replied us with messages that include short forms which most humans are doing nowadays. Furthermore, it should have stated that it does not know the answer when it cannot give us the correct answers instead of changing the topics of the conversation . (CHIN JUN WAI)

3. ChatBot should follow the trend and understand the “Internet Language” and short forms. (Kong MJ)

4. The chatbot will feel the emotion like a human. Some chatbot will respond and care about us like a friend. For example, when the users tell ELIZA about today is sad, then ELIZA will reply “Did you come to me because you are sad?”.(KOW YEE HUI)

5. The ChatBot replies too quickly, which is not possible for a human to do so I try to let her define robots. She literally replied to me 6 lines of definition in a millisecond. Also to some extent she tends to use repetitive sentences in some responses, I send mimicking human emotion symbols like “<3” and “XD” she replies to me with “Are you artistic? That looks like a heart.” and “Are you artistic? That looks like someone laughing.” (Lee Jun Xian)

6. The information provided by the ChatBot should be accurate and timely. For example, I asked about the information for who is the richest person in the world currently, then it replied Bill Gates but currently the founder of Amazon Jeff Bezos is the richest person in the world. This has proven that the information retrieved by the ChatBot should be improved to be more accurate as and information available people would like to know the current news. (Chong Jia Loong)

7. The chatbot should need some memory space to keep our conservation record like a human brain so it will not forget if we accidentally close the chat. Therefore, we do not need to inform the chatbot again. (Lai Xin Yi)

8. Chatbot does not entirely feel like chatting in front of a human because they cannot simplify that what we do usually using those abbreviation verbs such as expo stands for exposition. Hence, they will reply with something irrelevant to your topic. (Lee Chun Xian)

9. The chatbot should have many types of responses when chatting with real humans and respond to them depending on humans' emotional state. For example, chatbot would like add some kinds of advertisement in case of its owner is happy because when human is happy, he/she would buy some kind of things that he/she wants very bad for example, the things/commodities he/she searches many times either in the shopping app or google.(Lee Kah Wei)

10.The chatbot should have logical responses. To ensure this human-like interaction, it is necessary to use a particular tone or dialect when chatting. Human should feel like they are having a real world conversation. The chatbot should ask human questions that are relevant to their preferences and needs. (Ko Zhi Xin)

Principal Component Analysis (PCA) can be used to identify two significant components

Tutorial 3

1. Discuss the advantages and disadvantages of breadth-first search and depth-first search.

CHIN JUN WAI

	breadth-first search	**depth-first search**
Advantages	**Completeness** It guarantees a solution. **Optimality** Always find the shortest path between initial state and goal state with least number of steps. ✓	**Completeness** It guarantees a solution. **Space complexity (Advantage)** Low consumption in memory because it stores the parent nodes and finds the goal from its children. When the goal is not found, it can trace back their parents and continue exploring until the goal is found.
Disadvantages	**Time complexity** Need longer time to get the goal if the goal state is on the far deeper level. **Space complexity** Need to expand all branches before proceed to next level and the memory consumption is high ✓	**Time complexity** Depends on the depth level but if the depth level for both of the search is the same, depth-first search is faster **Optimality (Disadvantage)** Always expand nodes to a deeper level until a goal is found. If it reaches the dead end, it will trace back to the previous node and continue expanding. The goal it returns might be not the shortest path.

2. Figure 1 below shows a puzzle problem, which requires rearrangement of the tiles to transform the order from start state to goal state. One is only permitted to slide a tile **left, right, up or down** into the blank square.

1	**2**			**2**	**3**
	3	**Start**	**Goal**	**1**	

Figure 1: The puzzle problem

(i) Provide the goal formulation of the puzzle problem above.

CHONG JIA LOONG

The goal is to move the tile's state with the least amount of steps from state[1,2,0,3] to goal[2,3,1,0].
✓

Principal Component Analysis (PCA) can be used to identify two significant components

Optimal solution? Reach a solution with the least amount of steps.
Abstraction? By moving the tiles left, right, up or down.

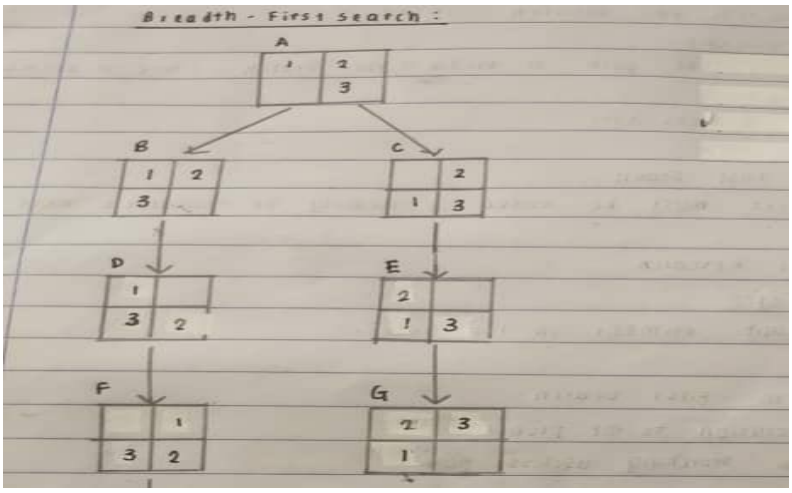
(ii) Formulate the puzzle problem above by specifying the initial state, successor functions, goal test, step cost, and path cost.

KONG MUN JUN

Initial State = [1,2,0,3]
Successor Func. = Move[1, down]; Move[2, left], Move[3, Up]
Goal Test = goal[2,3,1,0] a function to test whether a state is a goal state. If it is the goal, then return path, else continue
Step Cost = 1
Path Cost: 3 summation of the step cost of a path
:Move[1, down]; Move[2, left], Move[3, Up]

(iii) Perform **breadth-first search** and **depth-first search** on the puzzle problem above. Draw the resulting trees for both.

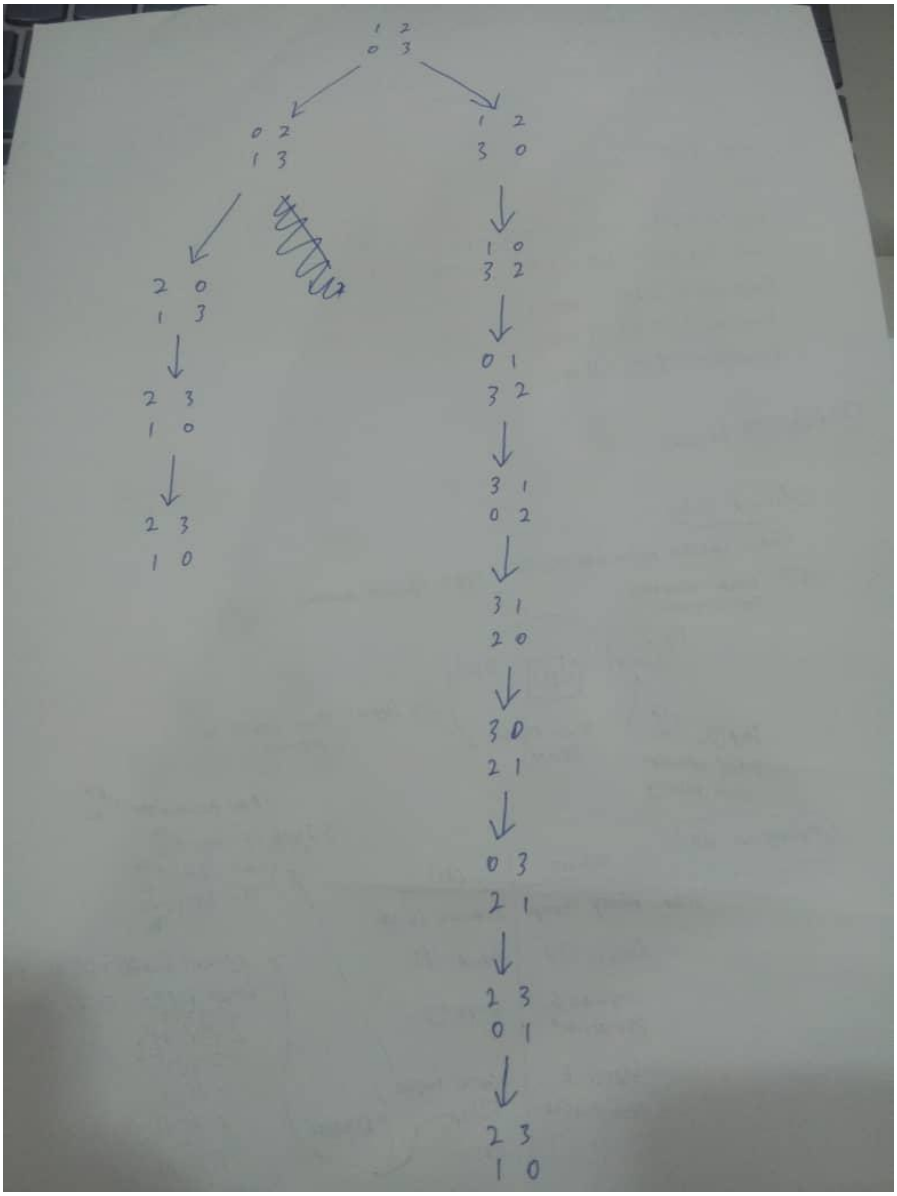
KOW YEE HUI (breadth-first)

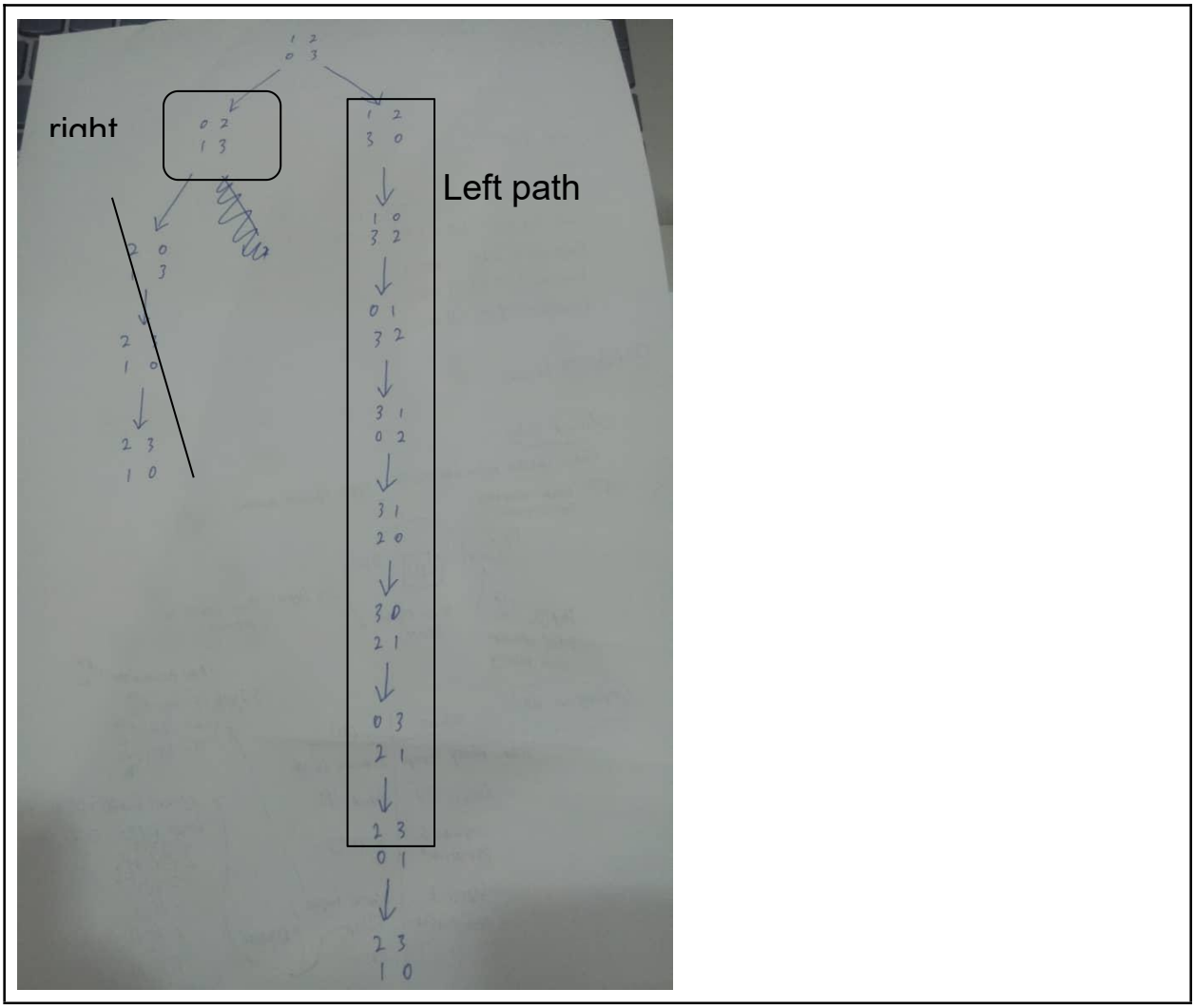


Breadth-first search usually stops once it found the goal
So ms, it is until f and g only ? yes
Okay thank you ms, now the above one is true arldy?✓ok,thx miss

LIM MING JUN (depth-first)

Principal Component Analysis (PCA) can be used to identify two significant components





3. Figure 1 below shows an 8-puzzle problem, which requires rearrangement of the tiles to transform the order from start state to goal state. One is only permitted to slide a tile **up, down, left or right** into the blank square.

1	**2**	**3**
	5	**6**
4	**7**	**8**

Start

Goal

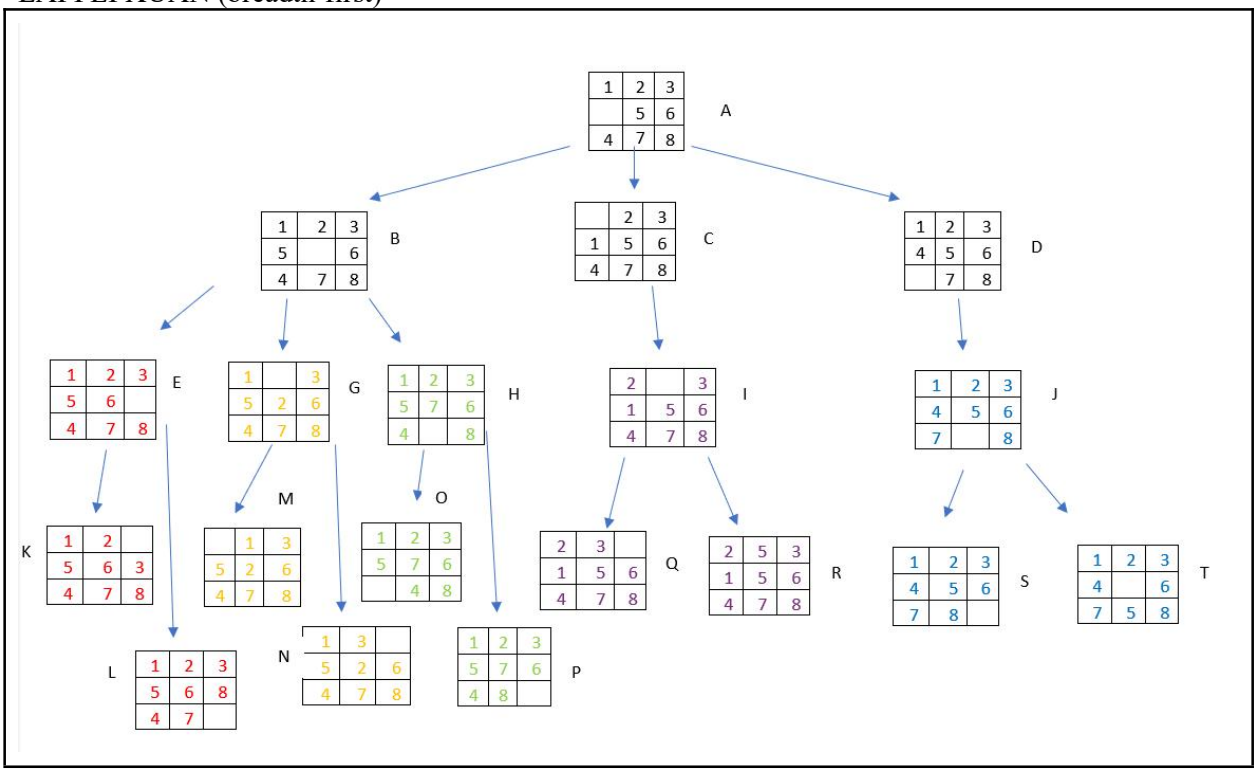
1	**2**	**3**
4	**5**	**6**
7	**8**	

Figure 2: The 8-puzzle problem

Principal Component Analysis (PCA) can be used to identify two significant components

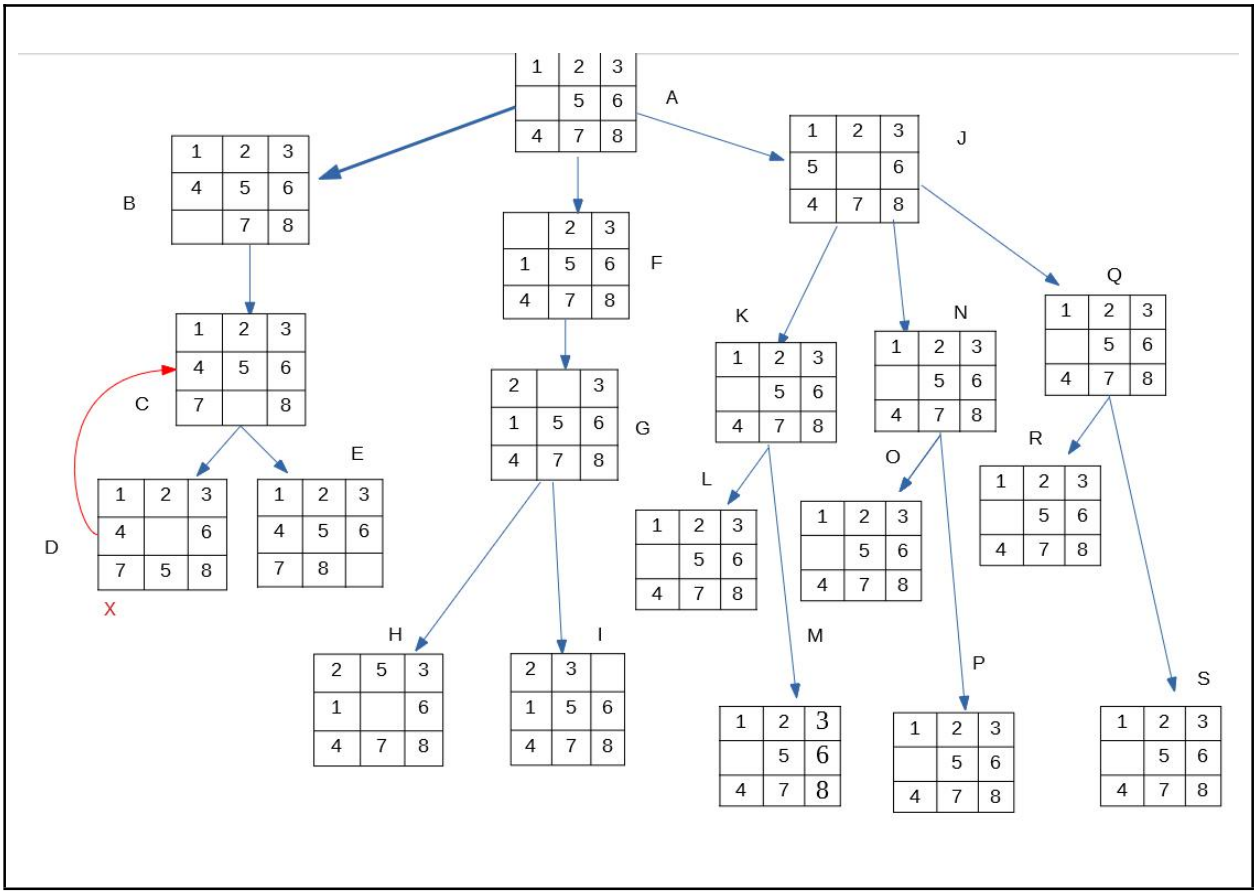
(i) Perform **breadth-first search** and **depth-first search** on the 8-puzzle problem above. Draw the resulting trees for both. (Remark: you may stop the search at level 4)

LAI PEI XUAN (breadth-first)



LIM YIH FENG (depth-first)

Principal Component Analysis (PCA) can be used to identify two significant components



(ii) Evaluate the efficiency of **breadth-first search** and **depth-first search** in terms of completeness, optimality, time efficiency and space efficiency in solving the problem above.

LAI XIN YI

Breadth-first search

- ***Completeness*** : BFS will provide a solution
- ***Optimality*** : BFS will provide the optimal solution which requires the least number of steps.
- ***Time efficiency*** : BFS needs a lots of time if the solution is far away from the root node
- ***Space efficiency*** : It requires a lots of memory since each level of the tree must be saved into memory to expand the next level

✓

Depth-first search

- ***Completeness :*** DFS will find the solution in end
- ***Optimality :*** DFS is non-optimal, as it may generate a large number of steps or high cost to reach to the goal node.

Principal Component Analysis (PCA) can be used to identify two significant components

- ***Time efficiency :*** DFS takes less time to reach to the goal node than BFS
- ***Space efficiency :*** DFS requires very less memory as it only needs to store a stack of the nodes on the path from root node to the current node

✓

4. In the family river-crossing problem, two parents are with their two children - a son and a daughter - came to a wide river. The only way to get to the other side was to ask a fisherman if he could lend them his boat. However, the boat could carry only two persons. For safety reason, no child should be left alone without the supervision of at least one parent. The family must get to the other side and finally returns the boat to the fisherman, assuming only the fisherman and the two parents know how to row the boat.

   Based on the family river-crossing problem, answer the following questions.

   (a) Describe the goal formulation and problem formulation.

LEE CHUN XIAN (goal formulation)

Goal - The entire family across the river and return the boat to the fisherman.
There are 3 items in a goal formulation
Optimal solution - Least number of trip to across the river
abstraction - A boat can carry only two persons

ONG T'NSAM (problem formulation)

How can the family get to the other side and return the boat to the fisherman?

Correction:
Initial state:
-the number of parents on the right
-the number of children on the right
-the side of the boat is on the right

Successor function:
-the possible moves for the family to get to the other side

Goal test:
-To check whether the family get to the other side and finally return the boat to the fisherman, and no child should be left alone without the supervision of at least one parent

Cost:
Step cost = 1
Path cost = total number of travel

Principal Component Analysis (PCA) can be used to identify two significant components

(b) (i) Suggest a simple representation of the initial state. You must briefly explain the representation.

LEE JUN XIAN

Representation: State(Num of Parents on Left,Num of Children on Left ,Num of Fisherman on Left, Direction, Num of Parents on Right, Num of Children on Right ,Num of Fisherman on Right)
Initial State: State(0,0,0,R,2,2,1)✓

For the state function, the first three parameters represent the number of peoples on the left side of the river, the fourth parameter represent which side the boat is at and the last three parameters represent the number of peoples on the left.

TAN KAI YUAN

[FL,PL,CL,B,FR,PR,CR]
state([0,0,0,right,1,2,2])
F represents Fisherman, P represents parents and C represents children
B represents the riverside boat is on while L and R represents Left and Right
✓

(ii) Draw the depth-first search tree to show how all the states are being traversed.

LEE KAH WEI

5. Figure 2 below shows a directed graph. Assume that the traversal would start from **Vertex 0 to Vertex 7**. All vertices to be visited in **ascending order** (i.e. from smaller number to bigger number).

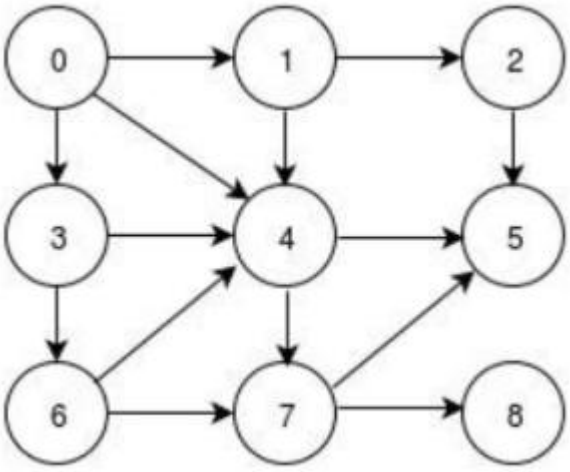
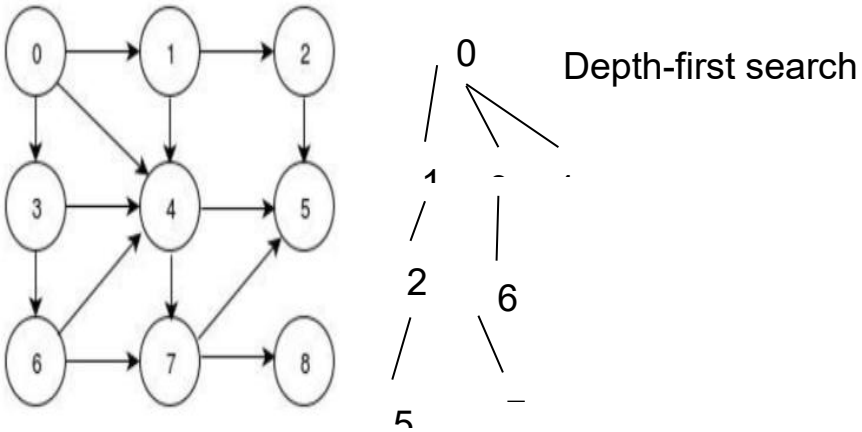


Figure 2. The directed graph

(i) Perform a **depth-first search** on the directed graph to traverse from Vertex 0 to Vertex 7. Draw the resulting tree and list the returned path.

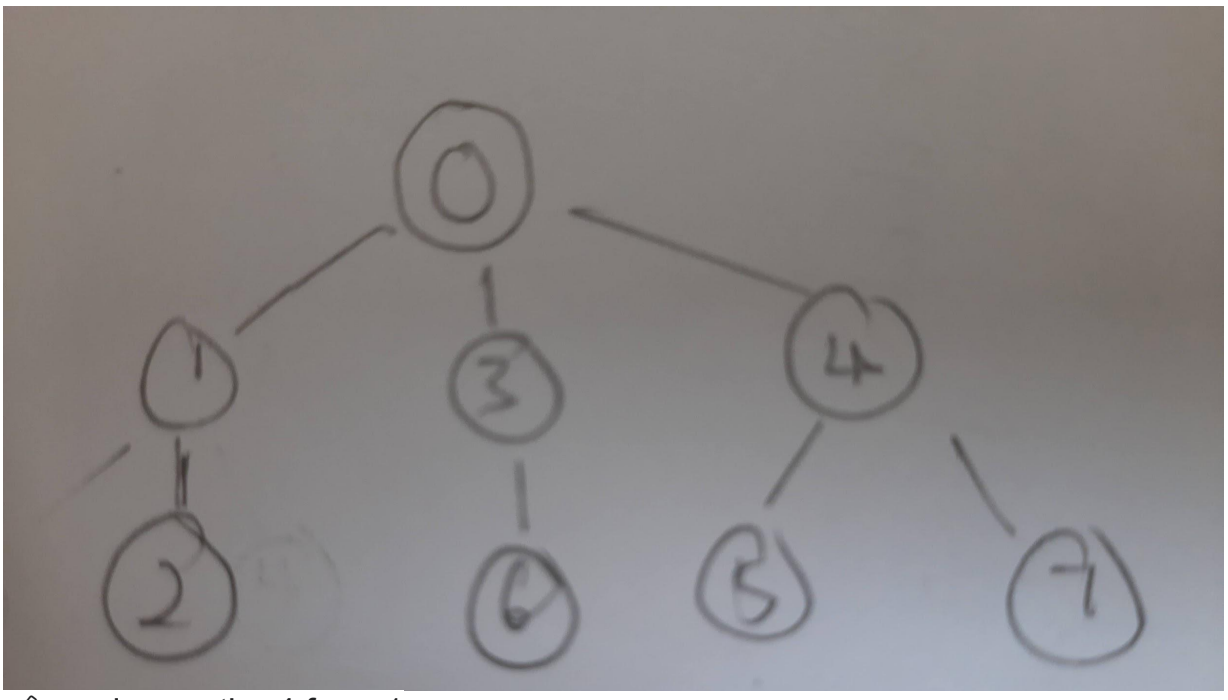
LEONG YIT WEE



(ii) Perform a **breadth-first search** on the directed graph to traverse from Vertex 0 to Vertex 7. Draw the resulting tree and list the returned path. Avoid repeated state.

LIM CHIA CHUNG

Principal Component Analysis (PCA) can be used to identify two significant components



✓ can ignore the 4 from 1

(iii) Evaluate the efficiency of **breadth-first search** and **depth-first search** in terms of completeness, optimality, time efficiency and space efficiency in solving the problem above.

LIM JUN RONG

- Breadth-first search (BFS)
  - Completeness : Solution is found in the end.
  - Optimality : BFS will always look for the optimal solution by identifying the shortest path between the initial state and the solution/goal, with the least number of steps needed.
  - Time efficiency : BFS takes longer time to reach the solution/goal.
  - Space efficiency : BFS consumes large memory space as it needs to expand all the branches before going to the next level.

✓

- Depth-First search (DFS)
  - Completeness : Solution is found in the end.
  - Optimality : DFS may not find the optimal solution to the problem,
  - Time efficiency : Less time is needed to reach the solution/goal
  - Space efficiency : DFS consumes very less memory space

✓

Principal Component Analysis (PCA) can be used to identify two significant components

Tutorial 4

Q1. Figure 1.1 below shows a Block World Problem. A robot will move the blocks one by one from initial state S1 to reach the goal state S19. Figure 1.2 shows the state space of the Block World Problem and the heuristic costs for each state are shown in parentheses next to their respective nodes.

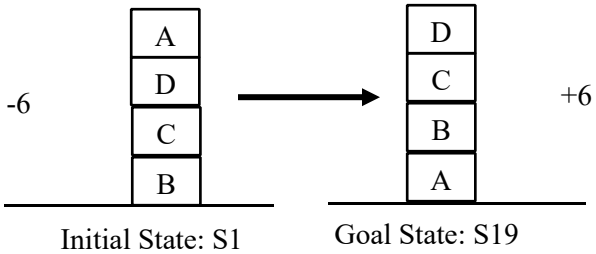


Figure 1.1

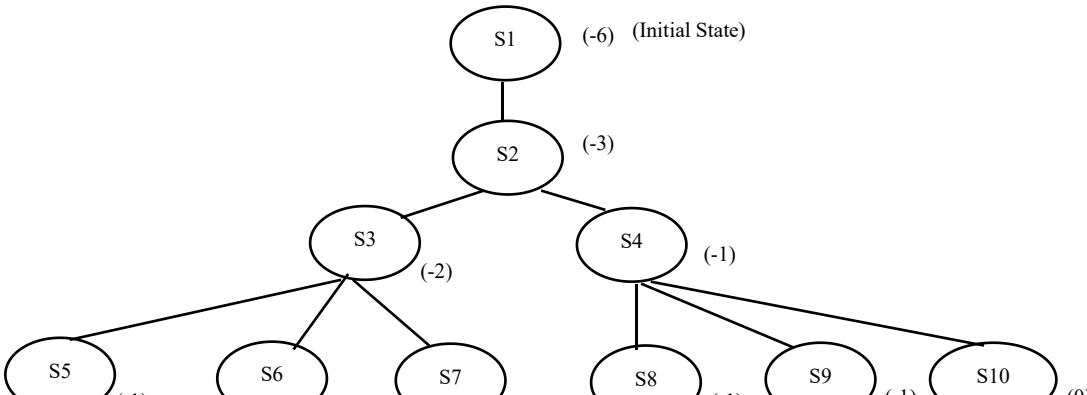


Figure 1.2

(a) Explain **step cost** used in problem formulation. Specify the value of the step cost for the problem above.

CHIN JUN WAI

Step cost = move one block from one location to another (it is either stacked on the other blocks or on an empty platform)

The value of the step cost is 1

(b) Hill climbing search is unable to guarantee completeness and optimality as it may be trapped into local maximum.

(i) Explain local maximum.

CHONG JIA LOONG

A state that child nodes are unable to explore further and the child nodes are non better than the parent nodes

(ii) Discuss why hill climbing search always leads to a local maximum.

LEE JUN XIAN

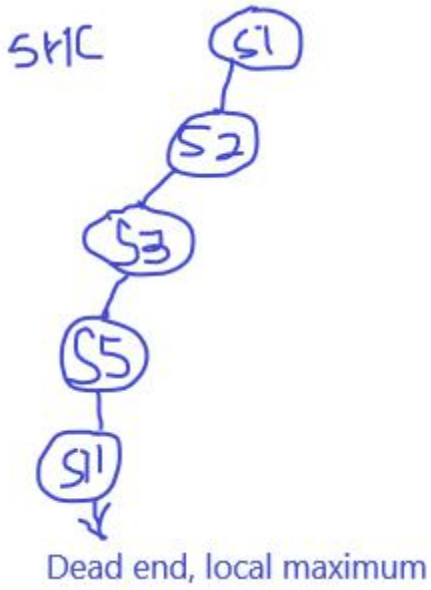
Principal Component Analysis (PCA) can be used to identify two significant components

-The hill climbing search always finds a state which is best but it ends in local maximum due to the neighbouring states having worse values compared to the current state.
-It does not allow backtracking.

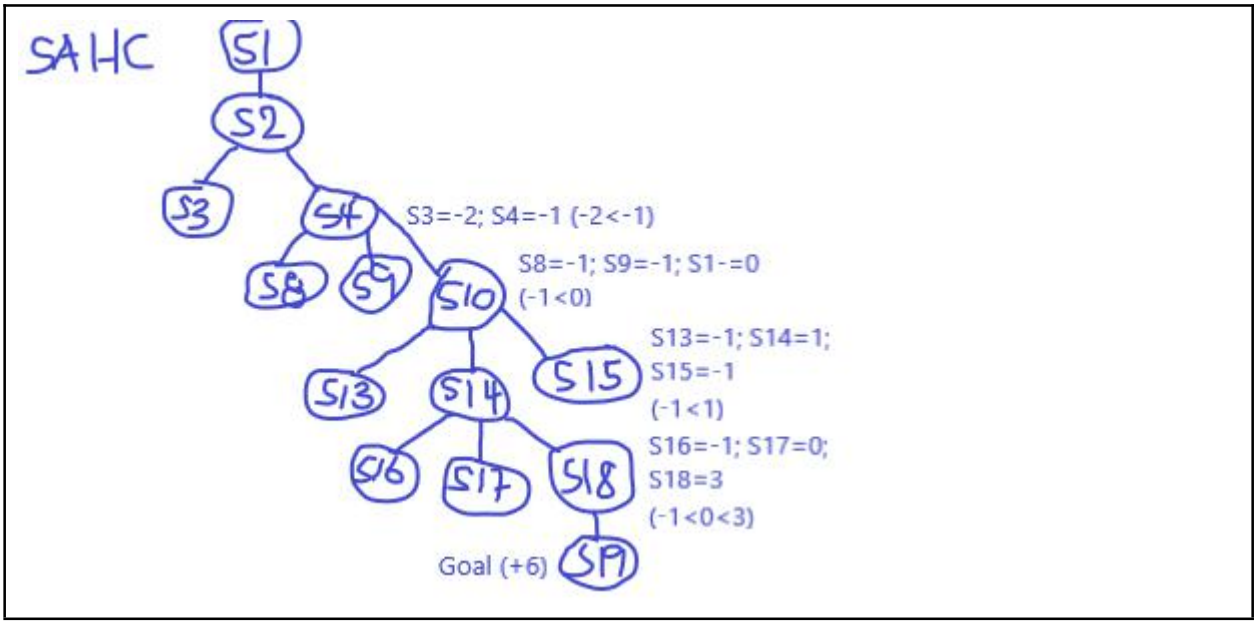
(iii) Use simple hill-climbing and steepest-ascent hill-climbing to search for the best path from S1 to S19 on the state space shown in Figure 1.2. Then for each search technique, **draw the resulting search tree** that shows the visited nodes. **Show that hill-climbing technique can be trapped into a local maximum.**

KONG MUN JUN

Best path {S1, S2,S4,S10,S14,S18,S19}



Principal Component Analysis (PCA) can be used to identify two significant components



(c) A search technique can be evaluated based on four criteria: completeness, optimality, time complexity and space complexity. Evaluate the efficiency of breadth-first search and steepest-ascent hill-climbing. Conclude which technique is better to solve the Block World Problem mentioned in Figure 1.1.

LAI XIN YI

Breadth-first search (BFS)

- Completeness : BFS guarantees a solution in the end
- Optimality : BFS will find the shortest path with the least number of the steps
- Time complexity : BFS will take longer to reach the goal
- Space complexity : BFS needs large memory space because it needs to expand all the states before going to the next state.

✓

Steepest ascent hill climbing (SAHC)

- Completeness : SAHC will not guarantees that it will find the solution in end
- Optimality : SAHC will lead to a suboptimal solution (longer path)
- Time complexity : SAHC will take shorter time compare with BFS because it does not need to explore all the path
- Space complexity : SAHC use less memory space than BFS

Breadth first solution is the better technique.

✓

Q2. Figure 2 below shows an 8-puzzle problem, which requires rearrangement of the tiles to transform the order from start state to goal state. One is only permitted to slide a tile **up, down, left or right** into the blank square.

Principal Component Analysis (PCA) can be used to identify two significant components

1	**2**	**3**
	4	**5**
7	**8**	**6**

Start

Goal

1	**2**	**3**
4	**5**	**6**
7	**8**	

Figure 2: The 8-puzzle problem

(i) Suggest a heuristic function to produce a heuristic cost for a state. Demonstrate how such heuristic cost can be computed on the **start state.** Then perform best-first search.

LEE KAH WEI

h(n)= number of tile(s) at the right place

(ii) Evaluate the efficiency of **breadth-first search** and **best-first search** in terms of completeness, optimality, time efficiency and space efficiency in solving the problem above.

LIM CHIA CHUNG

- Breadth-first search
  - Completeness: It will always reach a goal.

- ➢ Optimality: It will always find the less optimal solution between initial state and goal state, with the least number of steps.
- ➢ Time efficiency: It takes longer to reach the goal.
- ➢ Space efficiency: It keeps nodes in memory to expand all the branches before proceeding to the next level.

❖ Best-first search

- ➢ Completeness: It always selects the solution which appears at that moment.
- ➢ Optimality: It will always find the optimal solution depending on heuristic function and counting steps.
- ➢ Time efficiency: It takes less time to reach the goal.
- ➢ Space efficiency: It uses less memory space because it does not explore all paths.

(iii) Explain why hill-climbing will fail in this problem.

LIM MING JUN

This is because we will encounter local maximum if we using hill-climbing to search.
Cannot backtracking

Q3. The following graph in Figure 1.1 shows all the nodes in a telecommunication network. The distance (in km) from one node to another is shown on the arc.

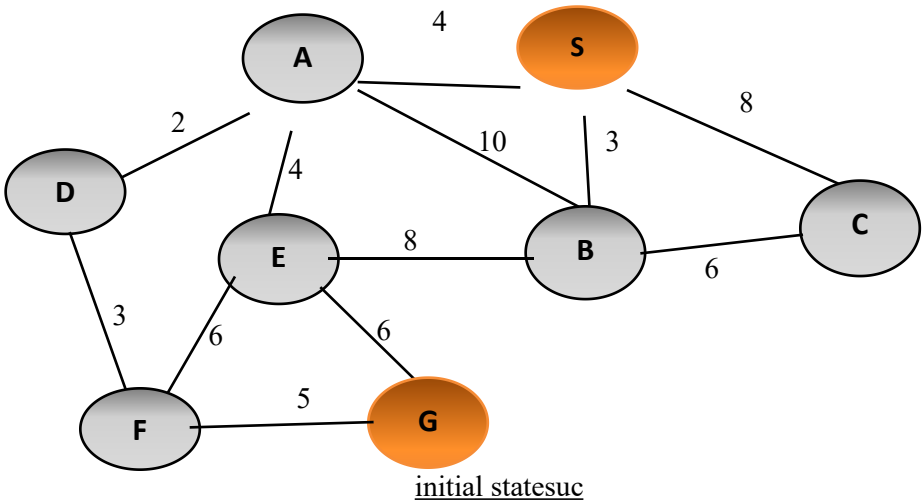


Figure 3 the search graph of a new LRT network

The Euclidean distance (in km), which is used as the heuristic cost (h) for different node, is provided in Table 1 below.

Table 1: The heuristic costs for different nodes

S	A	B	C	D	E	F	G

Principal Component Analysis (PCA) can be used to identify two significant components

14	10	13	8	8	11	5	0

(i) Assume that some data are to be sent from node S to node G using the shortest route. Describe the goal formulation and problem formulation.

LIM YIH FENG

Goal - From node S to node G
Optimal solution - Shortest distance to reach node G
abstraction -can't find answer

Initial state:
-heuristic cost = h(14)

Successor function:
-heuristic cost - distance

Goal test:
-h(0)

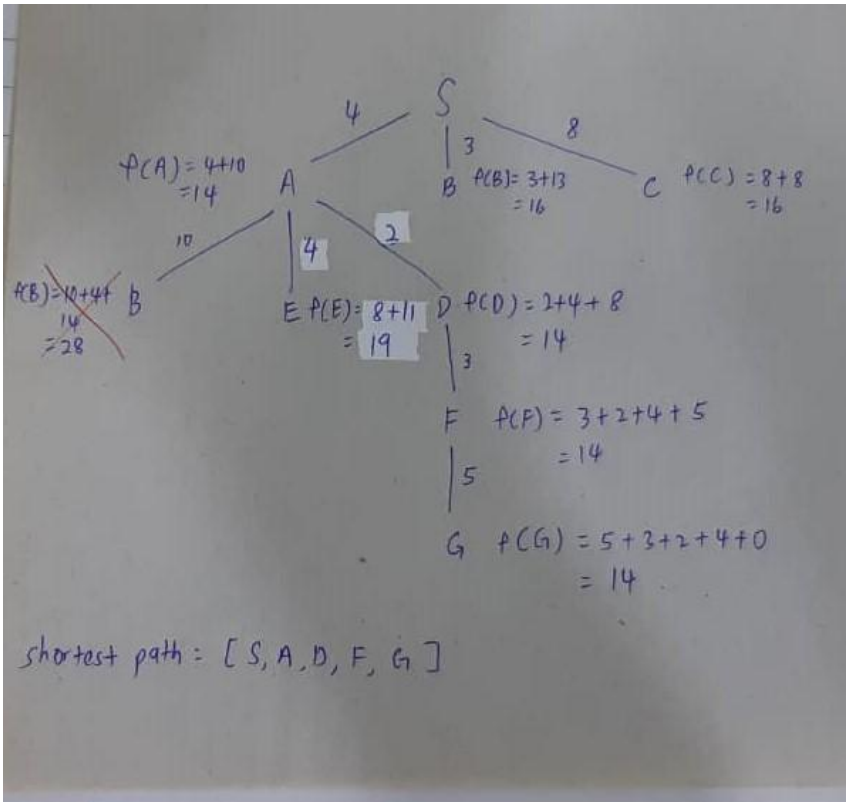
Cost:
Step cost = heuristic cost per step
Path cost =total distance travelled

This question is similar to Tutorial 3, Chapter 2

(ii) **Show the resulting search tree** of A* search to find the shortest path from S to G. **State the shortest path**.
(Remark: Ignore repeated nodes that have been visited previously)

ONG T'NSAM

Principal Component Analysis (PCA) can be used to identify two significant components



Shortest path : [S,A,D,F,G]

(iii) Evaluate the efficiency of A* search in solving the path-finding problem above.

TAN KAI YUAN

A* efficiency is high because the step cost is different, therefore the solution it returns will be the best route with shortest distance

Answer this based on the 4 criteria

Q4. Consider 2 heuristic h_1 and h_2 of A* for the puzzle problem are defined as:

$h_1(n)$= number of misplaced tiles
$h_2(n)$= total Manhattan distance

3	1

1	2

Principal Component Analysis (PCA) can be used to identify two significant components

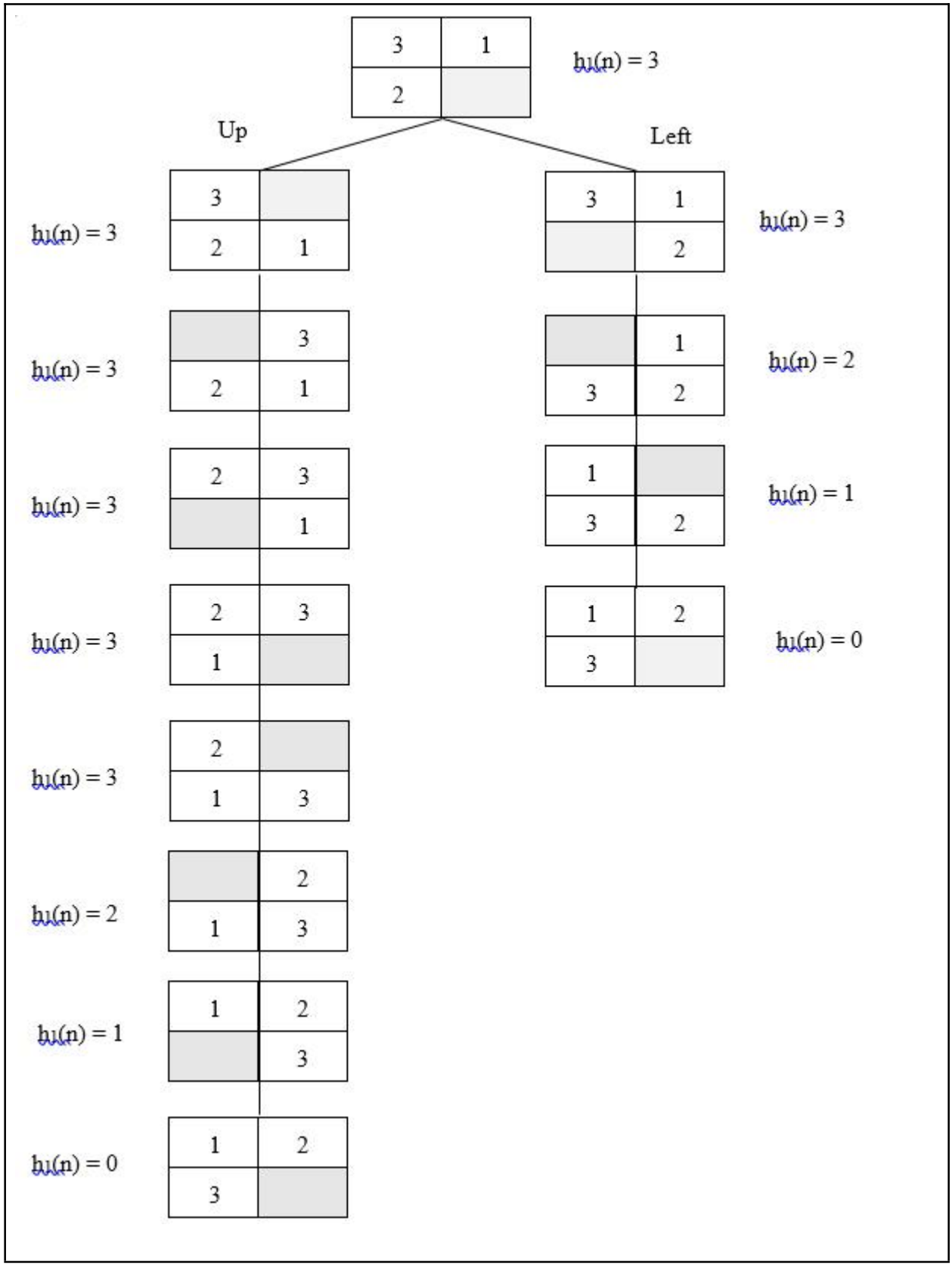
2		Start State	Goal State	3	

(i) Illustrate the **state space** of the puzzle to reach the goal state based on:

- $h_1(n)$
- $h_2(n)$

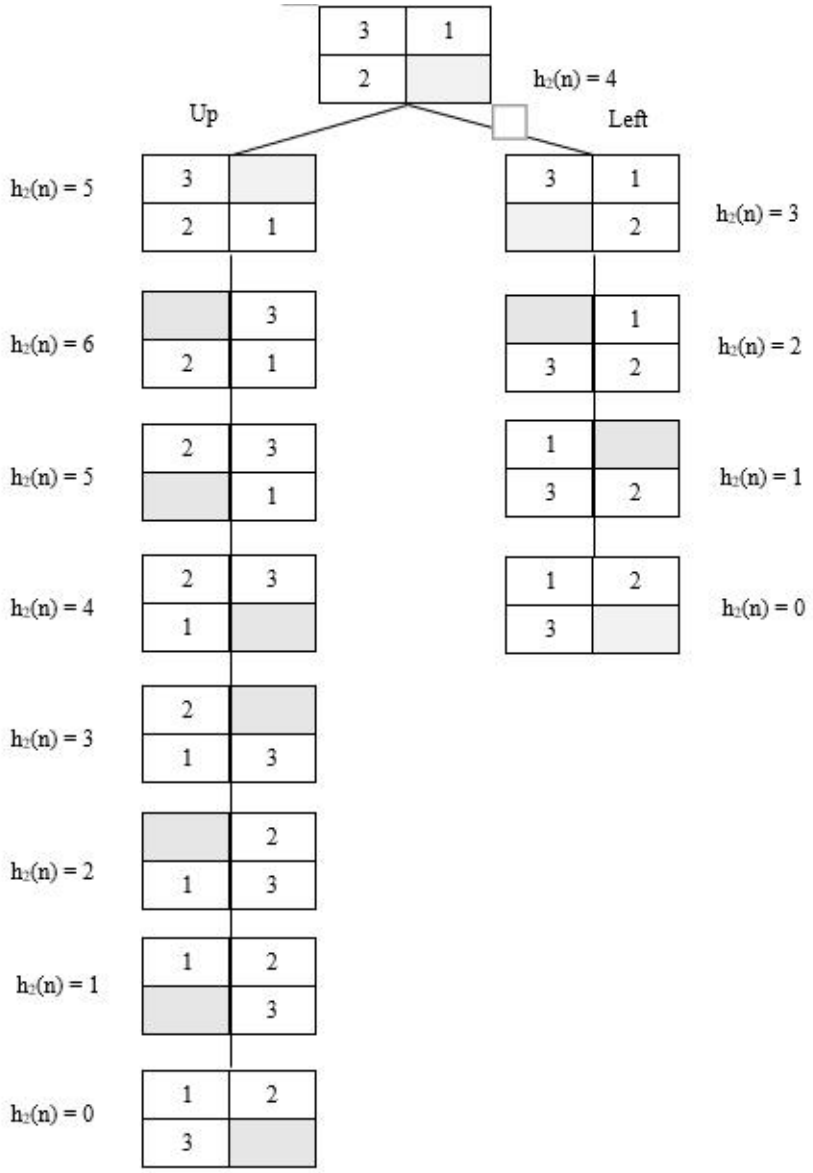
LIM JUN RONG (h1(n))

Principal Component Analysis (PCA) can be used to identify two significant components



Principal Component Analysis (PCA) can be used to identify two significant components

LEE CHUN XIAN (h2(n))



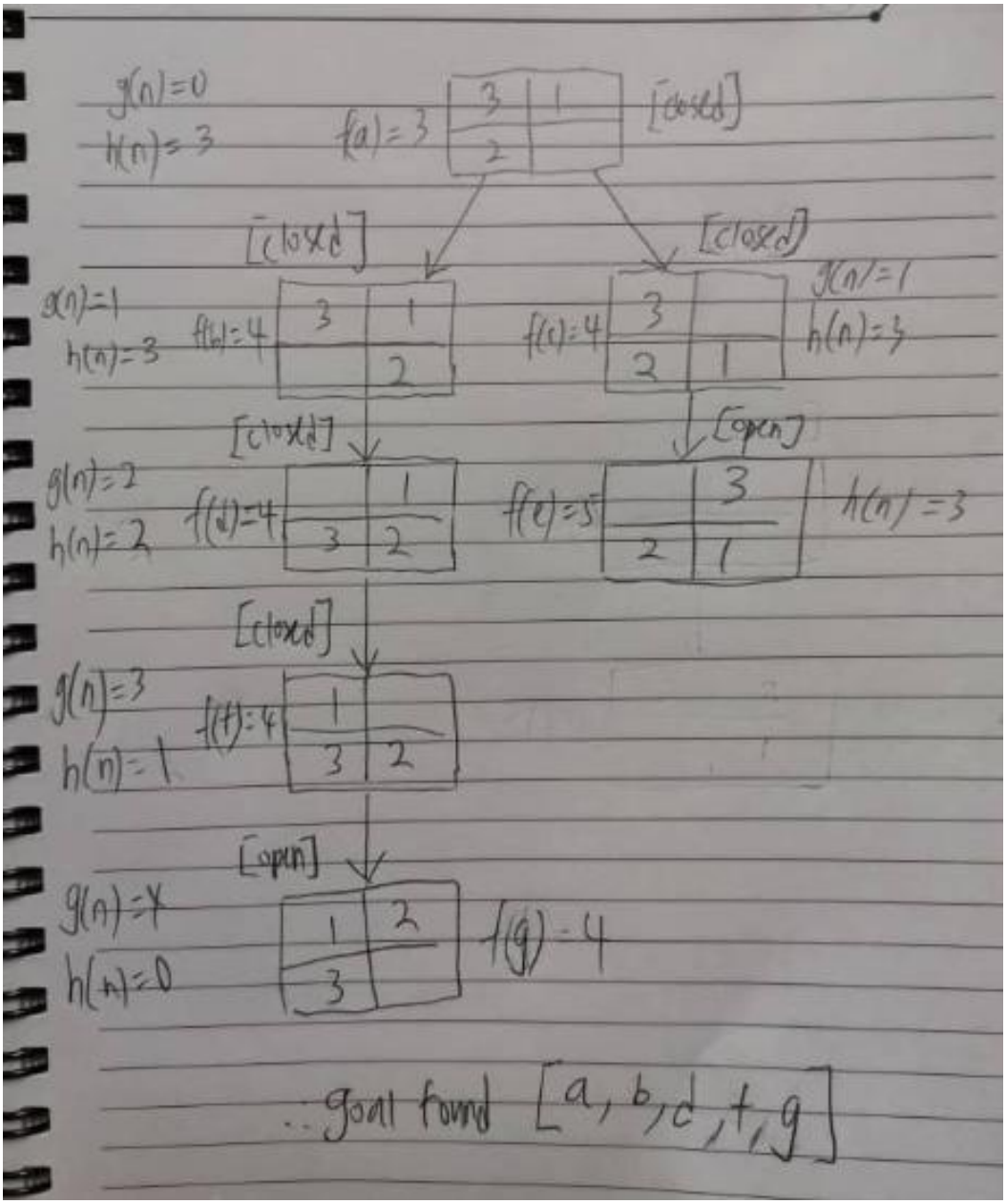
(ii) **Show the resulting search trees** of A* search to find the shortest path using the heuristic functions of:

- $h1(n)$
- $h2(n)$

Principal Component Analysis (PCA) can be used to identify two significant components

You must clearly show the function cost, given that:
$f(n) = h(n) + g(n)$, where $g(n)$ is the path cost.

LEONG YIT WEE (h1(n))



KOW YEE HUI (h2(n))

Principal Component Analysis (PCA) can be used to identify two significant components