University of Toronto at Mississauga
Department of Mathematics and Computer Science

CSC 384 H5F: NTRODUCTION TO ARTIFICIAL INTELLIGENCE

Fall 2005 Final Exam

Examiner: Steve Engels
Duration: 3 Hours

Aids: This exam allows the use of the following:

- A single-sided 8½″x11″ handwritten study sheet
- a non-programmable calculator.

Use of any other computing and/or communicating devices is NOT permitted.

Other Instructions:

Do not remove any sheets from this test book. Answer all questions in the space provided. No additional sheets are permitted.

Do not begin writing until told by your instructor. Keep all bags and notes far from the desk, with only the permitted aids visible.

Write your name and student number in the space below, and at the top of each sheet of this exam book. Make sure to always write legibly.

Name: ________________________________
(Underline last name)

Student Number: ________________________________

Q1. __________ / 24
Q2. __________ / 8
Q3. __________ / 10
Q4. __________ / 12
Q5. __________ / 10
Q6. __________ / 6

Q7. __________ / 11
Q8. __________ / 9
Q9. __________ / 10
Bonus. ________ / 3

Total ☐ /100

Question 1. (24 marks). *General*.

Answer the following questions by providing a **very brief and direct** answer.

1) **(2 marks)** A cleaning agent goes from room to room, making sure not to clean the same room twice. What type of agent is this?

Model-based agent

2) **(2 marks)** True or False? The number of possible states in the logistics domain is O(2^n), where `n` in the number of packages.

[✓ True] False

3) **(2 marks)** If a multi-agent system has n agents, each of which has m possible moves, the search space increases to how many possible states?

a) m^n-mn [✓ b) m^n]

c) n^m-nm d) none of the above

4) **(2 marks)** True or False? Two agents work better than one.

True [✓ False]

5) **(2 marks)** True or False? When searching through the state space of a Minesweeper game, the solution can be found faster by using alpha-beta pruning.

True [✓ False]

6) **(2 marks)** Using big-oh notation, what is the time usage of the Viterbi algorithm for finding the most likely state sequence, given `N` possible states and `M` emissions?

$$O(M*N)$$

7) **(2 marks)** A game tree uses alpha-beta pruning and has a branching factor `b` and a search horizon depth `d`. Using big-oh notation, what is the space usage of this game tree?

$$O(\sqrt{b^d})$$

8) **(10 marks)** Match the following terms to the AI domain on the right by writing the domain’s letter in the blank next to the term.

7 Let’s Make A Deal

9 Viterbi algorithm

8 Anaphora

10 Overfitting

5 First-order logic

11 Perceptrons

2 Manhattan heuristic

4 Information gain

3 Nash equilibrium

6 STRIPS

1. Agents

2. Searching

3. Game Theory

4. Decision Trees

5. Reasoning

6. Planning

7. Probability

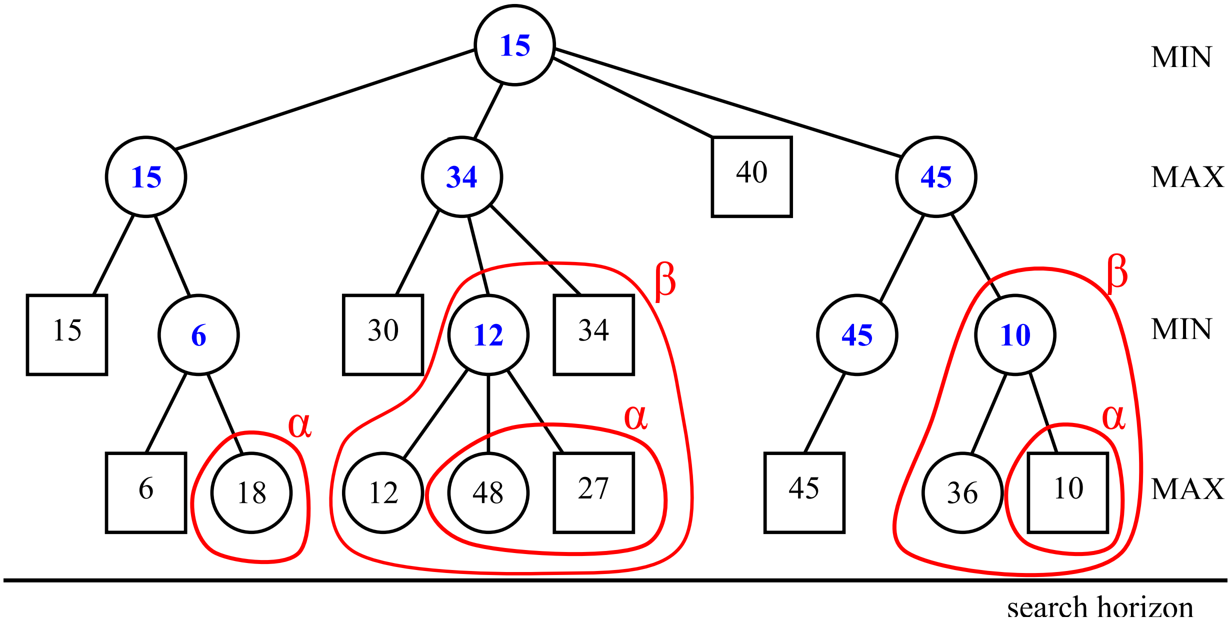
8. Natural Language

9. Hidden Markov Models

10. Learning

11. Neural Networks

Question 2. (8 marks). *Game Trees*.



a) (4 marks) Evaluate and fill the heuristic values for all the empty states in the game tree above. Assume that the minimax algorithm is being used, according to the labels on the right.

b) (3 marks) On the same diagram above, indicate which states will **not** be explored if alpha-beta pruning is used. Circle all unvisited subtrees, and indicate next to them whether alpha-pruning or beta-pruning was used by writing '*α*' or '*β*' next to the state. You may assume that the branches are explored from left to right.

c) (1 mark) Assuming that the game tree is only capable of searching N states. If this game tree has a branching factor B and a search horizon of depth D without alpha-beta pruning, how deep could the search horizon be *with* alpha-beta pruning?

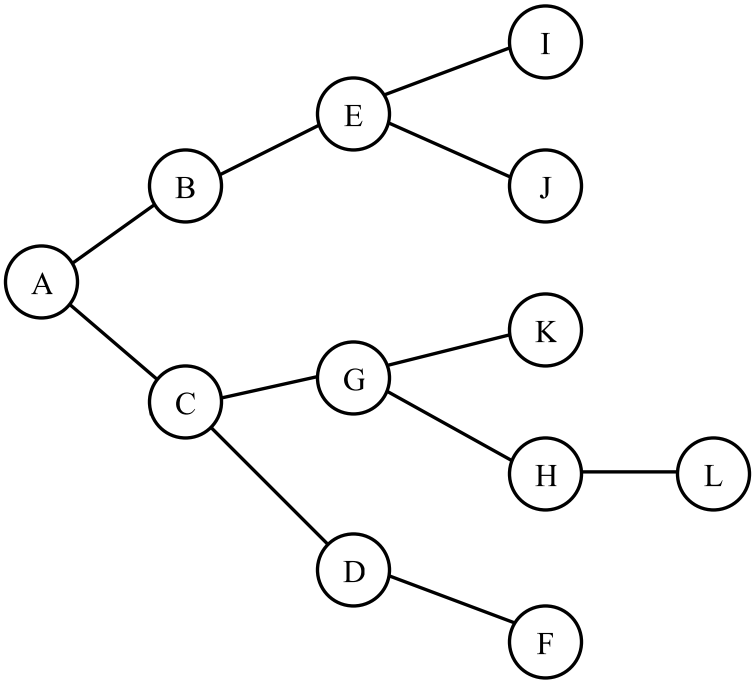
2*D

Question 3. (10 marks). *Searching*.

The table below lists the states in a search domain, along with the transitions available from that state. For this question, assume that the start state is always **`A`** and the goal state is always **`K`**.

A → B, C
B → E
C → D, G
D → F
E → I, J
F →
G → H, K
H → L
I →
J → K
K →
L →

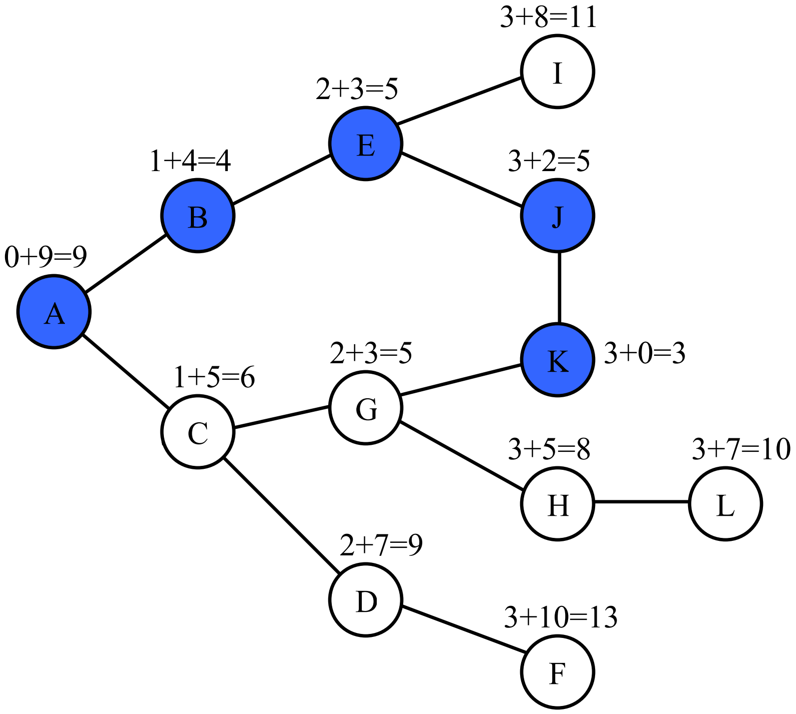
a) (4 marks) Assuming that there is a cost of 1 (unit cost) to move from one state to another, draw a diagram in the space below that illustrates the search tree of explored states, given a breadth-first search.



Question 3. (continued).

```
A:   9            G:   3
B:   4            H:   5
C:   5            I:   8
D:   7            J:   2
E:   3            K:   0
F:   10           L:   7
```

b) (5 marks) The heuristic values for each state are shown above. Given these values, draw a diagram in the space below that illustrates the search tree of explored states, given an A* search (the path cost is still a unit cost). Indicate the calculated cost at each node in the tree.



c) (1 mark) What would the **minimum** uniform path cost between states have to be for the heuristic shown above to be admissible?

3

Question 4. (12 marks). *Reasoning*.

a) (2 marks) Circle the two expressions below that are logically equivalent.

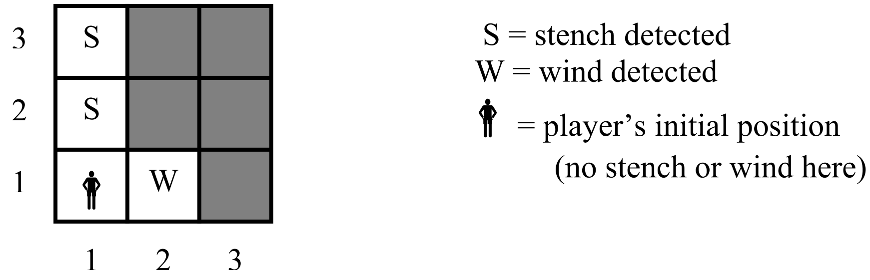
[✓ i) $(P \vee Q) \Rightarrow \neg R$]

ii) $(R \wedge P) \vee (R \wedge Q)$

iii) $\neg R \Rightarrow (P \vee Q)$

[✓ iv) $R \vee (\neg P \wedge \neg Q \wedge \neg R)$]

b) (6 marks) .Consider the Wumpus World domain, displayed in the diagram below:



The light squares have been explored by the player, the dark squares have not. Given that a stench indicates the presence of a Wumpus in an adjacent square (diagonal squares included) and a wind indicates the presence of a pit in an adjacent square, use first-order logic to do the following:

i) represent the knowledge indicated by the diagram
ii) express relations between stenches, wind, pits and Wumpi

i) $\text{Stench}(1,2) \wedge \text{Stench}(1,3) \wedge \text{Wind}(2,1)$

ii)
$$\begin{array}{lll}\text{Adjacent}(a,b,c,d) & \Rightarrow & a=b \wedge c=d-1 \vee \\ & & a=b \wedge c=d+1 \vee \\ & & a=b-1 \wedge c=d \vee \\ & & a=b+1 \wedge c=d \vee \\ & & a=b+1 \wedge c=d-1 \vee \\ & & a=b+1 \wedge c=d+1 \vee \\ & & a=b-1 \wedge c=d-1 \vee \\ & & a=b-1 \wedge c=d+1\end{array}$$

(more space provided on following page)

Question 4. (continued).

ii) *(cont'd)* $\text{Stench}(a,b) \Rightarrow \exists x,y \;\; \text{Wumpus}(x,y) \wedge \text{Adjacent}(a,b,x,y)$

$\text{Wind}(a,b) \Rightarrow \exists x,y \;\; \text{Pit}(x,y) \wedge \text{Adjacent}(a,b,x,y)$

c) (4 marks) What inferences can be made about the Wumpus world diagram on the previous page, given your first-order logic statements?

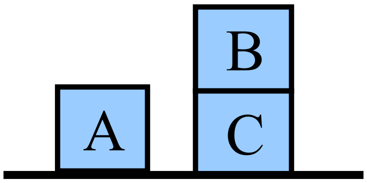
Wumpus(2,3)

or

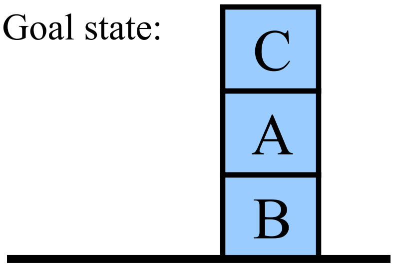
"There is a Wumpus at position (2,3)"

Question 5. (10 marks). *Planning*.

Start state:



Goal state:



a) (4 marks) Given the Blocks World diagrams above, what are the CNF representation for the start state and goal state, assuming the use of **`On(X,Y)`** literals?

Start state: $On(A,Table) \wedge On(B,C) \wedge On(C,Table)$

Goal state: $On(C,A) \wedge On(A,B) \wedge On(B,Table)$

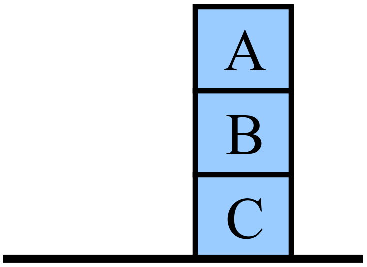
b) (4 marks) What are the CNF preconditions and postconditions of the **`Move(A,B)`** action, assuming that this action moves block `A` onto block `B`? Assume that blocks `A`, `B` and `C` are the only ones in this instance of Blocks World.

Preconditions: $\neg On(B,A) \wedge \neg On(C,A) \wedge \neg On(C,B)$

Postconditions: $On(A,B) \wedge \neg On(A,C) \wedge \neg On(A,Table)$

Question 5. (continued).

c) (2 marks) Given the start and goal states illustrated on the previous page, there is a move that would reduce the Hamming distance of the system. State this move, and draw a Blocks World diagram that would result from performing it.



Move(A,B)

Question 6. (6 marks). *Expert Systems*.

Destination	Climate?	Beaches?	Cost?	Book?
Cancun	Warm	Yes	High	No
Greece	Warm	Yes	High	Yes
Iceland	Cold	Yes	High	Yes
Nice	Warm	Yes	High	Yes
Rio de Janeiro	Warm	Yes	Low	No
Sweden	Cold	No	High	Yes
Tremblant	Cold	No	Low	No
Vancouver	Cold	No	High	No

Herman is trying to figure out a travel destination, to try to recover from the mental and emotional damage of the horrific CSC384 final exam. He has listed the travel destinations he considered in the past, along with his booking decision for each one.

a) (2 marks) Based on this booking history, what factor should Herman consider first when deciding on his travel destinations in the future? Make sure to base your selection on good decision tree principles.

Cost

b) (4 marks) Calculate the resulting entropy that results in each category, after splitting the data according to your decision boundary. Show your work.

Climate: [-1/2log(1/2) - 1/2log(1/2)] +
[-1/2log(1/2) - 1/2log(1/2)] = 0.602

Beaches: [-3/5log(3/5) - 2/5log(2/5)] +
[-1/3log(1/3) - 2/3log(2/3)] = 0.568

Cost: -1/3log(1/3) - 2/3log(2/3) = 0.276

Question 7. (11 marks). *Language*.

a) (5 marks) Write a grammatical sentence with an ambiguous meaning, and draw sensible parse trees for each meaning in the space below. The decompositions do not have to conform to a particular grammar or POS set, but they should show how the sentence decomposes into reasonable but different ways.

The solutions to this question will vary.

Question 7. (continued).

During a football game, you always hear the quarterback yell out three play numbers before hiking the ball. For example:

"... ready! ... 10! ... 28! ... 42! ... hike! ..."

One of these numbers is actually the play being used, but only his teammates know which one. The other numbers are just used to distract the opposing team. However, having observed this quarterback often in the past, you've discovered the following model for the quarterback's calls:

P(NotPlay|Ready) = 0.6
P(Play|Ready) = 0.4

P(NotPlay|Play) = 0.75
P(Play|Play) = 0.0
P(Hike|Play) = 0.25

P(Play|NotPlay) = 0.25
P(NotPlay|NotPlay) = 0.05
P(Hike|NotPlay) = 0.20

P("10"|Play) = 0.5
P("28"|Play) = 0.25
P("42"|Play) = 0.25
P("10"|NotPlay) = 0.2
P("28"|NotPlay) = 0.5
P("42"|NotPlay) = 0.3

b) (6 marks) Given the model for this quarterback's yells, determine which play number is probably being used, as well as the probability of this play, given the utterance sequence above. Give a brief explanation to justify your approach and your calculation.

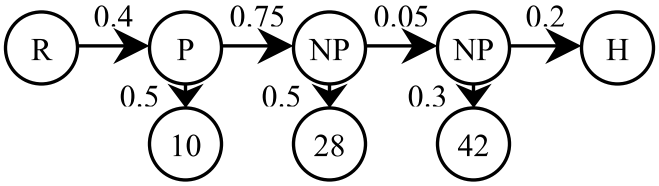
There are three possible sequences:

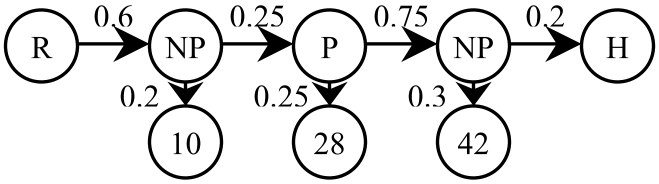
1. "10" is the play being used
2. "28" is the play being used
3. "42" is the play being used

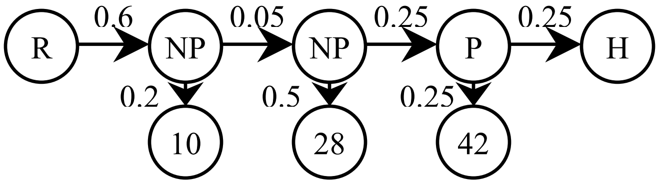
From this, we have 3 possible paths:

P(path_1) = 0.000225
P(path_2) = 0.0003375
P(path_3) = 0.000046875

Therefore, Path #2 has the highest likelihood, so the play is #28.

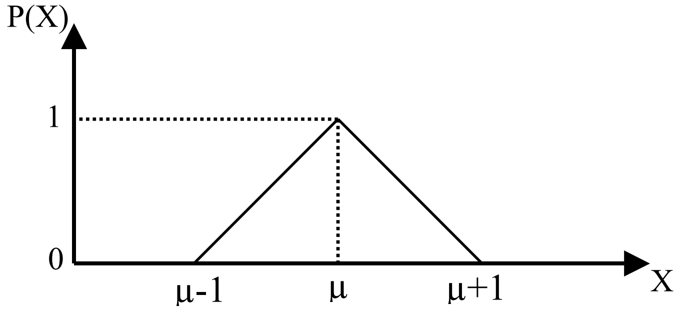






Question 8. (9 marks). *Learning.*

A *triangular distribution* is shown in the diagram below:



This probability distribution reaches a value of 1 at $x=\mu$, has a value of 0 for all $x>\mu$**+1** and $x<\mu$**-1**, and has a total area of 1 under the distribution.

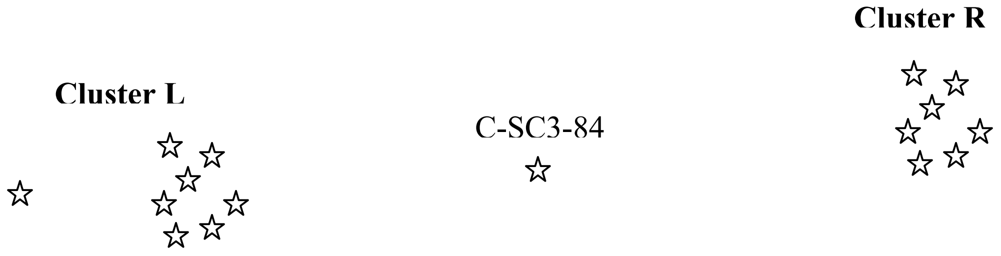
a) (5 marks) The following six points are a result of two triangular distributions in the range of x=0 to x=5. Given initial midpoints for two triangular distributions of $\mu_1=1$ and $\mu_2=4$, find the successive midpoints for these two distributions after one iteration of the EM algorithm. Briefly justify your answer.

```
A:  1        B: 3.5
C:  1.5      D: 3.5
E:  4        F: 2.5
```

μ_1 : 1.25 (average of 1 and 1.5)

μ_2 : 3.66 (average of 3.5, 3.5 and 4)

Question 8. (continued).



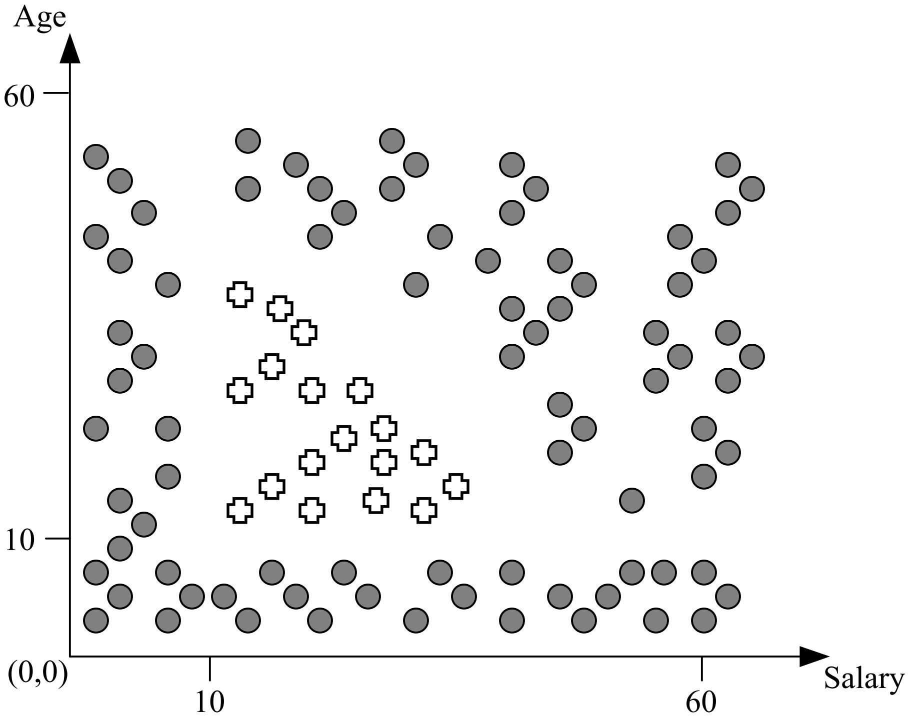
b) (4 marks) A new star, designation C-SC3-84 has just been discovered in the night sky, between existing star clusters L and R. Scientists all agree that this new star must belong to one of the existing clusters, but when they use their nearest-neighbour classifiers, they get conflicting results. Explain why the scientists are classifying this star in two possible ways.

There are two different kinds of nearest-neighbour classifiers:

1. Nearest-point classifier
2. Farthest-point classifier

If using the first classifier, Cluster L is the closest to the new star. If using the second, Cluster R is the closest cluster, since Cluster L has that one outlying star on the left.

Question 9. (10 marks). *Neural Networks.*



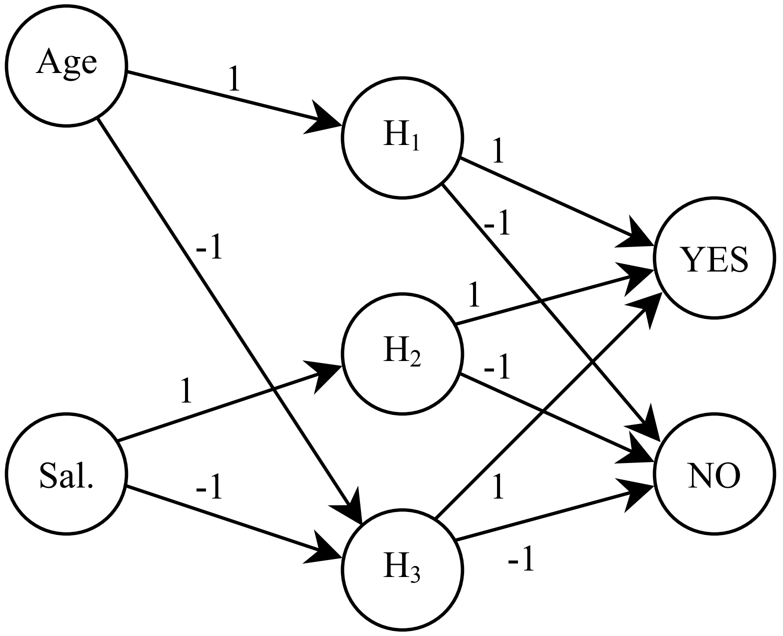
The diagram above classifies a group of people according to whether they enjoy playing arcade games (✙) or not (●). Each data point is a person surveyed, and the axes represent the age (in years) and the salary (in hourly wages) of each respondent.

a) (3 marks) Can this data be classified using a neural network with no hidden layer? Explain why or why not.

This data cannot be classified using a neural network with no hidden layer (a.k.a. a perceptron), because perceptrons are only able to classify data that has a linear separation boundary.

Question 9. (continued).

b) (7 marks) Based on the diagram from the previous page, draw the simplest neural network that can classify this data, given inputs of age and salary, and outputs of **YES** (people who like arcade games) and **NO** (people who don't like arcade games). Indicate the weights and activation functions for each node in the network.



Activation function for H_1: Unit step at x=10
Activation function for H_2: Unit step at x=10
Activation function for H_3: Unit step at x=-60

Activation function for YES: Unit step at x=3
Activation function for NO: Unit step at x=-3