Final Exam

CSCI 561 Spring 2024: Artificial Intelligence

Problems	**100 Percent total**
1- General AI Knowledge	10
2- Fuzzy Logic	10
3- Search, Planning, Markov Decision Process	20
4- Constraint Satisfaction Problem	15
5- Decision Tree	10
6- Bayesian Network Inference	20
7- Resolution proof	15

DO NOT OPEN EXAM UNTIL YOU ARE TOLD TO

Instructions:

1. Date: **Friday May 3rd, 2:00pm - 4:00pm**
2. Maximum credits/points/percentage for this final: 100
3. The percentages for each question are indicated in square brackets [ ] near the question.
4. **No books** (or any other material) are allowed.
5. **Write down your name, student ID and USC email address.**
6. **Your exam will be scanned and uploaded online.**
7. **Write within the boxes provided for your answers.**
8. **Do NOT write on the 2D barcode.**
9. **Do not write within less than 1” from the paper edges to avoid lost work during scanning.**
10. **The back of the pages will not be graded. You may use it for scratch paper.**
11. **The back of the pages will not be scanned. Do not write any answer there!**
12. No questions during the exam. **If something is unclear to you, write that in your exam.**
13. **Be brief: a few words are enough if using the correct vocabulary studied in class.**
14. When finished, raise completed exam sheets until approached by proctor.
15. **Adhere to the Academic Integrity code.**

1. [10%, 1% each, no partial]

General AI Knowledge

For each of the statements below, fill in the bubble **T** if the statement is **always and unconditionally true**, or fill in the bubble **F** if it is **always false, sometimes false, or just does not make sense**.

1.	T	F
2.	T	F
3.	T	F
4.	T	F
5.	T	F
6.	T	F
7.	T	F
8.	T	F
9.	T	F
10.	T	F

1. When solving a game problem, a heuristic evaluation function is always needed, even if you can run full exhaustive search on every move. (F)

2. How mutation is done in genetic algorithms only affects the convergence speed but not the final result. (F)

3. The environment of a contingency problem is nondeterministic and inaccessible. (T)

4. Forward and backward chaining are complete for general first-order logic. (F).

5. The tasks of a logical reasoning system only include adding new facts to KB and deciding whether a query is explicitly stored in the KB. (F, L11, P3)

6. In logical reasoning systems, FETCH and STORE actions always take at least O(n) time on n-element KB. (F, L11, P5-8)

7. Convolutional Neural Networks (CNN) extend traditional multi-layer Perceptrons, based on 3 ideas: local receptive fields, shared weights, and spatial/temporal sub-sampling. (T)

8. In reinforcement learning, we give correct answers after each action as the feedback. (F)

9. In natural language processing, people are always using bag of words to translate the discrete words into one-hot encodings. (F)

10. In Markov chain Monte Carlo, the samples generated are statistically independent. (F)

2. [10%] Fuzzy Logic

A smart thermostat uses fuzzy logic to control the heating in a house. The thermostat receives two inputs: the current temperature in Fahrenheit (°F) and the humidity. Based on these inputs, it controls the heating level as the speed of its fan, in RPM.

Temperature is measured in Fahrenheit (°F) and can be described as follows:

- **Cold**: ≤ 65°F
- **Hot**: ≥ 71°F

Humidity is measured as a percentage (%) and can be described as follows:

- **Dry**: ≤ 45%
- **Wet**: ≥ 55%

The heating output is described as follows:

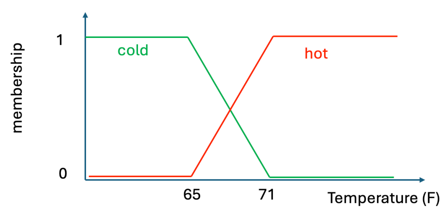
- **Weak:** fan speed ≤ 1000 rpm
- **Strong:** fan speed ≥ 2000 rpm

The rules in the thermostat's fuzzy logic system are:

1) **If the temperature is Cold and the humidity is Dry, then the heating should be Strong.**
2) **If the temperature is Hot or the humidity is Wet, then the heating should be Weak.**

2A. [3%, -0.5% if any one curve is incorrect.] Create the membership functions for temperature being Cold or Hot, humidity being Dry or Wet, and heating being Weak or Strong. Use a piecewise linear function for each fuzzy concept.

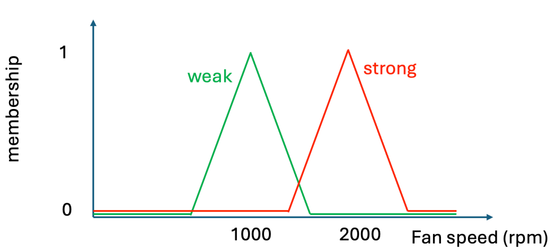
Temperature:



Exact location of inflexions (shown above at 65 and 71) is not critical. Shape of each curve is critical (cold starts high (y value near but not more than 1.0) for low temp, then decreases near 0.0 for high temp.

Similar plots for humidity (shows 2 curves, for Dry and Wet, x-axis is humidity %, inflexions around 45% and 55%) (see next question to look at the curves)

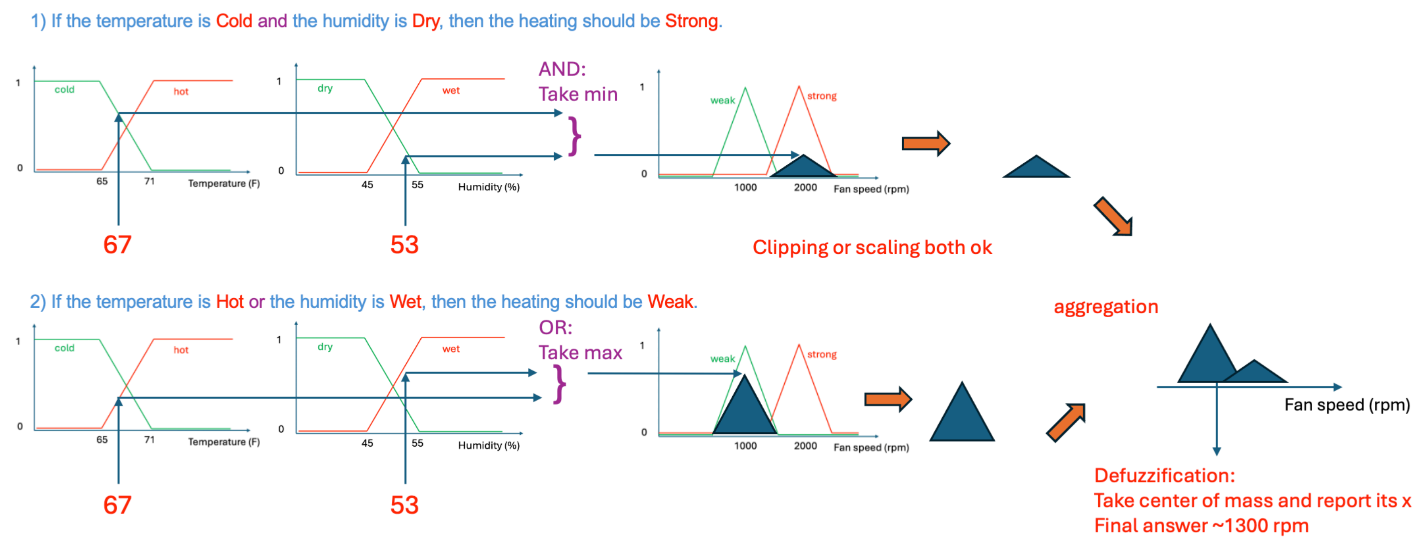
Heating: shows 2 triangle-shaped curves for Weak and Strong, x axis is fan speed in rpm, peaks of triangles are around 1000 for weak and 2000 for strong. Non-triangle shapes ok if they are bounded and roughly centered at 1000 and 2000:



2B. [7%] Using the provided rules, explain how the thermostat determines the fan speed. Demonstrate this process under the following conditions:

The temperature is 67°F and the humidity is 53%.

Detail all the steps of fuzzy inference, including fuzzification, application of the rules, aggregation, and defuzzification, to clearly explain how the system processes the inputs to compute the fan speed.



For rule 1 (top), 67F must project to Cold, while 53% must project to Dry, otherwise -1%.
For rule 2 (bottom), 67F must project to Hot, while 53% must project to Wet, otherwise -1%.

AND in rule 1 means take the min of the clipping levels, otherwise -1%
OR in rule 2 means take the max of the clipping levels, otherwise -1%

For rule 1, clipping level should clip (or scale, as shown here) Strong, otherwise -1%
For rule 2, clipping level should clip (or scale, as shown here) Weak, otherwise -1%

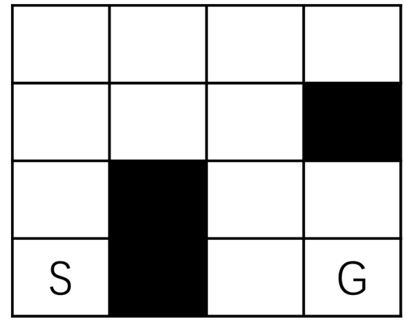
Aggregation merges the two clipped or scaled areas, otherwise -1%

Defuzzification takes center of mass of aggregated area and reports its x value on the Fan Speed axis, otherwise -1%.

Final answer must be a speed in RPM otherwise -1%, should be around ~1300 rpm, value not critical.

3. [20%] Search, Plan, MDP

Consider a 4x4 grid world as shown in the picture. We use a coordinate system with origin at the bottom-left corner to refer to specific cells. For example, the cell in the left-top corner with no content is referred to as cell **(0, 3)**. There is currently an agent at **(0,0)**, marked with "S" and the goal cell is **(3, 0)**, marked with "G". There are three cells with obstacles in the map that are colored in black.



Each time, the agent could choose to move to one of the 4-adjacent cells, i.e., choose to move up, down, left, or right.
The Manhattan distance between two cells $(x_1, y_1), (x_2, y_2)$ is defined as: $D = |x_1 - x_2| + |y_1 - y_2|$ *(hint: this definition ignores any obstacles).*

3A. [3%, action, precondition, effect worth 1% each.] We unify the definition of state as the current location of the agent: At(x,y), where x and y are the current coordinates of the agent. Please describe the **STRIPS action(s)** in this system, using the **graphical definition** of STRIPS actions studied in class.

Graphical like in lectures:

At(x,y) or At(x,y-1) or At(x,y+1) or At(x+1,y) or At(x,y+1)

MoveTo(x, y)

At(x,y)

Also ok if defined 4 actions for MoveUp, MoveLeft, MoveDown, MoveRight (preconditions and effects must make sense; eg, for MoveUp, precond: At(x,y) and effect: At(x,y+1). Ok to ignore that the space has boundaries.

The following questions will use search algorithms on this problem.

Step cost: We define the cost of each move (step cost) to be the Manhattan distance from the cell that we move into, to the goal cell. ***Note: This is highly unusual; an approximate cost to goal is here used in the step cost.*** For example: the cost to move from (2,3) to (1,3) given that the goal is at (3, 0) is $|1 - 3| + |3 - 0| = 5$

Heuristic function: for each cell, the heuristic value is defined as the Manhattan distance from this cell to the goal cell.

When all else is equal, expand the nodes in coordinate-numerical order, i.e., you first expand the node with smaller x, and, if x coordinates are equal, with smaller y.

3B. [6%, 3% for expansion order and 3% for solution path, no partial score] Apply A* and write down the order of expansions and the solution path.

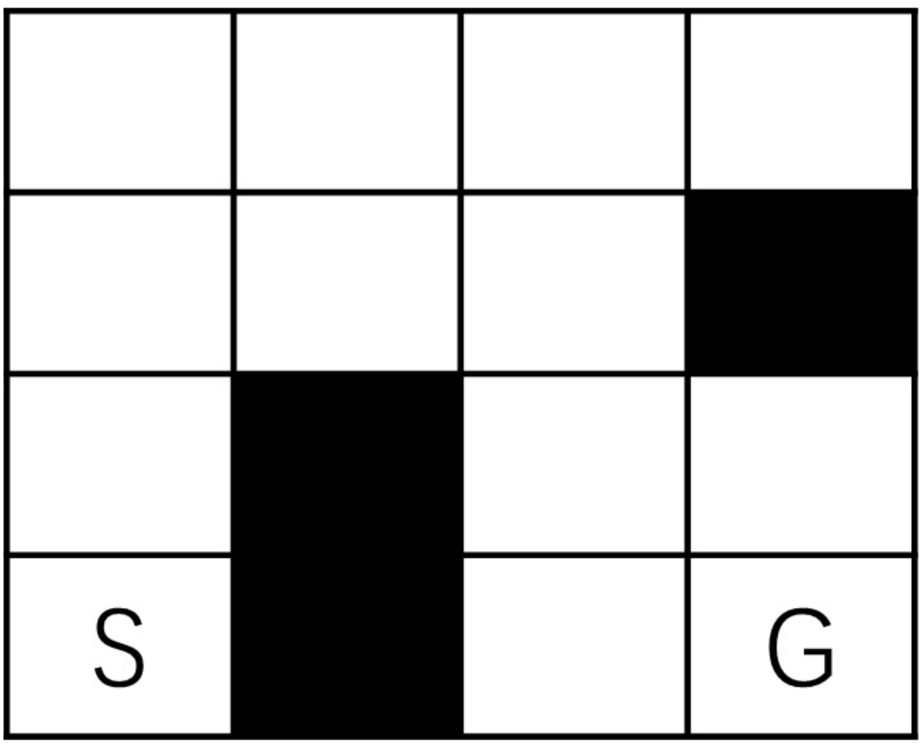
EXPANSIONS: ___(0,0),(0,1),(0,2),(1,2),(2,2),(2,1),(2,0),(3,0). where only (3,0) could be optional. (0,0) is not.___________________________
SOLUTION: __(0,0),(0,1),(0,2),(1,2),(2,2),(2,1),(2,0),(3,0)________________________

3C. [2%, no partial score] List all the cells that break the rule of admissible heuristic for the Manhattan Distance heuristic defined above. If none, write "NONE".

______NONE____

3D. [7%, -1% for each cell with incorrect value, no negative score.]. Now let's re-consider this problem as an MDP. Assume that the value function is initialized to 0 in every cell. Assume that the agent receives a reward of **256** for entering the goal G in (3,0). Also assume that once the agent is at the goal, it will always remain there. The discount factor is $\gamma = 0.5$.

Write down the value function for all cells after convergence of the value iteration algorithm.



8	16	32	16
16	32	64	
8		128	256
4		256	G

3E. [2%]. What is the optimal action for location (1,2) based on the numbers in 3F?

Moving right____

4. [15%] CSP

Relax and let's play the game Sudoku.
We use the same coordinate system
to refer to specific cells as in problem 3. For example, the cell in the left-top corner with number "2" is referred to as cell **(0, 8)**, and the left-bottom corner with number "1" is referred to as cell **(0,0)**. There are 9 sub-squares with size 3*3, shown with thicker borders. There should be exactly one of each digit from 1 to 9 in each sub-square, each row, and each column.

2		5			3	8	6	1
4	3	1	8	6				
	7	6	1		2		4	3
3	8	7		5		2	1	6
6	1	2			7	4		5
5	4	9	2	1	6	7		
7	6	3	5	2	4	1	8	9
	2	8	6	7	1			4
1		4		3		6		

4A. [1%]. If you use cells as variables, what is the domain? Enumerate the elements in the domain.

1-9

4B. [2%, no partial score]. Applying all the immediate constraints, what are the possible values for cell **(7,7) in current state**?

2,5,7,9

4C. [2%, no partial score]. Applying all the immediate constraints, what are the possible values for cell **(5,5) in current state**?

9

4D. [4%, no partial score]. What cells are the options if we use the **most constrained variable** heuristic (MRV heuristic)?

(5,5),(1,8),(4,6),(7,3),(8,3),(0,1),(3,0)

4E. [2%, no partial score]. Which cells are then the options
if we use the **most constraining**
variable heuristic (degree heuristic)
to break ties among the results of 4D?
[note: make sure your answers to 4D are correct, otherwise you may receive zero here].

____(3,0)____

4F. [4%, -0.5% for each incorrect cell result, 4% reduction max]. Fill in the blanks by running the AC-3 algorithm studied in class. In each cell below, write down all the remaining values in the reduced domain for that cell.

Answer: NOTE: blank cells below should instead have two possible values: 2,7

2	9	5	7	4	3	8	6	1
4	3	1	8	6	5	9		
8	7	6	1	9	2	5	4	3
3	8	7	4	5	9	2	1	6
6	1	2	3	8	7	4	9	5
5	4	9	2	1	6	7	3	8
7	6	3	5	2	4	1	8	9
9	2	8	6	7	1	3	5	4
1	5	4	9	3	8	6		

2		5			3	8	6	1
4	3	1	8	6				
	7	6	1		2		4	3
3	8	7		5		2	1	6
6	1	2			7	4		5
5	4	9	2	1	6	7		
7	6	3	5	2	4	1	8	9
	2	8	6	7	1			4
1		4		3		6		

5. [10%] Decision Trees

5A. [2%, no partial]. Select the correct answer for the following question:
The ID3 algorithm uses which of the following methods to select the best attribute?

A. ___ Most-Even
B. ___ Least-Value
C. ___ Most-Value
D. ___ Max-Gain

5B. Let's consider the following dataset for determining whether one wants to buy a specific cloth based on their material, price, color, and size.

Instance	*Attributes*				*Target* Buy
	Woolen	*Price*	*Color*	*Size*	
I_1	*T*	*H*	*R*	*L*	*T*
I_2	*T*	*M*	*G*	*S*	*F*
I_3	*T*	*H*	*R*	*L*	*T*
I_4	*T*	*M*	*G*	*S*	*T*
I_5	*F*	*H*	*R*	*L*	*T*
I_6	*F*	*L*	*R*	*L*	*F*
I_7	*F*	*H*	*G*	*L*	*T*
I_8	*F*	*L*	*G*	*L*	*F*

(a) [4%, 2% for Woolen and 2% for Size, no partial score.] Calculate the information gain *IG* in **bits** for splitting on Woolen and Size. You do not need to answer fully computed numbers: an expression like "0.5 $\log_2(0.5)$" is acceptable as long as it evaluates to the correct answer.

E(original)=-3/8*log2(3/8) – 5/8*log2(5/8)

E(wood)=1/2(-3/4(log2(3/4))-1/2*log2(1/2)-1/2*log2(1/2)-1/4(log2(1/4)))

E(size)= -1/2*log2(1/2)-1/2*log2(1/2)-2/3*log2(2/3)-1/3*log2(1/3)

IG(wood)=E(original)-E(wood)= -5/8 log2(5) + 3/4

IG(size)= E(original)-E(Size) = -39/40log2(3)-1/8log2(5)+3.6

(b) [4%, 2% for correct calculation on IG or entropy for Price and Color, 1% for the final answer, and 1% for doing the comparison among the values calculated as reason.] Which one of the four attributes is the best choice for the root of the decision tree using the ID3 algorithm? Explain why.

E(color)=1/2(-3/4(log(3/4))-1/2*log(1/2)-1/2*log(1/2)-1/4(log(1/4)))

E(Price)=1/3(-1/2*log(1/2)-1/2*log(1/2))

Comparing between the information gain, Price is clearly dominating because its entropy is the biggest.

You may use the space below for rough work. It will **not** be graded.

6. [20%] Bayes Inference

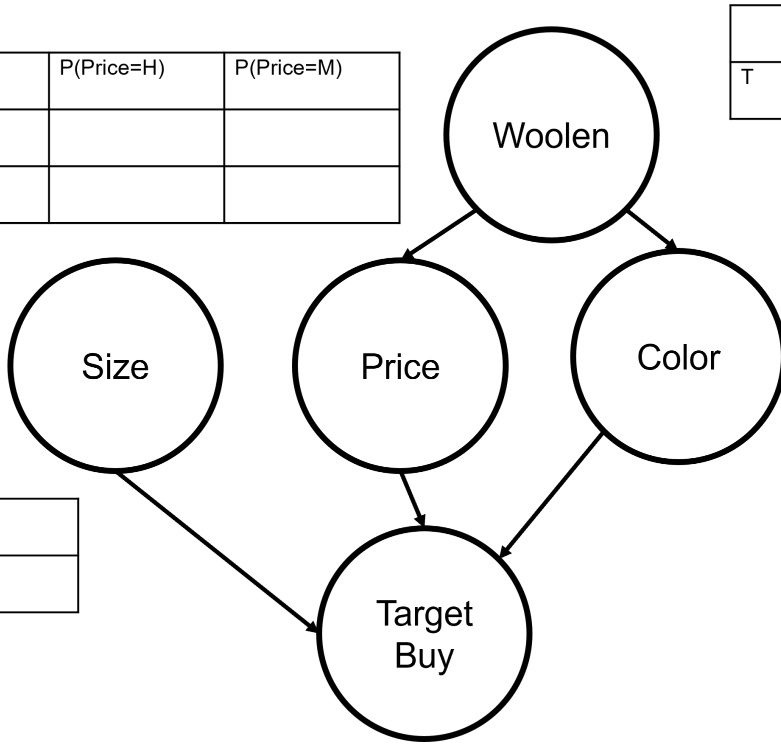
Let's consider the same dataset
as in the previous question 5.2:

Instance	*Attributes*				*Target* Buy
	Woolen	*Price*	*Color*	*Size*	
I_1	T	H	R	L	T
I_2	T	M	G	S	F
I_3	T	H	R	L	T
I_4	T	M	G	S	T
I_5	F	H	R	L	T
I_6	F	L	R	L	F
I_7	F	H	G	L	T
I_8	F	L	G	L	F

And given the following Bayesian network that describes the relationship between different attributes, with blank conditional probability tables (CPTs) here only shown for Price and Color:

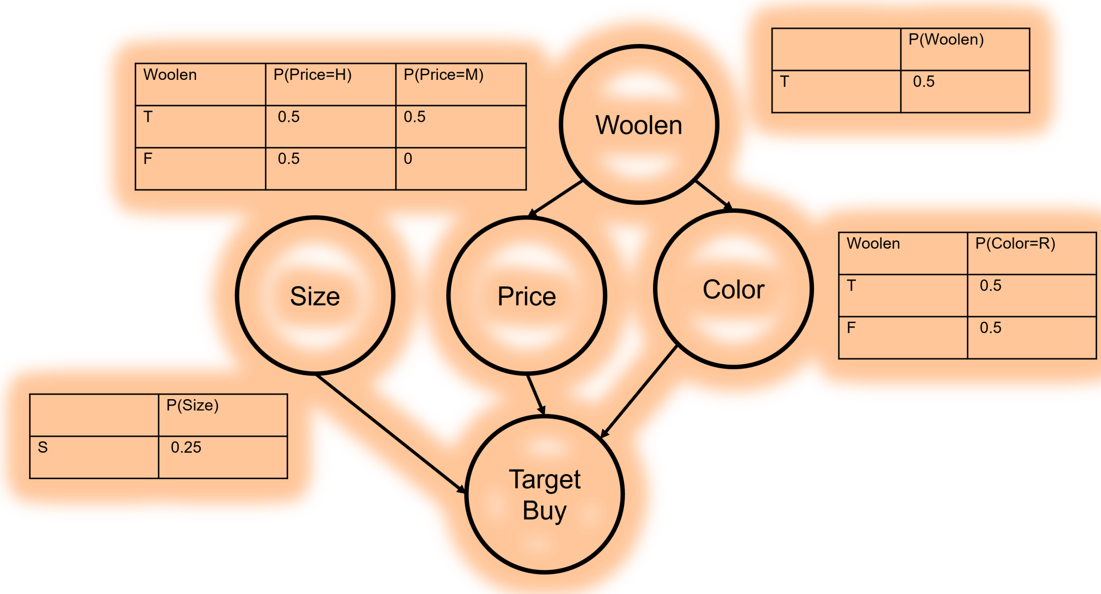
Woolen	P(Price=H)	P(Price=M)
T		
F		

	P(Woolen)
T	



Woolen	P(Color=R)
T	
F	

	P(Size)
S	



6A. [8%, 1% for each number, no further partial score] Using the provided dataset table, fill in the CPT tables in the picture above.

6B. [1%] Write down the statement
corresponding to the following sentence:

"People are more likely to buy **a woolen cloth with price L than a cloth with color R and size S.**"

P(Target Buy=T|Woolen=T, Price=L) > P(Target Buy=T|Color = R, size=S)

Suppose the CPT table for Target Buy is as follows, which may be conflict with the values from the above dataset, but please use the numbers from this table onward:

Price	Color	Size	P(Target Buy=T | Size, Color, Price)
H	R	L	0.3
H	R	S	0.3
H	G	L	0.1
H	G	S	0.2
M	R	L	0.6
M	R	S	0.3
M	G	L	0.2
M	G	S	0.2
L	R	L	0.8
L	R	S	0.3
L	G	L	0.1
L	G	S	0.5

6C. [5%, 1% for final conclusion, 2% for each probability] Prove whether the above statement in 6B is true or false based on the CPT above and on the CPTs that you filled in question 6A.

P(Target Buy=T|Woolen=T, Price=L)=0.25*0.4+0.75*0.45=0.4375

P(Target Buy=T|Color = R, size=S)=0.3

It is true.

6D. [6%, 2% for each, no partial score] Calculate the following probabilities based on the CPTs.

(a). $P(Target\ Buy = T \mid Woolen = F,\ size = L)$:

0.325

(b). $P(Woolen = T \mid Target\ Buy = F,\ Size = S)$:

150/285=30/57=0.526

(c). $P(Woolen =\ T \mid Target\ Buy = F, Size = S, Price = H)$:

0.5

7. [15%] Resolution proof

Consider the following axioms:

1. Anyone with allergies will sneeze
2. Anyone who is allergic to cats has allergies
3. Felix is a cat
4. Liz is allergic to Felix

We want to prove:

5. Liz will sneeze

7A [5%; 1% per sentence, no partials] Convert the axioms and the query that we want to prove to CNF. Use the following predicates:

- allergies(X) is true iff X has allergies
- sneeze(X) is true iff X will sneeze
- cat(X) is true iff X is a cat
- allergic-to(X, Y) is true iff X is allergic to Y

1:

2:

3:

4:

5:

1. allergies(X) → sneeze(X)
2. cat(Y) ∧ allergic-to(X, Y) → allergies(X)
3. cat(Felix)
4. allergic-to(Liz, Felix)
5. sneeze(Liz)

Equivalently:

1. ~allergies(X) v sneeze(X)
2. ~cat(Y) v ~allergic-to(X, Y) v allergies(X)

7B. [10%, -1% each missing substitution, -2% each incorrect resolution step]. On the following page, draw a proof by refutation to prove sentence 5 using the resolution algorithm. **Only use the FOL resolution inference rule,** or you will lose points. Be sure to indicate substitutions used next to each resolution step.

ANSWER: (others are possible)