Name _____

ID# _____

CSCI 5582 Final Exam

I. The following questions address the topic of Search. **(15 points)**

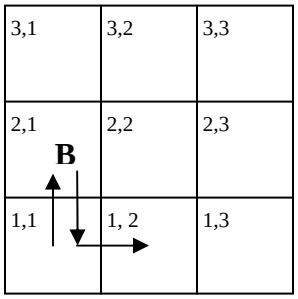
1. **(5 points)** Which of the following search algorithms are optimal?
   - Uniform cost
   - A*
   - Greedy
   - IDA*

2. **(5 points)** True or False: A search algorithm that is *complete* needs to examine all possible paths to a goal state.

3. **(5 points)** What problem with A* is IDA* designed to solve and how does it solve it?

II. The following questions address the topic of Logic. **(25 points)**

1. Recall the following Wumpus world rule from the text. When the agent detects a Breeze in a location, it means that there is at least one Pit in an adjacent location (up, down, left, or right; for some reason the Breeze doesn't go diagonally). Assume that our agent travels the path shown in the following figure.



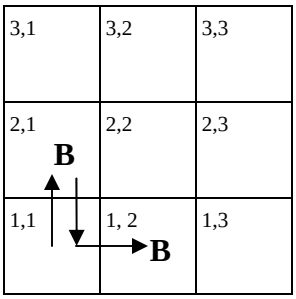
   a. **(10 points)** What is the state of the agent's knowledge base with respect to the location of any pits at the point where it has moved from 1,1 to 2,1 and detects the Breeze there? More specifically, show this using propositions and show all the logical inference steps needed to get these propositions.

b. **(10 points)** Show the state of the agent's knowledge base at the point where it has moved back through 1,1 to 1,2 and detects NO Breeze in 1,2. Again show all the needed inferences.

2. **(5 points)** Translate the following sentences to First Order Logic.
    a. All dogs are mammals.
    b. Mammals are warm-blooded.
    c. Not all dogs like the snow.

III. The following questions address the topic of uncertainty. **(45 points)**

1. Consider the following Wumpus situation, which is similar to the one described in Question 1. The only difference is that after traveling to position [1,2], a second breeze is encountered.



   a. **(5 points)** What can the agent logically conclude about the state of the board at this point?

   b. **(10 points)** Now assume that you are told that there are only two pits on this board, and that your agent now has the ability to reason probabilistically. What cell should your agent visit next? Why? Show your work with appropriate probabilistic notation.

2. Professor Solomon has noticed that on days when it snows two of his students have a tendency to skip class. Sally likes to go skiing and Tommy stays home since he hates to drive in the snow. After investigating, Professor Solomon determines that when it snows Sally goes skiing 60% of the time, whereas she only skips class to go skiing 10% of the time when it doesn't snow. Tommy skips class to stay home 90% of the time when it snows and stays home from class 20% of the time when it doesn't snow. Finally, he figures out that it snows 20% of the time.

    a. **(5 points)** Draw a Bayesian belief net that corresponds to the above facts (include all the appropriate tables).

    b. **(10 points)** Without knowing anything about the weather, what should Professor Solomon conclude about the chances that Tommy will stay home from class?

c. **(10 points)** Professor Solomon begins to teach and notes that Tommy is definitely not present. What should he conclude (without looking outside) about the chances that it is snowing?

d. **(10 points)** What should he conclude (without looking) about the chances that Sally is in attendance?

IV. The following questions address the topic of Machine Learning. **(25 points)**

1. **(10 points)** Given the data in the following table what attribute would be selected by the decision tree algorithm for placement at the top of the tree?

#	Credit History	Income	Debt	Approved
1	Good	Low	Low	Yes
2	Good	Medium	Low	Yes
3	OK	High	High	Yes
4	OK	High	High	Yes
5	OK	Medium	Low	Yes
6	OK	High	Low	Yes
7	Bad	High	Low	No
8	Bad	Medium	Low	No
9	OK	Low	High	No
10	OK	Low	High	No
11	OK	Low	Low	No
12	OK	Low	Low	No

2. Assume that some dumb summer intern decided to include the applicant's Social Security Number (assume they are unique) as an additional feature for all the objects in the training set (as shown here for one example).

#	Credit History	Income	Debt	SSN	Approved
1	Good	Medium	Low	329-456-0891	Yes

a. **(5 points)** What effect would including this feature have on the decision tree's information gain criteria?

b. **(10 points)** Generalize from this extreme example and point out a potential flaw in the information gain criteria.

V. The following questions address the topics of Speech and Language processing. **(15 points)**

1. **(5 points)** Match the following linguistic terms (levels) with their English equivalent.

Syntax	Sounds people make
Phonology	What things mean
Semantics	How language is used
Pragmatics	Order and grouping of words

2. **(10 points)** Text-to-speech systems are designed to take sequences of words in as input and output a spoken signal. One of the problems these systems face is the fact many written words have multiple pronunciations. For example, words like *bow, read, and bass* each have two basic pronunciations (ignoring dialect and context which may give rise to more). You are asked to design a system that pre-processes the inputs to a text-to-speech system so that words like bow are replaced with forms like bow1 and bow2 that the system knows how to pronounce. For example an input like *I will read about bow ties* turns into *I will read1 about bow2 ties*. Describe the design of a system that does this. Assume that you have access to a list of such words and to a corpus of correctly tagged words in context, but that you don't have access to a dictionary, parser, or any other syntactic or semantic information.