Name: Connor Johnson

CS1571 Introduction to Artificial Intelligence
Final 2020

You have 24 hours to complete this test. It is due Thursday, Dec. 3 at 7:59AM. The test has 6 pages. There are 13 questions, worth 90 points total.

The test should be submitted on **Gradescope**. Upon submission, please use the Gradescope interface to indicate where your responses to each question are on the page.

This test should represent your **individual work**. You may not discuss any parts of this test with your classmates or other people. You may not post test questions to Q&A sites like StackOverflow to receive answers during the testing period. Consultation on test questions constitutes an Academic Integrity Violation.

This test is **open book** and **open notes**. You may use any course resources you like to answer questions on this test. You may also use other resources **already available** on the internet, as long as you cite them below. Do not list all resources viewed while searching; just focus on the ones that directly contributed to the answer. Failure to cite resources you have used to compose your answer consists of an Academic Integrity Violation.

Certifications

I certify that this test represents my own individual work, and I have not consulted with others in taking the test.

Connor Johnson
Signature

The following are any resources other than the course textbook or course materials I used while taking the test in order to compose my answers.

URL or Resource Name	**Question Number**
https://my.eng.utah.edu/~mccully/cs5300lw/	7
https://www.cs.toronto.edu/~sheila/384/w11/Lectures/csc384w11-KR-tutorial.pdf	2
<<add more rows as necessary>>	

I certify that the above list represents the external resources that influenced my answers on the test.

_______Connor Johnson__________
Signature

A. Logic (12 pts)

1. (4 pts) For the following statements about inference in propositional logic, indicate whether they apply to model checking (MC), theorem proving with inference (TP-I), and theorem proving with resolution (TP-R) by placing an 'X' in the appropriate cells. Some sentences might not apply to any of the approaches, and some might apply to more than one.

Statement	**MC**	**TP-I**	**TP-R**
It uses a search algorithm.			X
It requires storing a truth table.	X		
It can be used on any propositional logic *KB*.	X	X	
It is only complete with a proof by refutation.			X

2. (8 pts) Translate the following English sentences into first-order logic Horn Clauses. As with converting to conjunctive normal form, assume that all variables are universally quantified. You will need to remove existential quantifiers through skolemization, and standardize your variables.

   All people who are curious enjoy learning some subject. There is a person who enjoys learning about every subject.

∀x(Person(x)→ ∃y (Subject(y) ^ Likes(x,y)))
∀x(~Person(x) v ∃y (Subject(y) ^ Likes(x,y)))
∀x(~Person(x) v (Subject(F(x)) ^ Likes(x,F(x))))
~Person(x) v (Subject(F(x)) ^ Likes(x,F(x)))

∃x(Person(x) → ∀y(Subject(y) → Likes(x,y)))
∃x(~Person(x) v ∀y(~Subject(y) v Likes(x,y)))
∃x∀y (~Person(x) v (~Subject(y) v Likes(x,y)))
∀y (~Person(G(y)) v (~Subject(y) v Likes(G(y),y)))
~Person(G(y)) v ~Subject(y) v Likes(G(y),y)

Hints:
- Take care to capitalize the first letter of your constant symbols but not your variables (e.g., "Jake" is a constant, "jake" is a variable).
- There are two sentences, but your final response could have more than two clauses.

B. Inference (8 pts)

The next two questions relate to the following *KB*.

Person(x) ^ Course(y) ^ Take(x,y) -> WriteFinalExam(x,y)

Person(Linda)
Course(Chemistry)
Take(Linda,Chemistry)

One new fact that you could infer from this *KB* is *WriteFinalExam(Linda, Chemistry)*.

3. (4 pts) What is one unification and one substitution that would need to be performed to make that inference?

UNIFY(Person(x), Person(Linda)) = {Person(x) ^ Course(y) ^ Take(x,y), Person(Linda) ^ Course(Chemistry) ^ Take(Linda,Chemistry) = {x/Linda, y/Chemistry}

SUBST({x/Linda,y/Chemistry}, WriteFinalExam(x,y)) = WriteFinalExam(Linda, Chemistry)

4. (4 pts) What technique could be used to infer *WriteFinalExam(Linda, Chemistry)* from the above KB? Check all that apply.

_X_ Model checking

_X_ Resolution

___ Forward chaining

___ Backward chaining

C. Planning (12 pts)

5. (12 pts) Consider the following spare tire problem described in the textbook. Walk through how a solution could be discovered using backward planning. For all solution steps, list the action that you are applying and the new subgoal that would be created as part of the backward planning search.

1. The goal is At(Spare, Axle)). Backtracking, we find that this is the effect of the PutOn action. Using substitution, the action is PutOn(Spare, Axle). Looking at the preconditions, Tire(Spare) and ~At(Spare, Axle) are satisfied as preconditions. At(Spare, Ground)and ~At(Flat,Axle) are the new subgoals and PutOn(Spare, Axle) is added as an action.
2. For At(Spare, Ground), this is the effect of the action Remove(Spare, Axle). The precondition for this action is At(Spare, Trunk) which is an initial precondition. So, this action is added and At(Spare, Ground is removed from the subgoals.
3. For ~At(Flat, Axle), this is an effect for Remove(Flat, Axle) The precondition for this action is At(Flat,Axle) which is one of the initial preconditions. The last subgoal is satisfied and the search is complete.

The solution is Remove(Flat, Axle), Remove(Spare, Trunk) PutOn(Spare, Axle).

$Init(Tire(Flat) \wedge Tire(Spare) \wedge At(Flat, Axle) \wedge At(Spare, Trunk))$
$Goal(At(Spare, Axle))$
$Action(Remove(obj, loc),$
  PRECOND: $At(obj, loc)$
  EFFECT: $\neg At(obj, loc) \wedge At(obj, Ground))$
$Action(PutOn(t, Axle),$
  PRECOND: $Tire(t) \wedge At(t, Ground) \wedge \neg At(Flat, Axle) \wedge \neg At(Spare, Axle)$
  EFFECT: $\neg At(t, Ground) \wedge At(t, Axle))$
$Action(LeaveOvernight,$
  PRECOND:
  EFFECT: $\neg At(Spare, Ground) \wedge \neg At(Spare, Axle) \wedge \neg At(Spare, Trunk)$
  $\wedge \neg At(Flat, Ground) \wedge \neg At(Flat, Axle) \wedge \neg At(Flat, Trunk))$

Figure 11.2 The simple spare tire problem.

D. Bayesian Inference (16 pts)

6. (12 pts) Consider a modification of the following scenario discussed in class, where Marcella is likely to be on time for her morning class if she did not oversleep and if she does not eat breakfast. Marcella is likely to eat breakfast if she is hungry. She is likely not to oversleep if she went to bed early the night before. The following are the conditional probability tables for this scenario.

P(BedEarly)
.60

P(Hungry)
.80

BedEarly	P(Oversleep)
T	.20
F	.40

Hungry	P(EatBreakfast)
T	.90
F	.30

Oversleep	EatBreakfast	P(OnTime)
T	T	.30
T	F	.50
F	T	.75
F	F	.90

What is the exact probability of Marcella being on time, given that she went to bed early and she was hungry in the morning?

Pr(Oversleep = T and EatBreakfast = T) = .2 * .9 = .18
Pr(Oversleep = T and EatBreakfast = F) = .2 * .1 = .02
Pr(Oversleep = F and EatBreakfast = T) = .8 * .9 = .72
Pr(Oversleep = F and EatBreakfast = F) = .8 * .1 = .08

.18*.3 + .02*.5 + .72*.75 + .08*.9 = .054 + .01 + .54 + .072 = 0.676

7. (4 pts) Assume that you would like to answer a query about a Bayes net in less than a minute, and this requires that you perform fewer than 10^8 operations. Say you had a Bayes net with 20 variables, each with three different possible values. In this course, we discussed two methods of answering queries from Bayes Nets: a) Inference by enumeration and b) Monte Carlo sampling with likelihood weighting. Assume all 20 of the variables are relevant to answering your query, and for your Monte Carlo sampling you generate 10,000 samples. Which methods could be used in this situation:

   a) Both inference by enumeration and likelihood weighting

   b) Only inference by enumeration

   c) Only likelihood weighting

   d) Neither method would be fast enough

E. Machine Learning & Naïve Bayes (12 pts)

8. (8 pts) Imagine you were building an application for a smart watch that used supervised learning to predict the number of steps a user was going to take by the end of the day based on its access to the user's calendar, location, environmental factors such as the weather, and previous month's step history. Fill in the following table in order to explain how this approach might be used. The first row is filled in for you.

Supervised Learning Concept	**How it Applies to the Example**
The outcome y	*Number of steps the user takes in a day*
An example feature x_j of each training data instance d_i	The inputs like calendar, location, etc.
A feature engineering method you might use involving x_j to create a potentially more effective feature.	This would be combining inputs to create a new one like combining the weather with location to show how unusual it is
The type of model you could use to predict y from d_i	Linear Regression Model

The error function you would try to minimize in your prediction.	Function that shows how far off the prediction was from the actual steps

9. (4 pts) Let's say that instead of treating the above supervised learning problem in question #8 as a regression problem, you treat it as a classification problem. You have 20 multinomial features in each training data instance, and you would like to predict whether each user met their step goal ($y=1$) or did not meet their step goal (y=0) for the day. You decided to use a Naïve Bayes approach, and are computing the probabilities for P(x_i | y) for all 20 features. If you are using Laplace smoothing on your data with $\alpha = 1$, which of the following is true:

_T_ The technique corrects for cases where a given feature x_i is not observed in your training data.

_F_ The technique will function properly in cases where a given class (e.g., $y=0$) is not observed in your training data.

_F_ You will add 15 to the denominator and 1 to the numerator when computing the probability for the fifteenth feature $P(x_{15} \mid y=1)$.

_T_ The technique will decrease your confidence in the accuracy of a predicted class for a test data instance, because you are manipulating the values of probabilities derived from your training data.

F. Application Areas (24 pts)

10. (8 pts) In his guest lecture, Dr. Harrison discussed being able to use low-cost sensors in the home as part of machine learning in order to detect what a user is doing at any given moment. For example, activities with similar acoustic properties might get clustered together. These activities then get given a label by the user, and then future activities can be assigned a cluster, and thus a label. For example, all activities where the user is washing dishes have a similar acoustic pattern, different from all activities where the user is making coffee. The next time the user makes coffee, the system can recognize that this is the activity the user is doing because it sounds similar to the other making coffee instances.

    Assuming you want your approach to generalize to new users, how should you **train your model?** Explain in reference to **what** data is collected, **when** it is collected, **where** it is collected, and **who** it is collected with. Explain why your approach will **prevent overfitting** and **reduce algorithmic bias.** A response that earns full marks will cover all of these points.

The model would be trained by collecting data on the sounds being made by test users. The data that would be collected is the frequency of the sounds, the length of them, and the loudness. This data would be collected in people's homes that are randomly selected to participate. The users would tell the AI what they are currently doing that is making noise. The AI could use this info to train itself. Since the test users are randomly selected, there would be no bias toward a specific group of people and it would be more accurate because the training is done in people's homes.

11. (4 pts) The Dr. Harrison guest lecture also touched on multiple ways to preserve the privacy of users. Assume microphones are being used to collect data in the home. What is one good method for preserving the privacy of your users in this context?

Once the data is used, it can be deleted because it isn't needed after it has already been used to train the AI. This would help protect the user's privacy.

12. (12 pts) Intelligent tutoring systems typically have a model of how to solve problems in a domain, so that the student solutions can be compared to a correct solution, often on a step by step level. There are several different approaches for developing this model using AI, including:
    a) Creating a knowledge base (founded on logic) that consists of expert-defined rules for solving these problems. For example, in solving the equation "3x+2 = 9" for x, you would have rules for identifying and subtracting the constant term (in this case, 2), or identifying when there is a single variable term on a side and dividing by the coefficient (in this case, 3). The rules would include information about what the sequence of actions should be (e.g., you should remove the constant term before you remove the coefficient).
    b) Encoding these problems as planning problems, and then allowing the AI to discover the sequence of steps for solving these problems. For example, in solving the equation "3x+2 = 9" for x, you would have *actions* representing different steps that could be taken (e.g., removeConstantTerm, removeCoefficient), and the system would try to find a plan using those actions that take you to the goal state.
    c) Using reinforcement learning by having actions represent different steps that can be taken in a problem. For example, when solving "3x+2 = 9" for x, actions might include removeConstantTerm and removeCoefficient. The AI can observe what state it is in after executing an action, and a user "teaches" the AI how to solve the problem by assigning a reward after each action. The goal state would be a terminal state, and the AI would go through several training iterations until it has discovered an optimal policy.

    Your task is to generate three example problems, one for each approach (expert rules, planning, reinforcement learning), where it would be better to use that approach to model the correct solution than the other two. Each problem can be from any learning domain (e.g., math, chemistry, English, history, engineering, etc.). Provide a justification for each example.

a) A knowledge base would be best for problems like multiple choice or matching that are based on facts. A knowledge base could be developed on the information that the student needs to know. This would be better because the other 2 Ais would overcomplicate a simple solution.
b) Planning problems would be best for solving equations like the example given. There are very specific rules for solving these problems that need to be followed. These steps could be defined as actions with preconditions effects, and a goal state. Since there are often many steps to the solution, the KB would be very inefficient. Reinforcement learning would also be unnecessarily inefficient because the actions that need to be taken are well defined and purely based on simple portions of the equation.
c) Reinforcement learning would be best for infinite series problems in calculus. There are many different rules that can show whether converges to a value of diverges to infinity. It isn’t always clear which one is best or should be tried first. There is no way to simply give an AI rules to follow. Because of this, the other two models would not work well. Reinforcement would train itself based on how people would approach these problems so it could then choose the most likely best options based on the equations given.

G. Reflection (6 pts)

13. Which question on the exam do you feel like you learned the most from trying to answer, and why? (6 pts)

Part E was the part where I think I learned the most. I didn’t fully understand Naïve Bayes and Machine learning with regression at first. I did more research on it in the notes and the study guide and my understanding improved from it.