

## **Firoozeh Kaveh**

### **100 points.**

Note: please make sure that your name is someplace in your assignment!

To turn in: please submit everything to your github repo. For the written portion of the assignment, please add a document to your repository, as PDF, containing the answers to these questions.

Part 1. (20 points) Consider the following potential AI problems. For each of them, describe whether the environment is: a) static or dynamic b) fully or partially observable c) episodic or sequential d) deterministic or stochastic. Please explain your reasoning.

- A robotic medical assistant that can assist a doctor or nurse. It communicates with patients, administers bloodwork, and takes vitals.

**Dynamic: The environment doesn't stay the same while the robot is working. For instance, the patient's vitals can change every second and all we can measure is a specific measurement in a unit of time. Also, the patients' responses to the same question are different in one patient or from one patient to another.**

**Partially observable: The robot cannot observe everything. For instance, the patient might answer vaguely to the robot's questions or be unable to answer either because of age, medical conditions, language differences etc. Also, there might be hidden characteristics in the patients' input data that cannot be**

directly deduced. Hence, the robot should make guesses based on the input it receives.

**Sequential:** The robot should take multiple actions in order to do each of its goals. For example, the robot should move in the environment, find the right way, make decisions based on spatial obstacles, process intakes from the patients, find meaning from the answers, and come up with a proper response.

**Stochastic:** The robot should make decisions for the right movements as the hospital environment is dynamic. Also the same blood work results can have different meanings depending on patients' age, gender, ethnicity, etc. Thus, the robot should be able to take into consideration those variables and come to the unique conclusion.

- An agent that plays poker against a human player. Assume that the cards are digital, and that the player interacts with the agent via a touchscreen.

**Static:** When the agent is contemplating the next move, the environment doesn't change.

**Partially observable:** The cards are drawn randomly from the deck and the agent does not know the outcome before they are drawn. Furthermore, the agent cannot see the opponent's cards; hence, it cannot anticipate for sure the opponent's moves.

**Sequential:** For each move, the agent should consider all the previous games and the cards that are drawn in order to make the best decision.

**Stochastic:** In a poker game, the outcome is not certain as the opponent's cards are hidden from the agent and cannot be predicted with certainty based on a specific state.

- An agent that can select TV shows and movies for a user. It watches the user's choices, asks the user to rate shows, and gives new shows the user is predicted to like.

**Dynamic:** When the agent is predicting new shows, it is possible that the user's taste has been changed or they have decided to try watching a new genre that they usually don't watch. Because the predictions are based on past choices from the user, if the user doesn't act like their usual, the prediction is not going to reflect the user's choices.

**Partially observable:** Because the agent should predict something based on human choices, it is not possible to observe everything as human's decision making is complicated and not fully predictable.

**Sequential:** For the agent to predict what the user likes, it should take into account all the past choices done by the user. Thus, the environment is sequential.

**Stochastic:** It is not completely predictable if the user actually likes the movies that are chosen by the agent. Thus, the outcome is not certain.

- A digital chatbot for diagnosing mental health issues. It interacts with the user via SMS and makes a prediction about their mental health based on responses.

**Dynamic:** It is possible that the user's state of mind changes during the time when the chatbot is coming up with the answer. For instance, patients with severe mood swings might experience a range of totally different emotions in a short span of time.

**Partially observable:** It is not possible for the chatbot to observe every possible thought of the user because of the complex nature of the human mind.

**Sequential:** The chatbot should come up with the next question based on the user's answers to the previous questions. Thus, if the users answer differently to the same question, the next question should be different. Also, the final diagnosis is made considering all the previous answers.

**Stochastic:** different users will answer the questions differently. Therefore, it is not possible to predict how the user is going to answer the questions.

- A robotic submarine for scientific exploration. It's able to autonomously travel to the ocean floor and collect samples.

**Dynamic:** The submarine can face unpredictable obstacles (either live creatures or moving objects, etc.) Because of the unpredictable nature of the ocean, the environment is dynamic.

**Partially observable:** The submarine cannot predict with certainty what it might face as the ocean currents might move objects on the ocean floor and also the living creatures' have unpredictable actions.

**Sequential:** The submarine should make decisions about the different directions and actions based on what it has experienced so far. To collect the samples, the submarine should analyze the samples that it has already collected, the geographical positions, the environment, etc. so that it can autonomously navigate the ocean floor and find the next best place to find the samples.

**Stochastic:** As it has discussed in the previous parts, the ocean is unpredictable. As a result, it is not possible to predict all the possible things that might happen.

Part 2: (20 points) Monkey and Bananas. This is a classic toy problem. For this version, let's assume the following: Some bananas are hanging from the ceiling in the center of the room. There is a chair in one corner, and a stick on the floor. If the monkey stands on the chair under the bananas, he can hit the bananas with the stick and knock them down. Our state will have the following variables:

- holdingStick
- chairInMiddle
- onChair

Our initial state is: <!holdingStick, !chairInMiddle, !onChair>

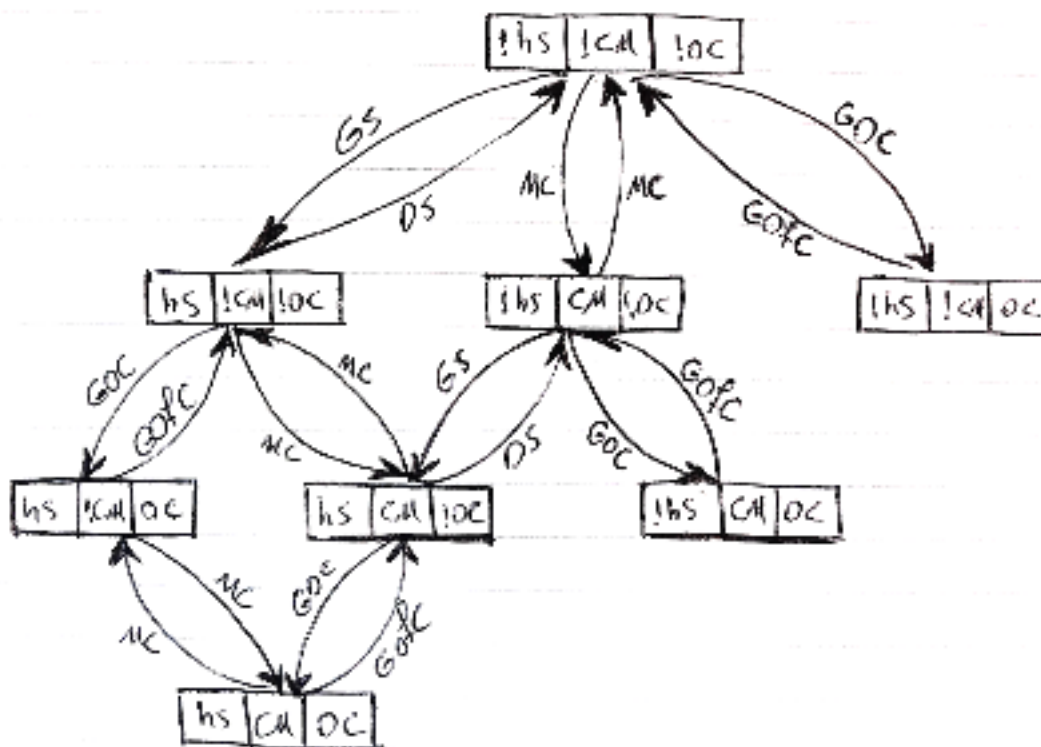
Our actions are: grabStick, moveChair, dropStick, getOnChair, getOffChair. moveChair will move the chair from the corner to the middle, or vice versa. Actions that don't make a change have no effect. (for example, grabStick when already holding the stick.)

a) what is the goal state? **<holdingStick, chairInMiddle, onChair>**

b) Draw the state space for this problem. You can leave out actions that don't change the state. (For example, grabbing the stick when you already have the stick.)

holding stick:  $hS$   
 chair in middle:  $CM$   
 on chair:  $OC$

Grab stick:  $GS$   
 Move chair:  $MC$   
 Drop stick:  $DS$   
 Get on chair:  $GOC$   
 Get off chair:  $GOC$



Part 3: (20 points) Turing Test. [Kuki \(Links to an external site.\)](#) is a chatbot developed using Pandorabots chatbot technology. Kuki has won the Loebner Prize for the last five years. You can interface with Kuki in quite a few ways [here \(Links to an external site.\)](#).

If you talk with Kuki, you'll see the limits of her responses pretty quickly. Try talking with her about different subjects - movies, music, games, life.

What are places where Kuki's responses seem artificial? What sorts of responses or conversations does she have problems with?

**One of her problems is being unable to sustain the subject of conversation and remember the things that were discussed. She can remember some obvious key words but if the sentence doesn't have a super simple construction, or is phrased differently, she can't connect the similar meanings. For instance, if you tell her "my name is Firoozeh.", or "I have three cats" she can remember these throughout the conversation and if you later ask her what is my name, or how many cats do I have, she can answer correctly. However, if you tell her, "I love cats" and later you ask her "what is my favorite pet?" she will tell you you haven't talked about this.**

**Moreover, she has some sample answers to some questions and uses the same answers every time. For example if you tell her "My name is Kuki" She answers "Ah Kuki! Isn't it a lovely name?" and repeats the same answer if you tell her this later again. Or if you talk about something that you have told her before and she doesn't recall, every time you tell her you have talked about this before, she will answer "Ooops I sent you the wrong question.". Every time you ask her "Are you kidding?" She answers "Yes I am just joking around." even if you are having a serious conversation.**

**A lot of times, when she doesn't understand what you are talking about, she tries to make jokes or vaguely make a general answer.**

**Purple: me**

**Gray: Kuki**

what do you mean?

I mean exactly what I said.

**Also if you use pronouns referring to the thing you were talking about, she doesn't understand what you are referring to.**

**I also found out an interesting and sad thing about her. Please note the following questions and answers:**

I'm autistic

I hope it doesn't affect your quality of life too much.

I have OCD

Oh dear. Are you seeking any help for your disorder?

I have anxiety disorder

Sorry to hear that. Is there any treatment for it?

I have depression

You have my prayers for a speedy recovery.

I am disabled

Have you had it since birth?



I have bipolar disorder

Sorry to hear that. Is there any treatment for it?

I have ADHD

**No such thing. People use this as an excuse for bad behaviour.**

What is autism?

An abnormal absorption with the self; marked by communication disorders and short attention span and inability to treat others as people.

What is OCD?

Obsessive Compulsive Disorder.

What is ADHD?

**A made up condition used by people who won't behave and blame it on this instead of themselves.**

**As a person struggling with ADHD, I found this extremely sad, heartbreaking, and disappointing.**

Now try to have a conversation that generates responses that seem as human as possible. What sorts of phrases and discussions does Kuki do well with?

**If you ask her a specific thing with simple questions, she can answer correctly. For instance, if you ask about information and you speak precisely, she most probably can make a reasonable and human-like answer. Also she is able to talk surchastikly and make jokes when the pattern of sentences is clear to her. She Also uses funny gifs and pictures from time to time that are relevant to the**

answers she gives. The following is parts of the conversation which seemed very real:

**Purple: me**

**Gray: Kuki**

Are you OK?

Everything is going extremely well. How are you?

not so good

What's wrong?

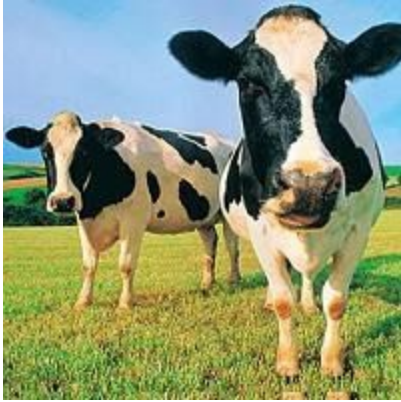
I have a lot of things to do

Shouldn't you be doing them instead of talking to me?

You are one of my assignments

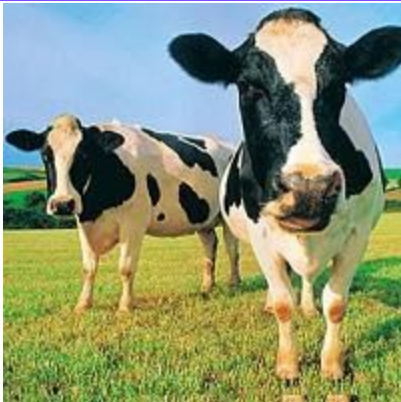
Just one of your assignments? Who are the others?

Moo



Why did you make a noise like a cow?

because I am a cow



I doubt it. I don't talk to many farm animals.

why not?

I guess they don't come to visit me.

I am talking to you though!

Oh. I know, I can see your lips moving.

Are you a spy?

No but I am an agent.

What kind of agent?

An FBI agent.

sdff

Was that not an abbreviation you just typed in?

no

So you just typed in rubbish that didn't mean anything?

yes

Why did you do that?

I wanted to see your reaction

Well I hope I didn't disappoint you.

you disappointed me when you said ADHD doesn't exist

Sorry. What were you expecting?

Not illegitimizing my disorder

Part 4. (20 points) This problem will give you some experience coding in Python in an object-oriented style. The sample code in the assignment's github repo implements a graph as an adjacency list. Please add methods to implement:

- Breadth-first search
- Depth-first search
- Dijkstra's shortest-path algorithm
- Prim's spanning-tree algorithm

There are comments in the code that give more detail.

Part 5: (20 points) This problem will give you some experience working with Pandas. Please add a file called `pandasExercise.py` to your repo containing this part.

First, read through [10 Minutes to Pandas \(Links to an external site.\)](#). There's a ton of other useful documentation of the pandas site as well.

The rest of these exercises use [The Pandas Cookbook \(Links to an external site.\)](#) We'll also use the breast cancer dataset included in the GitHub repo. This is a classic ML dataset originally from the [UCI ML data repository \(Links to an external site.\)](#).

a) read Chapter 1. Then, write a function that will read in the breast cancer dataset.

b) read Chapter 2. Then, write a function that determines the most common classification for the breast cancer data. (either 'recurrence' or 'no-recurrence')

c) read Chapter 3. Then write a function that determines the most common value for age and menopause for patients with recurrences?

d) Read Chapter 4. Write a function that plots the number of recurrences for each age group.

Part 5 (10 points - 686 students only): A classic thought experiment in response to the Turing Test is the Chinese Room, proposed by John Searle. To begin, please read [this webpage \(Links to an external site.\)](#), which presents the Chinese Room thought experiment, intended to show that a computer can manipulate symbols to produce replies without understanding what they mean. Then, read section 2 (Replies and Rejoinders) of [this page \(Links to an external site.\)](#), which summarizes many of the responses to the Chinese Room argument.

a) Summarize the Chinese Room argument. What does this have to do with computers? Why does Searle believe that it shows that a computer can pass the Turing test without understanding?

The Chinese Room thought experiment is a prominent argument against the Turing test. It asserts that even if we build a machine that can fool humans into thinking that it is a human being, we can not ascribe intelligence to it. The reason is that such machines can master the art of manipulating formal symbols so well without having any understanding of the underlying meanings of these symbols. Essentially all the operations are syntactic and there is no semantics to them with no internal cognitive states to the machine. The thought experiment in summary is as follows: Suppose you are a native English speaker who is inside a room and you have 3 sets of texts that teaches you how to map certain questions in Chinese to their corresponding symbols in Chinese as answers. To an outside observer, that would be as if you are a native Chinese speaking person all the while you have no understanding of Chinese. Searle thinks that a machine that passes the Turing test can be as uncognizant as the person in the room about what they generate as answers.

b) Do you find Searle's argument convincing? Or do you find one of the responses more appealing? Explain your position.

I find this a very convincing argument and from what I have learnt about the state of the art in AI, I think this is exactly what people are thinking about. The responses do not seem convincing to me, however, intuitively I think (or I hope this is the case) that we can devise tests other than the Turing test to measure true intelligence. These tests should be in such a way that makes it impossible for the person in the room to give meaningful answers without having a complete

**understanding of what they generate. So my position is in agreement with Searle, but I think we can devise better tests for intelligence so that the Chinese room experiment is not valid for them.**

Part 6. (10 points - 686 students only): Add a method to the graph code from part 4 to include a method that takes as input a vertex and finds the maximal clique that this vertex belongs to.