15-110 Check6-1 - Written Portion

Name:

AndrewID:

#1 - Files in Code - 4pts

We've written some code to count the number of times the string `"Hello World"` occurs within a file:

```
f = open("sample.txt", "r")
print(f.count("Hello World"))
f.close()
```

There is a bug in this code that will make it raise a runtime error, even if sample.txt is located in the same directory as the code. What is the bug, and how could you fix it?

#2 - Data Formats - 4pts

For each of the following data formats, identify whether it should be interpreted as a CSV, a JSON file, or as plaintext.

```
{ "restaurants": [
      { "restaurant" : "Chipotle",
        "date" : "11/01/19",
        "menu" :
            { "lunch" :
                  { "burrito" : 7.99,
                    "tacos" : 6.99,
                    "bowl" : 8.99 },
              "dinner" :
                  { "burrito" : 8.99,
                    "tacos" : 7.99,
                    "bowl" : 9.99 } } },
     { "restaurant" : "Sushi Fuku",
       "date" : "11/01/19",
       "menu":
            { "lunch" :
                  { "bowl" : 8.99,
                    "sushi" : 10.99 },
              "dinner":
                  { "bowl" : 10.99,
                    "sushi" : 12.99 } } } ]
}
```

- ☐ CSV
- ☐ JSON
- ☐ Plaintext

```
City,Longitude,Latitude
Los Angeles,34°03′N,118°15′W
New York City,40°42′46″N,74°00′21″W
Paris,48°51′24″N,2°21′03″E
```

- ☐ CSV
- ☐ JSON
- ☐ Plaintext

#3 - Parsing Data - 7pts

You have been given a set of data about CMU classes in the following format (whitespace added for clarity):

```
Professor,      ClassNum,  Days,      Time
Cortina,        15104,     MTWF,      09:20-10:10
Rivers,         15110,     MWThF,     16:00-16:50
Touretzky,      15110,     MWThF,     17:20-18:10
Touretzky,      15394,     MW,        19:00-20:20
```

Assume you've already split the string on `"\n"` and used the variable `row` to iterate through each class one line at a time.

How would you determine which department each class is in? Recall that the first two digits of the class number indicate the department.

- ☐ `row.split(',')[1][0:1]`
- ☐ `row.split(',')[1][0:2]`
- ☐ `row.split(',')[1][2:]`

How would you determine the start time and end time of a class and set those times as strings in the variables `start` and `end`? **Select all lines that are needed.** Assume that the code is run from the top selected line to the bottom selected line.

- ☐ `times = row.split(',')[0]`
- ☐ `times = row.split(',')[3]`

- ☐ `start = row.split('-')[0]`
- ☐ `start = times.find('-') - 1`
- ☐ `start = times.split('-')[0]`

- ☐ `end = times.find('-') + 1`
- ☐ `end = times.split('-')[1]`
- ☐ `end = start + "1:00"`

#4 - Components vs. Rules - 5pts

Let's say we want to design a simulation that determines how many students will sign up for a course during registration week. The simulation's time loop will loop over each sign-up time slot in order.

We need to design the model for this simulation. For each of the following values, would this value work better as a **component** of the model, or as a **rule** of the model?

Current length of the course's waitlist

- ☐ Component
- ☐ Rule

Students are more likely to sign up if a class is required for their major

- ☐ Component
- ☐ Rule

Number of students who are required to take this class, and haven't taken it yet, organized by sign-up timeslot

- ☐ Component
- ☐ Rule

Students are less likely to sign up for a class if the waitlist is long

- ☐ Component
- ☐ Rule

Information on whether or not the course will be offered again in the following semester

- ☐ Component
- ☐ Rule

#5 - Simulation Code - 10pts

We want to write code for a simulation that moves a circle from the left side of the screen to the right side of the screen in a 400px x 400px window. When the user clicks on the circle or presses 'Enter', the circle moves back to the left side of the screen.

For each part of the simulation (the Model, the View, the Time Rules, and the Event Rules), select the line of code that needs to be included in that part.

Hint: if you're not sure, try implementing this using the simulation starter code!

Which line of code should be included in the **model**, in `makeModel(data)`?

- ☐ `x = 5`
- ☐ `data["left"] = 5`
- ☐ `canvas.create_oval(x - 20, y - 20, x + 20, y + 20)`

Which line of code should be included in the **view**, in `makeView(data, canvas)`?

- ☐ `data["left"] = data["left"] + 5`
- ☐ `canvas.create_oval(200 - 20, 200 - 20, 200 + 20, 200 + 20)`
- ☐ `canvas.create_oval(data["left"]-20, 200-20, data["left"]+20, 200+20)`

Which line of code should be included in the **time rules**, in `runRules(data, call)`?

- ☐ `data["left"] = 5`
- ☐ `data["left"] = data["left"] + 5`
- ☐ `x = data["left"] + 5`

How would you check if the user clicked in the circle in `mousePressed(data, event)`?

- ☐ `((data["left"] - data["x"])**2 + (200 - data["y"])**2)**0.5 <= 20`
- ☐ `((data["left"] - event.x)**2 + (200 - event.y)**2)**0.5 <= 20`
- ☐ `(data["x"]**2 + data["y"]**2)**0.5 <= 20`
- ☐ `(event.x**2 + event.y**2)**0.5 <= 20`

How would you check if the user pressed "Enter" in `keyPressed(data, event)`?

- ☐ `if (data["char"] == "Return"):`
- ☐ `if (data["keysym"] == "Return"):`
- ☐ `if (event.char == "Return"):`
- ☐ `if (event.keysym == "Return"):`

#6 - Machine Learning Categories - 10pts

For each of the following prompts, fill in the blanks with the type of **learning algorithm** that should be used and/or the type of **reasoning algorithm** that should be used.

You have a dataset that consists of student grades from past semesters of 15-110, including final grades. Use _____ learning to predict a student's **numerical** final grade based on their **numerical** quiz scores with a _____ algorithm.

You have a dataset of weather patterns in different major cities around the world. Use _____ learning to propose **new groupings** of cities based on the **categorical** weather patterns with a _____ algorithm.

You have a dataset of athletes' descriptions (age, height, weight, etc) and the sport that they play. Use _____ learning to predict an athlete's **categorical** sport based on their **numerical** age, height, weight, etc. with a _____ algorithm.

To train a robot how to throw a basketball through a hoop through repeated practice and feedback, you'd want to use _____ learning.

To identify previously-unknown market trends based on stocks that go up and down together at similar times, you'd want to use _____ learning.

#7 - Machine Learning Process - 5pts

Imagine a scenario where Bill wants to train a machine learning algorithm to identify which pictures on the internet have cats in them. He downloads 1,000 pictures of cats and other animals from the internet, decides to use a basic image recognition algorithm which will identify important features, trains on all 1,000 pictures, then tests his on a quarter of that dataset (250 pictures). He finds that his algorithm has a 97% success rate, which he publishes on his blog.

Bill made a few mistakes in this process. **What was his <u>biggest</u> mistake?**