Assignment 1

Benjamin Sorenson

September 23, 2015

1 Written Questions

1. [20 points]Performance measures for an agent could be designed based on the effects the agent has on the environment or according to the behaviors of the agent. Explain briefly what is the difference, provide an example for each, and explain if one of the two choices is best and why.

   Answer: Taking two examples from the book, in the case of the vacuum cleaner agent, one could define a performance measure by the amount of dirt cleaned up in a single 8 hour shift, awarding more points for more dirt cleaned—a performance measure according to the behaviours of the agent. Alternatively, one could define a performance measure by awarding one point for each clean square at each time step—a performance measure according to the effects the agent has on the environment.

   In the case of a performance measure designed according to the behaviours of the agent, we reward behavior that we think will lead to the desired effects on the environment rather than reward for the desired effects directly. This can allow a rational agent to maximize its performance measure without the desired effects to the environment. For example, as the book points out, a rational vacuum cleaner agent that is maximizing the amount of dirt cleaned in an 8 hour period could maximize its performance measure by cleaning up the dirt, dumping it back on the floor and cleaning it again. For this reason, it is better to define a performance measure based on the desired effects to the environment rather than the behaviors of the agent.

2. [30 points]You are given the following problem: Given a 5-gallon jug filled with water and an empty 2-gallon jug how can you have precisely 1 gallon of water in the 2-gallon jug? Assume you can fill the jugs with water as many times as desired, but you cannot measure how much water is in each jug. When you move water out of a jug you can either fill up the other jug or dump the water.

   You are to formulate the problem using a state-space search representation.

   Describe (precisely):

   (a) what is the initial state

   The initial state is a 5-gallon jug (referred to as "5G" for the remained or the description) filled with water and an empty 2-gallon jug (referred to as "2G" for the remainder of

the description). Note: for the remainder of this description a state will be denoted as $\{Jug1 : NumberofGallons, Jug2 : NumberofGallons\}$. For example, the initial state wwould be $\{5G : 5, 2G : 0\}$.

(b) the goal test

Precisely 1 gallon of water in the 2G.

(c) the actions (called in the textbook successor function)

i. `pour_into(PouringJug, ReceivingJug)`
Definition: if the amount of water in `PouringJug` ($g1$) plus the amount of water in `ReceivingJug` ($g2$) is $\leq$ to the capacity of `ReceivingJug` ($G2$), the amount of water in `ReceivingJug` is set to $g1 + g2$, otherwise, fill `RecievingJug` to $G2$ and set the level of `PouringJug` to $g1 + g2 - G2$
Example:

$$RESULT(\{5G : 5, 2G : 0\}, pour\_into(5G, 2G)) = \{5G : 3, 2G : 2\}$$

ii. `empty(Jug)`
Definition: set the level of `Jug` to 0
Example:

$$RESULT(\{5G : 5, 2G : 0\}, empty(5G)) = \{5G : 0, 2G : 0\}$$

iii. `fill(Jug)`
Definition: set the level of `Jug` to its maximum capacity
Example:

$$RESULT(\{5G : 5, 2G : 0\}, fill(2G)) = \{5G : 5, 2G : 2\}$$

(d) the path cost

Each step costs 1 so the path cost is the number of steps in the path

(e) the state-space for the problem

The state-space is any state of fullness of the two-gallon and five-gallon jugs reachable by the actions `pour_into(Jug1, Jug2)`, `fill(Jug)`, and `empty(Jug)` performed on the initial state and any resultant states. Though other states are possible, any state can be trivially converted to one of the states reachable by the initial state in this problem definition in at most two actions.

(f) is the state-space a tree or a graph?

The sate space is a graph since there are repeated states.

(g) what search algorithm would you use and why?

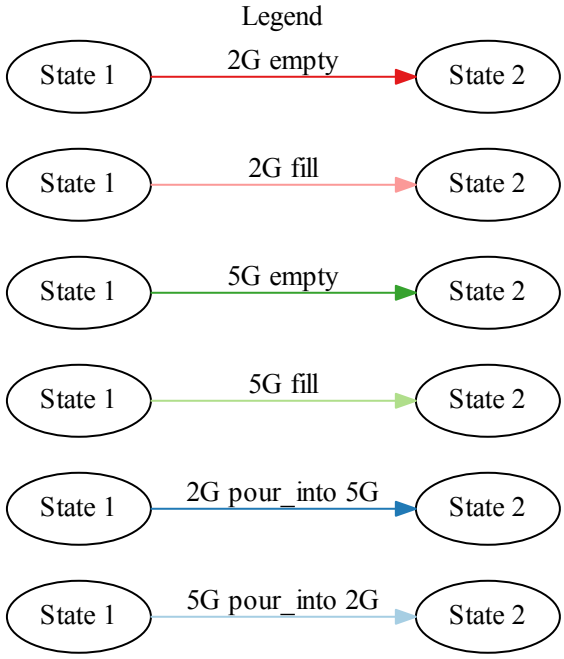
I would use a uniform-cost search because the search space is small, and it will always find an optimal solution.

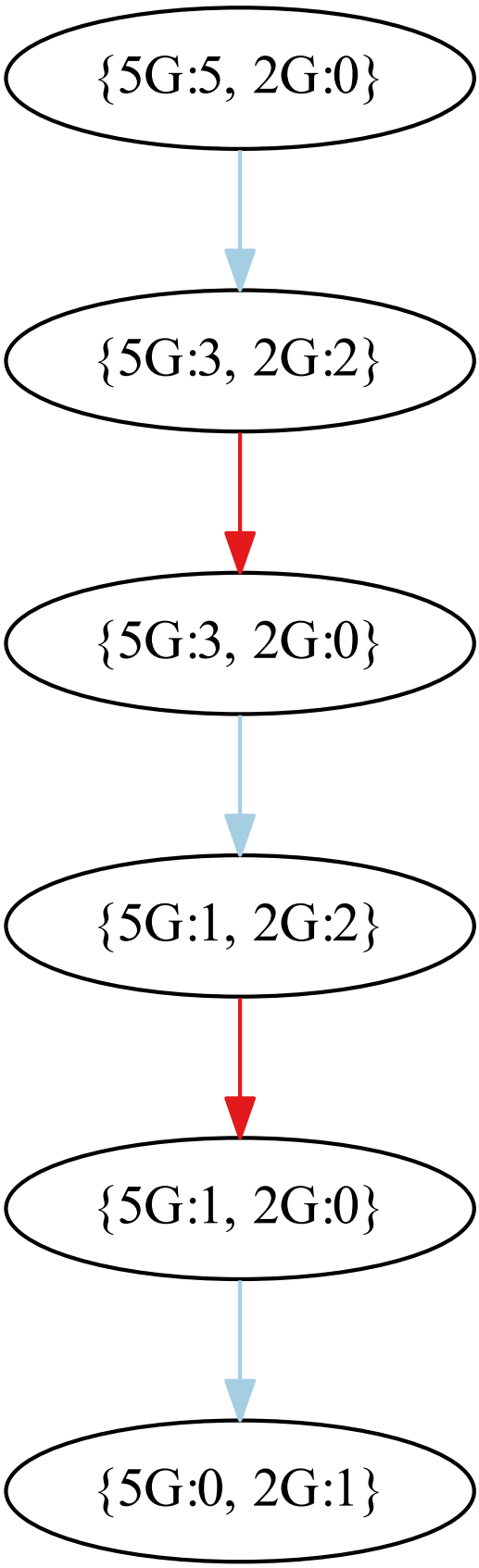
(h) show graphically the search space explored and the solution (there might be more than one solution)

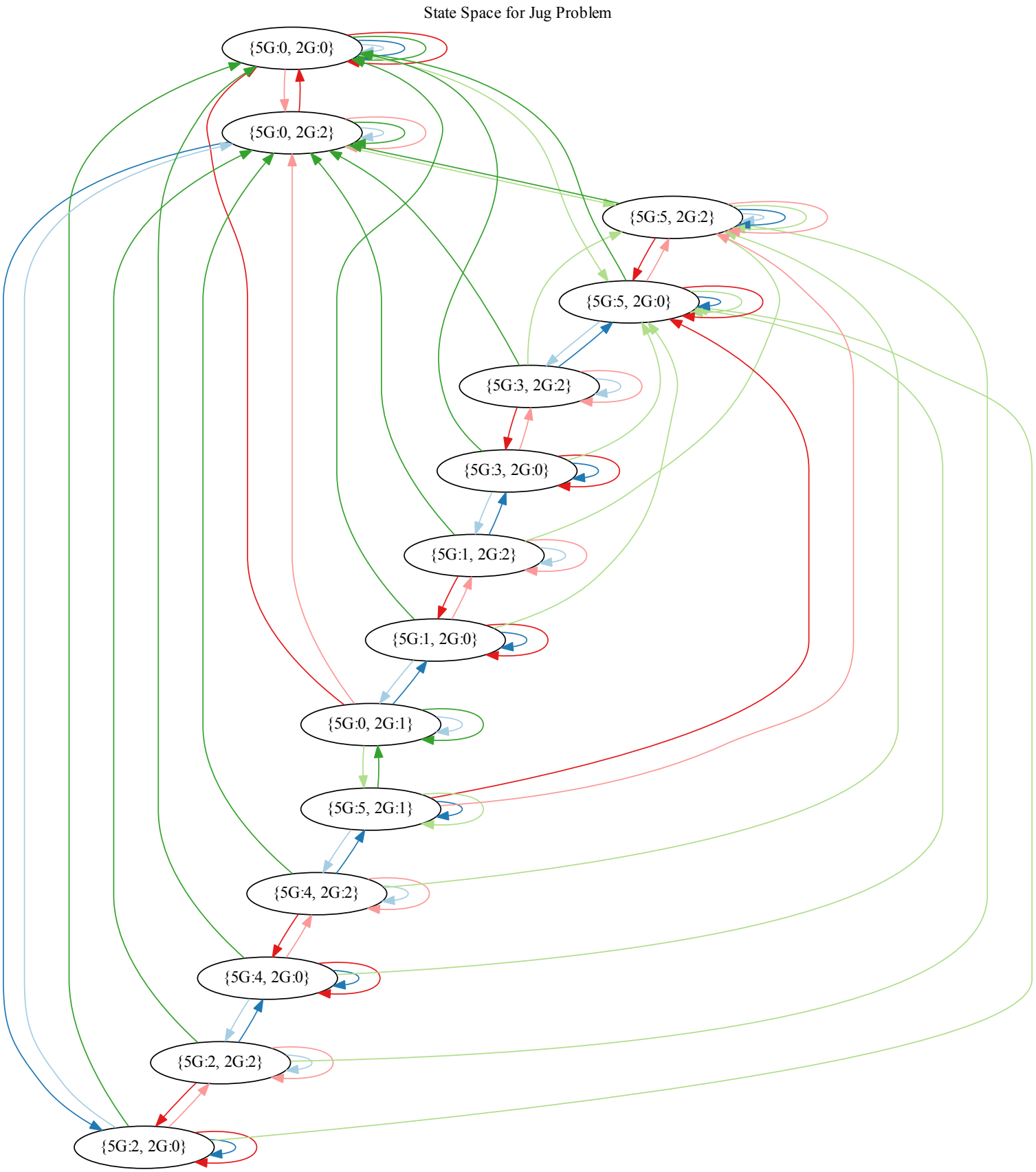
Next two pages show

i. The search space explored
ii. The optimal soluion

The color of the arrows indicate the action according to the following legend




Optimal Solution
{5G:5, 2G:0}
{5G:3, 2G:2}
{5G:3, 2G:0}
{5G:1, 2G:2}
{5G:1, 2G:0}
{5G:0, 2G:1}


State Space for Jug Problem
{5G:0, 2G:0}
{5G:0, 2G:2}
{5G:5, 2G:2}
{5G:5, 2G:0}
{5G:3, 2G:2}
{5G:3, 2G:0}
{5G:1, 2G:2}
{5G:1, 2G:0}
{5G:0, 2G:1}
{5G:5, 2G:1}
{5G:4, 2G:2}
{5G:4, 2G:0}
{5G:2, 2G:2}
{5G:2, 2G:0}

3. [20 points]There has been a lot of discussion recently in the news and on the web on the dangers of AI, started by Stephen Hawking and Elon Musk

.

(a) Search for a few on line writings on the controversy and summarize (in 1/2 to 1 page) the main arguments made.

(b) Read the white paper by Eric Horvitz outlining the project One Hundred Year Study on Artificial Intelligence. Can you think of some additional topics not listed there? or could you add more questions to one of the topics listed? Summarize your thoughts (in a few paragraphs).

1.1 Ethics

I wish I could spend more time thinking about the topics covered. In each topic, the central concern seems to be to one degree or another "What are the consequences of turning over traditionally human decisions over to machines?" and, in the other direction "What can we accomplish (for good for bad) with the aid intelligent machines?" Personally, I'm most looking forward to autonomous assistants that don't take any direction—they observe your behaviour and try to help out. For the first few weeks or months, they may do almost nothing, and suddenly rudimentary tasks begin to get completed on their own. Of course, this comes with its own set of ethical concerns. What if you're a serial killer, and you now have a serial killer assisstant—even if checks are put in place so that the machine can't assist killing or some other hard safe-guard, real life is fuzzy. Where doee the threshold lie between contributing to nefarious activity and not? Who decides?

1.2 Educations

I saw an info-graphic (I can't remember where, now), that put education near the bottom of fields most likely impacted by "Data Science." I tend to instead agree with the idea that education is one of the key areas of opportunity for AI. I've heard (source?) that the most difficult part of teaching is figuring out what when wrong when a student comes up with an incorrect solution—this difficulty is then multiplied by the number of students that a teacher has to observe. A machince that could perfectly and objectively observed, memorize, and perform rigorous analyses to assist a teacher in discovering these paths leading to incorrect understanding of the material, I would have to beleive, would be invaluable to both teachers and students. Maybe this would take the form of a study tool where not only is the student shown the correct answer when, but also given instruction specific to the deficiency in their understanding of the material.

1.3 Augmentation

The first concern is addressed most directly in law and ethics. Who is responsible when a machine makes the "wrong" decision? Is it the programmer, the owner? Will we see a new standard of right and wrong as rigorous analysis and unbiased memory picks informs our actions?

With regard to economics, we're already seeing impacts of smart, targeted advertising. There's the famous story of Target figuring out a teenage girl was pregnant before her parents. I don'tk