

## Problem Set 1

Dan McQuillan

*Handed In: February 5, 2014*

## 1. Solution to problem 1

(a) (i) depth-first

<b>n</b>	<b>q</b>
-	(A)
A	(B, C, D)
B	(E, F, C, D)
E	(K, F, C, D)
K	(F, C, D)

(ii) breadth-first

<b>n</b>	<b>q</b>
-	(A)
A	(B, C, D)
B	(C, D, E, F)
C	(D, E, F, G, H)
D	(E, F, G, H, I, J)
E	(F, G, H, I, J, K)
F	(G, H, I, J, K, L)
G	(H, I, J, K, L)

(iii) uniform-cost

<b>n</b>	<b>q</b>
-	(A)
A	(B, C, D)
B	(C, E, F, D)
C	(E, F, H, D, G)
E	(F, H, D, G, K)
F	(H, L, D, G, K)
H	(L, D, N, G, M, K)
L	(D, N, G, M, K)
D	(N, G, M, I, K, J)
N	(G, M, I, K, J)

(iv) greedy/best-first

<b>n</b>	<b>q</b>
-	(A)
A	(C, D, B)
C	(G, H, D, B)
G	(H, D, B)

(v)  $A^*$ 

<b>n</b>	<b>q</b>
-	(A)
A	(C, B, D)
C	(B, H, D, G)
B	(H, D, G, E, F)
H	(D, G, N, E, F, M)
D	(G, N, E, F, I, J, M)
G	(N, E, F, I, J, M)

- (b) (i) Depth first is not admissible since it does not take weights or distance into consideration and could easily not choose the cheapest goal.
- (ii) Breadth first search is not admissible for the same reason as depth first search since it does not consider weights or distance when finding a goal state.
- (iii) Uniform cost search is admissible since it is using the actual cost to the node and recalculates and sorts the order of the queue based upon the total cost from the root node.
- (iv) Greedy / Best-First search is not admissible since it does not take the cost of getting from  $n \rightarrow n'$  into consideration and it only depends on how close the node is estimated to be from the goal.
- (v)  $A^*$  search is admissible because when it terminates its search, it has found a path whose cost is lower than the estimated cost of any path through any open node. Although, since those estimates are optimistic,  $A^*$  can safely ignore those nodes. Therefore,  $A^*$  will never overlook the possibility of a lower cost and is admissible.
- (c) We know that for all heuristic functions, it is consistent if for every node  $n$  and every successor  $n'$  of  $n$  generated by any action  $a$ , the estimated cost of reaching the goal from  $n$  is no greater than the step cost of getting to  $n'$  plus the estimated cost of reaching the goal from  $n'$ . Or more formally:

$$h(n) \leq c(n, a, n') + h(n')$$

**Suppose:**

B is an unreachable node and therefore the cost from any node,  $n \in nodes$ , the cost will be infinite.

**Then:**

$$h(n) \leq c(n, a, B) + h(B)$$

$h(B)$  must therefore also be infinite since it is an unreachable and infinitely far

away from any goal node

$\therefore h(n) \leq \infty + c$  where  $c$  is some constant  
 $h(n) \leq \infty \rightarrow$  the heuristic is consistent

## 2. Solution to problem 2

- (a) (i) A data structure to represent this structure could be a Pair of the current node and the state of the box. Each state would be a tuple3, where it would contain three slots which would be filled with an enum value where the enumeration is as follows:

```
StateAttributeType {
    Empty,
    Box,
    Agent,
    Agent_With_Box
}
```

The pair would be defined as:

```
Pair {
    Tuple3 {
        StateAttributeType,
        StateAttributeType,
        StateAttributeType
    },
    Node
}
```

A goal state for this data structure would be defined as when the current node has the corresponding tuple3:

```
Tuple3 {
    StateAttributeType.Box,
    StateAttributeType.Box,
    StateAttributeType.Agent
}
```

- (ii) Pickup and Drop Preconditions and Effects

**Pickup Operator** preconditions:

- Must be a valid box
- Agent must not be holding a box
- Agent must be in a square with a box

effects:

- Must be a valid box
- Agent now has a box.
- Box is removed from square

**Drop Operator** preconditions:

- Agent must be holding a box

effects:

- Box replaces empty state
- Agent state replaces agent with box state

- (b) (i) 22 nodes must be visited at minimum for breadth first to find a goal. It is guaranteed to find a goal state since it will expand every level of the tree. However, an optimal solution would not be found since the cost of going from state to state is not considered. This is because breadth first is not an admissible search algorithm.
- (ii) The fewest number of nodes that will be visited is 5 nodes. The search is not always guaranteed to reach the goal since the algorithm is not admissible and could for example end up stuck in a infinite depth caused by the switching between two states. For the same reason as Breadth first search an optimal solution would not be found.
- (iii) The fewest number of nodes that will be visited is 22 nodes. We know that the uniform cost search is admissible and by definition would reach a goal state of which is optimal.

### 3. Solution to problem 3

(a)

$$f(x, y, z) = x \ln x + y \ln y + z \ln z + \alpha(x + y + z - 1)$$

$$= x \ln x + y \ln y + z \ln z + \alpha x + \alpha y + \alpha z - \alpha$$

$$\nabla f = \frac{\delta f}{\delta x} \vec{i} + \frac{\delta f}{\delta y} \vec{j} + \frac{\delta f}{\delta z} \vec{k}$$

$$\frac{\delta f}{\delta x} = \ln x + \frac{x}{x} + \alpha = \ln x + \alpha + 1$$

$$\frac{\delta f}{\delta y} = \ln y + \frac{y}{y} + \alpha = \ln y + \alpha + 1$$

$$\frac{\delta f}{\delta z} = \ln z + \frac{z}{z} + \alpha = \ln z + \alpha + 1$$

$$\nabla f = (\ln x + \alpha + 1) \vec{i} + (\ln y + \alpha + 1) \vec{j} + (\ln z + \alpha + 1) \vec{k}$$

(b) **let**  $\alpha = 1$

$$\nabla f = (\ln x + 2) \vec{i} + (\ln y + 2) \vec{j} + (\ln z + 2) \vec{k}$$

$$\begin{aligned}\nabla f(0.9, 0.4, 0.01) &= [\ln(0.9) + 2]\vec{i} + [\ln(0.4) + 2]\vec{j} + [\ln(0.01) + 2]\vec{k} \\ \nabla f(0.9, 0.4, 0.01) &= 1.89464\vec{i} + 1.08371\vec{j} - 2.60517\vec{k}\end{aligned}$$

$$\begin{aligned}\nabla f(0.7, 0.2, 0.04) &= [\ln(0.7) + 2]\vec{i} + [\ln(0.2) + 2]\vec{j} + [\ln(0.04) + 2]\vec{k} \\ \nabla f(0.7, 0.2, 0.04) &= 1.64333\vec{i} + 0.390562\vec{j} - 1.21888\vec{k}\end{aligned}$$

$$\begin{aligned}\nabla f(0.1, 0.12, 0.14) &= [\ln(0.1) + 2]\vec{i} + [\ln(0.12) + 2]\vec{j} + [\ln(0.14) + 2]\vec{k} \\ \nabla f(0.1, 0.12, 0.14) &= -0.302585\vec{i} - 0.120264\vec{j} + 0.0338871\vec{k}\end{aligned}$$

(c) Compute the Hessian

$$H = \begin{bmatrix} \frac{\delta^2 f}{\delta x x} & \frac{\delta^2 f}{\delta x y} & \frac{\delta^2 f}{\delta x z} \\ \frac{\delta^2 f}{\delta y x} & \frac{\delta^2 f}{\delta y y} & \frac{\delta^2 f}{\delta y z} \\ \frac{\delta^2 f}{\delta z x} & \frac{\delta^2 f}{\delta z y} & \frac{\delta^2 f}{\delta z z} \end{bmatrix}$$

$$\frac{\delta^2 f}{\delta x} = \ln x + \alpha + 1$$

$$\frac{\delta^2 f}{\delta x x} = \frac{1}{x} \quad \frac{\delta^2 f}{\delta x y} = 0 \quad \frac{\delta^2 f}{\delta x z} = 0$$

$$\frac{\delta^2 f}{\delta y} = \ln y + \alpha + 1$$

$$\frac{\delta^2 f}{\delta y x} = 0 \quad \frac{\delta^2 f}{\delta y y} = \frac{1}{y} \quad \frac{\delta^2 f}{\delta y z} = 0$$

$$\frac{\delta^2 f}{\delta z} = \ln z + \alpha + 1$$

$$\frac{\delta^2 f}{\delta z x} = 0 \quad \frac{\delta^2 f}{\delta z y} = 0 \quad \frac{\delta^2 f}{\delta z z} = \frac{1}{z}$$

$$H = \begin{bmatrix} \frac{1}{x} & 0 & 0 \\ 0 & \frac{1}{y} & 0 \\ 0 & 0 & \frac{1}{z} \end{bmatrix}$$

(d) Is the function convex, concave or neither?

We know that if the eigen values are positive over the interior of the function's domain then a matrix is positive semidefinite. The eigenvalues are:

$$\lambda_1 = \frac{1}{x}, \lambda_2 = \frac{1}{y}, \lambda_3 = \frac{1}{z}$$

Since all the eigen values are positive over the interior of the functions domain then matrix is positive semidefinite.

$\therefore$  the function is convex

**(e) Is the domain of the function convex?**

In order to prove that the domain of the function is convex we must prove that the function is defined and finite on the interval  $[0, 1]$ .

In this case this will be true if we define  $0 \ln 0 = 0$  which will limit the behavior of  $x \ln x$ ,  $y \ln y$  and  $z \ln z$ . The function is therefore defined and finite for all values  $x, y, z \in [0, 1]$  and we know that  $[0, 1]^3$  is the cartesian product of intervals.

And since we also know that the cartesian product of convex sets is also convex.

$\therefore$  the domain is convex

**(f) Compute the optima of the function and say, for each point, whether it is a minimum or a maximum.**

The Jacobian is:

$$\nabla f(x, y, z) = (\ln x + \alpha + 1)\vec{i} + (\ln y + \alpha + 1)\vec{j} + (\ln z + \alpha + 1)\vec{k}$$

Note, that since  $x, y, z = 0$  is not in the domain, but this function has a value in the limit:

$$\lim_{x \rightarrow 0^+} (x \ln x + \alpha x) = 0$$

$$\lim_{y \rightarrow 0^+} (x \ln y + \alpha y) = 0$$

$$\lim_{z \rightarrow 0^+} (x \ln z + \alpha z) = 0$$

We treat points where some or all of  $x, y, z = 0$  as though they are in the domain, because to exclude them means there is no optimum near  $0, 0, 0$  although the function is increasing as  $x, y, z \rightarrow 0^+$  and bounded above in the local region. This means that  $(0, 0, 0)$  is a local maximum: any point with  $x, y, z < 0$  does not exist, and any nearby point with  $x, y, z > 0$  takes a smaller value.

We have stationary points for all  $x, y, z \in \{0, \frac{1}{e^{\alpha+1}}\}$ . This means we have  $2^3 = 8$  stationary points to consider.

The Hessian is:

$$H = \begin{bmatrix} \frac{1}{x} & 0 & 0 \\ 0 & \frac{1}{y} & 0 \\ 0 & 0 & \frac{1}{z} \end{bmatrix}$$

We have determined we should treat  $(0, 0, 0)$  as a local maximizer. The Hessian is undefined here, but at a local maximizer it must be negative definite. For a coordinate with a zero value we treat that row of the Hessian as though it contains a negative value on the diagonal and is zero elsewhere.

Next, consider  $(\frac{1}{e^{\alpha+1}}, \frac{1}{e^{\alpha+1}}, \frac{1}{e^{\alpha+1}})$ . At this point the Hessian has positive values along the diagonal and is zero elsewhere it is positive definite. Thus the point is a local maximizer.

Consider the remaining optima, where at least one coordinate is 0 and at least one coordinate is  $\frac{1}{e^{\alpha+1}}$ . Here the Hessian has a positive value in the diagonal in the row corresponding to  $\frac{1}{e^{\alpha+1}}$  and an effectively negative value in the row corresponding to 0. Hence these are all saddle points since it is indefinite whether it is positive or negative definite.

Since the function is unbounded above as  $x, y, z \rightarrow \infty$  there is no global maximizer. But the function is bounded below, hence  $(\frac{1}{e^{\alpha+1}}, \frac{1}{e^{\alpha+1}}, \frac{1}{e^{\alpha+1}})$  is the global minimizer.