Assignment 2

1.1

用 A* 搜索算法求解初始状态和目标状态如下图所示的 15 数码问题，写出算法过程。（PS： 只用写前 5 步和后 5 步完整状态，中间过程省略）

5	1	2	4
9	6	3	8
13	15	10	11
14	0	7	12

Ans：

前五步:

[5, 1, 2, 4]
[9, 6, 3, 8]
[13, 0, 10, 11]
[14, 15, 7, 12]

[5, 1, 2, 4]
[9, 6, 3, 8]
[13, 10, 0, 11]
[14, 15, 7, 12]

[5, 1, 2, 4]
[9, 6, 3, 8]
[13, 10, 7, 11]
[14, 15, 0, 12]

[5, 1, 2, 4]
[9, 6, 3, 8]
[13, 10, 7, 11]
[14, 0, 15, 12]

[5, 1, 2, 4]
[9, 6, 3, 8]
[13, 10, 7, 11]
[0, 14, 15, 12]

后五步:

[1, 2, 0, 4]
[5, 6, 3, 8]
[9, 10, 7, 11]
[13, 14, 15, 12]

[1, 2, 3, 4]
[5, 6, 0, 8]
[9, 10, 7, 11]
[13, 14, 15, 12]

[1, 2, 3, 4]
[5, 6, 7, 8]
[9, 10, 0, 11]
[13, 14, 15, 12]

[1, 2, 3, 4]
[5, 6, 7, 8]
[9, 10, 11, 0]
[13, 14, 15, 12]

[1, 2, 3, 4]
[5, 6, 7, 8]
[9, 10, 11, 12]
[13, 14, 15, 0]

1.2

在下图所示的博弈树中，进行 α − β 剪枝搜索，写出算法过程。

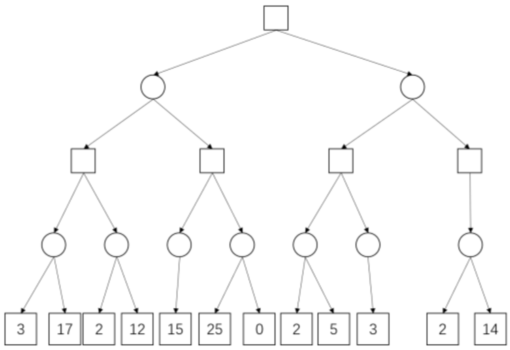
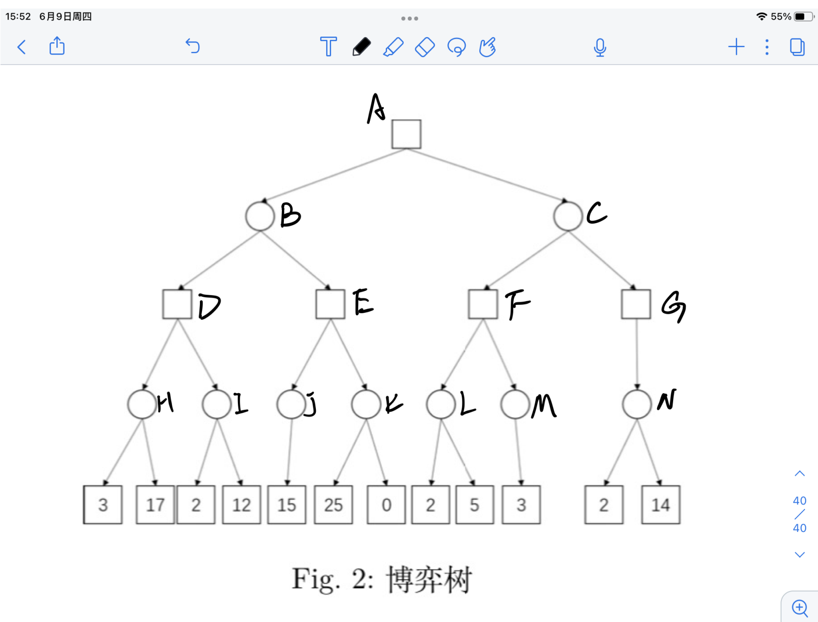


Fig. 2: 博弈树

Ans：



1. 在H点取min值，看其子节点，故返回b = 3，故D点的alpha为3

2. 看I点的子节点，为2，意味着I返回值不大于2，故将I的右儿子剪枝
3. 在D点取max值，故返回a = 3，故B点的beta为3
4. 看J点唯一的子节点，为15，故J返回值为15
5. E点取max值，故其返回值至少大于15，故将E的右儿子剪枝
6. 在B点取min值，故返回b = 3，故A点的alpha为3
7. 看L点的子节点，为2，意味着L返回值不大于2(小于alpha)，故将L的右儿子剪枝
8. L点返回b = 2， 看M点唯一的子节点，为3，故M点返回b = 3
9. F点取max值，故返回a = 3
10. 看N点的子节点，为2，小于3，故将N的右儿子剪枝
11. C点取min值，故返回b = 2
12. A点取max值，故返回3

1.3

如图 3 所示多层感知机模型，第一层是输入层，包含两个神经元：x1=0.08，x2=0.12 和偏置 b1；第二层是隐藏层，包含两个神经元：h1，h2 和偏置项 b2；第三层是输出：y1，y2。每条 线上标的 wi,j 是第 i 层第 j 个权重参数，激活函数是 sigmoid 函数（h 神经元之后），Loss 函数使用 MSE（均方误差）函数，真实标签 Label1 = 0.05, Label2 = 0.95，学习率 α = 0.5，求 在经过一次反向传播后所有权重参数的值（写出计算过程）。

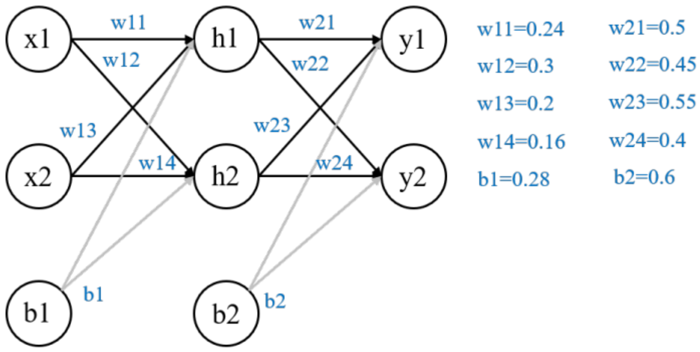


Fig. 3: MLP

Ans：

前向传播：

net_h1 = x1*w11 + x2*w13 + b1 = 0.3232

net_h2 = x1*w12 + x2*w14 + b1 = 0.3232

out_h1 = $\frac{1}{1+e^{-net_{h1}}}$ = 0.5801039

out_h2 = 0.5801039

net_y1 = out_h1*w21 + out_h2*w23 + b2 = 1.209109

net_y2 = out_h1*w22 + out_h2*w24 + b2 = 1.0930883

out_y1 = $\frac{1}{1+e^{-net_{y1}}}$ = 0.7701413

out_y2 = 0.7489628

反向传播：

MSE_1 = $\frac{1}{2}(label_1 - out_{y1})^2$ = 0.2593017

MSE_2 = $\frac{1}{2}(label_2 - out_{y2})^2$ = 0.0202079

MSE_total = MSE_1 + MSE_2 = 0.2795096

$\frac{\partial MSE}{\partial w_5}$ = abs($\frac{\partial MSE}{\partial out_{y1}}$ $\frac{\partial out_{y1}}{\partial net_{y1}}$ $\frac{\partial net_{y1}}{\partial w_5}$) = (label1 - out_y1)(out_y1)(1-out_y1)(out_h1)

= 0.0739528

w_{21}^+ = w21 - $\alpha\frac{\partial MSE}{\partial w_{21}}$ = 0.463023

同理可得

w_{22}^+ = 0.4390364

w_{23}^+ = 0.513023

w_{24}^+ = 0.3890364

$\frac{\partial MSE}{\partial w_{11}}$ = $\frac{\partial MSE}{\partial out_{h1}}\frac{\partial out_{h1}}{\partial net_{h1}}\frac{\partial net_{h1}}{\partial w_{11}}$ = $(\frac{\partial MSE_1}{\partial out_{h1}} + \frac{\partial MSE_2}{\partial out_{h1}})\frac{\partial out_{h1}}{\partial net_{h1}}\frac{\partial net_{h1}}{\partial w_{11}}$

= ($\frac{1}{2}(1-e^{-net_{y1}})^{-2}(-e^{-net_{y1}})w_{21}$+$\frac{1}{2}(1-e^{-net_{y2}})^{-2}(-e^{-net_{y2}})w_{22}$)(out_h1)(1-out_h1)(x1)

=0.0062793

故w_{11}^+ = w11 - $\alpha\frac{\partial MSE}{\partial w_{11}}$ = 0.2368603

同理可得

$\frac{\partial MSE}{\partial w_{12}}$ = 0.006205

$\frac{\partial MSE}{\partial w_{13}}$ = 0.009419

$\frac{\partial MSE}{\partial w_{14}}$ = 0.009308

w_{12}^+ = 0.2968973

w_{13}^+ = 0.1952904

w_{14}^+ = 0.1553459

最终参数为

w_{11}^+ = 0.2368603

w_{12}^+ = 0.2968973

w_{13}^+ = 0.1952904

w_{14}^+ = 0.1553459

w_{21}^+ = 0.463023

w_{22}^+ = 0.4390364

w_{23}^+ = 0.513023

w_{24}^+ = 0.3890364