

武汉大学计算机学院

2018-2019 学年度第一学期 2018 级

《认知过程的信息处理》期末考试试卷 (B) 答案 (开卷)

试题 1 解答:

由条件 $AB//CD$ 有, $Y_2=y_3$

由条件 $AC // BD$ 有: $(x_3-x_1)y_2=x_2y_3$

由条件 $AE // DE$ 有 $X_4y_3=x_3y_4$

由条件 $CB//EB$ 有 $Y_2x_4+y_4x_1=y_2x_1+y_4x_2$

求证目标: $4(x_4^2+y_4^2)=x_3^2+y_3^2$

试题 2 解答:

“吴方法”的基本思想是将几何问题代数化,即用多项式来表达几何问题的条件以及结论,通过证明条件所组成的多项式交集的零点是结论对应多项式的子集来完成证明。主要步骤如下:

- 将条件和结论用代数多项式表达;
- 确定自由变元和约束变元,对约束变元排序,确定消元的次序;
- 将条件所对应的多项式三角化,设三角化后的多项式为 $F_1, F_2, \dots, F_k, \dots, F_n$, 保证 F_k 只包含前面 k 个约束变元;
- 将结论所对应的多项式从三角形底部到顶部依次消除最后的约束变元,如果最后所得到的剩余多项式为 0 则表明命题为真, 否则命题为假。

试题 3 解答:

x_1-x_2 乘以 F_1 , x_3 乘以 F_2 , 再相加可得

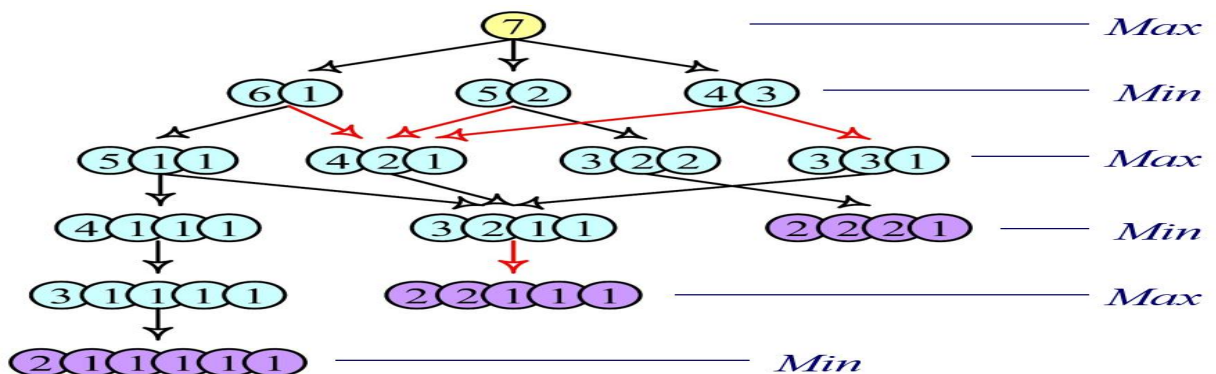
$$(x_1-x_2)x_4y_3+y_2x_3x_4-x_3y_2x_1=0$$

标准化后可得如下关于 x_4 的一元多项式:

$$X_4(x_1y_3-x_2y_3+y_2x_3)-x_3y_2x_1=0$$

试题 4 解答:

假设博弈者表示为 $Player \in \{Max, Min\}$, 开局为 $\{Max, 7\}$, 博弈树如下图所示.



试题 5 解答：

(1) 设定种群规模, 编码染色体, 产生初始种群。

将种群规模设定为 4；用 5 位二进制数编码染色体；取下列个体组成初始种群 S_1

$$s_1 = 13 \text{ (01101)}, \quad s_2 = 24 \text{ (11000)}$$

$$s_3 = 8 \text{ (01000)}, \quad s_4 = 19 \text{ (10011)}$$

(2) 定义适应度函数, 取适应度函数

$$f(x) = x^2$$

(3) 计算各代种群中的各个体的**适应度**，并对其染色体进行遗传操作, 直到适应度最高的个体，即 31 (11111) 出现为止。

● 选择：

$$P(x_i) = \frac{f(x_i)}{\sum_{j=1}^N f(x_j)}$$

$$q_i = \sum_{j=1}^i P(x_j)$$

选择-复制：设从区间 [0, 1] 中产生 4 个随机数：

$$\begin{aligned} r_1 &= 0.450126, & r_2 &= 0.110347 \\ r_3 &= 0.572496, & r_4 &= 0.98503 \end{aligned}$$

染色体	适应度	选择概率	积累概率	选中次数
$s_1=01101$	169	0.14	0.14	1
$s_2=11000$	576	0.49	0.63	2
$s_3=01000$	64	0.06	0.69	0
$s_4=10011$	361	0.31	1.00	1

于是，经复制得群体：

$$\begin{aligned} s_1' &= 11000 \text{ (24)}, & s_2' &= 01101 \text{ (13)} \\ s_3' &= 11000 \text{ (24)}, & s_4' &= 10011 \text{ (19)} \end{aligned}$$

● 交叉

设交叉率 $p_c=100\%$ ，即 S_1 中的全体染色体都参加交叉运算。

假设 s_1' 与 s_2' 配对， s_3' 与 s_4' 配对。分别交换后两位基因，得新染色体：

$$\begin{aligned} s_1'' &= 11001 \text{ (25)}, & s_2'' &= 01100 \text{ (12)} \\ s_3'' &= 11011 \text{ (27)}, & s_4'' &= 10000 \text{ (16)} \end{aligned}$$

- 变异

设变异率 $p_m=0.01$, 则上述群体 S_t 中共有 $5 \times 4 \times 0.01 = 0.2$ 位基因可以变异, 故本轮遗传操作不进行变异。

于是, 得到第二代种群 S_2 :

$$s_1=11001 (25), \quad s_2=01100 (12)$$

$$s_3=11011 (27), \quad s_4=10000 (16)$$

算法循环, 直至满足终止条件.

试题 6 解答:

Google Duplex 三大主要模块为自动语言识别系统 (ASR)、循环神经网络 (RNN)、文本转语音系统 (TTS)。

自动语言识别系统: 该模块功能为将听到的声音转换为文本信息。

循环神经网络: 该模块功能为理解输入的文本信息, 并产生对话内容(回答)。

文本转语音系统: 将文本信息转换为语音, 并决定语音在语调、语气以及一些语言习惯上的特征, 使之更加自然。

试题 7 解答:

DeepQA 的核心设计准则是: (1) 大规模并行处理; (2) 概率问题和内容分析的整合; (3) 可行度评估; (4) 浅层与深层知识的整合。

试题 8 解答:

策略网络是一个 13 层的深度卷积神经网络, 它从当前已有的棋局中进行学习, 根据输入的棋局状态, 预测所有合法的下一步的概率, 选择概率较大的位置落子, 模拟人类棋手的落子棋感, 属于局部行为。估值网络是一个 12 层的深度卷积神经网络, 它侧重于全局形势判断, 模拟人类棋手的胜负棋感, 输出单个预测结果。

试题 9 解答:

深度学习与浅层学习的区别是网络的深度不同, 或者网络的层数不同, 深度学习极端情况下可达到上百层。对于简单的模式识别问题, 浅层学习的分类工具就足够了, 但模式变得非常复杂时, 就需要深度学习来实现。深度学习区别于传统机器学习主要在以下四个方面:

- 强调 ANN 模型结构的深度, 与浅层学习相比使用了更多的隐藏层。
- 突出特征学习的重要性, 通过逐层特征转换, 将数据在原始空间的特征表示变换到一个新的特征空间, 使得分类或预测变得更容易而且精度更高。
- 深度学习来源于人工神经网络的发展, 但是训练的方式与传统人工神经网络不同, 采用逐层训练的方式, 然后对网络参数进行微调。
- 深度学习利用大数据来学习特征, 而浅层学习一般不需要。

试题 10 解答:

语音分析: 根据音位规则, 从语音流中区分出一个个独立的音素, 再根据音位形态规则找出一个个音节及其对应的词素或词。

词法分析: 找出词汇的各个词素, 从中获得语言学的信息。

句法分析: 对句子和短语的结构进行分析, 找出词、短语等的相互关系以及各自在语句中的作用。

语义分析: 找出词义、结构意义及其结合意义, 从而确定语言所表达的真正含义或概念。

语用分析：研究语言所存在的外界环境对语言使用者所产生的影响。

试题 11 解答：

知识图谱（Knowledge Graph）的概念由谷歌 2012 年提出，旨在实现更智能的搜索引擎。知识图谱本质上是一种叫做语义网络（semantic network）的知识库，即具有有向图结构的一个知识库，其中图的结点代表实体或者属性，而图的边代表实体与实体或者实体与属性之间的各种语义关系。

基于知识图谱的检索和基于关键词检索的主要区别在于：找到最想要的信息；提供全面的摘要；让搜索更有深度和广度。例如，当用户输入“Jackie Chan”，基于知识图谱的搜索引擎可以识别出 Jackie Chan 其实就是成龙，而且会给出成龙的各种属性信息，比如说出生日期、国籍、配偶等。基于关键词的检索则不能做到这些。