**姓名：**  **学号：** **任课教师:**

一、论文《Mathematical Analysis of Algorithm》

1. 论文中的In Situ Permutation算法最坏情形的时间复杂度是多少？对应的输入是怎样的？（6分）
2. 如果按论文中所述增加了tally变量，其最坏情形的时间复杂度是怎样的？对应的输入又是怎样的？（6分）
3. 在增加了tally变量后，第6行代码循环次数平均减少多少次？（注：论文中的结论有误）（8分）
4. **论文《PRIMES is in P》**
5. 请说明文章标题“PRIMES is in P”中的“PRIMES”指什么？ （3分）
6. 文章摘要：“We present an unconditional deterministic polynomial-time algorithm determines whether an input number is prime or composite.” 请分别、简单解释 “unconditional”，“deterministic”，“polynomial-time” 这三个词在此处的含义。（6分）
7. 给定及假设

证明： （3分）

1. 某同学将论文（Sec. 4）中AKS算法里的第2-4步作了如下改动（保留其他步不变）：

2. Find the smallest such that either

or

and .

3. If output PRIME.

4. If , output COMPOSITE.

请判断：修改后的算法是否具有与AKS算法同样的效果，即是否仍然是一个unconditional deterministic polynomial-time素性测试算法？（只给出判断“是”或“否”，无需说明理由。）（3分）

1. 数学家陶哲轩教授写了一篇博客文章The AKS primality test（https://terrytao.wordpress.com /2009/08/11/the-aks-primality-test/）解释AKS素性测试。下面这段话是署名Jonathan Vos Post的留言，请将其翻译成中文。（5分）

“Thank you for the clarity in showing that PRIME is in P. Now, is SEMIPRIME in P? By definition, a semiprime is a product of exactly two primes, not necessarily distinct. They matter because of an important class of cryptosystems, in a multibillion dollar industry. The complication: there exist specific semiprimes in the literature that have been proven to be semiprimes, without any prime factorization know. We don’t know the complexity class for prime factorization. So what is the complexity class for determining if a given integer is or is not a semiprime? Several experts in number theory and quantum computing have told me that mine is an interesting question, but probably very difficult. I mention quantum computing because SEMIPRIME might be in a qc complexity class of interest. What do you think?”

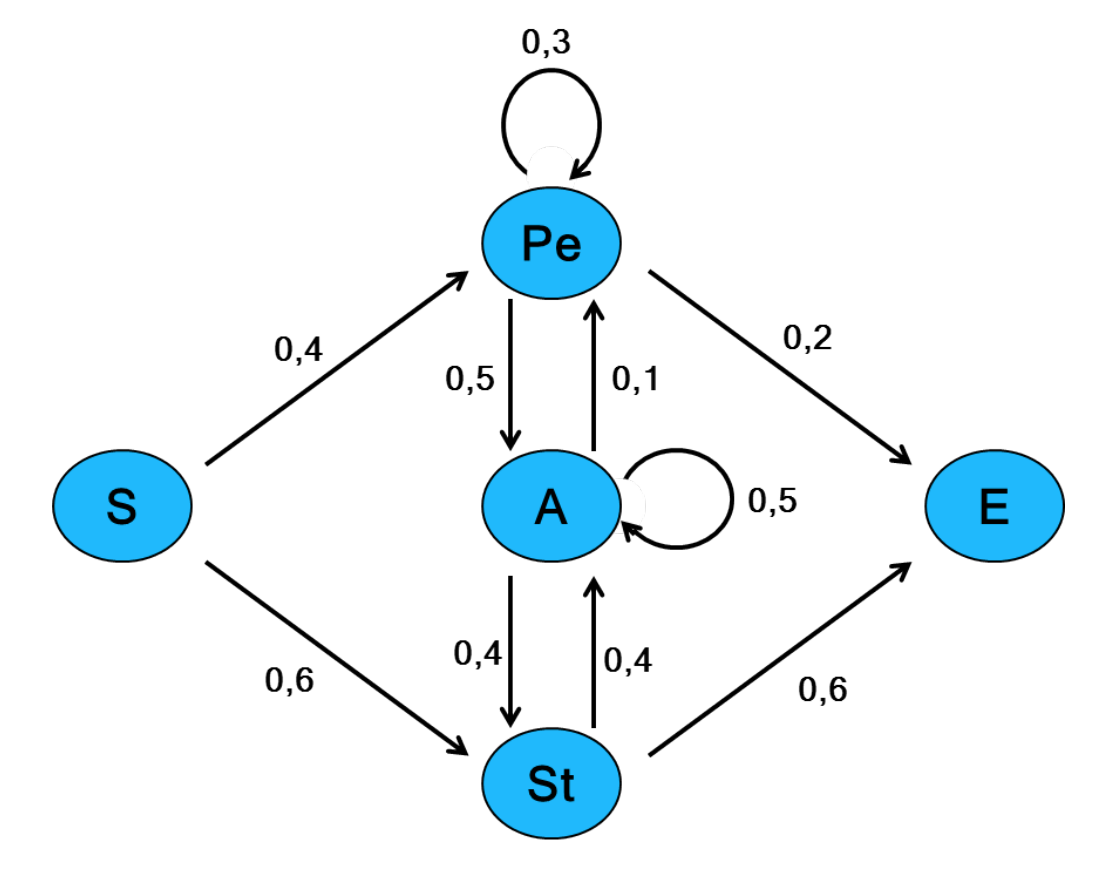
三、隐马尔科夫模型（Hidden Markov Model）可以被用来识别文本中的实体类型。

(1) 我们考虑两种实体类型People(Pe)和Cities(St)。假设下面给出了已经标注出两种实体类型的训练数据，其中，$代表句子的开始，€代表句子的结束：

所有没有被标注实体类型的单词，可以看成属于类别A（=other words）。

请画出上述问题的隐马尔科夫模型的状态转移图，图中需包含开始状态S（代表句子开始）、结束状态E（代表句子结束）和隐藏状态Pe、St、A. 根据训练数据，标出状态转移图中每条边的概率。然后写出隐藏状态Pe的所有输出概率(Emission Probilities)。（10分）

(2) 对于上述问题，假设给定其隐马尔科夫模型如下图所示：（10分）



同时我们给出每个状态的输出概率如下：

请用维特比(Viterbi)算法计算出以下句子的最有可能的标注序列：Paris lives in Denver

四、论文《Word Image Matching Using Dynamic Time Warping》

(1) 抛开具体的任务，DTW要解决什么问题，试举例说明传统的欧式距离有什么问题？（5分）

1. 给定两个序列X = {x1, x2, x3 …, xn}， Y = {y1, y2, y3 …, ym}, xi和yj的相似度为d(xi, yj)。形式化描述DTW动态规划算法，计算算法的计算复杂度，并描述状态转移方程中每一步决策的意义。(5分)
2. 本篇论文在动作规划的基础上加入了什么限制？如何体现在状态转移方程上？时间复杂度有什么变化？(3分)

(4) 文中是如何把图片信号转化为最后DTW使用的特征的，用了哪些方法，目的分别是什么？（7分）

五、论文《QUIC-SVD：Fast SVD using cosine trees》

(1) 假设，SVD全分解算法的复杂度是多少？ （3分）

1. 论文中QUIC-SVD快速算法的复杂度是多少？（3分）
2. QUIC-SVD算法中每一次迭代在V中增加几个向量？（3分）
3. 为什么可以采用cosine函数衡量向量之间的距离？（3分）
4. 为什么要用蒙特卡罗方法估计误差？（4分）
5. 简述QUIC-SVD算法的思路。（4分）