

一、选择或填空题 (每空2分, 共30分)

1. 下列陈述错误的是 c

(a) 数值概率算法一般是求数值计算问题的近似解 ✓

(b) Monte Carlo总能求得问题的一个解, 但该解未必正确 ✓

(c) Las Vegas算法的一定能求出问题的正确解 决不返回错误的结果, 即返回的结果必正确,

(d) Sherwood算法的主要作用是减少或是消除好的和坏的实例之间的差别 ✓ 但有时根本找不到答案

2. 下列陈述错误的是 d

(a) 概率算法的期望执行时间是指反复解同一个输入实例所花的平均执行时间 ✓

(b) 概率算法的平均期望时间是指所有输入实例上的平均期望执行时间 ✓

(c) 概率算法的最坏期望时间是指最坏输入实例上的期望执行时间 ✓

(d) 概率算法的期望执行时间是指所有输入实例上的所花的平均执行时间

3. 下述算法是求有限集 X 的势 $n=|X|$ 。请选择正确语句填空, 算法的时间复杂度是 $n^{\frac{1}{2}}$ 量级。

b SetCount(X) ($k:=0$; $S:=\emptyset$;

$a:=\text{uniform}(X)$;

do { $k:=k+1$; $S:=S \cup \{a\}$; $a:=\text{uniform}(X)$;

$a \notin S$ }while($a \notin S$);

return $2k \cdot \pi$;

// $\pi=3.14159\dots$

}

(a) $S:=S \cup \{a\}$

(b) a 属于 S

(c) a 不属于 S

(d) $a:=\text{uniform}(X)$

(e) $S:=\{a\}$

(f) n 量级

(g) n^2 量级

(h) $n^{1/2}$ 量级

(i) $\lg n$ 量级

(j) 常量级

4. Sherwood算法中随机预处理提供了某种加密计算 $f(x)$ 的可能, 其步骤是

(a) 使用函数 u 将 x 加密为某一随机实例 y

(b) 将 y 提交给 f 计算出 $f(y)$ 的值

(c) 使用函数 v 转换为 $f(x)$

5. Las Vegas算法的一般形式为obstinate(x){repeat LV(x,y,succes) until
sucess; return y};
当用他来解8皇后问题时, 设LV成功的概率 $p=$

二, 简要回答下述问题 (每题8分, 共32分)

1. 若要将一个偏y的, 55%—正确的, 一致的MC算法改进到95%—正确的算法, 需要重复调用MC算法多少次? 并给出推导过程。

$$(1 - (1 - 0.55)^x) \geq 95\% \rightarrow 0.45^x \leq 0.05 \rightarrow x \geq \lg 0.05 / \lg 0.45 = 3.75 = 4$$

2. 在分布式算法中, bit复杂性是指算法发送的所有消息中bit的总数; 消息链复杂性是指算法的任何执行中最长消息链的长度, 若某消息链是 m^1, m^2, \dots, m^k , 则消息 m^i 在因果关系上领先于消息 $m^{(i-1)}$, 该消息链的长度为k。请问这两种复杂度应分别属于通信复杂性和时间复杂性中的哪一种? 并简述其理由。

3. 在分布式算法的时间复杂性和ont-time复杂性中, 一个msg的延迟分别假定为至多1个时间单位和恰好1个时间单位, 但有时后者是前者的一个下界。为什么? 举例说明。

◇ 时间复杂性

① 一个分布式算法的时间复杂性是满足下述两个假定的一个计算所耗费的最大时间

T1: 一个进程在零时间内可计算任何有限数目的事件

T2: 一个msg的发送和接受之间的时间至多为1个时间单位

缺点: 针对一算法的所有计算, 其结果可能是极不可能发生的计算。

② 一个分布式算法的one-time复杂性是满足下述假定的一个计算的最大时间

O1: 同T1

O2: 发送和接收一个msg之间的时间恰好是1个单位时间

缺点: 某些计算可能被忽略, 而其中可能有极其耗时的计算

表面上, 1-time复杂性至少等于时间复杂性, 因为T2假定下的最坏时间不会高于O2假定下的时间。但事实并非如此, 而往往O1和O2假定之下的1-time复杂性是前一种时间复杂性的一个下界。

例如: 在echo算法里1-time复杂性是 $O(D)$, 时间复杂性是 $\Theta(N)$, 即使直径为1的网

络。

4. 对于同步环, 在一个均匀的leader选举算法中, 为什么一个id为i的msg是以 2^i 速率被转发的? 其目的是什么?

三, 算法设计题:

1. 量子运动的随机聚集过程可用量子赌博来描述。其规则是:

- (1) 开始时, A和B的赌本分别为x和y;
- (2) 每次通过掷一枚神奇的硬币来决定输赢。设正面A赢, 反面B赢, 但每次仍出硬币的正反面的概率正比于A和B当前的赌本;
- (3) 每次的输家将按固定的比例k从自己的赌本中付给赢家;
- (4) 设最小的赌本单位为1, 若输家当前的赌本小于等于1, 他付出自己的赌本后, 游戏结束。

例如: 设x和y的初值分别为20分和80分, $k=10\%$, 则第一次硬币仍出正面和反面的概率分别是20%和80%, 若扔出的是正面, 则B要付8分给A; 第二次赌博时, $x=28$, $y=72$, 硬币扔出正面和反面的概率将分别是28%和72%。赌博依此规则进行, 直至一方赌光为止。

要求:

- 1——写一算法实现赌博游戏: (15分)
 - 2——A和B最终输赢取决于什么? (3分)
 - 3——请分析A、B最终输赢的概率。(5分)
- (1)

变量说明:

a, b, p 整型

k 浮点小数

Gamble(A的赌本a, B的赌本b)

```
{
    While(a>0 && b>0)
    {
        p = uniform(0, a+b)
```

```

If(p<a) //0~a-1正面, a~a+b-1背面
{
    If(b>1)
    {
        a = a + K * b
        b = (1 - k) * b
    }
    Else
    {
        b = -1
    }
}
Else
{
    If(a>1)
    {
        b = b + k * a
        a = (1 - k) * a
    }
    Else
    {
        a = -1
    }
}
}

If(a<0) a输
Else b输
}

```

(2) 取决于赌本

(3) $A(a_0, b_0) = a_0 / (a_0 + b_0) \times A(a_0 + b_0 * k, (1 - k) * b_0) + b_0 / (a_0 + b_0) \times A((1 - k) * a_0, b_0 + a_0 * k)$

2. 设集合S和T中各有n个互不相同元素, 要求:

1——写一Monte Carlo算法判定S和T是否相等：(10分)

2——分析算法出错的概率：(3分)

3——算法是否有偏，若有偏，偏什么？(2分)

(1)

```
MC(S,T) {  
    i = uniform(1..n)  
    for j = 1 to n do {  
        if S[i]=T[j] return true  
    }  
    Return false  
}
```

(2)

(3)偏假



中国科学技术大学计算机学院
2013 级研究生《算法设计与分析》期终考试试题

学号或单位:

姓名:

成绩:

注意: 试卷须交回, 否则无分。

一. 单项选择题 (每空 3 分, 共 30 分)

1. 下列陈述错误的是_____。

- (a) 数值概率算法一般是求数值计算问题的近似解
- (b) Monte Carlo 总能求得问题的一个解, 但该解未必正确
- (c) Las Vegas 算法的一定能求出问题的正确解
- (d) Sherwood 算法的主要作用是减少或是消除好的和坏的实例之间的差别

2. 下述算法的估计值是_____。

```
Darts(n){  
    k:=0;  
    for i:=1 to n do {  
        x:=uniform(0,1); y:=x;  
        if ( $x^2+y^2 \leq 1$ ) k++;  
    }  
    return 4k/n;  
}
```

- (a) π (b) $\sqrt{2}$ (c) $2\sqrt{2}$ (d) $1/\sqrt{2}$

3. 若 A 是一个偏真的 MC 算法, 则下述陈述正确的是_____。

- (a) 只有 A 返回 true 时解正确;
- (b) A 以较大的概率返回 true;
- (c) A 返回 true 时解必正确, A 返回 false 时解必错误;
- (d) A 返回 true 时解必正确, A 返回 false 时有可能产生错误的解。

4. 重复调用一个一致的、p-正确的、偏真的 MC 算法 k 次, 可以得到一个_____的算法。

- (a) $(1-p)^k$ -正确 (b) $(1-p)^k$ -正确 (c) $(1-(1-p)^k)$ -正确 (d) 正确概率不能确定

5. 用 Las Vegas 算法求解 n 皇后问题的形式为:

```
obstinate(T, n) { //求出的解存放在 T[1..n]中  
    repeat  
        QueensLV(T, n, success);  
    until success;  
    return T; //返回解  
};
```

若 obstinate 陷入无限循环, 则可以断定_____成立。

- (a) $n < 4$ (b) $n = 4$ (c) $n = 8$ (d) n 太大

6. 在下述因素中, 已知有 3 个阻碍分布式系统了解系统的全局状态, 与全局状态无关的是_____。

- (a) 非即时的通信 (b) 相对性影响 (c) 中断 (d) 算法的正确性

7. 下列说法错误的是_____。

- (a) 在一个精确的异步系统中，一个算法的执行只取决于初始配置
- (b) 在一个精确的同步系统中，一个算法的执行只取决于初始配置
- (c) 在一个精确的异步系统中，对于相同的初始配置，一个算法可能有不同的执行
- (d) 异步系统中的消息延迟是不确定的

8. 在异步系统中，leader 选举算法的消息复杂性下界是_____。

- (a) $O(\log n)$
- (b) $O(n)$
- (c) $O(n \log n)$
- (d) $O(n^2)$

9. 已知事件 e_1 , e_2 , e_3 和 e_4 的向量时间戳分别为 $(4,0,0,0)$, $(3,5,0,0)$, $(3,4,1,2)$, $(3,6,4,3)$ ，与 e_3 有因果关系的是_____。

- (a) e_1
- (b) e_2
- (c) e_4
- (d) 都无关系

10. 下列陈述错误的是_____。

- (a) P 类问题可用多项式时间的确定性算法求解。
- (b) NP 类问题可用多项式时间的非确定性算法求解。
- (c) 所有需要指数时间求解的问题均属于 NP 类问题。
- (d) NP 完全问题也是 NP-hard 问题。

二. 简要回答下述问题 (每题 8 分, 共 40 分)

1. 在分布式算法中, bit 复杂性是指算法发送的所有消息中 bit 的总数; 消息链复杂性是指算法的任何执行中最长消息链的长度, 若某消息链是 m_1, m_2, \dots, m_k , 则 m_i 在因果关系上领先于 m_{i+1} , 该消息链的长度为 k 。请问这两种复杂性应分别属于通信复杂性和时间复杂性中的哪一种? 并简述其理由。
2. 已知事件 e_1 , e_2 , e_3 和 e_4 的向量时间戳分别为 $(1,0,0,0)$, $(3,5,0,0)$, $(0,0,1,2)$, $(3,6,4,3)$ ，与 e_3 有因果关系的是哪个事件? 若该事件发生在 e_3 之前, 则会发生什么情况?
3. 对于优化问题 Π , 最佳可达性能比 $R_{\text{min}}(\Pi)$ (定义如下) 分别为何值时, 问题 Π 易于近似和难于近似?

$$R_{\text{min}}(\Pi) = \inf\{r \geq 1 \mid \exists \Pi \text{ 的多项式时间算法 } A \text{ 使 } R_A^* \leq r\}$$

4. 对于一个优化问题, 什么情况下其近似算法的绝对性能比和渐近性能比相同?
5. 装箱问题是将 n 件物品放入尽可能少的若干个箱子中, 不妨设每个箱子的容量均为 1, 物品 $1, 1, \dots, j, \dots, n$ ($n=6$) 的大小依次为: 0.5, 0.6, 0.3, 0.7, 0.5, 0.4。请给出其最优解, 以及采用首次适应 (First Fit) 策略得到的近似解。这里, 解是指使用了几个箱子, 每个箱子中放了哪些物品。

三. 算法题 (共 30 分)

1. 设一个同步匿名的单向环有 n 个结点, 每个结点均知道 n , 每个结点的初始均状态相同, 每个结点上的程序相同且开始于同一时刻。

- (1) 请问是否存在一个确定的算法选出一个 leader? 请简述理由。(5 分)
- (2) 试设计一个概率的 leader 选举算法。提示: 算法由若干个 phase 构成, 每个 phase 包括 n 轮, 可用 phase 和轮控制算法流程, 每个结点可以设置一个随机数发生器 $\text{uniform}(1..m)$, 这里 m 是局部变量, 初值等于 n 。(20 分)
- (3) 请问你设计的概率算法属于哪一类算法? (5 分)

算法分析

一、单选(11*3)

1、下列描述正确的是 A

- A、概率算法的期望执行时间是指反复解同一输入实例所花的平均执行时间
- B、概率算法的期望执行时间是指所有输入实例上所花的平均执行时间
- C、概率算法的平均期望时间是指算法执行时间的上界
- D、概率算法的最坏期望时间是指算法执行时间的上界

2、当问题只有一个正确的解，不存在近似解时，某概率算法总是给出一个未必正确的解，但是随着调用该算法次数的增加，可将错误的概率控制在任意给定的范围，该算法属于 C

- A、数字概率算法
- B、Las Vegas 算法
- C、Monte Carlo 算法
- D、Sherwood 算法

3、Las Vegas 算法的一般形式是_____

```
Obstinate(x){
    Repeat
        LV(x,y,success)
    Until success;
    Return y
}
```

设 $p(x)$ 是 LV 成功的概率， $s(x)$ 和 $e(x)$ 分别是 LV 成功和失败的期望时间， $t(x)$ 是算法 obstinate 得到一个正确解的期望时间，则 $t(x)$ 的表达式应该是 B

- A、 $t(x) = s(x) + e(x)(1-p(x))/p(x)$
- B、 $t(x) = p(x)t(x) + (1-p(x))(e(x) + t(x))$ $t(x) = p(x)s(x) + (1-p(x))(e(x) + t(x))$
- C、 $t(x) = p(x)s(x) + (1-p(x))(e(x) + s(x))$
- D、 $t(x) = p(x)s(x) + (1-p(x))(t(x) + s(x))$

4、若一个一致的、 p -正确的 MC 算法是有偏的，则 p 至少应该满足 B

- A、 $p < 0$
- B、 $p > 0$
- C、 $p \geq 1/2$
- D、 $p > 1/2$

5、若 A 是一个偏真的 MC 算法，则下列陈述正确的是 D

- A、只有 A 返回 true 时解正确
- B、A 以较大的概率返回 true
- C、A 返回 true 时解必正确，A 返回 false 时解必错误
- D、A 返回 true 时解必正确，A 返回 false 时有可能产生错误的解。

6、用 Las Vegas 算法求解某问题，已知 obstinate(x) 找到正确的解的期望时间是 288。其中 LV 成功的概率为 $p(x) = 0.2$ ，成功时的期望 $s(x)$ 是 8，则失败的期望时间 $e(x)$ 是 A

- A、70
- B、102
- C、210
- D、280

7、一个 MC 算法是一致的、 $3/5$ -正确的，偏 y_0 的，若要求出错概率不超过 ϵ ，则重复调用 MC 至少为 109

A、

B、

$$(1 - \frac{3}{5})^k = \epsilon$$

$$k = \log_{\frac{2}{5}} \epsilon$$

C、
D、

8、若两个环 x_0, x_1, \dots, x_{n-1} 和 y_0, y_1, \dots, y_{n-1} 是序等价的，则通常是指_____

- A、若对每个 $i \in [0, n-1]$ ，均有 x_i 和 y_i 匹配
- B、若对每个 $i \in [0, n-1]$ ，均有 x_i 和 y_i 匹配
- C、若 $i < j$ ，则有 $x_i < x_j$ ，且 $y_i < y_j$
- D、要求 x_0, x_1, \dots, x_{n-1} 和 y_0, y_1, \dots, y_{n-1} 均是有序序列

9、在异步环上，一个 $O(n^2)$ 的 leader 选举算法按顺时针单向发送消息，假设只有最大标识符的结点可以当选为 leader，则当环上标识符次序为_____时该算法发送的消息数量最多。

- A、逆时针 0, 1, 2, ..., n-1
- B、逆时针 n-1, n-2, ..., 0
- C、顺时针 0, 1, 2, ..., n-1
- D、顺时针 n-1, n-2, ..., 0

10、下列序列代表的环中，没有空隙的环是_____

- A、10, 30, 20, 40, 60, 90, 80, 100
- B、10, 20, 30, 40, 50, 60, 70, 80
- C、1, 9, 30, 40, 50, 60, 70, 80
- D、其他序列

11、设正整数 d_1, d_2, \dots, d_n 是 n 个结点的标识符集合， $x = \min\{d_1, d_2, \dots, d_n\}$ ， $y = \max\{d_1, d_2, \dots, d_n\}$ ，则同步环上非均匀的 leader 选举算法的时间复杂度是_____

- A、 $O(n)$
- B、 $O(xn)$
- C、 $O(yn)$
- D、 $O(n \log n)$

二、简答题 (4*8) 元素问题

1. 设 $F(x)$ 是一个布尔函数，若 $F(x) = 1$ 的概率为 $\frac{1}{2}$ ，且 $F(x)$ 是随机的，则 $F(x)$ 是随机的还是偏真的？请分析 $F(x)$ 出错的概率是多少？

$F(x)$;

```
if F(x) then
    return true
else
    return F(x);
```

偏真

$F(x)$ 返回 true 概率 $> \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$
则出错概率小于 $\frac{1}{4}$ 。

2. 已知事件 e_1, e_2, e_3 和 m_1 的时间戳分别为 $(1, 0, 0, 0)$, $(2, 5, 0, 0)$, $(0, 0, 1, 2)$, $(3, 6, 4, 3)$ ，请列举出所有并发事件，以及所有因果相关事件。

3、对于同步环，一个均匀的 leader 的选举算法的消息复杂度是多少？算法中一个 id 为 i 的 msg 以 2^i 的速率被转发的目的是什么？简述原因，算法的时间复杂度是多少？

4、试举例说明 Causal Msg Delivery 算法可能出现的死锁情况。并分析为什么该算法通常被应用与织播通信的一部分？

三、算法题 (35)

1、设网络的生成树已经建立，各个节点 P_i 的 id 为 i ，并持有初值 x_i ，且 id 和持有的初值均互不相同，试写一个分布式算法使得根节点知道书中持有初值最大的节点，以及持有初值最小的节点。

2、设集合 S 和 T 中各有 n 个互不相同的元素，要求：

- 写一 Monte Carlo 算法判定 S 和 T 是否相等
- 分析算法出错的概率
- 算法是否有偏，若有偏，偏什么？

偏假

```

Isequal(S, T)
{
    i ← uniform(1..n)
    for j ← 1 to n do
        if S[i] = T[j] then
            return true;
    return false;
}
    
```

返回 false 一定有 $S \neq T$.

返回 true S 可能等于 T ，可能不等于 T

若 $S = T$ ，则无论 i 为何值， $S[i] = T[i]$

若 $S \neq T$ ， $S[i] \neq T[i]$ ，即 S 与 T 不相等，且 $i \neq k$ 时出错，算法误判相等。

出错概率 $\leq \frac{1}{n}$