计算机科学与技术学院
中国科学技术大学

随机算法习题

算法设计与分析

秋季 2023

截止时间：2023 年 12 月 29 日 18:00

题目 1 *(Markov 不等式的一个应用：把 Las Vegas 算法变成 Monte Carlo 算法) 25 分*

令 $\mathcal{A}$ 为某个判定问题 P 的随机算法（例如，判定图是否连通）。假设对于任何给定大小为 n 的 P 的实例，$\mathcal{A}$ 运行时间的期望为 $T(n)$，并且始终输出正确的答案。
使用 $\mathcal{A}$ 为问题 P 提供新的随机算法 NewAlg，使得 NewAlg 总是在 $100T(n)$ 时间内运行完毕，并且

- 如果输入是"是"实例[1], 则它将以至少 $\frac{9}{10}$ 概率被接受;
- 如果输入是"否"实例，则它将总是被拒绝。

描述你的算法，并证明其正确性和运行时间。

提示：你可以使用 Markov 不等式。

题目 2 *15 分*
假设 R 为在街上随机找的人的智商。请问找到的人智商至少为 150 的概率是多少？假设人群平均智商为 100，R 的方差为 80。

题目 3 *35 分*

假设每个 CPU 在任何时候最多可以执行一个进程。考虑以下实验，该实验进行一系列轮。对于第一轮，我们有 n 个进程，这些进程被独立且均匀地随机分配给 n 个 CPU。对于任意的第 $i \geq 1$ 轮，在第 i 轮之后，对于进程 p，如果 p 是已分配给 C 的唯一进程，我们就在第 i 轮中移除所有此类进程（因为它们将被执行）。剩余进程被保留到第 $i+1$ 轮，然后再被独立且均匀地随机分配给 n 个 CPU。

(a) 如果一轮开始时有 b 个进程，那么下一轮开始时进程数的期望是多少？

(b) 假设每一轮移除的进程数恰好是移除进程数的期望值。请证明所有进程将在 $O(\log \log n)$ 轮内被移除。

提示：如果 x_j 是 j 轮后留下进程数量的期望值，可以证明并使用 $x_{j+1} \leq x_j^2/n$。你还可以使用事实 $1 - kx \leq (1-x)^k$，对于 $0 < x < 1$ 和 $k \leq \frac{1}{x}$ 成立。

---

[1]如果问题的正确答案是"是"，则实例是"是"实例。否则（正确答案是"否"），则为"否"实例。

题目 4 *25 分*

设 X 和 Y 为有限集，并设 Y^X 表示所有函数从 X 到 Y。我们将这些函数视为“哈希”函数。一个族 $\mathcal{H} \subseteq Y^X$ 被称为强 2-全域的，如果以下性质成立，其中 $h \in \mathcal{H}$ 均匀随机抽取：

$$\forall x, x' \in X \; \forall y, y' \in Y \left(x \neq x' \Rightarrow \Pr_h[h(x) = y \wedge h(x') = y'] = \frac{1}{|Y|^2} \right) .$$

我们给定一个由 X 中元素构成的流 $\mathcal{S}$，其中包含最多 s 个不同的元素。假设 $\mathcal{H} \subseteq Y^X$ 是一个强 2-全域哈希族，其中 $|Y| = cs^2$，$c > 0$ 是某个常数。假设我们使用一个随机函数 $h \in \mathcal{H}$ 进行哈希。试证明碰撞（即 $\mathcal{S}$ 中两个不同元素哈希到相同位置的事件）的概率至多为 $1/(2c)$。
