Computer Science 341
Discrete Mathematics

Midterm Exam

Due at the beginning of class on Wednesday, November 9, 2005

Problem 1

Express each of the following summations as $\Theta(f(n))$ where $f(n)$ is an appropriate function of n in closed form.

a) $\sum_{k=1}^{n}(1.01)^k$.

b) $\sum_{k=1}^{n}\frac{1}{\sqrt{k}}$.

c) $\sum_{k=2}^{n}\frac{1}{k\log k}$.

d) $\sum_{k=1}^{n}\left(1+\frac{1}{k}\right)^k$.

Solution:

a)

$$\sum_{k=1}^{n}(1.01)^k = \frac{1.01^{n+1}-1.01}{1.01-1} = 101\cdot 1.01^n - 101 = \Theta(1.01^n).$$

Answer: $\Theta(1.01^n)$.

We use the following theorem in our solution for parts (b) and (c):

Theorem. Let $f:[a,+\infty)\to\mathbb{R}$ be a nonincreasing integrable positive function (where a is a natural number). Then

$$\sum_{i=a}^{n} f(n) = \Theta\left(\int_a^n f(t)dt\right).$$

(see below for the proof)

b) Since the function $\frac{1}{\sqrt{k}}$ is nonincreasing

$$\sum_{i=1}^{n}\frac{1}{\sqrt{k}} = \Theta\left(\int_1^n \frac{dt}{\sqrt{t}}\right) = \Theta\left(2\sqrt{t}\Big|_1^n\right) = \Theta(2\sqrt{n}-2) = \Theta(\sqrt{n}).$$

Answer: $\Theta(\sqrt{n})$.

c) Assume that the base of the logarithm is e (note that since we are interested only in the asymptotic behavior, the base does not matter). Since the function $\frac{1}{k \log k}$ is nonincreasing

$$\sum_{i=2}^{n} \frac{1}{k \log k} = \Theta\left(\int_{2}^{n} \frac{1}{t \log t} dt\right)$$

Now,

$$\int_{2}^{n} \frac{1}{t \log t} dt = \int_{2}^{n} \frac{d(\log t)}{\log t} = \log \log t|_{2}^{n} = \Theta(\log \log n).$$

Answer: $\Theta(\log \log n)$.

d) Note that $2 \leq (1 + 1/k)^k < e$. Therefore,

$$\sum_{k=1}^{n} 2 \leq \sum_{k=1}^{n} (1 + 1/k)^k \leq \sum_{k=1}^{n} e.$$

And,

$$2n \leq \sum_{k=1}^{n} (1 + 1/k)^k \leq en.$$

We conclude that $\sum_{k=1}^{n} (1 + 1/k)^k = \Theta(n)$.

Answer: $\Theta(n)$.

Theorem. Let $f : [a, +\infty) \to \mathbb{R}$ be a nonincreasing integrable positive function (where a is a natural number). Then

$$\sum_{i=a}^{n} f(n) = \Theta\left(\int_{a}^{n} f(t) dt\right).$$

Proof: Let $F(n) = \int_{a}^{n} f(t)dt$. Since the function $f(t)$ is nonincreasing

$$\sum_{k=a}^{n-1} f(k) \geq \int_{a}^{n} f(t)dt = F(n) \geq \sum_{k=a+1}^{n} f(k).$$

Since $f(x)$ is a positive function, $F(n)$ is an increasing function. Therefore, for $n \geq a + 1$

$$\sum_{k=a}^{n} f(k) \leq f(a) + F(n) = \left(\frac{f(a)}{F(n)} + 1\right) F(n) \leq \left(\frac{f(a)}{F(a+1)} + 1\right) F(n) = O(F(n)).$$

On the other hand,

$$\sum_{k=a}^{n} f(k) \geq F(n) = \Omega(F(n)).$$

This concludes the proof.

Problem 2

The running time $T(n)$ of a newly developed algorithm satisfies the following recurrence relation:

$$T(n) = 4T(n/2) - 4T(n/4) + n \qquad T(1) = 1 \qquad T(2) = 4.$$

Find a closed form expression for the running time of this algorithm for instances of size $n = 2^k$, where k is an integer.

Solution:

Let $a_k = T(2^k)$ for non-negative integer k. Then

$$a_k = T(2^k) = 4T(2^k/2) - 4T(2^k/4) + 2^k = 4T(2^{k-1}) - 4T(2^{k-2}) + 2^k = 4a_{k-1} - 4a_{k-2} + 2^k,$$

$$a_0 = T(1) = 1 \qquad a_1 = T(2) = 4.$$

We got a nonhomogeneous recurrence relation for the sequence a_k. Represent a_k as a sum of homogeneous and particular solutions:

$$a_k = a_k^{(h)} + a_k^{(p)}.$$

First, find the homogeneous solution. We have

$$a_k^{(h)} = 4a_{k-1}^{(h)} - 4a_{k-2}^{(h)}.$$

The characteristic equation is $\lambda^2 - 4\lambda + 4 = 0$ has a double root $\lambda = 2$. Therefore, the general homogeneous solution is $a_k^{(h)} = \alpha 2^k + \beta k 2^k$. Now, the nonhomogeneous part of the recurrence is 2^k, since 2 is the root of the characteristic equation, there is a particular solution of the form $\gamma k^2 2^k$. We have

$$\gamma k^2 2^k - 4\gamma(k-1)^2 2^{k-1} + 4\gamma(k-2)^2 2^{k-2} = 2^k.$$

Simplifying the expression in the left hand side, we get

$$2 \cdot \gamma \cdot 2^k = 2^k \Leftrightarrow \gamma = \frac{1}{2}.$$

We have $a_n = \alpha 2^k + \beta k 2^k + \frac{1}{2} k^2 2^k$. We now find the coefficients α and β:

$$\begin{cases} a_0 = \alpha 2^0 + \beta 0 \cdot 2^0 + \frac{1}{2} \cdot 0^2 \cdot 2^0 = 1 \\ a_1 = \alpha 2^1 + \beta 1 \cdot 2^1 + \frac{1}{2} \cdot 1^2 \cdot 2^1 = 4 \end{cases} \Leftrightarrow \begin{cases} \alpha = 1 \\ \beta = \frac{1}{2} \end{cases}.$$

We conclude that $a_k = 2^k + \frac{1}{2} k 2^k + \frac{1}{2} k^2 2^k$. Finally, $T(n) = n + \frac{1}{2} n \log n + \frac{1}{2} n (\log n)^2$.

Answer: $T(n) = n + \frac{1}{2} n \log_2 n + \frac{1}{2} n (\log_2 n)^2$.

Problem 3

Pick a random natural number r between 0 and 1000000: $0 \le r < 1000000$. What is the probability that the sum of its digits is divisible by 10?

Solution:

Let $r = \overline{d_5 d_4 d_3 d_2 d_1 d_0} \stackrel{\text{def}}{=} 10^5 d_5 + 10^4 d_4 + 10^3 d_3 + 10^2 d_2 + 10 d_1 + d_0$. For fixed values of d_4, d_3, d_2, d_1, and d_0, there exists exactly one value of d_5 s.t. $d_5 + d_4 + d_3 + d_2 + d_1 + d_0$ is divisible by 10 (e.g. d_5 is equal to the remainder of division of $-(d_4 + d_3 + d_2 + d_1 + d_0)$ by 10). Therefore, each set of d_4, d_3, d_2, d_1, and d_0 determines uniquely one number with the sum of its digits divisible by 10. So there are 100 000 numbers with the sum of its digits divisible by 10. The desired probability is $\frac{100\,000}{1\,000\,000} = \frac{1}{10}$.

Answer: $\frac{1}{10}$.

Problem 4

Consider a sequence of n independent tosses of a fair coin. A *run* is defined to be a maximal sequence of contiguous tosses that are either all heads or all tails. e.g. the sequence HHTHHTTTHH has 5 runs of length 2,1,2,3 and 2 respectively.

a) Compute the probability that there are exactly k runs.

b) Compute the probability that there are exactly k runs and every run is of length at most 2.

Solution:

a) For each sequence of n tosses with k runs denote the length of the i^{th} run by a_i. Then every sequence of n tosses with k runs defines one sequence $a_1, a_2, \ldots, a_k$ s.t. $a_1 + a_2 + \cdots + a_k = n$ and $a_i \ge 1$ (for $1 \le i \le k$). On the other hand, each such sequence a_i corresponds to exactly two distinct sequences of tosses

$$\underbrace{T\ldots T}_{a_1}\underbrace{H\ldots H}_{a_2}\underbrace{T\ldots T}_{a_3}\cdots$$
$$\underbrace{H\ldots H}_{a_1}\underbrace{T\ldots T}_{a_2}\underbrace{H\ldots H}_{a_3}\cdots$$

Therefore, the number of sequences of n tosses with k runs is twice the number of sequences a_i. Now the number of such sequences a_i is $\binom{n-1}{k-1}$: We distribute n ones between k variables a_i with repetitions allowed; at first we set each a_i to 1, and then distribute the remaining $n-k$ ones between k variables. The number of ways to do this is $\binom{(n-k)+k-1}{k-1} = \binom{n-1}{k-1}$.

So the number of sequences of n tosses with k runs is $2\binom{n-1}{k-1}$. Finally, the total number of all sequences of n tosses is 2^n.We conclude that the desired probability is

$$\frac{2\binom{n-1}{k-1}}{2^n}.$$

Answer: $\frac{1}{2^{n-1}}\binom{n-1}{k-1}$.

b) First we compute the number of runs of length 1 and the number of runs of length 2. Denote the former number by r_1 and the latter number by r_2. Since the total number of runs is k, we have $r_1 + r_2 = k$. The total number of tosses is n, so $1 \cdot r_1 + 2 \cdot r_2 = n$. Solving these two equations, we get $r_1 = 2k - n$ and $r_2 = n - k$. (Notice, that these formulas imply that if $2k - n < 0$ then there are no such sequences of tosses.)

We get that there are totally k runs, and $2k - n$ of them are of length 1. So there are $\binom{k}{2k-n}$ ways to arrange the lengths of k runs. Finally, for each arrangement of run lengths there are two sequences of tosses (one starts with heads, and the other starts with tails). We conclude that the number of such sequences of tosses is $2\binom{k}{2k-n}$, and the desired probability is

$$\frac{2\binom{k}{2k-n}}{2^n} = \frac{\binom{k}{2k-n}}{2^{n-1}}.$$

Answer: $\frac{1}{2^{n-1}}\binom{k}{2k-n}$ (if $2k - n < 0$ the answer is 0).

Problem 5

Two coins are placed in a bag. One of them is a fair coin, i.e. it comes up heads with probability 1/2 and tails with probability 1/2. The other is a special coin with tails on both sides. One of the coins is picked from the bag at random and this coin is tossed n times.

Let A_i be the event that the coin comes up tails on the i^{th} toss.

a) Calculate $\Pr[A_i]$.

b) Calculate $\Pr[A_2|A_1]$.

c) Calculate $\Pr[A_k|A_1 \cap A_2 \cap \ldots \cap A_{k-1}]$.

Solution:

Denote the event that we picked the fair coin by F.

a)

$$\Pr[A_i] = \Pr[A_i|F]\Pr[F] + \Pr[A_i|\bar{F}]\Pr[\bar{F}] = \frac{1}{2}\cdot\frac{1}{2} + 1\cdot\frac{1}{2} = \frac{3}{4}.$$

Answer: $\frac{3}{4}$.

b)

$$\Pr[A_2|A_1] = \frac{\Pr[A_2 \cap A_1]}{\Pr[A_1]} = \frac{4}{3}\Pr[A_2 \cap A_1]$$
$$= \frac{4}{3}\left(\Pr[A_2 \cap A_1|F]\Pr[F] + \Pr[A_2 \cap A_1|\bar{F}]\Pr[\bar{F}]\right) = \frac{4}{3}\left(\frac{1}{4}\cdot\frac{1}{2} + 1\cdot\frac{1}{2}\right) = \frac{5}{6}.$$

Answer: $\frac{5}{6}$.

c)

$$\Pr[A_k | A_1 \cap A_2 \cap \ldots \cap A_{k-1}] = \frac{\Pr[A_1 \cap A_2 \cap \cdots \cap A_k]}{\Pr[A_1 \cap A_2 \cap \cdots \cap A_{k-1}]}.$$

Now,

$$\begin{aligned}\Pr[A_1 \cap A_2 \cap \cdots \cap A_k] &= \Pr[A_1 \cap A_2 \cap \cdots \cap A_k | F] \Pr[F] + \Pr[A_1 \cap A_2 \cap \cdots \cap A_k | \bar{F}] \Pr[\bar{F}] \\ &= \frac{1}{2^k} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{2^k + 1}{2^{k+1}}.\end{aligned}$$

Similarly,

$$\Pr[A_1 \cap A_2 \cap \cdots \cap A_{k-1}] = \frac{2^{k-1} + 1}{2^k}.$$

We have,

$$\Pr[A_k | A_1 \cap A_2 \cap \ldots \cap A_{k-1}] = \frac{2^k + 1}{2^{k+1}} \Bigg/ \frac{2^{k-1} + 1}{2^k} = \frac{2^k + 1}{2^k + 2}.$$

Answer: $\dfrac{2^k + 1}{2^k + 2}$.