CMPUT 204 (B1) – Assignment # 2

DUE DATE: July 28, 2000

Solutions

---

1. *Fibonacci Sequence*

   Let $\{F_n\}_{n \geq 0}$ be the **Fibonacci sequence**

$$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ \ldots$$

   that is, F_n is the unique solution to the discrete initial value problem

$$\begin{aligned} F_{n+2} &= F_{n+1} + F_n, \quad n \geq 0 \\ F_0 &= 0 \\ F_1 &= 1. \end{aligned}$$

   Does there exist, among the first $100,000,001$ terms of this sequence, a term whose decimal representation ends with four zeros?

   SOLUTION: In this problem we are only interested in the final four digits of the integers F_k in the Fibonacci sequence whenever these integers contain 5 or more digits, so we are dealing with a sequence of integers all of which are less than 10^4.

   Let $\{a_k\}_{k \geq 0}$ be the sequence of (four-digit) integers such that

$$a_k = F_k \mod 10^4$$

   for $k \geq 0$, that is, a_k is the four-digit integer whose digits are the final four digits of F_k.

   If we know the integers a_{k+1} and a_k, then it is easy to find a_{k-1}, since the four digits composing a_{k-1} will depend only on the corresponding four digits of a_{k+1} and a_k.

   Now, suppose that we can find two nonnegative integers k and n such that $a_k = a_{n+k}$ and $a_{k+1} = a_{n+k+1}$, then it follows that

$$\begin{aligned} a_{k-1} &= a_{n+k-1} \\ a_{k-2} &= a_{n+k-2} \\ &\vdots \\ a_1 &= a_{n+1} \\ a_0 &= a_n \end{aligned}$$

   since any for any $i \geq 1$ we have $F_{i-1} = F_{i+1} - F_i > 0$. However, since $a_0 = 0$, then it follows that $a_n = 0$ also, that is, a_n consists of four digits, all of which are 0. Thus, if we can find indices k and n, such that $a_k = a_{n+k}$ and $a_{k+1} = a_{n+k+1}$, then the n^{th} Fibonacci number F_n will end in four zeros.

   Therefore, we only have to show that among the $10^8 + 1$ pairs,

$$\begin{array}{ll} a_0 & a_1 \\ a_1 & a_2 \\ a_2 & a_3 \\ \vdots & \vdots \\ a_{10^8-1} & a_{10^8} \\ a_{10^8} & a_{10^8+1} \end{array}$$

   there are at least two identical pairs. However, since none of the integers $a_0, a_1, a_2, \ldots, a_{10^8+1}$ is greater than 10^4, and since from the 10^4 integers $0, 1, 2, 3, 4, \ldots, 9999$ we can only obtain $10^4 \cdot 10^4 = 10^8$ distinct pairs (the first number in the pair can have 10^4 distinct values, and for each of these the second number in the pair can have 10^4 distinct values), from the pigeon hole principle, all the pairs cannot be distinct.

   Note: It can be shown that the smallest Fibonacci number that ends in 4 zeros is F_{7500}.

2. *Sum of Reciprocals of Primes*

Consider the sum

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{5}+\frac{1}{7}+\frac{1}{11}+\frac{1}{13}+\frac{1}{17}+\frac{1}{19}+\frac{1}{23}+\cdots+\frac{1}{p}$$

in which the denominators run through the prime numbers from 2 to some prime number p.

(a) Show that this sum can be made greater than any preassigned real number $M > 0$, no matter how large, provided the prime p is sufficiently large, that is, show that the infinite series

$$\sum_{p \text{ prime}} \frac{1}{p}$$

diverges.

(b) Use the result from part (a) to show that there exist infinitely many primes.

Solution I:

(a) Let $\{p_1, p_2, p_3, \ldots, p_m, \ldots\}$, be the sequence of primes, where p_m is the m^{th} prime (the sequence could be finite or it could be infinite), and suppose that the series

$$\sum_{m\geq 1} \frac{1}{p_m}$$

converges, then there exists a positive integer k such that

$$\sum_{m\geq k+1} \frac{1}{p_m} < \frac{1}{2}.$$

Let $Q = p_1 \cdot p_2 \cdots p_k$, and consider the numbers $1 + nQ$ for $n = 1, 2, 3, \ldots$. None of these is divisible by any of the primes $p_1, p_2, \ldots, p_k$. Therefore, all of the prime factors of $1 + nQ$ occur among the primes $p_{k+1}, p_{k+2}, \ldots$. Thus, for each integer $r \geq 1$, we have

$$\sum_{n=1}^{r} \frac{1}{1+nQ} \leq \sum_{t=1}^{\infty} \left(\sum_{m\geq k+1} \frac{1}{p_m} \right)^t,$$

since the sum on the right includes among its terms all the terms on the left.

However, the sum on the right is dominated by a convergent geometric series, so we have

$$\sum_{n=1}^{r} \frac{1}{1+nQ} \leq \sum_{t=1}^{\infty} \left(\sum_{m\geq k+1} \frac{1}{p_m} \right)^t \leq \sum_{t=1}^{\infty} \left(\frac{1}{2}\right)^t = 1$$

for all $r \geq 1$.

Now note that for each $n \geq 1$ we have $1 + nQ < 2nQ$, so that

$$\sum_{n=1}^{r} \frac{1}{1+nQ} > \frac{1}{2Q} \sum_{n=1}^{r} \frac{1}{n} = \frac{1}{2Q} H_r$$

which can be made arbitrarily large, since the harmonic series diverges. This condradiction shows that our original assumption must have been incorrect, therefore the series $\displaystyle\sum_{p \text{ prime}} \frac{1}{p}$ diverges.

(b) If there were only finitely many primes, then the series above would converge, which is a contradiction.

Solution II:

(a) First we prove the following lemmas:

Lemma 1. If x is a real number with $x \geq 2$, then $\displaystyle\prod_{\substack{p \leq x \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right) < \frac{1}{\log x}$.

proof. For each prime $p \leq x$, we have $p > 1$, and the geometric series

$$\frac{1}{1 - 1/p} = 1 + \frac{1}{p} + \frac{1}{p^2} + \cdots$$

converges, and therefore

$$\prod_{\substack{p \leq x \\ p \text{ prime}}} \frac{1}{1 - 1/p} = \prod_{\substack{p \leq x \\ p \text{ prime}}} \left(1 + \frac{1}{p} + \frac{1}{p^2} + \cdots\right),$$

and by the unique factorization theorem, when the product on the right is multiplied out it gives the sums of the reciprocals of all integers having only primes not exceeding x as prime divisors. In particular, all positive integers less than or equal to x are of this form, so that

$$\prod_{\substack{p \leq x \\ p \text{ prime}}} \frac{1}{1 - 1/p} > \sum_{k=1}^{\lfloor x \rfloor} \frac{1}{k} > \int_1^{\lfloor x \rfloor + 1} \frac{1}{u}\, du > \log x.$$

□

Lemma 2. For any real number $t > -1, \quad t \neq 0$, we have $\displaystyle\frac{t}{1+t} < \log(1+t) < t$.

proof. Let $t > -1$, with $t \neq 0$, if we define the function $f(u) = \log(1+u)$ for $u > -1$, then the mean value theorem implies the existence of a real number T between 0 and t such that

$$f(t) - f(0) = \log(1+t) - \log(1) = t \cdot \frac{d}{du} \log(1+u)\bigg|_{u=T},$$

that is, $\log(1+t) = \dfrac{t}{1+T}$ for some T between 0 and t.

Now, since the function $g(u) = \dfrac{1}{1+u}$ is strictly decreasing on the interval $-1 < u < \infty$, if $t < T < 0$, then $\dfrac{1}{1+T} < \dfrac{1}{1+t}$, and since $t < 0$, then $\dfrac{t}{1+T} > \dfrac{t}{1+t}$. Similarly, if $0 < T < t$, then $\dfrac{1}{1+t} < \dfrac{1}{1+T}$, and since $t > 0$, then $\dfrac{t}{1+t} < \dfrac{t}{1+T}$. Thus, $\log(1+t) > \dfrac{t}{1+t}$ for all $t > -1$, with $t \neq 0$.

Also, since the function $f(u) = \log(1+u)$ is concave down on the interval $-1 < u < \infty$, the entire graph lies below the tangent line to the curve at $u = 0$, that is, $\log(1+t) < t$ for all $t > -1$, with $t \neq 0$. □

Theorem. The series $\displaystyle\sum_{p \text{ prime}} \frac{1}{p}$ diverges. In fact, $\displaystyle\sum_{\substack{p \leq x \\ p \text{ prime}}} \frac{1}{p} > \frac{1}{2} \log\log x$.

proof. From Lemma 1 we have

$$\log \prod_{\substack{p \leq x \\ p \text{ prime}}} \left(1 - \frac{1}{p}\right) = \sum_{\substack{p \leq x \\ p \text{ prime}}} \log\left(1 - \frac{1}{p}\right) < -\log\log x.$$

Now, since

$$\frac{t}{1+t} \geq 2t$$

for $0 > t \geq -\frac{1}{2}$, and since $p \geq 2$, then from the left-hand side of the inequality in Lemma 2 we have

$$-\frac{2}{p} < \log\left(1 - \frac{1}{p}\right)$$

for all primes p, and therefore

$$\sum_{\substack{p \leq x \\ p \text{ prime}}} \frac{2}{p} > -\sum_{\substack{p \leq x \\ p \text{ prime}}} \log\left(1 - \frac{1}{p}\right) > \log\log x.$$

$\square$

3. *Rooks on an $n \times n$ Chessboard.*

This problem involves configurations of chess pieces on a chessboard. We will consider not only the usual chessboard of 8 rows and 8 columns, but also an $n \times n$ chessboard, having n rows and n columns.

A **rook** controls all squares in its row and column, up to and including the first square occupied by another piece.

(a) What is the greatest number of rooks that can be placed on an $n \times n$ chessboard in such a way that none of them controls the square on which another lies? In how many ways can this be done?

(b) What is the smallest number of rooks which can be arranged on an $n \times n$ chessboard in such a way that every square of the board is controlled by at least one of them? In how many different ways can this be done?

SOLUTION:

(a) An $n \times n$ chessboard has n rows and n columns. For none of the rooks on the board to control the square on which another lies, it is necessary and sufficient that no two rooks lie in the same row or in the same column. Therefore, the total number of rooks cannot exceed n. On the other hand, it is possible for n rooks to be arranged on the board so that none of them controls the square on which another one lies; for example, put them on the n squares of one of the principal diagonals of the chessboard.

The rook in the first column can lie in any one of the n rows. For any choice of the row in which it lies, there are $n-1$ choices for the row in which the rook in the second column can lie. For each of the $n(n-1)$ choices of the rows for the rooks in the first and second columns, there are $n-2$ choices for the row in which the rook in the third column can lie. Similarly, for each of the $n(n-1)(n-2)$ choices of the rows for the rooks in the first, second and third columns, there are $n-3$ choices for the row in which the rook in the fourth column can lie. Continuing in this manner, we see that the number of different arrangements of all n rooks so that none controls a square occupied by another rook is $n!$.

(b) It is impossible for fewer than n rooks to control all n^2 squares of an $n \times n$ chesboard. In fact, if there were less than n rooks on the chessboard, there would be a row on which there was no rook and a column on which there was no rook. The square common to this row and column would then not be controlled by any of the rooks. On the other hand, it is obviously possible to arrange n rooks on the board in such a way that they control every square of the board, for example, place them on the squares on one of the principal diagonals.

If n rooks on an $n \times n$ chessboard control every square of the board, then there is either one rook in each column or one rook in each row. Since otherwise there would be a row and a column neither of which contained any rooks, and the square common to this row and column would not be controlled by any of the rooks. Conversely, if there is either one rook in every row or one rook in every column, then these rooks will control the entire board.

The number of ways of arranging n rooks on an $n \times n$ chessboard, one in each column, is n^n (for each rook, we have n choices for the row in which to place it). Similarly, the number of ways of arranging n rooks on an $n \times n$ chessboard, one in each row, is n^n. Therefore, the number of ways of arranging n rooks so they control every square on the chessboard is the number of ways of arranging the rooks so that there is either one rook in each row or one rook in each column, so that the total number of such arrangements is $2n^n - n!$ (each of the arrangements where there is exactly one rook in each row and one rook in each column has been counted twice, so we have to subtract it).

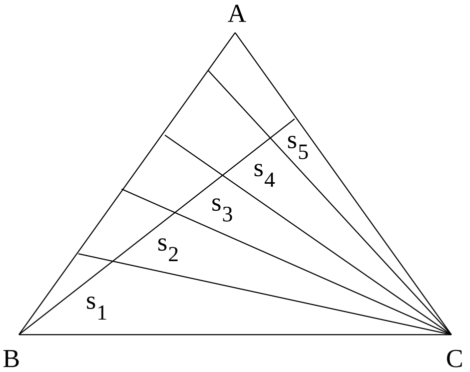
4. *Number of Regions in a Triangle.*

(a) Each of the vertices of the base of a triangle is connected by straight lines to n points on the side opposite it. Into how many parts do these $2n$ lines divide the interior of the triangle?

(b) Each of the three vertices of a triangle is joined by straight lines to n points on the opposite side of the triangle. Into how many parts do these $3n$ lines divide the interior of the triangle if no three of them pass through the same point?

SOLUTION:

(a) Given the traingle ABC, we first draw the lines joining C to the n points on side AB; these lines divide ABC into $n+1$ smaller triangles. Now draw a line S from B to one of the points on AC as in the figure.



This line consists of $n+1$ segments, $s_1, s_2, \ldots, s_{n+1}$, and each of these segments cuts one of the small triangles into two pieces. Therefore, drawing the line S has increased the number of regions into which triangle ABC was previously divided by $n+1$. The same argument applies to any of the n lines from B to AC; as each one is drawn, it increases the total number of regions by $n+1$. Thus, at the end of the process triangle ABC has been divided into $n+1+n(n+1) = (n+1)^2$ regions.

(b) The lines emanating from the vertices B and C divide the triangle into $(n+1)^2$ parts. Each of the lines emanating from the vertex A intersects all $2n$ of the lines which emanate from the vertices B and C (the points of intersection are all different, since no three of the lines are concurrent).

Therefore, each of the lines emanating from the vertex A is divided into $2n+1$ pieces by the other lines, and so increases the number of parts by $2n+1$. It follows that the triangle is divided into

$$(n+1)^2 + n(2n+1) = 3n^2 + 3n + 1$$

parts.

5. *Insertion Sort with a Twist.*

Consider the following variation of Insertion Sort:

For $2 \leq i \leq n$, to insert the key $L[i]$ among

$$L[1] \leq L[2] \leq \cdots \leq L[i-1],$$

do a binary search to find the correct position for $L[i]$.

(a) How many key comparisons would be done in the worst case?

(b) What is the total number of times keys are moved in the worst case?

(c) Using comparison of list entries as the basic operation, what is the order of the worst case running time?

(d) Can the number of moves be reduced by putting the keys in a linked list instead of an array? Explain.

SOLUTION:

(a) For $i = 2, 3, \ldots, n$, in the worst case we have to do

$$\lceil \log_2 i \rceil$$

key comparisons to find the place to insert $L[i]$. Thus, the total number of comparisons that must be made in the worst case is

$$W(n) = \sum_{i=2}^{n} \lceil \log_2 i \rceil.$$

(b) For each $i = 2, 3, \ldots, n$, in the worst case, after we find the position in which to insert $L[i]$, we have to move $i - 1$ keys, therefore, the total number of times keys are moved in the worst case is

$$\sum_{i=1}^{n} (i-1) = \frac{n(n-1)}{2}.$$

(c) Using comparisons of list entries as our basic operation, the worst case running time is

$$W(n) = \sum_{i=2}^{n} \lceil \log_2 i \rceil,$$

and since

$$\log_2 i \leq \lceil \log_2 i \rceil < \log_2 i + 1,$$

then

$$\sum_{i=2}^{n} \log_2 i \leq \sum_{i=2}^{n} \lceil \log_2 i \rceil < \sum_{i=2}^{n} \log_2 i + n - 1.$$

Therefore, $W(n) \in \Theta(\sum_{i=2}^{n} \log_2 i) = \Theta(\log_2 n!)$, but we saw in class that $\log_2 n! \in \Theta(n \log_2 n)$, so that $W(n) \in \Theta(n \log_2 n)$.

(d) The number of moves can be reduced to 0 by putting the keys in a linked list, since in that case we only need to reset the links instead of moving the keys. However, not much is accomplished, since implementing binary search with a linked list forces the list to be searched sequentially each time we want to find the middle key in a sublist.

6. *A Nonlinear Recurrence*

Let x_0 and x_1 be arbitrary positive numbers, and let

$$x_{n+1} = \frac{1}{2}\big(\sqrt{x_n} + \sqrt{x_{n-1}}\big), \qquad n = 1, 2, \ldots \tag{$*$}$$

show that $\gamma = \lim_{n\to\infty} x_n$ exists and does not depend on the choice of the starting values x_0, x_1.

SOLUTION: Note that if $\gamma = \lim_{n\to\infty} x_n$ exists, then $\gamma = \frac{1}{2}\big(\sqrt{\gamma} + \sqrt{\gamma}\big) = \sqrt{\gamma}$, and either $\gamma = 0$ or $\gamma = 1$. In order to show that $\gamma = 1$, let $x_0 > 0$ and $x_1 > 0$, and choose $0 < a < 1$, such that $x_0 > a$ and $x_1 > a$, then

$$\begin{aligned}
x_2 &= \tfrac{1}{2}\big(\sqrt{x_0} + \sqrt{x_1}\big) \ge \tfrac{1}{2}\big(\sqrt{a} + \sqrt{a}\big) = \sqrt{a} > a \\
x_3 &= \tfrac{1}{2}\big(\sqrt{x_1} + \sqrt{x_2}\big) \ge \tfrac{1}{2}\big(\sqrt{a} + \sqrt{a}\big) = \sqrt{a} > a \\
&\ \vdots \\
x_n &= \tfrac{1}{2}\big(\sqrt{x_{n-1}} + \sqrt{x_{n-2}}\big) \ge \tfrac{1}{2}\big(\sqrt{a} + \sqrt{a}\big) = \sqrt{a} > a \\
&\ \vdots
\end{aligned}$$

and an easy induction argument shows that $x_n > a$ for $n = 0, 1, 2, \ldots$.

From $(*)$ we have

$$x_{n+1} = \tfrac{1}{2}\big(\sqrt{x_n} + \sqrt{x_{n-1}}\big)$$

for all $n \ge 1$, and therefore

$$\begin{aligned}
x_{n+1} - 1 &= \tfrac{1}{2}\big(\sqrt{x_n} - 1\big) + \tfrac{1}{2}\big(\sqrt{x_{n-1}} - 1\big) \\
&= \tfrac{1}{2}\frac{x_n - 1}{1 + \sqrt{x_n}} + \tfrac{1}{2}\frac{x_{n-1} - 1}{1 + \sqrt{x_{n-1}}},
\end{aligned}$$

so that

$$|x_{n+1} - 1| \le \tfrac{1}{2}\frac{|x_n - 1|}{1 + \sqrt{x_n}} + \tfrac{1}{2}\frac{|x_{n-1} - 1|}{1 + \sqrt{x_{n-1}}},$$

for all $n \ge 1$. Since $\sqrt{x_n} > \sqrt{a} > a$, for all $n \ge 0$, then

$$|x_{n+1} - 1| \le \tfrac{1}{2}\frac{|x_n - 1|}{1 + \sqrt{a}} + \tfrac{1}{2}\frac{|x_{n-1} - 1|}{1 + \sqrt{a}}, \tag{$**$}$$

for all $n \ge 1$.

Now, $\{u_n\}_{n\ge 0}$ with $u_n = |x_n - 1|$, satisfies the inequalities

$$u_{n+1} \le \frac{1}{2(1+a)}u_n + \frac{1}{2(1+a)}u_{n-1} \tag{$***$}$$

where $u_0 > 0$ and $u_1 > 0$. Suppose that $\{b_n\}_{n\ge 0}$ satisfies the difference equation

$$\begin{aligned}
b_{n+1} &= \frac{1}{2(1+a)}b_n + \frac{1}{2(1+a)}b_{n-1} \\
b_0 &= u_0, \quad b_1 = u_1,
\end{aligned} \tag{$****$}$$

then

$$u_2 \leq \frac{1}{2(1+a)}u_1 + \frac{1}{2(1+a)}u_0 = \frac{1}{2(1+a)}b_1 + \frac{1}{2(1+a)}b_0 = b_2$$

$$u_3 \leq \frac{1}{2(1+a)}u_2 + \frac{1}{2(1+a)}u_1 \leq \frac{1}{2(1+a)}b_2 + \frac{1}{2(1+a)}b_1 = b_3$$

$$\vdots$$

$$u_n \leq \frac{1}{2(1+a)}u_{n-1} + \frac{1}{2(1+a)}u_{n-2} \leq \frac{1}{2(1+a)}b_{n-1} + \frac{1}{2(1+a)}b_{n-2} = b_n$$

$$\vdots$$

and an easy induction argument shows that

$$u_n \leq b_n \qquad (*****)$$

for all $n \geq 0$.

If we let $\epsilon = \dfrac{1}{4(1+a)}$, then the difference equation $(****)$ has characteristic equation $\lambda^2 = 2\epsilon\lambda + 2\epsilon$, with roots

$$\lambda_1 = \epsilon + \sqrt{\epsilon(2+\epsilon)}$$

$$\lambda_2 = \epsilon - \sqrt{\epsilon(2+\epsilon)}$$

both of which are less than 1 in absolute value since $\epsilon < \frac{1}{4}$. The solution to $(****)$ is

$$b_n = A \cdot \lambda_1^n + B \cdot \lambda_2^n$$

where

$$A = \frac{u_1 - \lambda_2 u_0}{\lambda_1 - \lambda_2},$$

and

$$B = \frac{\lambda_1 u_0 - u_1}{\lambda_1 - \lambda_2}.$$

From $(**)$ and $(*****)$ we have

$$|x_n - 1| = u_n \leq b_n = A \cdot \lambda_1^n + B \cdot \lambda_2^n$$

for all $n \geq 0$, and since $|\lambda_1| < 1$ and $|\lambda_2| < 1$, then $\lim\limits_{n\to\infty} x_n = 1$.