Midterm Exam – 3 hrs

PLEASE write **CLEARLY**, show **ALL** of your work and **EXPLAIN** everything in reasonable detail

---

The Master Theorem:

Let a, b be constants such that $a \geq 1$ and $b > 1$, let $f(n)$ be a function and let $T(n)$ be defined by the recurrence $T(n) = a\, T(\lfloor n/b \rfloor) + f(n)$. Then $T(n)$ can be estimated asymptotically as follows :

1) If $f(n) = O(n^{\log_b a - \varepsilon})$ for some $\varepsilon > 0$, then $T(n) = \Theta(n^{\log_b a})$

2) If $f(n) = \Theta(n^{\log_b a})$ then $T(n) = \Theta(n^{\log_b a} \log n)$

3) If $f(n) = \Omega(n^{\log_b a + \varepsilon})$ for some $\varepsilon > 0$, and if $a\, f(n/b) < c\, f(n)$ for some constant $c < 1$ all $n \geq n_0$ for some sufficiently large n_0, then $T(n) = \Theta(f(n))$.

---

1. In each case determine the exact asymptotic growth rate of the solution $T(n)$ to the recurrence.

    a. **[4pt]** $T(n) = 16\, T(n/4) + n\sqrt{n}$
    b. **[7pt]** $T(n) = T(3n/8) + (n+1)/n$
    c. **[7pt]** $T(n) = 3T(n/2) + n^{1+10/n}$
    d. **[12pt]** $T(n) = \sqrt{n}\, T(\sqrt{n}) + n\, \log_2 \sqrt{n}$

2. With the large amount of genomic DNA sequence data being made available, it is becoming more important to find genes (parts of the genomic DNA which are responsible for the synthesis of proteins) in these sequences. Each gene is composed of several pieces, called ***exons***, of regions coding proteins, plus "junk DNA" which separates the exons in the genomic sequence. It is known that the order of the exons is maintained in the protein synthesis process, but the number of exons and their lengths can be arbitrary. Prior to protein synthesis disjoint sequence of exons is chosen so that there are no complete exons in the remainder of the sequence and such remainder is junk which is removed. However, exons can overlap, in which case only one of the overlapping exons is chosen. Thus, a sequence might look as follows (please turn the page):

ATGCAATGGCAATTGGCACATTGCATTACCGATCCTTGAGAAGGC

with exons indicated by intervals, and with one choice of exons indicated by a thicker line.

   a. **[25pt]** Design an algorithm which given a sequence with exons indicated as above, finds a gene comprising of a largest number of disjoint exons;
   b. **[35pt]** Design an algorithm which given a sequence with exons indicated as above, finds a gene comprising of the largest number of bases (letters).

3. You are given a map of cities with roads connecting some of them, as well as the length of each road. You have to choose cities in which you will build schools, such that the distance from any city to the nearest city with a school is at most 10 miles, and so that the total number of schools is as small as possible.
   a. [**30pt**] Design a greedy algorithm which attempts to choose minimal number of cities to build schools in.
   b. **[30pt]** Give an example of a map which on which, unfortunately, such greedy algorithm might fail to produce the minimal number of schools to be build.

4. **[30pt]** Design an algorithm which multiplies any polynomial of degree 16 with any polynomial of degree 8 using only 25 multiplications in which both operands can be arbitrarily large.

5. [**20pt**] Design an algorithm which, given a set of positive integers X, determines if the equation $x^5 + x\,y - y^2 = y^3$ has a solution with both x and y belonging to X.

EXTRA PROBLEM FOR THE EXTENDED CLASSES 3821 & 9801 ONLY

6. **[30pt]** Design an algorithm which uses one fair die to generate 5 equally likely outcomes, and analyze the expected run time of your algorithm.

GOOD LUCK!!!