Assignment 6 **CAS705 Fall 2005** **Due on November 30th**

The work you submit must be your own. *You may discuss problems with each other; however, you should prepare written solutions alone. In particular, you should not leave with any written notes from such discussions. The style and clarity of your answers will be an important factor in the grade.*

This assignment has two questions, each worth 50%.

1. In this problem you will design a randomized algorithm for pattern matching. Consider the set of strings over $\{0,1\}$, and let $M : \{0,1\} \longrightarrow M_{2\times 2}(\mathbb{Z})$, that is, M is a map from strings to 2×2 matrices over the integers ($\mathbb{Z}$) defined as follows:

$$M(\varepsilon) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
$$M(0) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$
$$M(1) = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

and for strings $x, y \in \{0,1\}^*$, $M(xy) = M(x)M(y)$, where the operation on the LHS is concatenation of strings, and the operation on the RHS is multiplication of matrices. Show the following:

   (a) $M(x)$ is well-defined.

   (b) $M(x) = M(y)$ implies that $x = y$ (i.e., the map M is 1-1)

   (c) For $x \in \{0,1\}^n$, the entries of $M(x)$ are bounded by Fibonacci number F_n.

   (d) By considering the matrices $M(x)$ modulo a suitable prime p, show how would you peform efficient randomized pattern matching (i.e., how would you tell whether string x is a sub-string of string y).

2. Let $G = (V, E)$ be a connected undirected regular graph of degree d, meaning that for each $u \in V$, $|\{(u,v) :$ for some $v \in V$ s.t. $(u,v) \in E\ \}| = d$.

   Assume that for any given u we have ordered the edges incident upon u in some way. A string $U = l_1 l_2 \ldots l_m \in \{1, 2, \ldots, d\}^*$ and a node u induce a traversal of the graph as follows: we start at u, and we take edge l_1 out of node u and arrive at node u_1. Then we take edge l_2 out of node u_1 and arrive at node u_2, etc. (We may visit a node more than once).

   We say that U traverses G if it visits every node of G. Use a probabilistic argument to show that for each n there is a traversal sequence of length $O(d \cdot n^4)$ that works for all graphs with n nodes and degree d.

   What does this tell you about the complexity of undirected reachability?