cpts515 Midterm Exam

Let C be a finite set of colors and G be a directed graph where v_0 is a designated start node. Each edge of G is labeled with a color in C and multiple edges can share the same color. From v_0, one may have infinitely many walks α since there is no upper bound on the length of walks and there could be cycles in G. For each such walk α (that must start with v_0), we may collect the sequence $c(\alpha)$ of colors on the edges on the walk. We use P to denote the set of all walks α starting from v_0 and use $\mathcal{C}$ to denote all the resulting color sequences:

$$\mathcal{C} = \{c(\alpha) : \alpha \in P\}.$$

Clearly, $\mathcal{C}$ may be an infinite set (of color sequences).

1. (40pts) Let u be a node in G. From this u, one may have multiple outgoing edges, say $\langle u, v_1\rangle$, $\langle u, v_2\rangle$, $\cdots$, $\langle u, v_k\rangle$, for some $k \geq 2$, whose colors are all the same. One can understand the same color as a triggering event that leads from current node u to a "next" node chosen nodeterministically from $v_1, \cdots, v_k$. Clearly, if I show you a walk α in P, then there could be many edges on the walk that are the result of many nondeterministic choices. Please develop and justify a metric M_1 which is a function of G that measures the nondeterminism on all walks in P. (Your M_1 is high when nondeterminism is high.) You need also show me an algorithm in computing such M_1. In case when efficient algorithm is hard to obtain, please also provide an approximation algorithm.

2. (40pts) Multiple walks can share the same color sequence. Hence, if we measure nondeterminism from the angle of color sequences in $\mathcal{C}$, the metric would be very different. Color sequences in $\mathcal{C}$ may also have "nondeterminism" which can be understood in the following way. Consider two color sequences in $\mathcal{C}$, say

$$red, yellow, green, green, red, ...$$

and

$$red, yellow, blue, green, yellow....$$

The first two colors in the sequences are the same: *red* followed by *yellow*. However, the third colors are different (*green* in the first sequence and *blue* in the second). One may say that there is a nondeterministic choice of the next color right after the first two colors. Please develop and justify a metric M_2 which is a function of G that measures the nondeterminism on all color sequences in $\mathcal{C}$. You need also show me an algorithm in computing such M_2. In case when efficient algorithm is hard to obtain, please also provide an approximation algorithm.

3. (40pts) In good old days, a program is understood as a functional unit so that testing refers to figuring out its input/output relation. However, nowadays, a program (such as embedded systems code) is often reactive: it constantly interact with its environment so that when you test it, it will not just give you one pair of input/output; instead, it will give you a sequence of pairs of input/output:

$$(i_1, o_1), \cdots, (i_n, o_n)$$

for some n. Notice that the input sequence $i_1, \cdots, i_n$ is called a test case (provided by a test engineer) and the output sequence $o_1, \cdots, o_n$ is the test result. Now, we assume that each of i_j and o_j is a color in C. The program can be nondeterministic; i.e., one test case may have multiple test results. We now assume that the program is a blackbox (whose code is not available; not even assembly nor machine code nor design nor requirements. Do not use any code analysis techniques, design analysis techniques, requirements analysis techniques here since they are not applicable and you will get zero.). Sadly, researchers know little about testing a nondeterministic blackbox program! Good news is one can use the most stupid approach in testing such a program: try many many test cases and then collect many many test results. Our experiences are very intuitive: a more highly nondeterministic program is also harder to test. Please develop an algorithm to estimate, from those many many tests, how high the nondeterminism of the blackbox program under test is.