CS/ECE 374 A (Spring 2020)
Homework 1 (due Jan 30 Thursday at 10am)

Instructions: Carefully read `http://engr.course.illinois.edu/cs374/sp2020/A/hw_policies.html` and `http://engr.course.illinois.edu/cs374/sp2020/A/integrity.html`.

- **Groups of up to three people may submit joint solutions.** Each problem should be submitted by exactly one person, and the beginning of the homework should clearly state the Gradescope names and email addresses of each group member. In addition, whoever submits the homework must tell Gradescope who their other group members are.
- **Submit your solutions electronically on the course Gradescope site as PDF files.** Submit a separate PDF file for each numbered problem. If you plan to typeset your solutions, please use the LaTeX solution template on the course web site. If you must submit scanned handwritten solutions, please use a black pen on blank white paper and a high-quality scanner app (or an actual scanner, not just a phone camera).
- If you are not using your real name and your `illinois.edu` email address on Gradescope, you will need to fill in the form linked in the course web page.
- **You may use any source at your disposal**—paper, electronic, or human—but you *must* cite *every* source that you use, and you *must* write everything yourself in your own words.
- For any homework problem or subproblem (except for bonus questions), you can write **IDK** ("I Don't Know") and get 25% of the points for that problem or subproblem. You would get the points if and only if IDK is the only content of your answer. Any answer containing "IDK" anywhere and **any** additional text would immediately get a zero.
- **Avoid the Three Deadly Sins.** Any homework or exam solution that breaks any of the following rules may be given an *automatic zero.*
  - Always give complete solutions, not just examples.
  - Always declare all your variables, in English. In particular, always describe the specific problem your algorithm is supposed to solve.
  - Always avoid unnecessarily long answers.

Problem 1.1: Let $L \subseteq \{0,1\}^*$ be the language defined recursively as follows:

- The empty string ε is in L.
- For any string x in L, the string $1x$ is also in L.
- For any strings x, y, z in L, the string $0x0y0z$ is also in L.
- The only strings in L are those that can be obtained by the above rules.

(a) Prove that $L \subseteq \{x \in \{0,1\}^* : \#_0(x) \text{ is divisible by } 3\}$, by using induction (for example, on the length of the string). Here, $\#_0(x)$ denotes the number of 0's in the string x. (As always, use *strong* induction.)

(b) Prove that $\{x \in \{0,1\}^* : \#_0(x) \text{ is divisible by } 3\} \subseteq L$, by using induction.

Problem 1.2: A string $z \in \{0,1\}^*$ is said to be *balanced* if it has exactly the same number of 0's and 1's. Let L be the language consisting of all strings $x \in \{0,1\}^*$ that contain at least one balanced substring with length at least 6. [Note: the definition would be far less interesting if 6 is changed to 2 :-).]

For example, $11011011101 \notin L$ but $11\mathbf{101100}110 \in L$ (because of the bold substring).

Given a long string x of length n, deciding whether $x \in L$ may at first seem to require a lot of effort, as there are $\Theta(n^2)$ substrings to check. But it turns out we only need to check $\Theta(n)$ substrings. . .

More precisely, for $x \in \{0,1\}^*$, prove the following fact: $x \in L$ if and only if x contains a balanced substring of length between 6 and 11.

[Hint: in the more difficult direction, assuming that x does not contain a balanced substring of length between 6 and 11, prove that every substring z with length at least 6 is unbalanced, by induction on $|z|$.]

[More Hint: it may be helpful to first prove the following observation: if x is a balanced string that starts with a 0 and ends with a 0, then we can decompose x as uv for shorter strings u and v where both u and v are balanced. To prove this observation, let $\#_0(i)$ and $\#_1(i)$ denote the number of 0's and 1's in the first i symbols of x, and apply the "intermediate value theorem" to the function $\#_0(i) - \#_1(i)$ (over $i = 1, \ldots, |x| - 1$). . .]

Problem 1.3: Consider the following recurrence:

$$T(n) \;=\; \begin{cases} 8\,T(\lfloor n/3 \rfloor) + 9\,T(\lfloor n/9 \rfloor) + 2020 & \text{if } n \geq 9 \\ 374 & \text{if } 1 \leq n \leq 8 \end{cases}$$

Prove that $T(n) = O(n^2)$ by induction.

[Hint: you may need to use an induction hypothesis of the form $T(n) \leq cn^2 - c'$ for some constants c and c' to be determined.]