Homework 1: Strings and languages

CSE 30151 Spring 2018

Due: 2018/01/25 10:00pm

Instructions

- Create a PDF file (or files) containing your solutions. You can write your solutions by hand, but please scan them in the library or using a smartphone to get them into PDF.
- Please name your PDF file(s) as follows to ensure that the graders give you credit for all of your work:
  - If you're making a complete submission, name it `netid-hw1.pdf`, where *netid* is replaced with your NetID.
  - If you're submitting some problems now and want to submit other problems later, name it `netid-hw1-123.pdf`, where `123` is replaced with the problem numbers you are submitting at this time.
- Submit your PDF file(s) in Sakai. Don't forget to click the Submit button!

Problems

Each problem is worth 10 points.

1. **Strings and languages.** How would you represent each of the following sets as a formal language? (There are many possible right answers.) For each set, write what the alphabet would be. Describe informally how to encode an element as a string. Give an example of a string belonging to the language and a string not belonging to the language. You don't need to explain how you would actually decide whether a string belongs to the language.

   (a) The set of syntactically correct C programs.

   (b) The set of all solvable Sudoku boards.

   (c) The set of all ways to walk from North Dining Hall to DeBartolo Hall.

2. **String homomorphisms.** If Σ and Γ are finite alphabets, a *string homomorphism* is a function $\phi : \Sigma^* \to \Gamma^*$ that has the property that for any $u, v \in \Sigma^*$, $\phi(uv) = \phi(u)\phi(v)$.

   Intuitively, a string homomorphism does a "search and replace" where each symbol is replaced with a (possibly empty) string. For example, the function $\phi : \{\texttt{0}, \ldots, \texttt{9}, \texttt{A}, \ldots, \texttt{F}\}^* \to \{\texttt{0}, \texttt{1}\}^*$ that converts hexadecimal numbers (including ε) to binary is a string homomorphism because each hex digit is replaced with four bits ($\texttt{0} \mapsto \texttt{0000}, \texttt{1} \mapsto \texttt{0001}, \texttt{2} \mapsto \texttt{0010}, \ldots, \texttt{F} \mapsto \texttt{1111}$).

   Prove the above intuition more formally. That is, prove that if ϕ is a string homomorphism, then for any $w = w_1 \cdots w_n$ (where $n \geq 0$ and $w_i \in \Sigma$ for $1 \leq i \leq n$), we have $\phi(w) = \phi(w_1) \cdots \phi(w_n)$. Use induction, as follows:

   (a) Basis: Show that $\phi(\varepsilon) = \varepsilon$.

   (b) Induction step: Assuming that, for any $w = w_1 \cdots w_k$, we have $\phi(w) = \phi(w_1) \cdots \phi(w_k)$, show that, for any $w = w_1 \cdots w_{k+1}$, we have $\phi(w) = \phi(w_1) \cdots \phi(w_{k+1})$.

3. **Language classes.** Recall that a language class is a set of languages. In this course, we'll study several language classes, and as a warm-up to this concept, we'll think about the class of *finite languages*.

   Assume that Σ is a finite, nonempty alphabet. Let FINITE be the class of all finite languages over Σ, and let

$$\mathsf{coFINITE} = \{L \mid \overline{L} \in \mathsf{FINITE}\},$$

   where, for any language L over Σ, $\overline{L}$ is the complement of L, that is, $\Sigma^* \setminus L$.

   (a) Are there any languages in $\mathsf{FINITE} \cap \mathsf{coFINITE}$? Give an example language or briefly prove that there is none.

   (b) Are there any languages *not* in $\mathsf{FINITE} \cup \mathsf{coFINITE}$? Give an example language or briefly prove that there is none.