# Extra Credit

Note: All problems on this homework set are extra-credit. You may turn in solutions for up to four of the problems below. Turning in a solution for any part of a problem counts as attempting the entire problem.

I originally had not planned to assign homework during the week of the midterm. The idea of having an extra credit problem set was very popular in class, and I'm happy to do so. If I do this for both midterms, it creates the possibility that some people may get a total homework score greater than 100% – overall, this class has been doing very well on the homework so far. If at the end of the term, someone has a total homework score of $(100 + p)$% (with $p > 0$), I will count it as $100 + p/2$ when computing course grades. This means that you still benefit from the extra credit, but hopefully I'll avoid creating grade inflation in the process.

**Have fun!**

1. (**10 points**) Let $A$ be the language of all strings whose length is a power of 2:

$$A \quad = \quad \{w \mid \exists k \in \mathbb{N}.\ |w| = 2^k\}.$$

   Prove that $A$ is not regular.

2. (**15 points**) Each of the three languages described below is regular. For one language, show a DFA; for another language, show an NFA; and for the remaining language, show a regular expression.

   (a) (**5 points**) $A_1 = \{w \in \{\mathtt{a}, \mathtt{b}\}^* \mid \#a(w) - \#b(w) \text{ is divisible by three}\}$, where $\#a(w)$ denotes the number of $\mathtt{a}$'s in $w$ and likewise for $\#b(w)$.

   (b) (**5 points**) $A_2 = \{w \in \{\mathtt{a}, \mathtt{b}\}^* \mid \#ab(w) = \#ba(w)\}$ where $\#ab(w)$ denotes the number of occurrence of the string $\mathtt{ab}$ in $w$, and likewise for $\#ba(w)$. For example $\#ab(\mathtt{ababba}) = 2$ and $\#ba(\mathtt{ababba}) = 2$; therefore, $\mathtt{ababba} \in A_2$ .

   (c) (**5 points**) $A_3 = \{w \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^* \mid 2^{|\#a(w)|} - 1 \text{ is divisible by three or } 2^{|\#b(w)|} + 1 \text{ is divisible by five}\}$.

3. (**20 points**) Let $\alpha$, $\beta$ and $\gamma$ be regular expressions.

   (a) (**10 points**) Prove that $(\alpha \cup \beta)\gamma \ = \ \alpha\gamma \cup \beta\gamma$.

   (b) (**10 points**) Prove that $(\alpha^*\beta)^*\alpha^* \ = \ (\alpha \cup \beta)^*$.

4. (**20 points**) Let $A$ be a language. Define

$$\mathit{sameLength}(A) \quad = \quad \{w \mid \exists x \in A.\ |x| = |w|\}$$

   Show that if $A$ is regular, then $\mathit{sameLength}(A)$ is regular as well.

5. (**25 points**, problem 1.38 from Sipser)
   An *all*-NFA is a 5-tuple $N_\forall = (Q, \Sigma, \delta, q_0, F)$, that accepts $w \in \Sigma^*$ iff *every* possible state that $N_\forall$ could be in after reading input $w$ is in $F$. Prove that all-NFAs recognize the class of regular languages.

6. (**25 points**, problem 1.42 from Sipser)
   Let $A$ and $B$ be languages. Define

$$\mathit{shuffle}(A, B) \quad = \quad \{w \mid \exists x_1 \cdot x_2 \cdots x_k \in A.\ \exists y_1 \cdot y_2 \cdots y_k \in B.\ w = x_1 \cdot y_1 \cdot x_2 \cdot y_2 \cdots x_k \cdot y_k\}$$

   Note that this says that the *concatenation* of $x_1$ through $x_k$ produces a string in $A$; the indivual $x_i$ strings might or might not be strings in $A$. Likewise for the $y_i$'s.

   Show that the class of regular languages is closed under shuffle.

7. (**30 points**) Let $A$ be a regular language. Let

$$\begin{aligned} A_{-\frac{1}{3}-} &= \{w \mid \exists x, y, z.\ (|x| = |y| = |z|) \wedge (w = y) \wedge (xyz \in A)\} \\ A_{\frac{1}{3}-\frac{1}{3}} &= \{w \mid \exists x, y, z.\ (|x| = |y| = |z|) \wedge (w = xz) \wedge (xyz \in A)\} \end{aligned}$$

One of these languages is regular for any regular language $A$, and one is non-regular for some choices of $A$ (obviously, both $A_{-\frac{1}{3}-}$ and $A_{\frac{1}{3}-\frac{1}{3}}$ are regular if $A = \emptyset$ or $A = \Sigma^*$). Determine which is which and give proofs.

8. (**30 points**) Let $A$ be a regular language. Let

$$\begin{aligned} A_{exp_2} &= \{w \mid \exists x \in A.\ |x| = 2^{|w|}\} \\ A_{\log_2} &= \{w \mid \exists x \in A.\ |w| = 2^{|x|}\} \end{aligned}$$

One of these languages is regular for any regular language $A$, and one is non-regular for some choices of $A$ Determine which is which and give proofs.

9. (**30 points**) Let $A$ be a regular language with alphabet $\{\mathsf{a}, \mathsf{b}\}$. Let

$$\begin{aligned} A_{=a\wedge=b} &= \{w \mid \exists v \in A.\ (\#a(v) = \#a(w)) \wedge (\#b(v) = \#b(w))\} \\ A_{=a\vee=b} &= \{w \mid \exists v \in A.\ (\#a(v) = \#a(w)) \vee (\#b(v) = \#b(w))\} \end{aligned}$$

Where $\#a(x)$ denotes the number of $\mathsf{a}$'s in $x$, and $\#b(x)$ denotes the number of $\mathsf{b}$'s in $x$. One of these languages is regular for any regular language $A$, and one is non-regular for some choices of $A$. Determine which is which and give proofs.

10. (**40 points**)

    (a) (**10 points**) Prove the following more general version of the pumping lemma:

    If $A$ is regular, then there is a constant $p > 0$ such that for any string $xyz \in A$ with $|y| \geq p$, then there exist strings $u$, $v$, and $w$ such that:

    - $|v| > 1$; and
    - $\forall i \geq 0.\ xuv^iwz \in A$.

    (b) (**10 points**, see Sipser exercise 1.54)
    Consider the language $A = \{a^i b^j c^k \mid k \geq 0 \text{ and if } i = 1 \text{ then } j = k\}$. In class (see the Sept. 29 slides), we showed that $A$ satisfies the simple version of the pumping lemma that we've been using, and we showed that $A$ is not regular.

    Show that $A$ does not satisfy the new improved version of the pumping lemma that you proved in part (a).

    (c) (**10 points**, from Kozen)
    Consider the language with alphabet $\Sigma = \{\mathsf{a}, \mathsf{b}, \mathsf{c}\}$:

    $$B = (\mathsf{a}^+\mathsf{c})^n(\mathsf{b}^+c)^n \cup \Sigma^* cc \Sigma^*$$

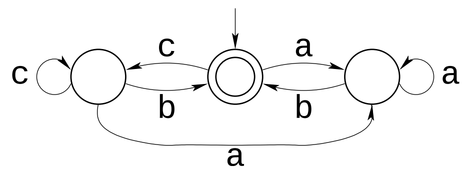
    Show that $B$ satisfies the conditions of the new improved pumping lemma from part (a).

    (d) (**10 points**) Prove that $B$ as defined in part (c) above is not regular.

11. (**45 points**) Let $\Sigma^\omega$ denote the set of all strings of *infinite* length composed of symbols from $\Sigma$. A *Non-deterministic Büchi Automaton* (NBA) is a five-tuple, $(Q, \Sigma, \delta, q_0, F)$ with $Q$, $\Sigma$, $\delta$, and $F$ defined as for an NFA. Let $N$ be a NBA and $w \in \Sigma^\omega$ be a string. Let $w(i)$ denote the $i^{th}$ symbol of $w$. A *run* of $N$ on $w$ is a sequence of states, $q : \mathbb{N} \rightarrow Q$ such that:

$$\forall i.\ q(i+1) \in \delta(q(i), w(i)),$$

where $q(i)$ denotes the $i^{th}$ state in the sequence $q$. We say that a run is accepting iff there are an infinite number of distinct choices for $i$ such that $q(i) \in F$. We say that $w$ is in the language of $N$, $w \in L(N)$ iff there is an accepting run for $N$ with input $w$. For example, the NBA shown below




accepts all strings that have an infinite number of b's and never have an a followed immediately by a c.

Let $\mathcal{B}$ denote the class of all languages that are recognized by NBAs. The class $\mathcal{B}$ is known as the $\omega$-regular languages.

(a) (**10 points**) Prove that $\mathcal{B}$ is closed under union. In other words, if $A_1$ and $A_2$ are languages in $\mathcal{B}$, then $A_1 \cup A_2$ is in $\mathcal{B}$ as well.

(b) (**15 points**) Prove that $\mathcal{B}$ is closed under intersection.

(c) (**20 points**) Deterministic Büchi Automata are defined in a manner analagous to ther non-deterministic cousins. In particular, a *Deterministic Büchi Automaton* (DBA) is a five-tuple, $(Q, \Sigma, \delta, q_0, F)$ with $Q$, $\Sigma$, $\delta$, and $F$ defined as for DFAs. Let $D$ be a DBA and $w \in \Sigma^\omega$ be a string. A *run* of $D$ on $w$ is a sequence of states, $q : \mathbb{N} \rightarrow Q$ such that:

$$\forall i.\ q(i+1) = \delta(q(i), w(i)),$$

We say that a run is accepting iff there are an infinite number of distinct choices for $i$ such that $q(i) \in F$. Let $\mathcal{D}$ denote the class of all languages that are recongnized by DBAs.

Prove that $\mathcal{D}$ is a proper subset of $\mathcal{B}$. In other words, you must show that there is at least one language recognized by an NBA that is not recognized by any DBA.