CSE596 **Prelim I Answer Key** **Fall 2018**

Open book, open notes, closed neighbors, 48 minutes. The exam totals 80 pts., subdivided as shown. Do *all three problems* in the exam book provided—there is no "choice" option. *Show your work*—this may help for partial credit.

(1) (5×4 = 20 pts. total) *True/False*.

Please write out the words *true* and *false* in full. Brief justifications are not necessary but may help for partial credit. Given sets A and B, the difference $A \setminus B$ is the same as $A \cap \tilde{B}$.

(a) If A and B are regular, then $A \setminus B$ is always regular.

(b) If A and B are decidable, then $A \setminus B$ is always decidable.

(c) If A and B are computably enumerable, then $A \setminus B$ is always computably enumerable.

(d) If A is regular, then A^* is decidable in linear time by a single-tape Turing machine.

(e) Every non-regular language is decidable.

Answers: (a) and (b) are true because both the regular languages and the decidable languages are closed under complements and intersection, which imply closure under all Boolean operations. But (c) is false: if you take $A = \Sigma^*$ and $B = K_{TM}$ then $A \setminus B$ is essentially D_{TM} which is not c.e. More simply put, closure under $\setminus$ entails closure under complements because $\Sigma^* \setminus B$ is the same as $\tilde{B}$, but RE is not closed under complements. Question (d), the trickiest, is true because A^* is also regular and hence has a DFA, which "Is-A" Turing machine with one tape that runs in time $n + 1$. Item (e) is false because if you take any undecidable language, such as D_{TM}, it must be non-regular because all regular languages are decidable.

(2) (18 + 12 = 30 pts.)

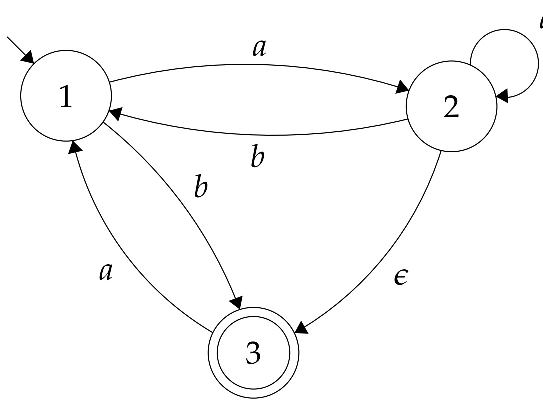
Consider the following nondeterministic finite automaton $N = (Q, \Sigma, \delta, s, F)$ where $Q = \{1, 2, 3\}$, $\Sigma = \{a, b\}$, $s = 1$, $F = \{3\}$, and $\delta = \{(1, a, 2), (1, b, 3), (2, a, 2), (2, b, 1), (2, \epsilon, 3), (3, a, 1)\}$. Convert N into a DFA M such that $L(M) = L(N)$ (18 pts.). *Answer:* Here is the equivalent DFA alongside the originally-given diagram of N. The ϵ-arc means "whenever 2, then also 3," and this helps build up the "delta-underbar" function as:

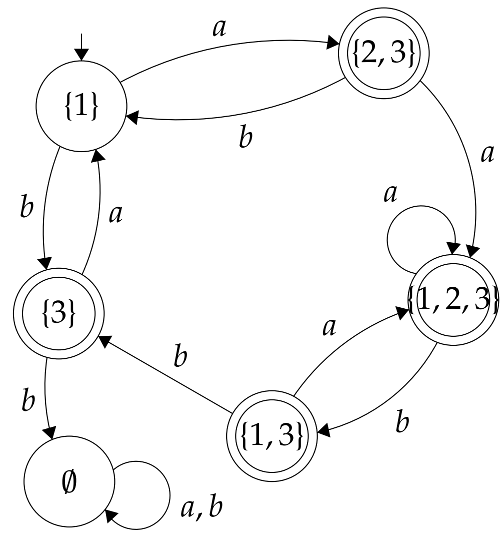
$$\left|\begin{array}{rcr} \bar{\delta}(1,a) & = & \{2,3\} \\ \bar{\delta}(2,a) & = & \{2,3\} \\ \bar{\delta}(3,a) & = & \{1\} \end{array}\right\|\begin{array}{rcr} \bar{\delta}(1,b) & = & \{3\} \\ \bar{\delta}(2,b) & = & \{1\} \\ \bar{\delta}(3,b) & = & \emptyset \end{array}\right|$$

The start state is just $\{1\}$, then $\Delta(\{1\}, a) = \bar{\delta}(1, a) = \{2, 3\}$ and $\Delta(\{1\}, b) = \bar{\delta}(1, b) = \{3\}$ to get things rolling. We also get

$$\Delta(\{2,3\}, a) = \bar{\delta}(2, a) \cup \bar{\delta}(3, a) = \{2, 3\} \cup \{1\} = \{1, 2, 3\},$$

which intuitively "turns all three lights on."





Also answer the following questions (3 pts. each).

(a) Is there a string u such that for each of its states q, N can process u from 1 to q? Give a shortest such string if so.

(b) Is there a string v that N cannot process starting from state 1 at all? Again give a shortest such string if so.

(c) Is there a string w such that for all $y \in \Sigma^*$, $wy \in L(N)$? Again give a shortest w if so.

(d) Does $L(N)$ include $b(aaabb)^*$? Briefly justify from your M.

Answers to (a)–(d): The fact that M goes into the "omni" state $\{1,2,3\}$ on the input $u = aa$ means that N can process aa from 1 to any of its states. Here aa is the string processed on a (the) shortest path to $\{1,2,3\}$. The string $v = bb$ on the other hanbd goes from $\{1\}$ to the dead state so it cannot be processed from start in N. The answer to (c) is *no* because if there were any such string it would be $w = u = aa$, but the followup string $y = bbb$ kills it. The answer to (d) is *yes* because the initial b takes the converted DFA M to the accepting state $\{3\}$, whereupon $aaabb$ goes a grand cycle back to state $\{3\}$. (It was also possible to answer (d) from N by noting the computation path $(\{1\}, b, \{3\})$ then $(\{3\}, a, \{1\}, a, \{2,3\}, a, \{1,2,3\}, b, \{1,3\}, b, \{3\})$, but especially in a "bushier" NFA, one could get lost trying other possibilities, whereas M shows you just the one path. To be sure, in problems with more states in N, the M can be enormous.)

(3) (8 + 4 + 18 = 30 pts.)

Define L to be the language of strings x such that $|x|$ is even and the second half of x contains at least one '1.' For instance 010100 is in L but 01010000 is not, and 0100001 is not because its length is odd.

(a) Which of the following strings belong to L? Say yes/no for each.

$$(i)\ \epsilon \qquad (ii)\ 1 \qquad (iii)\ 01 \qquad (iv)\ 010.$$

(b) Is $L \cdot L \subseteq L$? Justify your answer briefly.

(c) Prove via the Myhill-Nerode Theorem that L is nonregular.

Answers: In part (a), (i) is no because there is no 1, (ii) is no because $|1|$ is odd, (iii) is yes, but (iv) is no again because of odd length.

Part (b) is *yes* because any '1' in the second half of y will become a '1' in the second half of xy for any even-length x—and all $x \in L$ have even length.

The pitfall in (c) to avoid is to take $S = 0^*$, and upon being given $x = 0^m$ and $y = 0^n$ with $m < n$ (wlog.) to take $z = 1 \cdot 0^{m+1}$. Then $xy \notin L$ and it looks like $yz \in L$ because z is longer, but the problem is that $|z|$ could be odd given the choice of S. It is possible to fix this by dividing into cases accoprding to whether m and n are odd or even and choosing z accordingly.

It is cleaner, however, to control the parities of m and n from the get-go by choosing $S = (00)^*$ instead (or $S = (00)^+$ to stay away from the ϵ "edge case" though this doesn't matter here). Clearly S is infinite. Let any $x, y \in S$, $x \neq y$, be given. Then x and y are equal to 0^m and 0^n with $m < n$ (wlog.) and m, n both *even* and ≥ 2. So take $z = 10^{m+1}$. We have $xz = 0^m 1 0^{m+1} \notin L$ because the '1' is at the end of the first half, but $yz = 0^n 1 0^{m+1} \in L$ because $|yz| = n + 1 + m + 2$ is even and guarantees the '1' is in the second half by $n > m$. Thus $L(xz) \neq L(yz)$, and since $x, y \in S$ are arbitrary, S is an infinite PD set for L, which makes L non-regular by the Myhill-Nerode Theorem.