Final #1

Mark all correct answers for each of the following questions.

Σ denotes an arbitrary alphabet and L an arbitrary language over Σ, unless otherwise specified.

1. Given a language L over the alphabet $\{a, b\}$, denote:

$$L' = \{w \in L : wa \in L\}.$$

   (a) If $L_1, L_2 \subseteq \{a, b\}^*$, then $(L_1 \cup L_2)' = L_1' \cup L_2'$.

   (b) If $L_1, L_2 \subseteq \{a, b\}^*$, then $(L_1 L_2)' = L_1 L_2'$. However, we do not necessarily have $(L_1 L_2)' = L_1' L_2'$.

   (c) If $L_1, L_2 \subseteq \{a, b\}^*$, and both languages L_1 and L_2 are non-empty, then $(L_1 L_2)' \neq L_1' L_2'$.

   (d) If $L \subseteq \{a, b\}^*$, then $(L^*)' = L^* L'$.

   (e) If $M = (Q, \{a, b\}, \delta, s, A)$ is a deterministic automaton accepting L, then there exists a set of states $A' \subseteq Q$ such that the deterministic automaton $M' = (Q, \{a, b\}, \delta, s, A')$ accepts L'.

   (f) There exist languages L for which there exists a regular expression α such that $L(\alpha) = L$ but no regular expression β such that $L(\beta) = L'$.

   (g) None of the above.

2. Given a subset D of the set $\mathbf{Z}_+$ of non-negative integers, the set of their expansions in base 10 forms a partial language of $\{0, 1, \ldots, 9\}^*$. Denote this language by $L(D)$. Also, for $D_1, D_2 \subseteq \mathbf{Z}_+$ denote $D_1 D_2 = \{d_1 d_2 : d_1 \in D_1,\, d_2 \in D_2\} \subseteq \mathbf{Z}_+$.

(a) If $D = \left\{10^{n^2} : n \geq 0\right\}$, then $L(D)$ is context-free but not regular.

(b) Let $(a_n)_{n=0}^{\infty}$ be any sequence such that $a_0 = 0$ and such that for each n we have either $a_{n+1} = a_n + 1$ or $a_{n+1} = a_n + 2$, and let $D = \{10^{a_n} : n \geq 0\}$. If for all sufficiently large n we have $a_{n+1} = a_n + 1$, then $L(D)$ is regular. Similarly, if for all sufficiently large n we have $a_{n+1} = a_n + 2$, then $L(D)$ is regular. However, if there exist both infinitely many n's for which $a_{n+1} = a_n + 1$ and infinitely many n's for which $a_{n+1} = a_n + 2$, then $L(D)$ cannot be regular, and not even context-free.

(c) If $D = \{37 \cdot (1 + 10 + 10^2 + \ldots + 10^n) : n \geq 0\}$, then $L(D)$ is not regular.

(d) If $D_1, D_2 \subseteq \mathbf{Z}_+$ are such that both $L(D_1)$ and $L(D_2)$ are regular, then (by the closure properties of the family of regular languages) the set $L(D_1 \cup D_2)$ is regular as well. An analogous statement holds if "regular" is replaced by "context-free".

(e) If $D \subseteq \mathbf{Z}_+$ is such that $L(D)$ is regular, then $L(\mathbf{Z}_+ - D) = \{0, 1, \ldots, 9\}^* - L(D)$, so that $L(\mathbf{Z}_+ - D)$ is regular as well. However, since the family of context-free languages is not closed under complementation, the analogous statement is false if "regular" is replaced by "context-free".

(f) If $D_1, D_2 \subseteq \mathbf{Z}_+$ are such that both $L(D_1)$ and $L(D_2)$ are regular, then $L(D_1 D_2)$ is regular as well. An analogous statement holds if "regular" is replaced by "context-free".

(g) None of the above.

3. Let $L_1 \subseteq \{a, b\}^*$ be the language consisting of all words w over $\{a, b\}$ possessing the properties that $|w|_a = 2|w|_b$ and that $|u|_a \geq 2|u|_b$ for every prefix u of w. Let $L_2 \subseteq \{a, b, c\}^*$ be the language consisting of all words w over $\{a, b, c\}$ possessing the properties that $|w|_a = |w|_b = |w|_c$ and that $|u|_a \geq |u|_b \geq |u|_c$ for every prefix u of w. Let $L_3 \subseteq \{a, b\}^*$ be the language consisting of all words w over $\{a, b\}$ possessing the properties that $|w|_a = |w|_b$ and that $|u|_a \geq |u|_b$ for every prefix u of w.

(a) If the grammar G_1 is given by the rules
$S \to \varepsilon \mid a^2Sb \mid SS$,
then $L(G_1) = L_1$.

(b) There exist pushdown automata accepting L_1. Any such pushdown automaton must have at least two states.

(c) If the grammar G_2 is given by the rules
$S \to \varepsilon \mid aSbSc \mid SS$,
then $L(G_2) = L_2$.

(d) There exists a pushdown automaton M with a single state such that $L(M) = L_2$.

(e) The language L_2 is not context-free.

(f) If the grammar G_3 is given by the rules
$S \to \varepsilon \mid abSba \mid SS$,
then $L(G_3) = \{w\bar{w}^R : w \in L_3\}$, where $\bar{w}$ is the word obtained from w upon replacing each occurrence of a by b and vice versa.

(g) None of the above.

4. (a) There exists a constant C such that for every context-free language L there exists a pushdown automaton M with at most C states such that $L(M) = L$.

(b) The family of all context-free languages, accepted by context-free grammars $G = (N, \Sigma, R, S)$ with $|R| = 1$, is finite.

(c) For every constant C there exists a constant K such that, if $M = (Q, \Sigma, \Gamma, \Delta, s, A)$ is a simple pushdown automaton with $|Q| \le C$ and $|\Gamma| \le C$, then there exists a context-free grammar $G = (N, \Sigma, R, S)$ with $|N| \le K$ for which $L(G) = L(M)$.

(d) For every constant C there exists a constant K such that, if $G = (N, \Sigma, R, S)$ is a context-free grammar with $|N| \le C$, then there exists a pushdown automaton $M = (Q, \Sigma, \Gamma, \Delta, s, A)$ with $|Q| \le K$ and $|\Gamma| \le K$ for which $L(M) = L(G)$.

(e) A context-free grammar $G = (N, \{a, b\}, R, S)$ satisfies $L(G) \subseteq \{w \in \{a, b\}^* : |w|_a = |w|_b\}$ if and only if for every $(A, v) \in R$ (where $v \in (N \cup \{a, b\})^*$) we have $|v|_a = |v|_b$.

(f) If $G = (N, \Sigma, R, S)$ is a context-free grammar with $R \subseteq \{S\} \times (\Sigma^* S \cup S\Sigma^* \cup \Sigma^*)$, then G is not necessarily a regular grammar, yet $L(G)$ is a regular language.

(g) None of the above.

5. (a) Let M be a Turing machine with input alphabet Σ containing the letter a. If $sa \vdash_M^* aha$, then $sa \vdash_{M^2}^* a^2 ha$.

(b) If there exist at most finitely input words for which M does not halt, then M halts for every input.

(c) If M_1 and M_2 are Turing machines with the same input alphabet Σ and tape alphabet Γ, and both of them decide the same language L, then the machine $M_1 M_2$ decides L as well.

(d) If M_1 and M_2 are Turing machines with the same input alphabet Σ and tape alphabet Γ, and at least one of them decides L, then $M_1 M_2$ decides L as well.

(e) Let M be a Turing machine deciding L. It may be the case that one of the two machines MT_R or MT_L is also a machine deciding L, but it is impossible for both of these machines to decide L.

(f) Let M be a Turing machine deciding L. It may be the case that the machine MT_R^2 decides L also, but the machine MT_L^2 cannot possible decide L.

(g) None of the above.

Solutions

1. If $L_1 = \{a^2\}^*$ and $L_2 = \{a\}L_1$, then $L_1' = L_2' = \emptyset$, but $(L_1 \cup L_2)' = (\{a\}^*)' = \{a\}^* \neq L_1' \cup L_2'$.

If $L_1 = \{a\}^*$ and $L_2 = \{\varepsilon\}$, then $(L_1 L_2)' = \{a\}^*$ while $L_1 L_2' = \emptyset$.

If $L_1 = L_2 = \{a\}^*$, then $(L_1 L_2)' = \{a\}^* = L_1' L_2'$.

If $L = \{a\}$, then $(L^*)' = \{a\}^*$, but $L^* L' = \emptyset$.

If $M = (Q, \{a,b\}, \delta, s, A)$ accepts L, then $M' = (Q, \{a,b\}, \delta, s, A')$, where $A' = A \cap \{q \in Q : \delta(q,a) \in A\}$, accepts L'. Hence, if L is regular, then so is L'.

Thus, only (e) is true.

2. The length of the base 10 expansion of the number 10^{n^2} is n^2+1. Hence the set of all lengths of words in $L(D)$ for $D = \left\{10^{n^2} : n \geq 0\right\}$ is $\{n^2+1 : n \geq 0\}$, which is of unbounded gaps. Thus $L(D)$ is not context-free.

   Let $a_0 = 0$, $a_{n+1} = a_n + 1$ for even n and $a_{n+1} = a_n + 2$ for odd n. Taking $D = \{10^{a_n} : n \geq 0\}$ we have $L(D) = L\left(1\,(0^3)^*\,(\phi^* \cup 0)\right)$, which is a regular language.

   For $D = \{37 \cdot (1 + 10 + 10^2 + \ldots + 10^n) : n \geq 0\}$ we have

$$D = \{37, 407, 4107, 41107, 411107, \ldots\},$$

   so that $L(D) = \{37\} \cup \{4\}\{1\}^*\{07\}$, which is regular.

   Obviously, $L(D_1 \cup D_2) = L(D_1) \cup L(D_2)$, so that if $L(D_1)$ and $L(D_2)$ are regular (or context-free), then so is $L(D_1 \cup D_2)$.

   Since an expansion of a non-negative integer does not start with a 0 (except for the expansion of 0 itself), the set of all expansions is $\{0\} \cup \{1,2,\ldots,9\}\{0,1,\ldots,9\}^*$, and therefore

$$L(\mathbf{Z}_+ - D) = \{0\} \cup \{1,2,\ldots,9\}\{0,1,\ldots,9\}^* - L(D).$$

   If $D_1 = D_2 = \{10^n + 1 : n \geq 1\}$, then $L(D_1) = L(D_2) = \{1\}\{0\}^*\{1\}$ is regular. However, $D_1D_2 = \{10^{m+n} + 10^n + 10^m + 1 : 1 \leq m < n\} \cup \{10^{2n} + 2 \cdot 10^n + 1\}$. Taking any word of the form $10^n20^n1 \in L(D_1D_2)$, it is clear that pumping any non-trivial part of the word yields a word outside $L(D_1D_2)$, so that the latter language is not regular.

   Thus, only (d) is true.

3. Clearly, $a^3bab \in L_1$. However, by induction we show that, if $S \xRightarrow[G_1]{*} w$, then any occurrence of a in w is either preceded or followed by that of another a. Hence $a^3bab \notin L(G_1)$.

   The automaton $M = (\{s\}, \{a, b\}, \{a\}, \Delta, s, \{s\})$, where

   $$\Delta = \{((s, a, \varepsilon), (s, a)),\ ((s, b, a^2), (s, \varepsilon)),$$

   accepts L_1.

   If L_2 is context-free, then so is its intersection with the regular language $L(a^*b^*c^*)$. However,

   $$L_2 \cap L(a^*b^*c^*) = \{a^n b^n c^n : n \geq 0\},$$

   which is known to be non-context-free.

   Since $S \xRightarrow[G_3]{} abSba \xRightarrow[G_3]{} abba$, we have $abba \in L(G_3)$, yet $abba$ is not of the form $w\bar{w}^R$.

   Thus, only (e) is true.

4. The algorithm for constructing a pushdown automaton accepting the same language as a given context-free grammar $G = (N, \Sigma, R, S)$ always provides an automaton with two states only. Moreover, as the stack alphabet Γ of this automaton is $N \cup \Sigma$, an upper bound on $|N|$ yields an upper bound on $|\Gamma|$ (since Σ is fixed). On the other hand, since the size of the alphabet of non-terminals of the context-free grammar constructed to accept the same language as a given simple pushdown automaton $M = (Q, \Sigma, \Gamma, \Delta, s, A)$ is $1 + |Q|^2(|\Gamma| + 1)$, an upper bound on $|Q|$ and $|\Gamma|$ yields an upper bound on the number of non-terminals.

   Every language consisting of a single word is accepted by a context-free grammar with a single rule. Hence the family of context-free languages, accepted by context-free grammars with a single rule, is infinite.

   The grammar given by the rules

   - $S \to Ab$,
   - $A \to a$,

Accepts the language $\{ab\}$. The only word in the language has an equal number of a's and of b's although the right hand side of no grammatical rule enjoys this property.

If $R \subseteq \{S\} \times (\Sigma^* S \cup S\Sigma^* \cup \Sigma^*)$, then the grammar rules are $S \to u_1 S \mid u_2 S \mid \ldots \mid u_k S \mid S v_1 \mid S v_2 \mid \ldots \mid S v_l \mid w_1 \mid w_2 \mid \ldots \mid w_m$, for certain words $u_1, u_2, \ldots, u_k, v_1, v_2, \ldots, v_l, w_1, w_2, \ldots, w_m \in \Sigma^*$. By induction one shows that $S \overset{*}{\underset{G}{\Longrightarrow}} w \notin \Sigma^*$ if and only if

$$w \in L\left((u_1 \cup u_2 \cup \ldots \cup u_k)^* S (v_1 \cup v_2 \cup \ldots \cup v_l)^*\right),$$

and therefore $S \overset{*}{\underset{G}{\Longrightarrow}} w \in \Sigma^*$ if and only if

$$w \in L\left((u_1 \cup u_2 \cup \ldots \cup u_k)^* ((w_1 \cup w_2 \cup \ldots \cup w_m) (v_1 \cup v_2 \cup \ldots \cup v_l)^*\right).$$

Hence $L(G)$ is regular.

Thus, (a), (c), (d) and (f) are true.

5. If $sa \vdash_M^* aha$, then, roughly speaking, when the machine M^2 has aa written on its tape by the end of the first stage, with the head at the second square, it will continue to operate just as M does on input a, except the additional a at the leftmost square will stay there for the rest of the computation. More precisely, induction shows that, if the machine M changes the configuration sa to uqv after k steps, then it will change the configuration asa to $auqv$ after k steps. In particular, $sa \vdash_{M^2}^* a^2 ha$.

   For any finite set of words it is possible to construct a machine accepting exactly these words. In fact, the machine is designed first to check if the input is any of these words, in which case the machine halts. If all these tests fail, the machine is made to enter an infinite loop.

   The fact that both M_1 and M_2 decide the same language L gives no information regarding the behavior of M_2 on the configurations $s_2 Ⓨ$ (or $Ⓨ s_2$) and $s_2 Ⓝ$ (or $Ⓝ s_2$), so there is no reason for $M_1 M_2$ to decide L (or even just to halt for every input).

   According to our definitions, if a Turing machine M decides L, then for every input it halts with the head at the leftmost or second leftmost square. The machine MT_R will decide L if and only if M always halts with the head at the leftmost square. The machine MT_L will decide L if and only if M always halts with the head at the second leftmost square. However, the machine MT_R^2 will always halt with the head at the third or fourth square, and the machine MT_L^2 will always hang due to attempting to move the head leaftwards when already at the leftmost square.

Thus, only (a) and (e) are true.