CSL 705: Theory of Computation

II semester 2011-12

Sat 24 Mar 2012 14:30-15:30 II Minor WS-213 Max Marks 40

1. *Please answer in the space provided on the question paper. The other sheets are only for rough work and will not be collected.*
2. *You may use any paper-based material including your class notes and any other text books.*
3. *You are not allowed to share reference material or rough pages during the exam.*
4. *You are not allowed to bring into the exam hall any electronic gadgets such as computers, mobile phones or calculators.*
5. *Please keep your identity card with you. You may be asked for it at any time for verification.*

---

1. **(8 marks)** Prove that for every NPDA $N = \langle Q, \Sigma, \Gamma, \Delta, q_0, \perp, F\rangle$ such that $\Sigma \cap \Gamma \neq \emptyset$ there exists an equivalent NPDA $N' = \langle Q, \Sigma, \Gamma', \Delta', q_0, \perp, F\rangle$ such that $\Sigma \cap \Gamma' = \emptyset$

   Solution. Let $\Gamma' = \Gamma \cup \{A_i \mid A_i \notin \Gamma \cup \Sigma, a_i \in \Sigma \cap \Gamma\}$. Let $h : \Gamma \rightarrow \Gamma'$ be a 1-1 correspondence such that $h(a_i) = A_i$ whenever $a_i \in \Gamma' - \Gamma$ and $h(B) = B$ for all $B \in \Gamma$. h may be extended point-wise to strings in Γ i.e. $h(\varepsilon) = \varepsilon$ and for each $\beta = b.\beta' \in \Gamma^+$, $h(\beta) = h(b).h(\beta')$. Define Δ' such that $((q, a, B), (q', \beta)) \in \Delta$ iff $((q, a, h(B)), (q', h(\beta))) \in \Delta'$. Further it follows that (q, x, α) is a configuration of N iff $(q, x, h(\alpha))$ is a configuration of N' and $(p, ay, A\beta) \longrightarrow_N (q, y, \alpha.\beta)$ iff $(p, ay, h(A\beta)) \longrightarrow_{N'} (q, y, h(\alpha).h(\beta))$. Hence $\mathcal{L}(N) = \mathcal{L}(N')$ by any of the acceptance criteria and $\Sigma \cap \Gamma' = \emptyset$.

2. **(16 marks)** Let $\Sigma = \{a, b, c\}$. For each of the following languages determine whether it is context-free and if so whether it is a deterministic context-free language and give a proof of your answer as justification.

   (a) $L_1 = \{a^m b^{m+p} c^p \mid m, p \geq 0\}$

   (b) $L_2 = \{xx \mid x \in L \in \mathscr{CF}_\Sigma\}$

   (c) $L_3 = \{a^m b^n c^p \mid n \geq m + p + 2, m, n, p > 0\}$

   (d) $L_4 = \{a^m b^{\lfloor m/2 \rfloor} \mid m \geq 0\}$

Solution.

(a) It is easy to see that $L_1 = L_{ab}.L_{bc}$ where for any two distinct letters $x, y \in \Sigma$, $L_{xy} = \{x^n y^n \mid n \geq 0\}$ and since the context-free languages are closed under concatenation, L_1 is a context-free language over $\Sigma = \{a, b, c\}$. Further each L_{xy} is deterministic and deterministic CFLs are closed under concatenation. The DPDA for L_1 is $P_{abc} = \langle Q_{abc}, \Sigma, \Gamma_{abc}, \delta_{abc}, q_{abc_0}, \{q_{abc_F}\}\rangle$ where $Q_{abc} = \{q_{abc_0}, q_{abc_a}, q_{abc_b}, q_{abc_c}, q_{abc_F}\}$, $\Gamma = \{\perp, A, B\}$ and δ_{abc} is defined as

$$\begin{array}{lllllllll}
\delta(q_{abc_0}, \varepsilon, \perp) & = & (q_{abc_F}, \varepsilon) & , & \delta(q_{abc_0}, a, \perp) & = & (q_{abc_a}, A\perp) & , & \delta(q_{abc_0}, b, \perp) = (q_{abc_b}, B\perp), \\
\delta(q_{abc_a}, a, A) & = & (q_{abc_a}, AA) & , & \delta(q_{abc_a}, b, A) & = & (q_{abc_b}, \varepsilon) & , & \delta(q_{abc_b}, b, A) = (q_{abc_b}, \varepsilon), \\
\delta(q_{abc_b}, b, \perp) & = & (q_{abc_b}, B\perp) & , & \delta(q_{abc_b}, b, B) & = & (q_{abc_b}, BB) & , & \delta(q_{abc_b}, c, B) = (q_{abc_c}, \varepsilon), \\
\delta(q_{abc_c}, \varepsilon, \perp) & = & (q_{abc_F}, \varepsilon).
\end{array}$$

We could have designed the automaton with a single state too.

(b) One may use the pumping lemma to prove that L_2 is not context-free. For any $m > 0$ choose any $w \in L$ such that $|w| \geq m$ and $z = ww$. For any decomposition of $z = ww = tuvyx$ such that $|uy| > 0$, it is clear that $tuvy \preceq w$ and $x = x'w$ for some x'. Then for all $k \neq 1$ clearly $z_k = tu^k vy^k x \notin L_2$.

(c) L_3 is a deterministic context-free language as defined by the following DPDA

$$P_3 = \langle\{q_0, q_1, q_2, q_F\}, \{a, b, c\}, \{\perp, A, B\}, \delta_3, q_0, \{q_F\}\rangle$$

Notice that $\{b^n \mid n \geq 2\}$ is a subset of L_4. We use the states q_1 and q_2 to verify the predicate "at least 2 occurrences of b over and above the numbers of aa and cs".

$$\begin{array}{lllllllll}
\delta(q_0, \varepsilon, \perp) & = & (q_F, \varepsilon) & , & \delta(q_0, a, \perp) & = & (q_0, A\perp) & , & \delta(q_0, a, A) = (q_0, AA), \\
\delta(q_0, b, A) & = & (q_0, \varepsilon) & , & \delta(q_0, b, \perp) & = & (q_0, B\perp) & , & \delta(q_0, b, B) = (q_0, BB), \\
\delta(q_0, c, B) & = & (q_0, \varepsilon) & , & \delta(q_0, \varepsilon, B) & = & (q_1, \varepsilon) & , & \delta(q_1, \varepsilon, B) = (q_2, \varepsilon), \\
\delta(q_2, \varepsilon, B) & = & (q_2, \varepsilon) & , & \delta(q_2, \varepsilon, \perp) & = & (q_F, \varepsilon).
\end{array}$$

(d) It is clear that for all $m \geq 0$, $m = 2\lfloor m/2 \rfloor + m\%2$. The following DPDA accepts the language. Let $P_4 = \langle\{q_0, q_1, q_F\}, \{a, b\}, \{\perp, B\}, \delta_4, q_0, \{q_F\}\rangle$ where

$$\begin{array}{lllllllll}
\delta_4(q_0, \varepsilon, \perp) & = & (q_F, \varepsilon) & , & \delta_4(q_0, a, \perp) & = & (q_1, \perp) & , & \delta_4(q_1, a, \perp) = (q_0, B\perp) \\
\delta_4(q_0, a, B) & = & (q_1, B) & , & \delta_4(q_1, a, B) & = & (q_1, BB) & , & \delta_4(q_0, b, B) = (q_0, \varepsilon) \\
\delta_4(q_1, b, B) & = & (q_0, \varepsilon).
\end{array}$$

3. **(8 marks)** Prove that the intersection of a regular language and a context-free language over the same alphabet is context-free.

Solution.

Case $\varepsilon \notin R \cap C$**.** Let $R, C \subseteq \Sigma^*$ be regular and context-free languages respectively. Let $D = \langle Q, \Sigma, \delta, q_0, F \rangle$ be a DFA such that $\mathcal{L}(D) = R - \{\varepsilon\}$ and let $G = \langle V, \Sigma, S, P \rangle$ be a positive context-free grammar with $\mathcal{L}(G) = C - \{\varepsilon\}$. For any pair of states $p, q \in Q$ and $a \in \Sigma$, let a^{pq} be a new state and let $G = \langle V_{R\cap C}, \Sigma, S_{R\cap C}, P_{R\cap C} \rangle$ be a new grammar with $V_{R\cap C} = \{S_{R\cap C}\} \cup \{X^{pq} \mid p, q \in Q, X \in V \cup \Sigma\}$. Define $P_{R\cap C} = \{S_{R\cap C} \to S\} \cup \{X^{pq} \to X_1^{pr_1} X_2^{r_1 r_2} \cdots X_n^{r_{n-1}q} \mid X \to X_1 X_2 \cdots X_n \in P\} \cup \{a^{pq} \to a \mid a \in \Sigma, \delta(p, a) = q\}$ Then $G_{R\cap C}$ is a positive CFG. It is easy to see that $\mathcal{L}(G_{R\cap C}) = R \cap C$ because for each $\varepsilon \neq u = a_1 \ldots a_m \in R \cap C$, we have $S \Rightarrow_G^* u$ and for some sequence $q_0, \ldots q_m$ where $q_m \in F$, $\delta^*(q_0, u) = q_m$. This implies there exists a derivation $S_{R\cap C} \Rightarrow_{G_{R\cap C}}^* X^{q_0 q_m} \Rightarrow_{G_{R\cap C}}^* a_1^{q_0 q_1} \ldots a_m^{q_{m-1} q_m} \Rightarrow_{G_{R\cap C}}^* a_1 \ldots a_m = u$. This proves that $R \cap C \subseteq \mathcal{L}(G_{R\cap C})$. We may then prove that for each $X^{pq} \in V_{R\cap C}$, $X^{pq} \Rightarrow_{G_{R\cap C}}^* u \in \Sigma^*$ implies there exists a derivation $X \Rightarrow_G^* u$ in G and $\delta^*(p, u) = q$ in D.

Case $\varepsilon \in R \cap C$**.** Then we have $q_0 \in F$ and $G_{R\cap C}$ may be modified to include a new start symbol $S'_{R\cap C}$ and new rules $S'_{R\cap C} \to \varepsilon | S_{R\cap C}$.

4. **(8 marks)** Let $G = \langle V, \Gamma, S, P\rangle$ be a context-free grammar in Chomsky Normal Form (CNF), with $P = N \cup T$ where N is the set of *nonterminal* productions of the form $A \to BC$ with $A, B, C \in V$ and T is the set of *terminal* productions of the form $A \to a$ with $A \in V$ and $a \in \Gamma$.

   For each production $\pi \in N$ assume a new unique pair of bracketing symbols "$[_\pi$" and "$]_\pi$". Let Π be the set of all such bracketing symbols such that $\Pi \cap \Gamma = \emptyset$. Let $G_S = \langle V, \Gamma \cup \Pi, S, P_S\rangle$ be the grammar with $N_S = \{A \to [_\pi B]_\pi C \mid \pi = A \to BC \in N\}$ and $P_S = N_S \cup T$.

   Prove that there exists a deterministic push-down automaton that accepts the language $\mathcal{L}(G_S)$.

   Solution. We perform the following steps.

   (a) We transform the grammar G_S into an equivalent grammar G'_S in Greibach normal form (GNF)

   (b) Using the theorem that for every grammar G in GNF, there is a NPDA N' which accepts the language generated by G we construct the NPDA N'_S which accepts $\mathcal{L}(G'_S) = \mathcal{L}(G_S)$.

   (c) We prove that N'_S is actually deterministic.

   (a) Since G is in CNF, $\varepsilon \notin \mathcal{L}(G_S)$ and there are no ε rules in G and hence G is positive. From this it follows by construction that G_S is also positive and has no ε rules. We transform each of the productions in N_S as follows. For each rule $A \to [_\pi B]_\pi C \in P_S$, let K_π be a new non-terminal symbol and let N'_S and T'_S be defined as

$$\begin{array}{rcl} N'_S & = & \{A \to [_\pi B K_\pi C \mid A \to [_\pi B]_\pi C \in P_S\} \\ T' & = & T \cup \{K_\pi \to]_\pi \mid \pi \in P\} \end{array}$$

   Let $P'_S = N'_S \cup T'$ and $G'_S = \langle V \cup \{K_\pi \mid \pi \in P\}, \Gamma \cup \Pi, S, P'_S\rangle$. Then by the factoring theorem $\mathcal{L}(G'_S) = \mathcal{L}(G_S)$. Further G'_S is in GNF.

   (b) There exists a NPDA $N' = \langle \{q_0\}, \Gamma \cup \Pi, \Delta, q_0, \emptyset\rangle$ where

$$((q_0, a, A), (q_0, \alpha)) \in \Delta \text{ \underline{iff} } A \to a\alpha \in P'_S \quad (1)$$

   such that $\mathcal{L}_E(N') = \mathcal{L}(G'_S) = \mathcal{L}(G_S)$.

   (c) Assume Δ in (1) is not deterministic. Then there exist two distinct transitions $((q_0, a, A), (q_0, \alpha)), ((q_0, a, A), (q_0, \beta)) \in \Delta$ such that $\alpha \neq \beta$.

   Case $\underline{\alpha = \varepsilon \neq \beta}$**.** Then by (1) $A \to a$ and $A \to [_\pi B K_\pi C$ are two productions in G'_S. But this implies $a = [_\pi$ which is impossible since there is no rule of the form $A \to [_\pi$ in G'_S.

   Case $\underline{\alpha \neq \varepsilon = \beta}$**.** Similar to the previous case.

   Case $\underline{\alpha \neq \varepsilon \neq \beta}$**.** This implies for two distinct $\pi, \pi' \in P$ we have $A \to [_\pi B K_\pi C$ and $A \to [_{\pi'} B' K_{\pi'} C'$ which is again impossible since it implies $a = [_\pi \neq [_{\pi'} = a$. Hence N' is deterministic.