

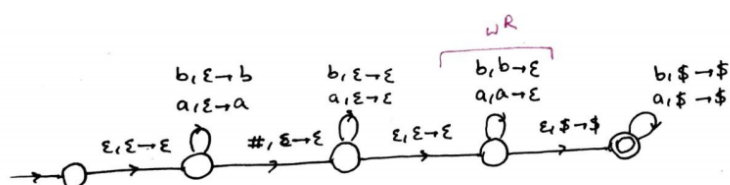


نظریه زبان‌ها و ماشین‌ها
پاسخ تکلیف چهارم
ترم دوم ۹۹-۰۰

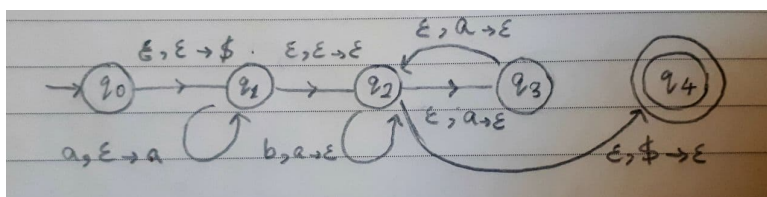
در همه بخش‌های تمامی سؤالات (به غیر از بخش‌هایی که مشخص شده است)، $\Sigma = \{a, b\}$ است.

۱- برای هریک از زبان‌های مستقل از متن معرفی شده یک PDA طراحی کنید.

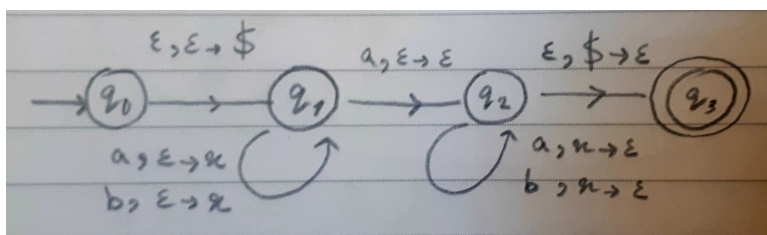
$$L_1 = \{ w \# x \mid w^R \text{ یک زیررشته برای } x \text{ است} \}$$



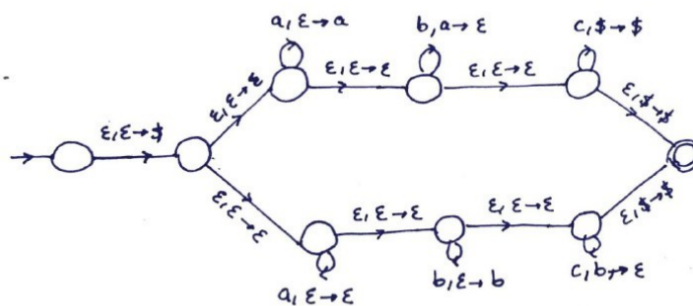
$$L_2 = \{ a^n b^m \mid n \geq m \text{ و } n - m \text{ زوج باشد} \}$$



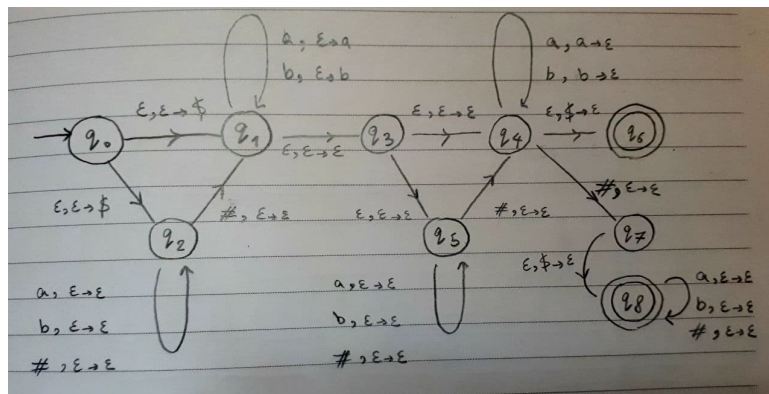
$$L_3 = \{ w \mid w \text{ فرد است و عضو وسط آن سیمبل } a \text{ باشد} \}$$



$$L_4 = \{ a^i b^j c^k \mid i = j \text{ یا } j = k \} \quad (\Sigma = \{a, b, c\} \text{ روی الفبای})$$



$$L_5 = \{ w_1 \# w_2 \# \dots \# w_k \mid k \geq 1, \forall i : w_i \in \Sigma^*, \exists i \neq j : w_i = w_j^R \}$$



۲ - اگر C یک زبان مستقل از متن و R یک زبان منظم باشد، ثابت کنید زبان‌های زیر مستقل از متن هستند.

$$L_1 = \{ xy \mid yx \in C \}$$

با داشتن یک PDA به نام P برای زبان C ؛ یک PDA به نام P' برای این زبان طراحی می‌کنیم. این عملیات را طی دو مرحله برای ساخت x و y انجام می‌دهیم:

۱- در ابتدا با شروع از یکی از استیت‌های P (q_x) شروع به ساخت قسمت x می‌کنیم. مشکلی که وجود دارد این است که در P قبل از این رسیدن به q_x در استک مقادیری ریخته شده، مقادیری که در مرحله قبل ساخته شدن رشته yx یعنی در قسمت ساخت y در استک push شده‌اند و P برای ساخت x از آن‌ها استفاده می‌کند. چون ما در PDA جدید به این مقادیر دسترسی نداریم؛ برای حل این مشکل می‌توانیم این مقادیر را پیش‌بینی کنیم و به جای pop کردن مقادیری که باید در استک وجود داشته باشند، مقادیر معادلی در استک push می‌کنیم (با non-determinism) و تا رسیدن به یک استیت اکسپت این روند را ادامه می‌دهیم.

۲- در مرحله دوم لازم است قسمت y را بسازیم؛ برای ساخت قسمت y از استیت آغازین P به ادامه عملیات با استکی که تا الان پر کرده‌ایم می‌پردازیم. در این مرحله هرجایی لازم به push کردن یک مقدار بود به جای این عمل متغیر نظیر آن که در مرحله قبل در استک ریخته‌ایم را pop می‌کنیم. در آخر آنقدر به این عملیات ادامه می‌دهیم تا به q_x برسیم و آن استیت را اکسپت می‌کنیم.

$$L_2 = \{ w \mid w^R \in C \}$$

گرامر زبان C را نوشته و آن را به فرم نرمال چامسکی تبدیل می‌کنیم در این صورت برای ساختن گرامری برای زبان L_2 کافیه همه‌ی rule‌هایی که به صورت $C \rightarrow AB$ در گرامر وجود دارند را با $C \rightarrow BA$ جایگزین کنیم.

$$L_3 = \{ w \mid \exists x \in R : wx \in C \}$$

زبان C یک CFL است پس برای آن یک PDA به نام N وجود دارد و همین‌طور زبان R یک زبان منظم است پس برای آن یک DFA به نام M می‌توانیم تعریف کنیم:

$$M = (Q_1, \Sigma, \delta_1, q_{01}, F_1)$$

$$N = (Q_2, \Sigma, \Gamma, \delta_2, q_{02}, F_2)$$

حال برای L_1 یک PDA جدید به نام N' طراحی می‌کنیم که به صورت زیر تعریف می‌شود:

$$N' = (Q_1 \times Q_2, \Sigma, \Gamma, \delta', q'_0, F')$$

$$q'_0 = (q_{01}, q_{02})$$

$$F' = (q_i, q_j) \text{ for all } q_i \in F_1 \text{ and } q_j \in F_2$$

$$\text{for all } \delta_2(q_{i1}, x, y) = (q_{i2}, z) \longrightarrow \delta'((q_{01}, q_{i1}), x, y) = ((q_{01}, q_{i2}), z)$$

$$\text{for all } \delta_2(q_{i1}, x, y) = (q_{i2}, z) \text{ and } \delta_1(q_{j1}, x) = q_{j2} \longrightarrow \delta'((q_{j1}, q_{i1}), \epsilon, y) = ((q_{i2}, q_{j2}), z)$$

۳ - با استفاده از الگوریتم CYK بررسی کنید که رشته‌های aaabb و aabab در زبان متناظر با گرامر زیر هستند یا خیر.

$$G : \begin{cases} S \longrightarrow AP \mid AB \\ E \longrightarrow AP \mid EB \mid b \\ P \longrightarrow EB \\ A \longrightarrow a \\ B \longrightarrow b \end{cases}$$

ϕ				
ϕ	ϕ			
ϕ	ϕ	$\{S, E\}$		
ϕ	ϕ	$\{S\}$	$\{E, P\}$	
$\{A\}$	$\{A\}$	$\{A\}$	$\{E, B\}$	$\{E, B\}$
swing	a	a	a	b

$S \notin \{\phi\}$ پس رشته aaabb عضو این زبان نیست.

ϕ				
ϕ	ϕ			
ϕ	ϕ	ϕ		
ϕ	$\{S\}$	ϕ	$\{S\}$	
$\{A\}$	$\{A\}$	$\{E, B\}$	$\{A\}$	$\{E, B\}$
ω	a	a	b	a

$S \notin \{\phi\}$ پس رشته aabab عضو این زبان نیست.

۴ - گرامرهای زیر را با حذف قواعد بی‌استفاده، اِسیلونی و یک‌ه تا حد ممکن ساده کنید.

$$G_1 : \begin{cases} S \longrightarrow aSab \mid bS \mid \epsilon \mid bA \mid aB \\ A \longrightarrow bB \mid B \\ B \longrightarrow B \mid AbB \\ C \longrightarrow A \mid aab \end{cases}$$

$$G_1 : \{ S \longrightarrow aSab \mid bS \mid \epsilon$$

$$G_2 : \begin{cases} S \rightarrow aCC \mid aAA \\ A \rightarrow aCC \mid a \\ B \rightarrow AaBc \mid ba \mid a \\ C \rightarrow cCC \mid A \end{cases}$$

$$G_2 : \begin{cases} S \rightarrow aCC \mid aAA \\ A \rightarrow aCC \mid a \\ C \rightarrow cCC \mid aCC \mid a \end{cases}$$

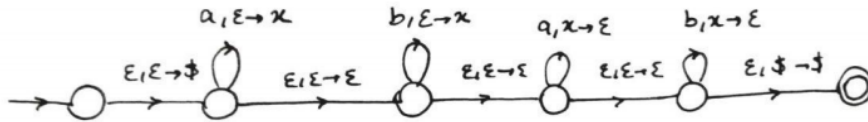
۵ - برای هر یک از موارد زیر، بررسی کنید زبان توصیف شده مستقل از متن است یا خیر. برای بخش‌های ۱ و ۳ روی الفبای $\Sigma = \{a, b, c\}$ بررسی شود.

$$L_1 = \{ w \mid n_a(w) < n_b(w) \cdot n_c(w) \}$$

زبان $L = \{ w \mid n_a(w) < n_c(w) \text{ and } n_a(w) < n_b(w) \}$ زیر مجموعه نامتناهی از L_1 است. در نتیجه با اثبات مستقل از متن نبودن L می توان ثابت کرد که L_1 نیز مستقل از متن نیست. حال اگر $s = a^p b^{p+1} c^{p+1}$ را در نظر بگیریم، با توجه به لم تزریق و شروط آن vy می تواند شامل a و b یا شامل b و c باشد. در حالت اول اگر $uv^2 xy^2 z$ را در نظر بگیریم در کمترین حالت تعداد a ها با c ها مساوی است و یا تعداد a ها بیشتر است. در حالت دوم اگر $uv^0 xy^0 z$ را در نظر بگیریم در بیشترین حالت تعداد b ها و یا c ها با تعداد a ها برابر می شود و یا کمتر. در نتیجه L و L_1 مستقل از متن نیستند.

$$L_2 = \{ a^n b^m a^j b^k \mid n + m = j + k \}$$

مستقل از متن است و PDA متناظر با آن به صورت زیر می باشد:



$$L_3 = L(a^* b^* c^*) - \{ a^n b^n c^n \mid n \geq 0 \}$$

زبان L_3 مستقل از متن است و به صورت زیر تعریف می کنیم:

$$G(L_3) : \begin{cases} S_0 \rightarrow S_{A>B} \mid S_{A>C} \mid S_{B>A} \mid S_{B>C} \mid S_{C>A} \mid S_{C>B} \\ S_{A>B} \rightarrow aS_{A=B}C \\ S_{B>A} \rightarrow S_{A=B}bC \\ S_{A=B} \rightarrow aS_{A=B}b \mid \varepsilon \\ C \rightarrow cC \mid \varepsilon \\ S_{A>C} \rightarrow aS_{A=C} \\ S_{C>A} \rightarrow S_{A=C}c \\ S_{A=C} \rightarrow aS_{A=C}c \mid B \\ B \rightarrow bB \mid \varepsilon \\ S_{B>C} \rightarrow AbS_{B=C} \\ S_{B<C} \rightarrow AS_{B=C}c \\ S_{B=C} \rightarrow bS_{B=C}c \mid \varepsilon \\ A \rightarrow aA \mid \varepsilon \end{cases}$$

$$L_4 = \{ w_1 \# w_2 \mid w_2 \text{ زیر رشته از } w_1 \text{ است} \}$$

زبان L_4 مستقل از متن نیست. فرض می کنیم طول لم تزریق p باشد. رشته $s = a^p b^p \# a^p b^p$ را در نظر می گیریم. اگر $s = uvxyz$ باشد. هیچ کدام از v و y نمی تواند شامل $\#$ باشند.

$$\begin{aligned} v, y \neq \varepsilon \text{ and } \# \in z &\Rightarrow uv^2xy^2z \notin L_4 \\ v, y \neq \varepsilon \text{ and } \# \in u &\Rightarrow uv^0xy^0z \notin L_4 \\ v = \varepsilon \text{ or } y = \varepsilon &\Rightarrow uv^0xy^0z \notin L_4 \\ v, y \neq \varepsilon \text{ and } \# \in x &\xrightarrow{|xyz| < p} v = b^+, y = a^+ \Rightarrow uv^2xy^2z \notin L_4 \end{aligned}$$

$$L_5 = \{ w \mid \text{حداقل یکی از } n_a(w) \text{ یا } n_b(w) \text{ عددی اول است} \}$$

زبان L_5 مستقل از متن نیست. با توجه به لم تزریق و شروط آن رشته ای به فرم $s = a^p$ را در نظر می گیریم که $s = uvxyz$ است.

$$\begin{aligned} v = a^q, y = a^t &\xrightarrow{\text{pumping lemma}} q + t > 0 \\ r = |uxz| &= p - q - t \\ |uv^rxy^rz| &= r + rq + rt = r(1 + q + t) \\ \text{if } r > 1 &\text{ then } r(1 + q + t) \text{ is not prime.} \\ \text{if } r = 0 &\text{ then } |uv^2xy^2z| = |v^2y^2| = 2p \text{ is not prime.} \\ \text{if } r = 1 &\text{ then } |uv^{p+1}xy^{p+1}z| = p^2 \text{ is not prime.} \end{aligned}$$

$$L_6 = \{ x_1 \# x_2 \# \dots \# x_k \mid \text{for } k \geq 0, \text{ each } x_i \in L(a^*), \text{ and } x_i \neq x_j \text{ for } i \neq j \} \quad (\Sigma = \{a, \#\})$$

زبان L_6 مستقل از متن نیست. با توجه به لم تزریق و شروط آن رشته ای به فرم $s = x^0 \# x^1 \# x^2 \# \dots \# x^p$ را در نظر می گیریم. طبق لم تزریق $|vy|$ نمی تواند صفر باشد. پس می توان گفت یکی از میان v و y حتماً متشکل از حروفی است. حال ممکن است ۳ حالت رخ دهد:

۱. فقط a دارد.
 ۲. فقط $\#$ دارد: با تزریق زیر رشته تهی تشکیل می شود و تکراری است.
 ۳. شامل a و $\#$ است.
- در حالت اول و سوم با تزریق و یا تخلیه زیر رشته تکراری تولید می شود.

موفق باشید :