

سوال ۱.

۱.۱

الف

با انتخاب یک رشته از مجموعه

$$L = \{a^m b^n c^p d^q \mid m + n = p + q\}$$

ما می‌خواهیم یک گرامر بسازیم که بتواند این رشته را تولید کند. بنابراین می‌توانیم گرامر زیر را بسازیم:

$$S \rightarrow aSd \mid A \mid B$$

$$A \rightarrow C \mid aAc$$

$$B \rightarrow C \mid bBd$$

$$C \rightarrow bCc \mid \epsilon$$

اگر به گرامر نگاه کنیم، می‌بینیم که در هر قاعده، تعداد حروف a و b همیشه با تعداد حروف c و d مطابقت دارد. بنابراین، با انتخاب انتخابات مناسب از قواعد، می‌توانیم هر رشته از L را تولید کنیم.

ب

برای این بخش، مجموعه رشته‌های ما به شکل زیر است:

$$L = \{uawb \mid u, w \in \{a, b\}^* \text{ and } |u| = |w|\}$$

برای این که این رشته‌ها را تولید کنیم، نیاز به گرامری داریم که بتواند حروف a و b را در مواقعی تولید کند که تعداد حروف قبل و بعد از آنها مساوی باشد. بنابراین، می‌توانیم گرامر زیر را بسازیم:

$$S \rightarrow Ab$$

$$A \rightarrow a \mid BAB$$

$$B \rightarrow a \mid b$$

این گرامر با توجه به تعداد مساوی حروف در رشته‌های u و w ، همیشه یک a را در میانه رشته و یک b را در انتهای رشته قرار می‌دهد.

۱.۲

الف

گرامر زیر را برای توصیف داریم:

$$(\circ^* \circ^* \circ^*) (\# (\circ^* \circ^* \circ^*))^*$$

در این گرامر، A می‌تواند یا یک \circ تولید کند یا با تولید A به حالت توصیف شده برسیم. علاوه بر این، عملگری است که رشته‌ها را در میانه جدا می‌کند.

ب

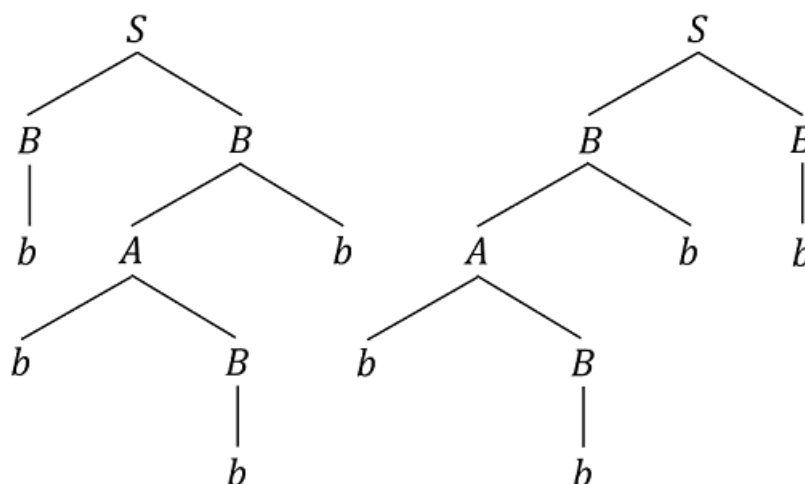
برای این قسمت داریم:

$$L = \{(a|b)uw \mid u, w \in \{a, b\}^* \text{ and } |u| = |w|\}$$

از آنجا که رشته یا با a یا با b شروع می‌شود و سپس دوتا دوتا ترکیب‌های یکسان از آنها تولید می‌شود، به توصیف مجموعه فوق می‌رسیم.

۱.۳

الف



برای نشان دادن اینکه گرامر ابهام دارد، می‌توانیم دو درخت مختلف برای یک رشته از گرامر بسازیم. این نشان می‌دهد که گرامر ابهام دارد زیرا برای یک رشته وجود چندین درخت اشتقاق ممکن است.

ب

برای تبدیل گرامر به فرم نرمال چامسکی، می‌توانیم قواعد زیر را داشته باشیم:

$$S \rightarrow AS' \mid BB$$

$$S \rightarrow AS' \mid BB$$

$$S' \rightarrow SB$$

$$A \rightarrow a \mid CB \mid A'D$$

$$A' \rightarrow SA$$

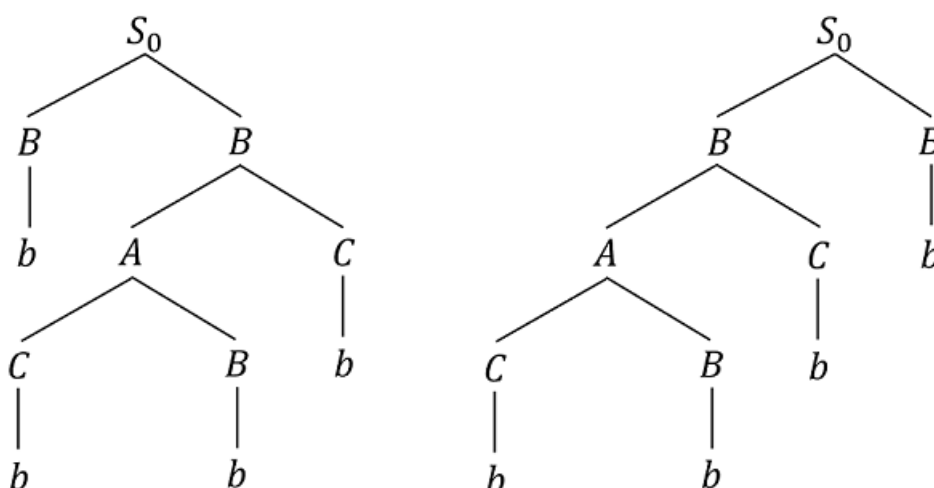
$$B \rightarrow b \mid AC \mid SD$$

$$C \rightarrow b$$

$$D \rightarrow a$$

ج

بله، گرامر هنوز ابهام دارد. همانطور که در شکل زیر می‌بینید، می‌توان برای یک رشته دو درخت اشتقاق مختلف ساخت.



سوال ۲.

۲.۱

الف

برای اثبات بستگی زبان‌های مستقل از متن در برابر عمل الحاق، دو گرامر زیر را در نظر می‌گیریم:

$$G_1 = (V_1, \Sigma_1, R_1, S_1) \quad G_2 = (V_2, \Sigma_2, R_2, S_2)$$

گرامر الحاق این دو به شکل زیر خواهد بود:

$$G_T = (V_1 \cup V_2 \cup \{S_1\}, \Sigma_1 \cup \Sigma_2, R_1 \cup R_2 \cup \{S_1 \rightarrow S_1 S_2\}, S_1)$$

که در آن یک نشان‌دهنده جدید، یعنی S_1 ، تعریف شده است و یک ارتباط جدید، $S_1 \rightarrow S_1 S_2$ ، برقرار شده است. این گرامر نشان‌دهنده عمل الحاق است، بنابراین زبان‌های مستقل از متن در برابر این عمل بسته هستند.

ب

حالا یک گرامر دلخواه، $G_1 = (V_1, \Sigma_1, R_1, S_1)$ ، را در نظر بگیرید. برای عملگر $*$ ، گرامری را می‌توان به شکل زیر در نظر گرفت:

$$G_* = (V_1 \cup \{S_*\}, \Sigma_1, R_1 \cup \{S_* \rightarrow S_1 S_* | \epsilon\}, S_*)$$

مانند بخش الف، یک نشان‌دهنده جدید، S_* ، در نظر گرفته شده است و با افزودن قاعده $S_* \rightarrow S_1 S_* | \epsilon$ به مجموعه قواعد، می‌توان عملگر $*$ را پیاده ساخت. پس زبان‌های مستقل از متن نسبت به این عملگر نیز بسته خواهند بود.

۲.۲

الف

با توجه به زبان مستقل از متن ارائه شده، زبان زیر را در نظر بگیرید:

$$L = \{a^m b^k c^n \mid m, n \geq 1 \text{ and } (m = k + n \text{ or } k = m + n \text{ or } n = k + m)\}$$

با توجه به بستگی زبان‌های مستقل از متن در برابر عمل اجتماع، می‌توان زبان بالا را به سه زبان مجزا تقسیم کرد و با اثبات مستقل بودن هر یک از آن‌ها، در نهایت به مستقل بودن زبان L می‌رسیم. بنابراین، زبان‌های زیر را در نظر بگیرید:

$$L_1 = \{a^m b^k c^n \mid m = k + n\}$$

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

$$L_3 = \{a^m b^k c^n \mid k = m + n\}$$

در ابتدا برای زبان L_1 ، گرامر مستقل از متن زیر را می‌نویسیم:

$$S \rightarrow aSc \mid X$$

$$X \rightarrow aXb \mid \epsilon$$

که در آن جمع تعداد تکرار حروف c و b با a برابر خواهد بود، چرا که برای هر a ، یک c یا یک b وجود دارد. بنابراین، گرامر زبان L_1 مستقل از متن است.

به همین ترتیب، برای زبان L_2 ، گرامر مستقل از متن زیر را در نظر بگیرید:

$$S \rightarrow aSc \mid X$$

$$X \rightarrow bXc \mid \epsilon$$

که در آن جمع تعداد تکرار a و b برابر با c خواهد بود. بنابراین، این زبان نیز یک زبان مستقل از متن است. در نهایت، برای زبان L_3 ، می‌توان دو زبان مستقل از متن زیر را الحاق کرد:

$$L_4 = (a^m b^k \mid m = k) \quad L_5 = (b^k c^n \mid k = n)$$

که می‌دانیم این دو زبان مستقل از متن هستند، پس الحاق آنها نیز مستقل از متن بوده و در نتیجه زبان L مستقل از متن خواهد بود.

برای بخش بعدی، زبان زیر را در نظر بگیرید:

$$L = \{a^{m_1} b^{k_1} c^{n_1} \dots a^{m_i} b^{k_i} c^{n_i} \mid i \geq 1 \text{ and } \forall j \leq i \ m_j, n_j \geq 1 \text{ and } k_j = 3m_j + 4n_j\}$$

از آنجایی که زبان‌های مستقل از متن نسبت به عمل $*$ بسته هستند، می‌توان زبان L را به این صورت تقلیل داد:

$$L = (L_1)^* \implies L_1 = (a^m b^k c^n \mid k = 3m + 4n)$$

بنابراین کافی است نشان دهیم زبان L_1 مستقل از متن است. برای اینکار، دو زبان زیر را با گرامر متناظر با آنها در نظر می‌گیریم:

$$L_2 = (a^m b^k \mid k = 3m) \quad S \rightarrow aSbbb \mid \epsilon$$

$$L_3 = (b^k c^n \mid k = 4n) \quad S \rightarrow bbbbSc \mid \epsilon$$

با الحاق این دو زبان، L_1 ساخته می‌شود. از آنجایی که هرکدام از آنها مستقل از متن هستند و عملگر الحاق نیز نسبت به این کار بسته است، در نهایت زبان L یک زبان مستقل از متن خواهد بود.

سوال ۳.

۳.۱

الف

زبان

$$L = \{w \mid w \in \{0, 1\}^* \text{ and } w = w^R\}$$

رشته‌هایی را شامل می‌شود که توالی آن‌ها با توالی معکوس آن‌ها برابر است. در عبارت دیگر، این زبان شامل رشته‌های پالیندروم است.

ب

برای اینکه ماشین بتواند رشته‌های با طول زوج را بپذیرد، باید از حالت q_1 بدون دریافت ورودی به q_2 بتوانیم بریم. برای اینکار، می‌توانیم یک ترانزیشن $\epsilon, \epsilon \rightarrow \epsilon$ بین این دو حالت اضافه کنیم. با این کار، ماشین می‌تواند زبان L را بپذیرد.

برای این قسمت زبانی که داریم به شکل زیر است:

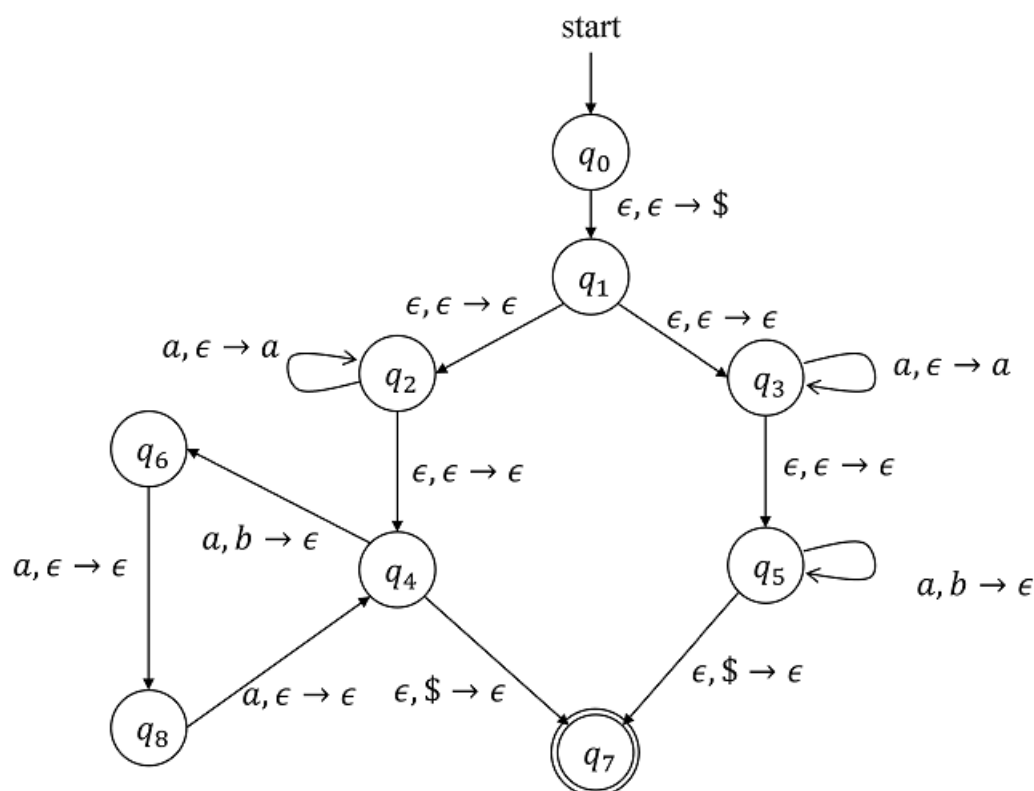
$$L = \{a^i b^j \mid i = j \text{ or } i = 3j\}$$

برای نشان دادن این زبان با یک ماشین پشته‌ای، می‌توانیم یک گرامر مستقل از متن به شکل زیر تعریف کنیم:

$$S \rightarrow A|B$$

$$A \rightarrow aAb|\epsilon$$

$$B \rightarrow aaBbb|\epsilon$$



ب

زبان بعدی به شکل زیر است:

$$L = \{a^i b^j c^k \mid k \geq \min(i, j)\}$$

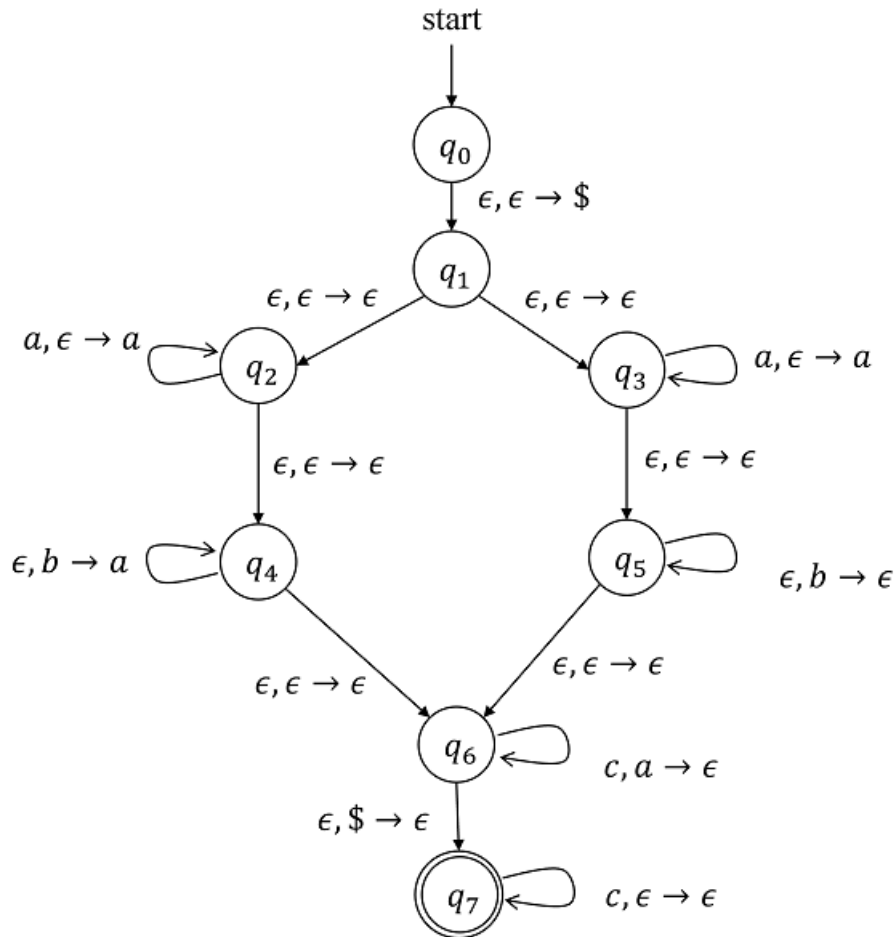
برای نشان دادن این زبان با یک ماشین پشته‌ای، می‌توانیم یک گرامر مستقل از متن به شکل زیر تعریف کنیم:

$$S \rightarrow AS|BS|T$$

$$A \rightarrow aAc|\epsilon$$

$$B \rightarrow bBc|\epsilon$$

$$T \rightarrow cT|\epsilon$$



ج

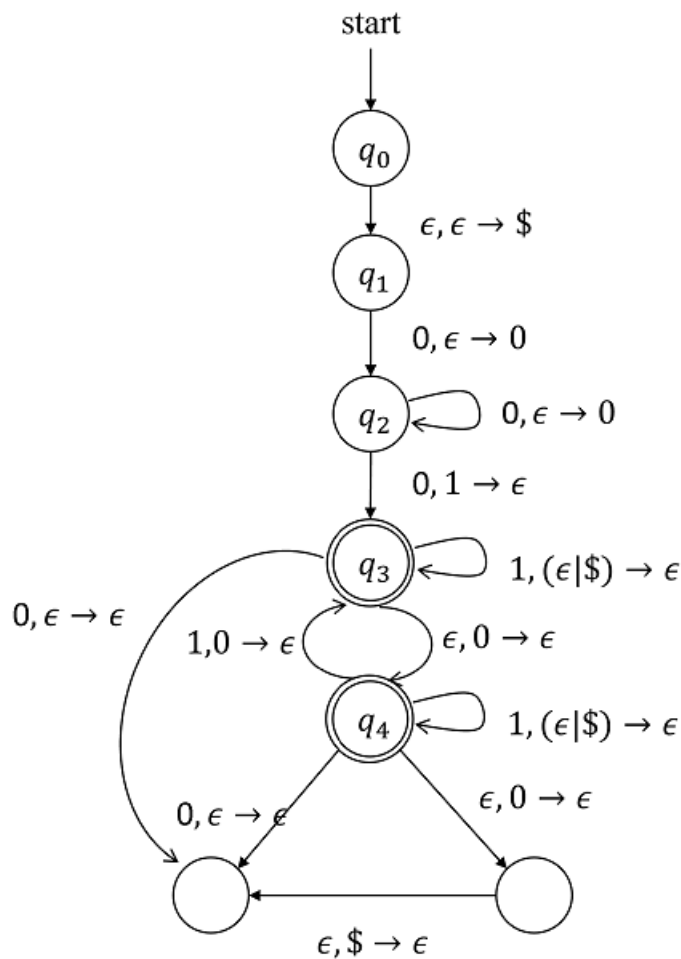
زبان بعدی به شکل زیر است:

$$L = \{a^n b^m \mid n \neq 2m + 1 \text{ and } n, m > 0\}$$

برای نشان دادن این زبان با یک ماشین پشته‌ای، می‌توانیم یک گرامر مستقل از متن به شکل زیر تعریف کنیم:

$$S \rightarrow aS \mid aaS \mid T$$

$$T \rightarrow aT \mid \epsilon$$



سوال ۴.

۴.۱

الف

به زبان L_1 بیان شده به صورت زیر توجه می‌کنیم:

$$L_1 = \{w\#x\#y\#z \mid w, x, y, z \in a, b^* \text{ and } |w| = |z|, |x| = |y|\}$$

برای گرامر این زبان خواهیم داشت.

$$S \rightarrow CSC \mid A$$

$$A \rightarrow \#B\#$$

$$B \rightarrow CBC \mid \#$$

$$C \rightarrow a \mid b$$

در اینجا می‌دانیم عبارت با حروف a و b شروع و خاتمه پیدا می‌کند. همچنین در بین علامت‌های $\#$ طوری گرامر را قرار داده‌ایم که اندازه دو دنباله وسطی با یکدیگر و دو دنباله ابتدایی و انتهایی با هم برابر باشند.

ب

در اینجا داریم.

$$L_2 = \{w\#x\#y\#z \mid w, x, y, z \in a, b^* \text{ and } |w| = |y|, |x| = |z|\}$$

نشان می‌دهیم برعکس قسمت بالا این زبان مستقل از متن نیست. طبق لم تزریق بدست می‌آوریم که:

$$\Rightarrow w\#x\#y\#z = (a|b)^p\#(a|b)^p\#(a|b)^p\#(a|b)^p = uvxyz$$

در اینجا دو بخش میانی یعنی y و v نمی‌توانند دارای علامت $\#$ باشند زیرا که با تزریق این قسمت بیش از سه هشتگ خواهیم داشت. پس این دو قسمت یا هر دو داخل یکی از زیر رشته‌های w, x, y و z قرار داشته و یا در دو زیررشته کناری قرار دارند.

اگر هر دو در زیررشته‌های کناری باشند، از آنجایی که اندازه دو دنباله x و z و دو دیگر دنباله باید برابر باشند، پامپ شدن این قسمت تساوی به هم خورده و رشته جدید تولید شده خارج از زبان L_2 قرار می‌گیرد.

اگر هر دو در یک زیررشته نیز باشند با اعمال تزریق تناسب زیررشته‌ها به هم خورده پس می‌توان نتیجه گرفت که براساس لم تزریق رشته‌های تولید شده خارج از زبان قرار گرفته و این زبان یک زبان مستقل از متن نخواهد بود.

۴.۲

الف

$$L_1 = \{a^n b^n c^n \mid n \geq 0\}$$

با توجه به لم تزریق رشته زیر را در نظر می‌گیریم.

$$\Rightarrow a^p b^p c^p = uvxyz$$

با توجه به رشته، هرکدام از رشته‌های v و y فقط شامل یک حرف خواهند بود زیرا در غیر این صورت با افزایش آنها تناسب رشته به هم می‌ریزد. حال سه حالت داریم.

- اگر هر دو یک حرف یکسان باشند، با انجام تزریق تعداد آن حرف بیشتر از مابقی شده و تناسب رشته به هم می‌ریزد.
- اگر v برای حرف a و y برای حرف b باشد. با پامپ کردن این دو مقدار نسبت حروف به حرف c به هم خورده و رشته‌ای خارج از زبان خروجی می‌دهد.
- رشته v برای حرف b و y برای حرف c باشد. همانند قسمت قبل تناسب رشته نسبت به حرف a به هم می‌ریزد.

در نتیجه در تمامی حالات رشته‌ای خارج از زبان L_1 تولید شده و این زبان مستقل از متن نیست.

ب

$$L_2 = \{a^{(n-1)(n+1)} \mid n > 0\}$$

همانند قسمت قبل داریم.

$$\implies a^{(p-1)(p+1)} = a^{p^2-1} = uvxyz$$

در اینجا دو رشته v و y تنها از حرف a تشکیل شده‌اند و برای آنها خواهیم داشت.

$$|uy| = t, \quad 0 < t \leq p$$

$$\implies |uv^t xy^t z| = p^2 - 1 + t$$

$$\implies p^2 - 1 < p^2 - 1 + t \leq p^2 + p - 1$$

$$\implies p^2 + p - 1 < p^2 + 2p$$

این نشان دهنده این است که طول رشته $uv^t xy^t z$ بین دو مقدار $p^2 - 1$ و $p^2 + 2p$ خواهد بود. در صورتی که در زبان L_2 این رشته با این طول پذیرفته نمی‌شود چون نمی‌تواند طولی بین این دو مقدار برای رشته تولید کند.

$$n = p \implies a^{(p-1)(p+1)} = a^{p^2-1}$$

$$n = p + 1 \implies a^{(p)(p+2)} = a^{p^2+2p}$$

بنابراین با استفاده از لم تزریق رشته‌ای تولید می‌شود که خارج از زبان است پس این زبان نمی‌تواند مستقل از متن باشد.

ج

$$L_3 = \{a^n b^n c^i \mid n \leq i \leq 2n\}$$

در اینجا داریم.

$$\implies a^p b^p c^{2p} = uvxyz$$

در اینجا دو رشته میانی v و y فقط می‌توانند شامل یک حرف باشند زیرا در غیر این صورت نظم رشته را از دست خواهیم داد.

حال اگر این دو رشته شامل یک حرف باشند با تکرار آنها یکی از عناصر a ، b و c بیشتر از بقیه شده و رشته‌ای خارج از زبان L_3 تولید می‌شود.

اگر v شامل b باشد با تکرار آنها تعداد حالت‌ها از a بیشتر شده و این حالت نیز قابل قبول نیست.

در حالت v شامل a باشد و y متشکل از b باشد بدست می‌آوریم:

$$\implies |v| = |y| = t \quad 1 \leq t \leq p$$

اگر دو رشته مدنظر را صفر بار تزریق کنیم خواهیم داشت.

$$\implies a^{p-t} b^{p-t} c^{2p} = uxz \implies 2p \leq 2p - 2t$$

در نتیجه با توجه به شرط زبان L_3 این رشته خارج از این زبان بوده و این یک زبان غیرمستقل از متن خواهد بود.

موفق باشید.