به نام خدا

نظریه زبان‌ها و ماشین‌ها- بهار ۱۴۰۱

پاسخ تمرین شماره ۷

دستیار آموزشی این مجموعه: پاریاب مرادی
paryabmoradi1378@gmail.com

---

1. با استفاده از لم تزریق (Pumping lemma)، اثبات کنید که زبان مربوطه، عضو کلاس زبان‌های منظم نیست: (20 نمره)

a. $L = \{ a^i b^j \mid j = i \, or \, j = 2i \}$

با برهان خلف فرض کنید زبان منظم باشد و طول پمپ را p در نظر بگیرید. لم تزریق را برای رشته‌ی $s = a^p b^p$ اجرا می‌کنیم. طبق لم تزریق داریم

$$s = xyz,\ |xy| \leqslant p,\ |y| > 0 \ \Rightarrow \ y = a^j$$

اما

$$xy^2z = a^{p+j} b^p \notin L$$

تناقض حاصله نشان می‌دهد زبان مورد نظر منظم نیست.

b. $L = \{ w \mid w \in \{a, b\}^* \, and \, n_a(w) < 2n_b(w) \}$

فرض کنید زبان مورد نظر منظم و طول پمپ p است. لم تزریق را روی رشته $s = a^{2p-1} b^p$ اجرا کنید. مانند قسمت قبلی داریم

$$s = xyz,\ |y| > 0,\ |xy| \leqslant p \ \Rightarrow \ y = a^j$$

اما داریم

$$xy^2z = a^{p+j} b^p \notin L$$

این تناقض نشان می‌دهد زبان مورد نظر ما منظم نیست.

c. $L = \{ a^{2^n} \mid n \geqslant 0 \}$

فرض کنید این زبان منظم باشد و طول پمپ را p در نظر بگیرید. لم تزریق را برای رشته a^{2^p} اجرا می‌کنیم (دقت کنید طول این رشته از p بیشتر است). پس داریم

$$s = xyz,\ |y| > 0,\ |xy| \leqslant p \ \Rightarrow\ y = a^j,\ 1 \leqslant j \leqslant p$$

که نتیجه می‌دهد

$$xy^2z = a^{2^p+j} \ \in\ L$$

اما دقت کنید عدد $2^p + j$ توانی از 2 نیست چون

$$2^p < 2^p + j < 2^p + p \leqslant 2^{p+1}$$

تناقض حاصله نشان می‌دهد L منظم نیست.

d. $L = \{ w \in \{a, b\}^* \mid w = w_1 b w_2 b \ldots b w_k,\ for\ k \geqslant 0,\ each\ w_i \in a^*,$
$and\ w_i \neq w_j\ for\ i \neq j \}$

فرض کنید زبان مورد نظر منظم باشد و طول پمپ p باشد. لم تزریق را برای رشته $w = a^p b a^{p+1} b \ldots a^{2p-1} b a^{2p}$ اجرا می‌کنیم. مانند قسمت‌های قبل داریم

$$w = xyz,\ |y| > 0,\ |xy| \leqslant p \ \Rightarrow\ y = a^j,\ 1 \leqslant j \leqslant p$$

حال دقت کنید

$$xy^2z = a^{p+j} b a^{p+1} b \ldots b a^{p+j} b \ldots b a^{2p}$$

که عضوی از زبان L نیست چون دو بخش a^{p+j} دارد. تناقض حاصله نشان می‌دهد این زبان منظم نیست.

2. **نسخه‌ی تغییریافته‌ای از لم تزریق را با تعریف زیر در نظر بگیرید:**
اگر L یک زبان منظم باشد، ثابت n را می‌توان یافت، به نحوی که به ازای هر z_1 و z_2 و z_3 که $z_1z_2z_3 \in L$ صدق کند و $|z_2| = n$، بتوانیم z_2 را به صورت uvw بازنویسی کنیم، طوری که $|v| \geqslant 1$ و به ازای هر $i \geqslant 0$ رشته‌ی $z_1uv^iwz_3$ عضو زبان L باشد. (15 نمره)

a. **لم جدید را اثبات کنید.**

n را برابر تعداد استیت‌های DFA این زبان در نظر می‌گیریم. فرض کنید پس از ورودی دادن رشته z_1 از استیت q به استیت q_0 از اتوماتا برسیم و دنباله استیت‌های پس از ورودی دادن z_2 برابر $q_0, q_1, \dots, q_n$ باشد. در انتها نیز با ورودی دادن z_3 از استیت q_n به استیت f می‌رود. طبق اصل لانه کبوتری حداقل دوتا از این استیت‌ها مانند q_i و q_j یکی هستند و $i < j$. فرض کنید قسمت‌هایی از رشته z_2 که برای قسمت‌های $q_0, \dots, q_i$، $q_i, \dots, q_j$ و $q_j, \dots, q_n$ هستند به ترتیب u، v و w باشند. حال دقت کنید که رشته‌های z_1u، v و wz_3 اتومات را به ترتیب از استیت‌های q، q_i و q_j به استیت‌های q_i، q_j و f می‌برند. با مطالب گفته شده به راحتی می‌توان دید که رشته $z_1uv^iwz_3$ اتومات را از استیت اولیه q به استیت نهایی f می‌برد که نشان می‌دهد عضوی از زبان L است.

b. **با به کارگیری لم مربوطه، نشان دهید که $L = \{ a^ib^jc^j \mid i, j \geqslant 1 \}$ زبانی نامنظم است.**

با برهان خلف فرض کنید این زبان منظم باشد و طول پمپ برابر p باشد. حال لم جدید را برای رشته $s = ab^pc^p$ (که $z_1 = a,\ z_2 = b^p,\ z_3 = c^p$) اجرا می‌کنیم.

$$z_2 = uvw,\ |v| > 0 \Rightarrow v = b^j \Rightarrow z_1uv^2wz_3 = a^pb^{p+j}c^p$$

که به وضوح عضوی از زبان L نیست. تناقض حاصله نشان می‌دهد این زبان منظم نیست.

3. **زبان‌های مستقل از متن قطعی نسبت به کدام یک از عملگرهای اجتماع (Union)، مکمل (Complement) و اتصال (Concatenation) بسته هستند؟ اثبات کنید. (15 نمره)**

در ابتدا اثبات می‌کنیم این زبان‌ها نسبت به مکمل بسته هستند. سپس با یک مثال نقض نشان می‌دهیم نسبت به اجتماع و اتصال بسته نیستند.

فرض کنید ماشین $DPDA = (Q, \Sigma, \Gamma, \delta, q_0, F)$ را داریم که زبان L را قبول می‌کند. در این صورت ماشین $DPDA' = (Q, \Sigma, \Gamma, \delta, q_0, F' = Q/F)$ مکمل زبان L را قبول می‌کند. زیرا به دلیل قطعی بودن ماشین با

ورودی دادن رشته w ماشین به یک استیت مانند q می‌رسد. حال اگر $q \in F$ باشد داریم $w \in L$ و $w \notin L(DPDA')$ و اگر $q \notin F$ داریم $w \notin L$ و $w \in L(DPDA')$. این نشان می‌دهد که $L(DPDA') = L^c$ و زبان‌های مستقل از متن قطعی نسبت به مکمل بسته هستند.

این زبان‌ها نسبت به اجتماع بسته نیستند زیرا می‌دانیم زبان $\{ a^n b^n \mid n \geqslant 0 \} \cup \{ a^n b^{2n} \mid n \geqslant 0 \}$ مستقل از متن قطعی نیست ولی هر کدام از $\{ a^n b^n \mid n \geqslant 0 \}$ و $\{ a^n b^{2n} \mid n \geqslant 0 \}$ مستقل از متن قطعی هستند.

دو زبان مستقل از متن قطعی c^* و $\{c\, a^n b^n \mid n \geqslant 0 \} \cup \{ a^n b^{2n} \mid n \geqslant 0 \}$ را در نظر بگیرید. زبان اول چون منظم است، مستقل از متن قطعی نیز هست. زبان دوم نیز می‌توان با بررسی حرف اول رشته فهمید در کدام یک از زبان‌های مستقل از متن قطعی $\{ a^n b^n \mid n \geqslant 0 \}$ یا $\{ a^n b^{2n} \mid n \geqslant 0 \}$ باید ادامه داد که باعث می‌شود مستقل از متن قطعی باشد. حال نشان می‌دهیم زبان حاصل از اتصال این دو زبان مستقل از متن قطعی نیست. با برهان خلف فرض کنید $c^*(\{c\, a^n b^n \mid n \geqslant 0 \} \cup \{ a^n b^{2n} \mid n \geqslant 0 \})$ مستقل از متن قطعی باشد. اگر این زبان را با ca^*b^* که زبانی منظم است اشتراک بگیریم به زبان مستقل از متن قطعی $\{c\, a^n b^n \mid n \geqslant 0 \} \cup \{ ca^n b^{2n} \mid n \geqslant 0 \}$ می‌رسیم. چون تمامی رشته‌های این زبان با حرف c شروع می‌شوند می‌توان گفت پس از حذف آن c هم باز زبان مستقل از متن قطعی می‌ماند اما می‌دانیم $\{ a^n b^n \mid n \geqslant 0 \} \cup \{ a^n b^{2n} \mid n \geqslant 0 \}$ مستقل از متن قطعی نیست. تناقل حاصله نشان می‌دهد کلاس زبان‌های مستقل از متن قطعی نسبت به اتصال بسته نیست.

4. **نشان دهید اگر L زبان مستقل از متن قطعی باشد و M یک زبان منظم باشد آنگاه زبان $L \cap M$ یک زبان مستقل از متن قطعی است. (10 نمره)**

فرض کنید $PDA = (Q_1, \Sigma, \Gamma, \delta_1, q_{01}, F_1)$ و $DFA = (Q_2, \Sigma, \delta_2, q_{02}, F_2)$ به ترتیب زبان‌های L و M را قبول کنند. در این صورت ماشین پشته‌ای قطعی

$$PDA' = (Q' = Q_1 \times Q_2,\ \Sigma,\ \delta',\ q_0 = (q_{01}, q_{02}),\ F = F_1 \times F_2)$$

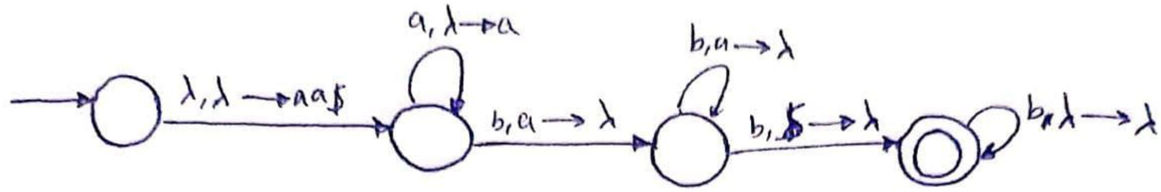
که δ' به صورت زیر تعریف شده است زبان $L \cap M$ را قبول می‌کند.

$$\forall q' = (q_1', q_2') \in Q',\ c \in \Sigma,\ \gamma \in \Gamma,\ \delta_1(q_1', c, \gamma) = (q_1^c, \gamma'):\ \delta'(q', c, \gamma) = ((q_1^c, \delta_2(q_2', c)), \gamma')$$

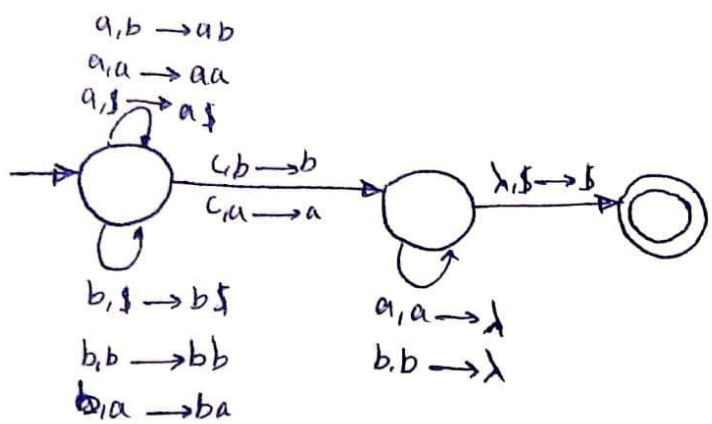
به صورت شهودی با این تعریف $delta'$ ما شبیه سازی اجرای رشته روی DFA و PDA را همزمان انجام می‌دهیم و حالت هر دو ماشین را با یک زوج مرتب نشان می‌دهیم. حال واضح است که رشته در هر دو زبان وجود دارد اگر استیت نهایی عضوی از $F_1 \times F_2$ باشد. پس ماشین ارائه شده زبان $L \cap M$ را می‌پذیرد.

5. برای زبان‌های زیر DPDA رسم کنید. ($\Sigma = \{a, b\}$) (20 نمره)

a. $L = \{ a^n b^m \mid m \geqslant n + 3 \}$



b. $L = \{ wcw^R \mid w \in \{a, b\}^* \}$



6. در خصوص ماشین‌های پشته‌ای دارای قطعیت (DPDA) به پرسش‌های زیر پاسخ دهید: (20 نمره)

a. نشان دهید که برای هر زبان منظم، یک ماشین پشته‌ای قطعی قابل رسم است که رشته‌های آن زبان را می‌پذیرد، تنها دو حالت دارد، هیچ گذار ε ندارد و در آن نمادها هیچ‌گاه از پشته حذف نمی‌شوند.

از آنجا که زبان‌های منظم زیرمجموعه‌ی محض زبان‌های مستقل از متن هستند، برای همه‌ی این زبان‌ها می‌توانیم یک ماشین پشته‌ای طراحی کنیم. در این سوال، هدف این است که نشان دهیم چگونه می‌توانیم عملکرد یک ماشین متنهای را به کمک یک ماشین پشته‌ای شبیه‌سازی کنیم.

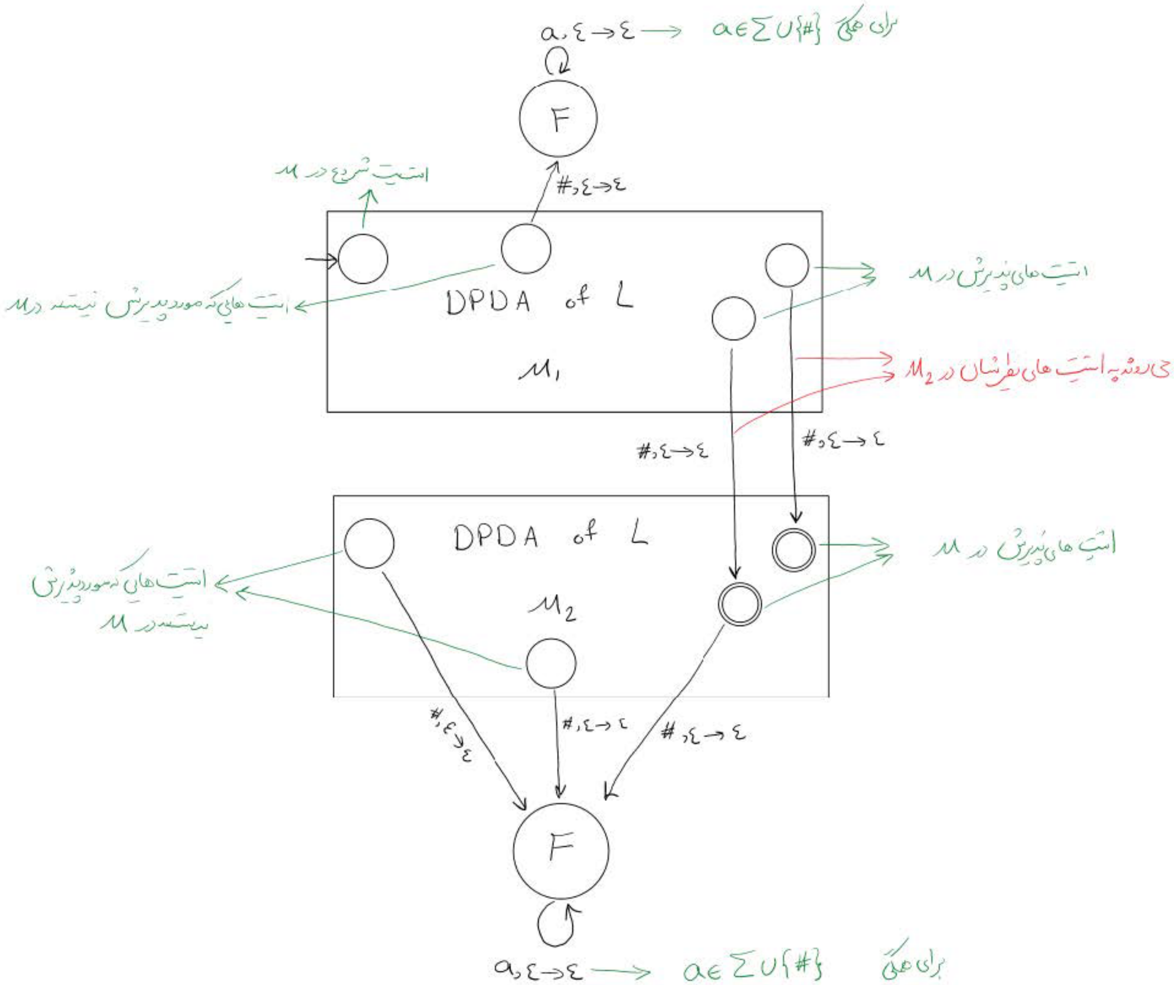
می‌دانیم که در ماشین‌های متناهی، صرفا داشتن نماد ورودی که هد خوانش روی آن قرار گرفته، و نیز حالت فعلی ماشین، برای تعیین حالت بعدی کفایت می‌کند. به همین جهت کافی است که از پشته، به همین منظور استفاده کنیم. یک روش آن است که به ازای هر حالت ماشین متناهی، یک نماد در نظر بگیریم. هر بار که ماشین متناهی به یک حالت جدید تغییر حالت می‌دهد، ماشین پشته‌ای که قرار است عملکرد آن را شبیه‌سازی کند، نماد مربوطه را بالای پشته push می‌کند. حال برای تعیین حالت بعدی، به حالت

قبلی که نماد آن در بالای پشته قرار گرفته دسترسی دارد و به همین جهت، می‌تواند برای نماد بعدی تصمیم‌گیری کند.

ماشین پشته‌ای مربوطه، بر حسب اینکه شیوه‌ی پذیرش در آن چگونه باشد، می‌تواند حداکثر با دو یا سه حالت پیاده‌سازی گردد.

b. نشان دهید که اگر برای L یک ماشین پشته‌ای قطعی وجود داشته باشد که رشته‌های این زبان را بپذیرد، یک ماشین پشته‌ای قطعی دیگر قابل تعریف است که رشته‌های زبان $\{\, x\#y \mid x \in L,\ xy \in L \,\}$ را بپذیرد. فرض می‌کنیم نماد # در هیچ‌یک از رشته‌های L ظاهر نمی‌شوند.

مطابق شکل پایین اگر داشته‌باشیم که M یک DPDA باشد که $L = L(M)$ آنگاه داریم که ما با استفاده از دو تا کپی از M با نام‌های M_1 و M_2 می‌توانیم که زبان خواسته شده را تولید کنیم به این صورت که هرگاه در ماشین M_1 هستیم و ورودی رشته ما # نباشد همان روند عادی M را ادامه می‌دهیم اما هر وقت # بیاید دو حالت دارد، حالت اول این است که ما در یک استیت پذیرش در M قرار داریم (در اصل آن استیتی که ما در M_1 قرار داریم خودش در ماشینی که تولید می‌کنیم پذیرش نیست اما استیت نظیرش در M پذیرش است) در این صورت ما به استیت نظیر آن در M_2 می‌رویم و ادامه می‌دهیم، در غیر این صورت به یک استیت F می‌رویم که در اصل استیتی است که نشان می‌دهد ما دیگر نمی‌توانیم به یک حالت پذیرش برسیم و از آن استیت دیگر خارج نمی‌شویم.
حال اگر در M_2 باشیم نیز اگر رشته ورودی ما # نباشد روال عادی M را طی می‌کنیم تا به یک استیت پذیرش برسیم. در هنگام پایان شدن رشته (استیت‌های پذیرش در M_2 همان استیت‌های پذیرش در M است) اما اگر رشته ورودی ما # باشد در هر حالتی به F می‌رویم.
توضیحاتی مختصر برای درک بیشتر در شکل آمده است.



7. (امتیازی) قضیه‌ی زیر معروف به لم اوگدن را در نظر گرفته، با استفاده از آن اثبات کنید که موارد a و b زبان‌های مستقل از متن نیستند. (20 نمره)

Ogden's Lemma: فرض کنید که L یک زبان مستقل از متن باشد. ثابت n وجود دارد، به نحوی که اگر w یک رشته‌ی دلخواه عضو زبان L باشد، و ما n تا، یا بیشتر، از نمادهای w را انتخاب کنیم، بتوانیم w را به صورت $uvxyz$ بنویسیم، به نحوی که سه شرط مقابل ارضاء شوند:

1. v و y **روی هم دست‌کم یک نماد منتخب داشته باشند**

2. vxy **حداکثر** n **نماد منتخب داشته باشد**

3. به ازای هر $i \geqslant 0$ **رشته‌ی** uv^ixy^iz **عضو زبان** L **باشد.**

a. $L = \{\, a^pb^qc^rd^s \mid p = 0 \ or\ q = r = s \,\}$

فرض کنید این زبان مستقل از متن باشد و طول پمپ را n بنامید. حال لم اوگدن را روی رشته $w = ab^nc^nd^n$ اجرا می‌کنیم. فرض کنید تمام حروف $b,\ c,\ d$ علامت‌گذاری شده باشند. در این

صورت داریم $w = uvxyz$ که تعداد حروف علامت‌دار vxy حداکثر n است و vy حداقل یک حرف علامت‌دار دارد. در این صورت مانند تمام مسائل لم تزریق به راحتی می‌توان ثابت کرد که v و y تنها شامل یک حرف هستند چون اگر حداقل دو حرف متفاوت داشته باشد فرم $a^*b^*c^*d^*$ در uv^2xy^2z از بین می‌رود. حال دقت کنید که تعداد حداکثر دو تا از حرف‌های b, c, d یکی می‌ماند که تناقض است. پس این زبان مستقل از متن نیست.

b. $L = \{ a^n b^n c^i \mid i \neq n \}$

فرض کنید این زبان مستقل از متن باشد و طول پمپ را n در نظر بگیرید. حال لم اوگدن را روی رشته $w = a^{n!+n} b^{n!+n} c^n$ اجرا می‌کنیم که تمام حروف c علامت‌گذاری شده‌اند. پس داریم $w = uvxyz$ که vxy حداکثر n حرف c دارد و vy نیز شامل حداقل یک حرف c هست و برای هر $i \geqslant 0$ رشته uv^ixy^iz نیز در زبان هست. دقیقا مانند قسمت قبلی می‌توان گفت که رشته‌های v و y شامل حداکثر یک حرف متفاوت هستند. در این‌صورت دو حالت ممکن برای آن‌ها وجود دارد:

حالت اول: $v = a^s \; or \; b^s \, , \; y = c^k$

در این حالت اگر به رشته uv^2xy^2z دقت کنید متوجه می‌شوید که تعداد a و b های آن متفاوت است چون نسبت به $uvxyz$ تعداد تنها یکی از این دو حرف تغییر کرده است. پس این رشته در L نیست که تناقض است.

حالت دوم: $v = c^s \, , \; y = c^k$

در این صورت با انتخاب $i = \frac{n!}{s+k}$ داریم $uv^ixy^iz = a^{n!+n} b^{n!+n} c^{n!+n} \in L$ که تناقض است. دقت کنید که انتخاب $i = \frac{n!}{s+k}$ قابل انجام است زیرا طبق فرض $|vxy| \leqslant n$ که نشان می‌دهد $s + k \leqslant n$ و تضمین می‌کند که i عددی طبیعی است.