نظریه زبان و ماشین - دکتر قوامی‌زاده

امیرحسین منصوری - ۹۹۲۴۳۰۶۹ - تمرین سری ۶

سوال ۱ - a)

با لم تزریق و برهان خلف، نامنظم بودن زبان را ثابت می‌کنیم. فرض می‌کنیم برای این زبان، ماشینی با k وضعیت وجود دارد. ثابت می‌شود که به ازای هر عدد طبیعی دلخواه k، می‌توان عدد اول p را پیدا کرد، به طوری که هیچ‌کدام از k عدد متوالی بعد از آن اول نباشد. با انتخاب چنین عدد p ای، رشته زیر را می‌سازیم:

$$z = b^{k}cd^{p-k}e$$

که با توجه به تعریف زبان، عضو این زبان است. بنا به لم تزریق رشته زیر نیز به ازای هر i باید عضو زبان باشد (با فرض $k = s + l + t$):

$$z' = b^{s}b^{i\times l}b^{t}cd^{p-k}e$$

با توجه به این که $1 \le l \le k$، داریم:

$$k = s + l + t$$
$$k + 1 \le s + 2 \times l + t \le 2k$$
$$p + 1 \le (p - k) + (s + 2 \times l + t) \le p + k$$
$$p + 1 \le n_{b}(z') + n_{d}(z') \le p + k$$

بنا به تعریف p در ابتدای اثبات، هیچ‌کدام از اعداد p+1 تا p+k اول نیستند. بنابراین z' در زبان نیست. این تناقض نشان می‌دهد که زبان منظم نیست.

سوال ۱ - b)

فرض می‌کنیم ماشین k وضعیتی برای این زبان وجود دارد. رشته زیر را در نظر می‌گیریم:

$$z = b^{k+2}a^{2}$$

که رشته‌ای از این زبان است. بنابراین بر اساس لم تزریق، اگر $k = s + l + t$، به ازای هر i ای باید رشته

$$z' = b^{k'+2}a^{2}$$

که در آن $k' = s + i(l) + t$ در زبان موجود باشد. داریم:

$$k + 2 \equiv 2 \ (mod\ (k + 2)^{2})$$
$$s + l + t + 2 \equiv 2 \ (mod\ (k + 2)^{2})$$
$$s + l + t \equiv 0 \ (mod\ (k + 2)^{2})$$
$$s + l + t \equiv k^{2} + 4k + 4 \ (mod\ (k + 2)^{2})$$

چون $1 \le l \le k$، به ازای $i = 0$ داریم:

$$s + t \equiv m \ (mod\ (k + 2)^{2})$$

که $k^{2} + 3k + 4 \le m \le k^{2} + 4k + 3$. در نتیجه $m \ne 0$ و $m \ne (k + 2)^{2}$. در نتیجه z' عضو زبان نیست و این تناقض، نامنظم بودن زبان را نشان می‌دهد.

سوال ۱ - c)

فرض می‌کنیم زبان منظم است و ماشینی k وضعیتی دارد. رشته زیر را در نظر می‌گیریم:

$$z = ba^{k}b^{k}$$

که در زبان است. بر اساس لم تزریق می‌توان نوشت $s = uvw = ba^{k-1}$ و در این صورت رشته

$$z' = uv^{i}wab^{k}$$

نیز در رشته است. برای u دو حالت وجود دارد:

۱- تهی باشد: در این صورت با انتخاب $i = 0$، رشته دیگر با b شروع نمی‌شود و در نتیجه z' در زبان نیست.

۲- با b آغاز شود: در این صورت باز با انتخاب $i = 0$، تعداد a ها کم می‌شود و در نتیجه رشته باز هم عضو زبان نیست.

بنابراین در هر صورت به تناقض بر می‌خوریم. در نتیجه زبان نامنظم است.

سوال ۱ - d)

زبان منظم است و عبارت منظم

(a|b)*a(a|b)*a(a|b)*a(a|b)*|(a|b)*b(a|b)*b(a|b)*b(a|b)*

را می‌توان برای آن نوشت.

در واقع این زبان مجموعه‌ی همه رشته‌هایی است که حداقل ۳ حرف یکسان داشته باشد. چرا که کافیست از بین این حداقل ۳ حرف یکسان، ۳ حرف را به دلخواه انتخاب کنیم، حرفی که بین ۲ حرف دیگر می‌آید را به عنوان رشته تک حرفی v در نظر بگیریم، همه کاراکترهای قبل آن را به عنوان رشته u، و همه کاراکترهای بعد آن را به عنوان رشته w در نظر بگیریم. درنهایت، شرط بیان شده در تعریف زبان، برای u و v و w برقرار خواهد بود.

(همچنین با بررسی محدود رشته‌هایی که حداکثر دو حرف یکسان دارند - مثلا baba - مشاهده می‌شود که برای هیچ‌کدام از این رشته‌ها نمی‌توان u و v و w به صورت گفته شده در سوال در نظر گرفت و درنتیجه عضو زبان نیستند)

سوال ۱ - e)

با استفاده از لم تزریق، نامنظم بودن زبان را ثابت می‌کنیم. ابتدا فرض می‌کنیم زبان منظم است و دارای ماشینی k وضعیتی است. رشته زیر را که در زبان است در نظر می‌گیریم:

$$z = a^k \# a^k \# a^k$$

بنا به لم تزریق، اگر $k = s + l + t$، رشته زیر باید به ازای هر i صحیح در زبان باشد:

$$z' = a^{k'} \# a^k \# a^k$$

که در آن $k' = s + i(l) + t$. اما به ازای $i = 0$، چون $l \geq 1$، خواهیم داشت $k' < k$ و در نتیجه z' عضو زبان نیست. این تناقض، نامنظم بودن زبان را نتیجه می‌دهد.

سوال ۱ - f)

با استفاده از لم تزریق، نامنظم بودن زبان را ثابت می‌کنیم. ابتدا فرض می‌کنیم زبان منظم است و دارای ماشینی k وضعیتی است. رشته زیر را که در زبان است در نظر می‌گیریم:

$$z = b^k c^{k!}$$

بنا به لم تزریق، اگر $k = s + l + t$، رشته زیر باید به ازای هر i صحیح در زبان باشد:

$$z' = b^{k'} c^{k!}$$

که در آن $k' = s + i(l) + t$. اما به ازای $i = 0$، چون $l \geq 1$، خواهیم داشت $k' \neq k$ و در نتیجه $(k')! \neq k!$. در نتیجه z' عضو زبان نیست و این تناقض، نامنظم بودن زبان را نتیجه می‌دهد.

سوال ۱ - i)

با استفاده از لم تزریق، نامنظم بودن زبان را ثابت می‌کنیم. ابتدا فرض می‌کنیم زبان منظم است و دارای ماشینی k وضعیتی است. رشته زیر را در نظر می‌گیریم:

$$z = b^k b^{2k+2} a^{3k}$$

که در زبان است. بنا به لم تزریق، اگر $k = s + l + t$، رشته زیر باید به ازای هر i صحیح در زبان باشد:

$$z' = b^{k'} b^{2k+2} a^{3k}$$

که در آن $k' = s + i(l) + t$. اما به ازای $i = 0$، چون $k \geq l \geq 1$، خواهیم داشت:

$$0 \leq k' \leq k - 1$$

$$2k + 2 \leq k' + 2k + 2 \leq 3k + 1$$

$$2k + 2 \leq n_b(z') \leq 3k + 1$$

همچنین می‌دانیم چون $3k + 2 \equiv 2 \ (mod\ 3k)$، بنابراین هیچ عدد طبیعی n که $2 < n < 3k + 2$ در رابطه $n \equiv 2 \ (mod\ 3k)$ صدق نمی‌کند. در نتیجه طبق نامساوی بالا:

$$n_b(z') \not\equiv 2 \ (mod\ n_a(z'))$$

این تناقض، نامنظم بودن L را نتیجه می‌دهد.

سوال ۱ - l)

زبان منظم است. زیرا می‌توان عبارت منظم

$$(a^3)^* a^2((b^3)^* b \mid (b^3)^* b^2)$$

را برای آن نوشت.

درواقع چون شرط $nm^2 \equiv 2\ (mod\ 3)$ وجود دارد، در نتیجه باقی‌مانده تقسیم n و m بر ۳ نباید صفر باشد. همچنین باقی‌مانده n بر ۳ نیز نباید ۱ باشد (با بررسی حالات مختلف باقی‌مانده m بر ۳ می‌توان دید که شرط مذکور برقرار نخواهد شد). در نهایت تنها حالت‌های مجاز n و m به صورت زیر است:

$$n \equiv 2\ (mod\ 3),\ m \equiv 1\ (mod\ 3)$$
$$n \equiv 2\ (mod\ 3),\ m \equiv 2\ (mod\ 3)$$

در نتیجه زبان را می‌توان با عبارت منظم مذکور توصیف کرد.

سوال ۱ - m)

با توجه به توصیف داده شده در سوال، می‌توان گفت هر رشته‌ای در L که با a^n آغاز شود، حداقل 2^n تا b دارد. با توجه به این موضوع و با برهان خلف، نامنظم بودن زبان را اثبات می‌کنیم. فرض می‌کنیم L منظم است و یک FSA با k وضعیت دارد. رشته زیر را در نظر می‌گیریم:

$$z = a^k b^{2^k}$$

که در زبان است. بنا به لم تزریق، اگر $k = s + l + t$، رشته زیر باید به ازای هر i صحیح در زبان باشد:

$$z' = a^{k'} b^{2^k}$$

که در آن $k' = s + i(l) + t$. اما به ازای $i = 2$، چون $l \geq 1$، $k < k'$ و بنابراین $2^k < 2^{k'}$. رشته بالا باید حداقل $2^{k'}$ تا b داشته باشد، اما فقط 2^k تا b دارد. این تناقض، نامنظم بودن L را نتیجه می‌دهد.

سوال ۲ - a)

با استفاده از برهان خلف، نامنظم بودن زبان را اثبات می‌کنیم. فرض می‌کنیم L یک DFA با وضعیت شروع q_0 و وضعیت‌های شناسایی F دارد. رشته $b^i d^{i^2}$ به ازای هر $i = 1, 2, 3, ...$ پذیرفته می‌شود. چون حالت‌های i نامتناهی است و تعداد وضعیت‌های DFA این زبان متناهی است، طبق اصل لانه‌کبوتری، وضعیتی مثل q در این DFA وجود دارد، به طوری که:

$$\delta^*(q_0,\ b^n) = \delta^*(q_0,\ b^m) = q,\ n \neq m$$

حال چون $b^n d^{n^2}$ پذیرفته می‌شود، در نتیجه:

$$\delta^*(q_0,\ b^n d^{n^2}) = \delta^*(\delta^*(q_0,\ a^n),\ d^{n^2}) = q_F \in F$$
$$\Rightarrow \delta^*(\delta^*(q_0,\ a^n),\ d^{n^2}) = \delta^*(\delta^*(q_0,\ a^m),\ d^{n^2}) = q_F \in F$$

که به این معنی است که رشته $a^m b^{n^2}$ نیز باید پذیرفته شود. اما این رشته طبق تعریف سوال، پذیرفته نمی‌شود (چون $m \neq n$). این تناقض، نامنظم بودن زبان را نتیجه می‌دهد.

سوال ۲ - b)

فرض می‌کنیم زبان L منظم است. بنابراین با اعمال تابع همریختی، باید منظم بماند. اما با تابع همریختی زیر:

$$h(a) = a,\ h(b) = a,\ h(c) = c,\ h(d) = c$$

زبان به صورت زیر در می‌آید:

$$\{a^{m+n} c^{m+n} \mid n, m \in N\}$$

که می‌دانیم منظم نیست. این تناقض نامنظم بودن زبان L را نتیجه می‌دهد.

سوال ۳ - a)
می‌دانیم زبان

$$L_U = \{a^* b^n c^m d^* \mid n \in N\}$$

منظم است. زیرا عبارت منظم آن برابر $a^* b^+ c^* d^*$ است. همچنین تعریف می‌کنیم:

$$L_C = \{a^* b^n c^m d^* \mid n \in N \wedge m = n!\}$$

حال داریم:

$$L_C = L_U - L$$

اگر L منظم باشد، با توجه به خواص بستاری زبان‌های منظم، L_C نیز باید منظم باشد. از سوال ۱ - f می‌دانیم که L_C نامنظم است. در نتیجه L نیز منظم نیست.

سوال ۳ - b)
اگر زبان L منظم باشد، پس باید با اعمال تابع همریختی نیز همچنان منظم باشد. با اعمال تابع همریختی زیر:

$$h(a) = a$$
$$h(b) = a$$
$$h(c) = b$$
$$h(d) = c$$
$$h(e) = c$$
$$h(f) = d$$

زبان به شکل زیر در می‌آید:

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

اما از سوال ۲-b می‌دانیم که زبان بالا منظم نیست. این تناقض، نامنظم بودن L را نتیجه می‌دهد.

سوال ۴ - الف)
می‌دانیم اگر زبان منظم باشد، عدد p ای وجود دارد که که به ازای آن:

$$s = ab^p c^p$$

را می‌توان به صورت $s = uxyzv$ نوشت و در این صورت رشته:

$$s' = uxy^i zv$$

نیز در زبان است. اما اگر در مثال بالا فرض کنیم

$$u = a$$
$$xyz = b^p$$
$$v = c^p$$

به ازای هر $i \neq 1$، s' دیگر در زبان نیست. این تناقض نامنظم بودن زبان را نشان می‌دهد. در نتیجه مثال داده شده درست است.

سوال ۴ - ب)
تفاوت تعمیم لم تزریق و لم تزریق معمولی این است که تعمیم لم تزریق به ما اجازه می‌دهد در هرجای رشته که خواستیم، رشته دیگری را pump کنیم. اما در لم تزریق معمولی، رشته pump شده حتما باید در p حرف اول رشته اصلی باشد. بنابراین تعمیم لم تزریق به ما آزادی بیشتری در اثبات نامنظم بودن بعضی از زبان‌ها می‌دهد.