



## مسئله‌ی ۱.

با توجه به گرامرهای داده شده، عبارت زیر را توکن بندی (tokenize) کنید.

**acccbabbbaacacaabaabbabc**

$$Q = (a + b)c^*$$

$$R = a^*b^+$$

$$S = a(ca)^*bc^*$$

پاسخ:

در توکن کردن یک رشته باید توجه داشته باشیم که با توجه به گرامرها، بتوانیم بزرگترین زیررشته هایی را انتخاب کنیم که پس از آن هم بتوانیم ادامه رشته را به زیررشته هایی که با گرامرها مطابقت داشته باشند، تقسیم کنیم. بنابراین توکن بندی عبارت داده شده به صورت زیر خواهد بود:

$$accc(Q) - b(Q) - abbb(R) - ac(Q) - ac(Q) - aab(R) - aabb(R) - abc(S)$$

## مسئله‌ی ۲.

با توجه به زبان  $L$  بر روی الفبای  $\Sigma = \{0, 1\}$  که با شرایط زیر تعریف شده است: هر رشته باید با ۱ شروع شود و پس از آن هر ترکیبی از ۰ها و ۱ها را دنبال کند. رشته باید حداقل یک زیررشته به فرم "۰۱۰" داشته باشد که بخشی از هیچ نمونه تداخلی "۰۱۰۱۰" نباشد.

الف. یک عبارت منظم را بنویسید که تمام رشته‌های درون این زبان را شامل شود.

ب. تأثیر گسترش الفبای  $\Sigma = \{0, 1\}$  برای شامل شدن یک کاراکتر سوم، مثلاً '۲'، بر عبارت منظم و زبان  $L$  تعریف شده را بررسی کنید. اضافه شدن '۲' چگونه بر ساختار رشته‌های درون  $L$  و خود عبارت منظم تأثیر می‌گذارد؟

پاسخ:

## الف

با توجه به اینکه رشته باید شامل ۰۱۰ باشد اما بخشی از ۰۱۰۱۰ نباشد یعنی قبل آن ۰۱ نیامده باشد (۱۱، ۰۰ یا ۱۰) آمده باشد و بعد آن هم ۱۰ نیامده باشد (۰۰، ۱۱ یا ۰۱) آمده باشد. همچنین باید حالات خاص که ۰۱۰ در اوایل یا اواخر رشته آمده نیز در نظر بگیریم. در نهایت داریم:

$$\begin{aligned} & 1(0+1)^*(00+11+10)010(11+00+01)(0+1)^* \\ & + 1(0+1)^*(00+11+10)010(0+1)? \\ & + 1(0+1)?010(11+00+01)(0+1)^* \end{aligned}$$

## ب

اگر فرض کنیم که با اضافه شدن حرف ۲ به مجموعه زبان ما زبان  $L$  میتواند شامل ۰ و ۱ و ۲ باشد به صورتی که شرط‌های قبلی را هم داشته باشد، عبارت منظم  $L$  به صورت زیر در می‌آید:

$$\begin{aligned} & 1(0+1+2)^*(00+11+10+2)010(11+00+01+2)(0+1+2)^* \\ & + 1(0+1+2)^*(00+11+10+2)010(0+1+2)? \\ & + 1(0+1+2)?010(11+00+01+2)(0+1+2)^* \end{aligned}$$

### مسئله ۳.

به سوالات زیر پاسخ بدهید:

۱. عبارت منظم توصیف کننده زبان های زیر را بنویسید.

(آ)  $L = \{ w \in \{a, b\}^* \mid \text{رشته } w \text{ با } a \text{ تمام بشود و شامل رشته } bb \text{ نباشد} \}$

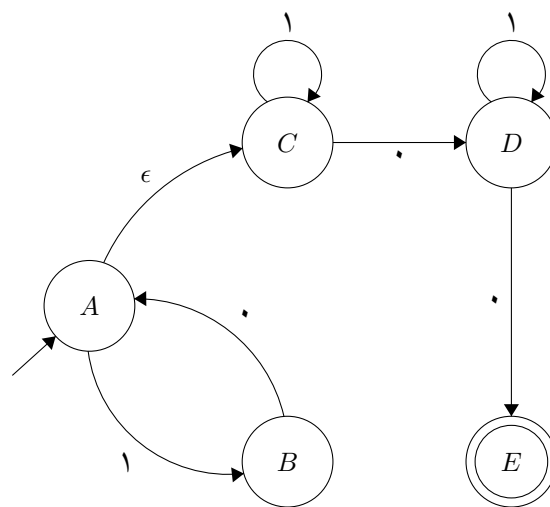
(ب)  $L = \{ w \in \{a, b\}^* \mid \text{رشته } w \text{ حداقل دو } b \text{ و یا دقیقاً دو } a \text{ داشته باشد} \}$

۲. برای عبارات منظم زیر،  $DFA$  رسم نمایید.

(آ)  $(a^*b^* + ba)^*$

(ب)  $011^* + 10(1 + 01)^*$

۳. (آ) ماشین  $NFA$  زیر را به  $DFA$  تبدیل نمایید.



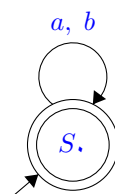
(ب) ماشین بالا چه زبان منظمی را تشخیص می دهد؟

پاسخ:

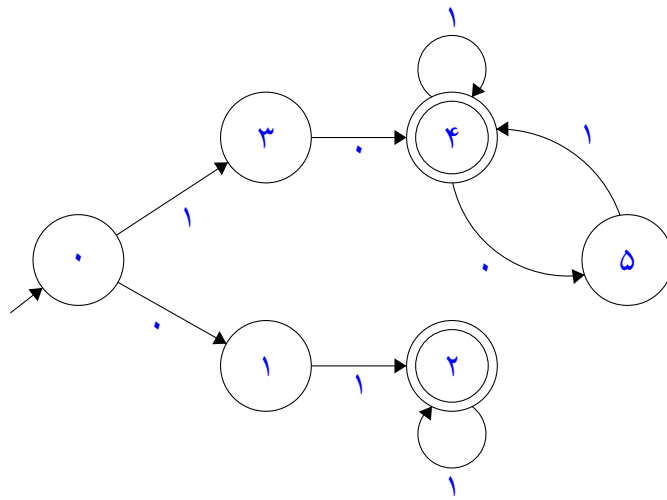
۱. (آ)  $(a + ba)^*$

(ب)  $b^*ab^*ab^* + (a + b)^*b(a + b)^*b(a + b)^*$

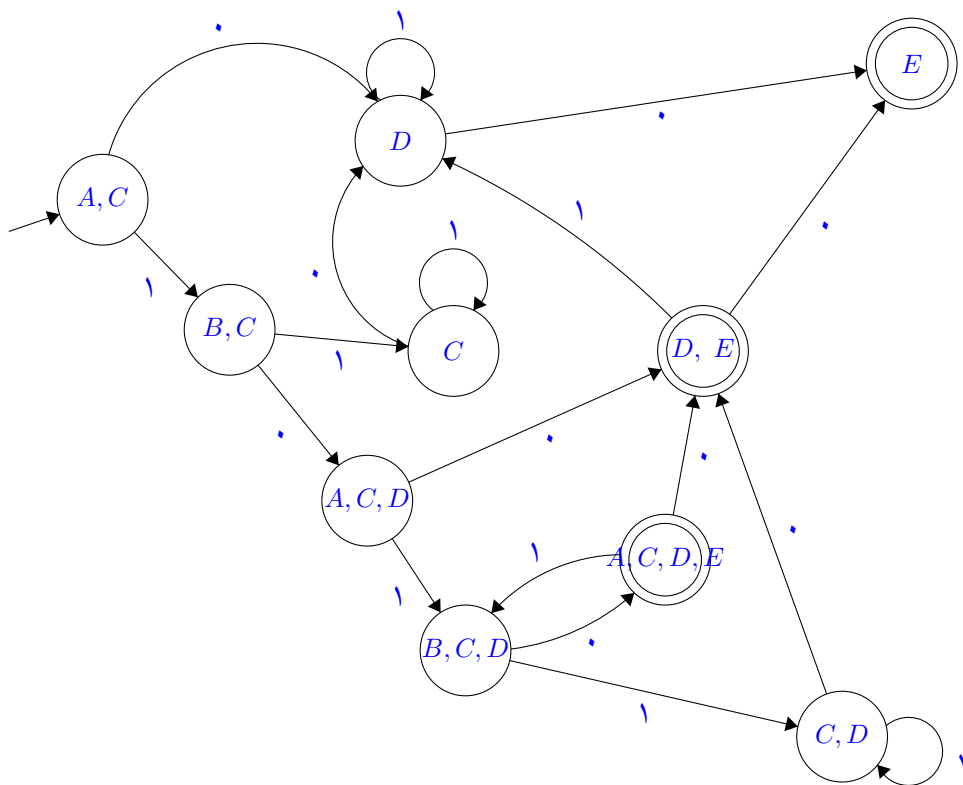
۲. (آ) دقت کنید ماشین رسم شده باید رشته  $\epsilon$  را نیز بپذیرد.



(ب) داریم:



۳. آ ماشین  $DFA$  معادل به شکل زیر می باشد:



ب)  $(10)^*1*01*0$

#### مسئله ۴.

با استفاده از گرامر مستقل از متن زیر، برای عبارت  $[a!b]@[a]$  یک left most derivation و یک right most derivation بنویسید و درخت parse مربوط به آن را رسم کنید.

$$E \rightarrow UV \mid EBE \mid V \mid [E]$$

$$V \rightarrow a \mid b$$

$$U \rightarrow < \mid >$$

$$B \rightarrow ? \mid ! \mid @$$

پاسخ:

ابتدا یک left most derivation برای عبارت می‌نویسیم، داریم:

$$\begin{aligned} E &\rightarrow EBE \rightarrow [E]BE \rightarrow [EBE]BE \rightarrow [VBE]BE \rightarrow [aBE]BE \rightarrow [a!E]BE \\ &\rightarrow [a!V]BE \rightarrow [a!b]BE \rightarrow [a!b]@E \rightarrow [a!b]@[E] \rightarrow [a!b]@[V] \rightarrow [a!b]@[a] \end{aligned}$$

در ادامه یک right most derivation برای عبارت می‌نویسیم، داریم:

$$\begin{aligned} E &\rightarrow EBE \rightarrow EB[E] \rightarrow EB[V] \rightarrow EB[a] \rightarrow E@[a] \rightarrow [E]@[a] \rightarrow [EBE]@[a] \\ &\rightarrow [EBV]@[a] \rightarrow [EBb]@[a] \rightarrow [E!b]@[a] \rightarrow [V!b]@[a] \rightarrow [a!b]@[a] \end{aligned}$$

درخت parse مربوط به هر دو derivation بالا در ادامه آمده است. (درخت هایشان یکسان است)

