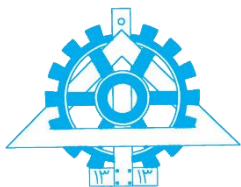


به نام خدا



نظریه زبان‌ها و ماشین‌ها - پاییز ۱۴۰۱

تمرین شماره ۵

دستیار آموزشی این مجموعه: محمد عظیم‌پور

azimpour102@ut.ac.ir



تاریخ تحویل: ۲۵ آبان

(۱) گرامر زیر را در نظر بگیرید: (۲۰ نمره)

$$S \rightarrow aBa|bDb|cAc|A|B$$

$$A \rightarrow aA|B|\varepsilon$$

$$B \rightarrow bB|D$$

$$C \rightarrow cC|\varepsilon$$

$$D \rightarrow cD|B$$

الف) قواعد اِپسیلون را از گرامر حذف کنید. (۵ نمره)

پاسخ:

$$C \rightarrow \varepsilon:$$

$$S \rightarrow aBa|bDb|cAc|A|B$$

$$A \rightarrow aA|B|\varepsilon$$

$$B \rightarrow bB|D$$

$$C \rightarrow cC|c$$

$$D \rightarrow cD|B$$

$$A \rightarrow \varepsilon:$$

$$S \rightarrow aBa|bDb|cAc|A|B|cc|\varepsilon$$

$$A \rightarrow aA|B|a$$

$$B \rightarrow bB|D$$

$$C \rightarrow cC|c$$

$$D \rightarrow cD|B$$

$$S \rightarrow \varepsilon:$$

$$S' \rightarrow S|\varepsilon$$

$$S \rightarrow aBa|bDb|cAc|A|B|cc$$

$$A \rightarrow aA|B|a$$

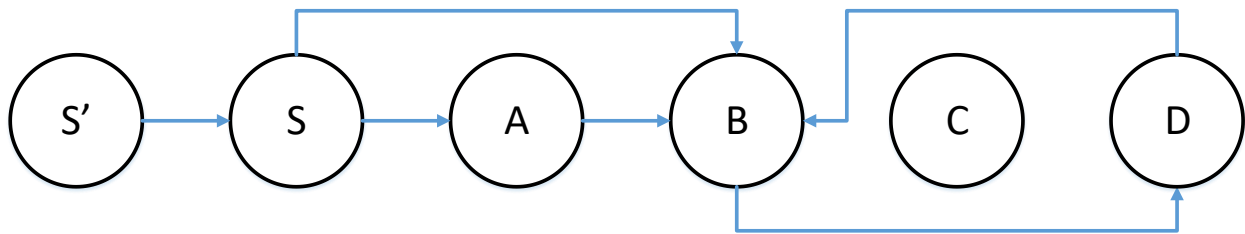
$$B \rightarrow bB|D$$

$$C \rightarrow cC|c$$

$$D \rightarrow cD|B$$

ب) گراف قواعد یک‌ه را رسم کنید. (۵ نمره)

پاسخ:



پ) قواعد یک‌ه را از گرامر حذف کنید. (۵ نمره)

پاسخ:

ابتدا حلقه $B \rightarrow D \rightarrow B$ را از بین می‌بریم:

$$S' \rightarrow S|\varepsilon$$

$$S \rightarrow aBa|bBb|cAc|A|B|cc$$

$$A \rightarrow aA|B|a$$

$$B \rightarrow bB|cB$$

$$C \rightarrow cC|c$$

سپس باقی قواعد یک‌ه را حذف می‌کنیم. گرامر نهایی به شکل زیر می‌شود:

$$S' \rightarrow aBa|bBb|cAc|aA|a|bB|cB|cc|\varepsilon$$

$$S \rightarrow aBa|bBb|cAc|aA|a|bB|cB|cc$$

$$A \rightarrow aA|bB|cB|a$$

$$B \rightarrow bB|cB$$

$$C \rightarrow cC|c$$

ت) قواعد بی کاربرد را از گرامر حذف کنید. (۵ نمره)

پاسخ:

از آن جا که متغیرهای S و C پس از تغییرات ایجاد شده در سمت راست هیچ رابطه‌ای قرار ندارد، قواعد مربوط به آن‌ها بی کاربرد خواهند بود. پس از حذف آن‌ها گرامر به صورت زیر درمی‌آید:

$$S' \rightarrow aBa|bBb|cAc|aA|a|bB|cB|cc|\varepsilon$$

$$A \rightarrow aA|bB|cB|a$$

$$B \rightarrow bB|cB$$

۲) گرامر زیر را به فرم نرمال چامسکی تبدیل کنید. (۱۵ نمره)

$$S \rightarrow aAabBb|cCcdDd$$

$$A \rightarrow aBa|aCa|\varepsilon$$

$$B \rightarrow bCb$$

$$C \rightarrow cDc|c$$

$$D \rightarrow dBd$$

پاسخ:

ابتدا باید قواعد اپسیلون و قواعد یک‌ه را از گرامر حذف کنیم. قواعد اپسیلون به صورت زیر حذف می‌شوند: (۵ نمره)

$$A \rightarrow \varepsilon:$$

$$S \rightarrow aAabBb|aabBb|cCcdDd$$

$$A \rightarrow aBa|aCa$$

$$B \rightarrow bCb$$

$$C \rightarrow cDc|c$$

$$D \rightarrow dBd$$

در گرامر حاصل قواعد یک‌ه وجود ندارد، در نتیجه می‌توانیم آن را به فرم نرمال چامسکی تبدیل کنیم. ابتدا متغیرهایی تعریف می‌کنیم که مستقیماً به پایانه‌ها تبدیل شوند:

$$A' \rightarrow a, B' \rightarrow b, C' \rightarrow c, D' \rightarrow d$$

سپس قواعدی از گرامر که طول طرف راستشان بزرگتر از ۲ است را با استفاده از این متغیرها بازنویسی می‌کنیم: (۵ نمره)

$$S \rightarrow A'AA'B'BB'|A'A'B'BB'|C'CC'D'DD'$$

$$A \rightarrow A'BA'|A'CA'$$

$$B \rightarrow B'CB'$$

$$C \rightarrow C'DC'|c$$

$$D \rightarrow D'BD'$$

$$A' \rightarrow a, B' \rightarrow b, C' \rightarrow c, D' \rightarrow d$$

اکنون باید با تعریف متغیرها و قواعد جدید و بازنویسی گرامر به کمک آن‌ها، قواعدی که طرف راستشان طول بیشتر از ۲ متغیر دارد از بین ببریم: (۵ نمره)

$$S \rightarrow A'E|A'E'|C'E''$$

$$E \rightarrow AE', E' \rightarrow A'F, E'' \rightarrow CF'$$

$$F \rightarrow B'G, F' \rightarrow C'G'$$

$$G \rightarrow BB', G' \rightarrow D'H$$

$$H \rightarrow DD'$$

$$A \rightarrow A'I|A'I'$$

$$I \rightarrow BA', I' \rightarrow CA'$$

$$B \rightarrow B'J$$

$$J \rightarrow CB'$$

$$C \rightarrow C'K|c$$

$$K \rightarrow DC'$$

$$D \rightarrow D'L$$

$$L \rightarrow BD'$$

$$A' \rightarrow a, B' \rightarrow b, C' \rightarrow c, D' \rightarrow d$$

۳) با استفاده از الگوریتم CYK، نشان دهید که رشته "aabaabaa" توسط زبان زیر پذیرفته می‌شود یا خیر. (۱۵ نمره)

$$S \rightarrow aAa|bBb$$

$$A \rightarrow aBa|b$$

$$B \rightarrow bAb|a$$

پاسخ:

ابتدا باید گرامر را به فرم نرمال چامسکی تبدیل کرد. چون قواعد اسیلون و یکه وجود ندارند نیاز نیست تا مراحل حذف آنها را انجام دهیم. پس از تبدیل گرامر به فرم نرمال چامسکی، به گرامر زیر می‌رسیم: (۵ نمره)

$$S \rightarrow A'C|B'C'$$

$$A \rightarrow A'D|b$$

$$B \rightarrow B'D'|a$$

$$A' \rightarrow a, B' \rightarrow b$$

$$C \rightarrow AA', C' \rightarrow BB'$$

$$D \rightarrow BA', D' \rightarrow AB'$$

حال می‌توان الگوریتم CYK را روی ورودی مورد نظر اجرا کرد: (۱۰ نمره)

S								
	C							
	A							
		D						
		B						
			D'					
	S		A		S			
D	C'	C	D	D	C'	C	D	
A',B	A',B	A,B'	A',B	A,B'	A',B	A,B'	A',B	A',B
a	a	b	a	a	a	b	a	a

با توجه به جدول، رشته $aabaaabaa$ توسط زبان تعریف شده پذیرفته می‌شود.

۴) چپ‌گردی موجود در گرامر زیر را به راست‌گردی تبدیل کنید. (۱۰ نمره)

$$S \rightarrow Aa|Bb|c$$

$$A \rightarrow aA|SA|a$$

$$B \rightarrow bB|SB|b$$

پاسخ:

ابتدا نیاز است تا چپ‌گردی‌های غیرآنی را از بین ببریم. مثلاً قواعد $S \rightarrow Aa$ و $A \rightarrow SA$ باعث ایجاد این چنین چپ‌گردی‌هایی می‌شوند: (۵ نمره)

$$S \rightarrow SAa|SBb|c$$

$$A \rightarrow aA|aSA|a$$

$$B \rightarrow bB|bSB|b$$

پس از آن می‌توانیم چپ‌گردی را به راست‌گردی تبدیل کنیم: (۵ نمره)

$$S \rightarrow cS'|c$$

$$S' \rightarrow AaS'|BbS'|Aa|Bb$$

$$A \rightarrow aA|aSA|a$$

$$B \rightarrow bB|bSB|b$$

(۵) گرامر زیر را به فرم نرمال گریباخ تبدیل کنید. (۲۰ نمره)

$$S \rightarrow SS|A$$

$$A \rightarrow Ba$$

$$B \rightarrow bB$$

پاسخ:

ابتدا باید چپ‌گردی موجود در گرامر را از بین ببریم: (۵ نمره)

$$S \rightarrow AS'|A$$

$$S' \rightarrow SS'|S$$

$$A \rightarrow Ba$$

$$B \rightarrow bB$$

سپس باید قواعدی که در آن‌ها متغیری در ابتدای طرف راست وجود دارد به قواعدی تبدیل کنیم که با یک پایانه آغاز می‌شوند:

(۱۰ نمره)

$$A \rightarrow Ba:$$

$$S \rightarrow AS'|A$$

$$S' \rightarrow SS'|S$$

$$A \rightarrow bBa$$

$$B \rightarrow bB$$

$$S \rightarrow AS' \text{ و } S \rightarrow A:$$

$$S \rightarrow bBaS'|bBa$$

$$S' \rightarrow SS'|S$$

$$A \rightarrow bBa$$

$$B \rightarrow bB$$

$$S' \rightarrow SS' \text{ و } S' \rightarrow S:$$

$$S \rightarrow bBaS'|bBa$$

$$S' \rightarrow bBaS'S'|bBaS'|bBaS'|bBa$$

$$A \rightarrow bBa$$

$$B \rightarrow bB$$

در انتها نیاز است تا قواعدی که طول طرف راست آن‌ها بیشتر از ۲ است و به غیر از خانه اول در میان آن‌ها همچنان پایانه مشاهده می‌شود حذف کنیم. برای این کار باید متغیرهای جدیدی تعریف کنیم که بتوانند به این پایانه‌ها تبدیل شوند و گرامر را با استفاده از این متغیرها بازنویسی کنیم: (۵ نمره)

$$S \rightarrow bBA'S'|bBA'$$

$$S' \rightarrow bBA'S'S'|bBA'S'|bBA'S'|bBA'$$

$$A \rightarrow bBA'$$

$$B \rightarrow bB$$

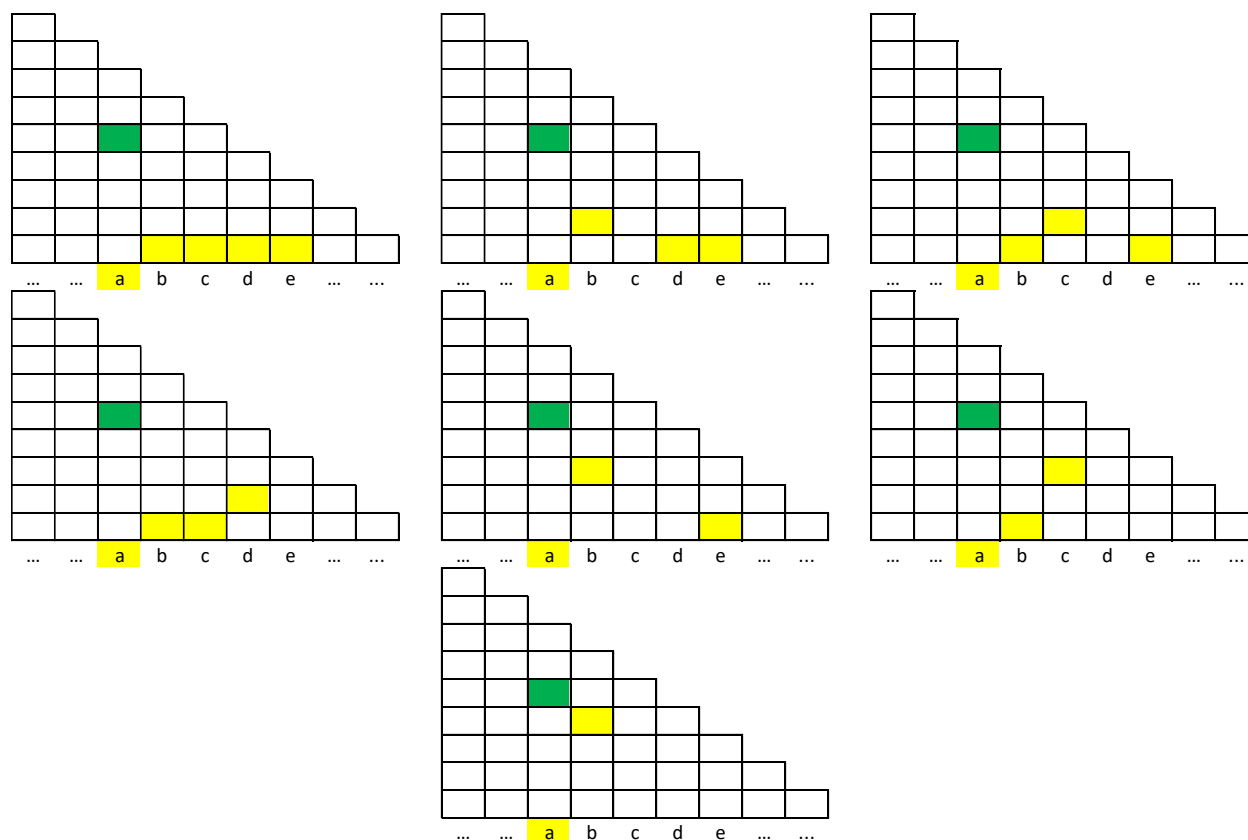
$$A' \rightarrow a$$

۶) الگوریتم CYK را به گونه‌ای تغییر دهید که برای گرامرهای به فرم نرمال گریباخ جوابگو باشد. در قیاس با حالت اصلی الگوریتم CYK، آیا استفاده از این الگوریتم جدید از نظر پیچیدگی زمانی به صرفه است؟ (۲۰ نمره)

پاسخ:

قواعد در فرم نرمال گریباخ به صورت $A \rightarrow \alpha A' A'' A''' \dots$ هستند که در آن α یک پایانه است و باقی حروف نشان دهنده متغیرها هستند. اگر بخواهیم الگوریتم CYK را برای گرامرهای به این فرم اجرا کنیم، هنگام پر کردن هر خانه از جدول باید تمام توالی‌های متغیرها که می‌تواند منجر به تولید زیررشته‌ای از رشته مورد نظر بشود را بررسی کنیم (بر خلاف الگوریتم CYK برای فرم نرمال چامسکی که صرفاً توالی‌های ممکن با طول ۲ مورد بررسی قرار می‌گیرند). همچنین، برای هر پر کردن خانه، قواعدی که به

دنبال آن‌ها می‌گردیم باید حتماً با اولین پایانه‌ی زیررشته‌ی متناظر آن خانه آغاز شوند. به عنوان مثال، برای خانه‌ای که در جدول به رنگ سبز مشخص شده، باید این مجموعه خانه‌ها بررسی شوند: (۱۰ نمره)



در الگوریتم CYK که برای گرامرهای به فرم نرمال چامسکی اجرا می‌شود، برای هر خانه در سطر m جدول، $m-1$ جفت خانه در سطر پایین باید بررسی شوند. در نتیجه پیچیدگی زمانی اجرای این حالت از الگوریتم به شکل زیر محاسبه می‌شود:

$$T = n + \sum_{m=2}^n m - 1 = 1 + \sum_{m=2}^n m = 1 + \frac{(n+2)(n-1)}{2} = \frac{n^2 + n}{2}$$

در الگوریتم جدید، به ازای هر خانه در سطر m جدول، یک خانه در سطر $m-1$ ، دو خانه در سطر $m-2$ ، ...، $m-2$ خانه در سطر دوم و یک حالت برای وقتی که تمام متغیرها از سطر اول انتخاب شوند باید بررسی شوند. پس نتیجه پیچیدگی زمانی اجرای این حالت از الگوریتم به شکل زیر قابل محاسبه است:

$$T = n + 1 + \sum_{m=2}^n (n-m)(m-1)$$

به وضوح پیچیدگی زمانی الگوریتم برای فرم نرمال گریباخ بیشتر از فرم نرمال چامسکی است. در نتیجه استفاده از این الگوریتم جدید از نظر پیچیدگی زمانی به صرفه نیست. (۱۰ نمره)

۷) (امتیازی) فرض کنید $G = (V, \Sigma, R, S)$ گرامری به فرم نرمال گریباخ باشد که در آن هیچ متغیری وجود ندارد که صرفاً در یک قانون شرکت کند و تبدیل به پایانه شود (مثل $A' \rightarrow \alpha$). اگر بخواهیم این گرامر را به فرم نرمال چامسکی تبدیل کنیم، تعداد قواعد مورد نیاز آن حداکثر چقدر بیشتر از تعداد قواعد G خواهد بود؟ (فرض کنید تابعی به نام L وجود دارد که ورودی آن یک قاعده و خروجی آن طول طرف راست قاعده ورودی است) (۱۰ نمره)

پاسخ:

قواعد در فرم نرمال گریباخ به صورت $A \rightarrow \alpha A' A'' A''' \dots$ هستند که در آن α یک پایانه است و باقی حروف نشان دهنده متغیرها هستند. از آنجا که در فرم نرمال چامسکی طرف راست قواعد نباید با پایانه ها شروع شود و طبق صورت سوال متغیرهایی نداریم که صرفاً قابلیت تبدیل به پایانه ها را داشته باشند، نیاز است تا به تعداد اعضای مجموعه الفبای G متغیر جدید اضافه کنیم که هرکدام صرفاً در یک قاعده شرکت دارند و در این قاعده مستقیماً به یک پایانه تبدیل می شوند.

پس از اضافه کردن این متغیرها و بازنویسی قواعد موجود در گرامر با استفاده از آن ها، باید قواعدی که در آن ها طول طرف راست بیشتر از ۲ است را از بین ببریم. برای این کار، باید به صورت متواتر متغیرهای جدیدی تعریف کنیم که هرکدام یکی از متغیرهای قاعده اصلی را در طرف راست تک قاعده ای که در آن شرکت می کند ظاهر کند. به عنوان مثال برای رابطه $A_0 \rightarrow A_1 A_2 \dots A_m$ با طول m ، متغیرهایی به این صورت ایجاد خواهد شد:

$$A_0 \rightarrow A_1 B_0$$

$$B_0 \rightarrow A_2 B_1$$

$$B_1 \rightarrow A_3 B_2$$

....

$$B_{m-4} \rightarrow A_{m-2} B_{m-3}$$

$$B_{m-3} \rightarrow A_{m-1} A_m$$

بنابراین طبق این فرآیند، به ازای یک قاعده با طول m ، $m - 2$ قاعده جدید ایجاد می شود. در نهایت، تعداد قواعد اضافه شده به این صورت قابل محاسبه است:

$$|R'| - |R| = |\Sigma| + \sum_{r \in R} (L(r) - 2) = |\Sigma| - 2|R| + \sum_{r \in R} L(r)$$