

نظریه زبان‌ها و ماشین‌ها - دکتر قوامی‌زاده

امیرحسین منصوری - ۹۹۲۴۳۰۶۹ - تمرین سری ۹

سوال ۱

ابتدا گرامر را ساده می‌کنیم. برای حذف قوانین تهی ابتدا متغیرهای میرا را پیدا می‌کنیم. به دلیل وجود قوانین $A \rightarrow \varepsilon$ و $B \rightarrow \varepsilon$ ، متغیرهای A و B میرا هستند. حال در همه قوانین که دارای A و B هستند، آن‌ها را تک تک با تهی جایگزین می‌کنیم (البته در قانون $B \rightarrow A$ ، جایگزینی A با تهی باعث تولید قانون تهی می‌شود، برای همین از این جایگزینی صرف نظر می‌کنیم):

$$S \rightarrow abAB \mid abB \mid abA \mid ab$$

$$A \rightarrow bAB \mid bB \mid bA \mid b$$

$$B \rightarrow A \mid BAa \mid Aa \mid Ba \mid a$$

برای حذف قوانین یکه، همه جفت متغیرهای u و v که $U \Rightarrow V$ را پیدا می‌کنیم. تنها جفت متغیر ممکن با این شرط، $B \Rightarrow A$ است. بنابراین $B \rightarrow A$ را با معادل آن جایگزین می‌کنیم:

$$S \rightarrow abAB \mid abB \mid abA \mid ab$$

$$A \rightarrow bAB \mid bB \mid bA \mid b$$

$$B \rightarrow BAa \mid Aa \mid Ba \mid a \mid bAB \mid bB \mid bA \mid b$$

در گرامر بالا، قانون بی‌فایده وجود ندارد. به عنوان مثال در اشتقاق زیر:

$$S \rightarrow abAB \rightarrow abbbB \rightarrow abbbb$$

همه قوانین استفاده شده‌اند.

حال گرامر ساده شده را به فرم نرمال چامسکی تبدیل می‌کنیم. ابتدا دو قانون زیر را تعریف می‌کنیم:

$$X_a \rightarrow a$$

$$X_b \rightarrow b$$

و در همه قوانین با بیشتر از یک پایانه یا ناپایانه در سمت راست، همه ناپایانه‌ها را با این دو قانون عوض می‌کنیم:

$$S \rightarrow X_a X_b AB \mid X_a X_b B \mid X_a X_b A \mid X_a X_b$$

$$A \rightarrow X_b AB \mid X_b B \mid X_b A \mid b$$

$$B \rightarrow BAX_a \mid AX_a \mid BX_a \mid a \mid X_b AB \mid X_b B \mid X_b A \mid b$$

حال در قوانین با بیشتر از ۲ ناپایانه در سمت راست، آن‌ها را به چند قانون با ۲ ناپایانه در سمت راست تبدیل می‌کنیم تا به فرم نرمال چامسکی برسیم:

$$S \rightarrow X_a D_1 \mid X_a D_3 \mid X_a D_4 \mid X_a X_b$$

$$D_1 \rightarrow X_b D_2$$

$$D_2 \rightarrow AB$$

$$D_3 \rightarrow X_b B$$

$$D_4 \rightarrow X_b A$$

$$A \rightarrow X_b D_2 \mid X_b B \mid X_b A \mid b$$

$$B \rightarrow BD_5 \mid AX_a \mid BX_a \mid a \mid X_b D_2 \mid X_b B \mid X_b A \mid b$$

$$D_5 \rightarrow AX_a$$

$$X_a \rightarrow a$$

$$X_b \rightarrow b$$

سوال ۲

خوشبختانه، گرامر داده شده دارای قوانین تهی و یکه و بی‌فایده نیست. برای تبدیل به نرمال گریباخ، ابتدا C را در $S \rightarrow CA$ جایگزین می‌کنیم. عملاً C نیز بی‌فایده می‌شود و می‌توانیم آن را حذف کنیم:

$$S \rightarrow bA \mid BB$$

$$A \rightarrow a$$

$$B \rightarrow b \mid SB$$

سپس در قاعده $S \rightarrow BB$ ، اولین B را جایگزین می‌کنیم:

$$S \rightarrow bA \mid bB \mid SBB$$

$$A \rightarrow a$$

$$B \rightarrow b \mid SB$$

سپس قاعده $S \rightarrow SBB$ که بازگشتی چپ است را به بازگشتی راست تبدیل می‌کنیم:

$$S \rightarrow bA \mid bB \mid bAS' \mid bBS'$$

$$S' \rightarrow BBS' \mid BB$$

$$A \rightarrow a$$

$$B \rightarrow b \mid SB$$

سپس در قاعده $B \rightarrow SB$ ، S را جایگزین می‌کنیم:

$$S \rightarrow bA \mid bB \mid bAS' \mid bBS'$$

$$S' \rightarrow BBS' \mid BB$$

$$A \rightarrow a$$

$$B \rightarrow b \mid bAB \mid bBB \mid bAS'B \mid bBS'B$$

در نهایت در قاعده‌های $S' \rightarrow BBS' \mid BB$ ، اولین B را جایگزین می‌کنیم تا نرمال گریباخ به دست آید:

$$S \rightarrow bA \mid bB \mid bAS' \mid bBS'$$

$$S' \rightarrow bBS' \mid bABBS' \mid bBBBS' \mid bAS'BBS' \mid bBS'BBS' \mid bB \mid bABB \mid bBBB \mid bAS'BB \mid bBS'BB$$

$$A \rightarrow a$$

$$B \rightarrow b \mid bAB \mid bBB \mid bAS'B \mid bBS'B$$

سوال ۳

$$S \rightarrow A \mid B$$

$$A \rightarrow 0A1 \mid 01$$

$$B \rightarrow 0B11 \mid 011$$

برای تبدیل به فرم نرمال چامسکی، ابتدا قواعد یکه را حذف می‌کنیم. قواعد یکه به صورت $S \Rightarrow A$ و $S \Rightarrow B$ هستند. با جایگزینی آن‌ها داریم:

$$S \rightarrow 0A1 \mid 01 \mid 0B11 \mid 011$$

$$A \rightarrow 0A1 \mid 01$$

$$B \rightarrow 0B11 \mid 011$$

با اضافه کردن دو قاعده $X_0 \rightarrow 0$ و $X_1 \rightarrow 1$ گرامر را تبدیل به فرم نرمال چامسکی می‌کنیم:

$$S \rightarrow X_0 D_1 \mid X_0 X_1 \mid X_0 D_2 \mid X_0 D_3$$

$$D_1 \rightarrow A X_1$$

$$D_2 \rightarrow B D_3$$

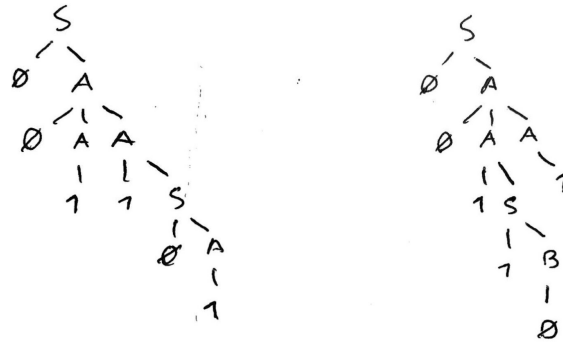
$$D_3 \rightarrow X_1 X_1$$

$$A \rightarrow X_0 D_1 \mid X_0 X_1$$

$$B \rightarrow X_0 D_2 \mid X_0 D_3$$

سوال ۴ - الف)

گرامر مبهم است. مثلاً برای رشته 001101 می‌توان دو درخت اشتقاق زیر را در نظر گرفت:



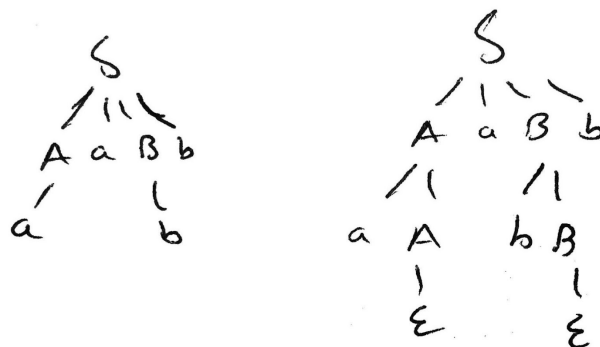
سوال ۴ - ب)

گرامر مبهم نیست. به ازای هر رشته تولید شده هنگام اشتقاق راست، تنها می‌توان یک قانون را برای به دست آوردن کاراکتر بعدی استفاده کرد. مثلاً برای رشته $aaa +^*$ فقط می‌توان به این شکل عمل کرد:

- $S \Rightarrow SS^*$
- $\Rightarrow SSS +^*$
- $\Rightarrow SSa +^*$
- $\Rightarrow Saa +^*$
- $\Rightarrow aaa +^*$

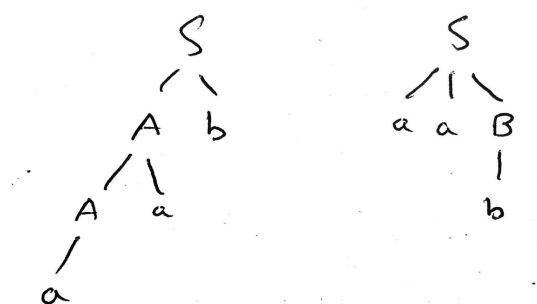
سوال ۴ - ج)

گرامر مبهم است. مثلاً برای رشته aabb می‌توان دو درخت اشتقاق زیر را در نظر گرفت:



سوال ۴ - د

گرامر مبهم است. مثلا برای رشته aab می‌توان دو درخت اشتقاق زیر را در نظر گرفت:



سوال ۵

در گرامر سوال ۴-ج، بعد از حذف قوانین تهی:

$$S \rightarrow AaBb \mid aBb \mid Aab \mid ab$$

$$A \rightarrow aA \mid bB \mid a \mid b$$

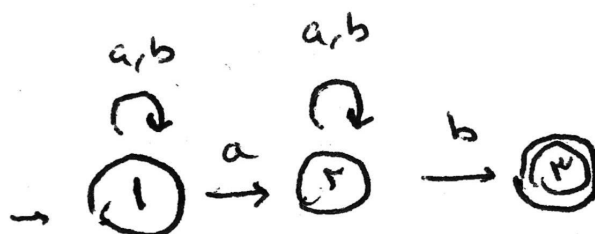
$$B \rightarrow aA \mid bB \mid a \mid b$$

و یکی کردن قوانین A و B (با توجه به این که معادل هستند):

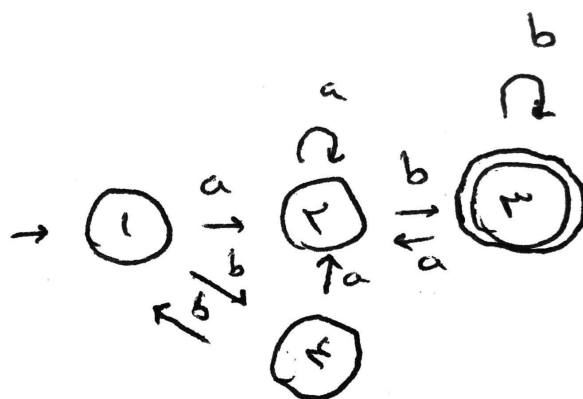
$$S \rightarrow AaAb \mid aAb \mid Aab \mid ab$$

$$A \rightarrow aA \mid bA \mid a \mid b$$

می‌توان مشاهده کرد که عملا زبان این گرامر منظم است و می‌توان با عبارت منظم $(a|b)^* a(a|b)^* b$ آن را نشان داد. با تبدیل این عبارت منظم به FSA:



و تبدیل به DFA:



می‌توان گرامر خطی راست برای آن به دست آورد، که مبهم نیست:

$$S_1 \rightarrow aS_2 \mid bS_4$$

$$S_2 \rightarrow aS_2 \mid bS_3 \mid b$$

$$S_3 \rightarrow bS_3 \mid aS_2 \mid b$$

$$S_4 \rightarrow aS_2 \mid bS_1$$

برای گرامر سوال ۴-د، تنها رشته‌ای که دو درخت اشتقاق دارد aab است. این دو اشتقاق به صورت زیر هستند:

$$S \Rightarrow aaB \Rightarrow aab$$

$$S \Rightarrow Ab \Rightarrow Aab \Rightarrow aab$$

با حذف قانون $S \rightarrow aaB$ ، رشته‌ی aab فقط با یک نوع اشتقاق چپ قابل تولید است. همچنین رشته‌های دیگر که قبلاً تولید می‌شدند نیز همچنان قابل تولید هستند. بنابراین گرامر زیر، معادل و نامبهم است:

$$S \rightarrow Ab$$

$$A \rightarrow Aa \mid a$$

همچنین قانون B غیرمفید خواهد بود و می‌توان آن را حذف کرد.