

فرادرس

فراتر از یک کلاس درس
www.faradars.org

نظریه زبان ها و ماشین ها

مدرس:

فرشید شیرافکن

دانشجوی دکتری دانشگاه تهران

(کارشناسی و کارشناسی ارشد : کامپیوتر نرم افزار) (دکتری: بیوانفورماتیک)

بخش دوم (قسمت دوم)

بسته بودن - ابهام - ساده سازی گرامر - فرم های نرمال - گرامر خطی

FaraDars.org

بسته بودن زبان های مستقل از متن :

خانواده زبان های مستقل از متن تحت اجتماع، الحاق، بستار ستاره ای، معکوس و هم ریختی بسته است و تحت اشتراک، مکمل گیری و تفاضل بسته نیست.

مثال

نشان دهید که خانواده زبان های مستقل از متن، تحت اشتراک بسته نیستند.

دو زبان L_1 و L_2 مستقل از متن هستند، چون برای آنها می توان گرامر مستقل از متن نوشت :

$$L_1 = \{a^n b^n c^m : n \geq 0, m \geq 0\}$$

$$L_2 = \{a^n b^m c^m : n \geq 0, m \geq 0\}$$

اما اشتراک این دو زبان یعنی $\{a^n b^n c^n : n \geq 0\}$ مستقل از متن نیست.

مثال

آیا زبان $L = \{a^n b^n : n \geq 0, n \neq 3\}$ مستقل از متن است؟

زبان L از اشتراک یک زبان مستقل از متن با زبان منظم تشکیل شده ، پس مستقل از متن است.

$$L = \{a^n b^n : n \geq 0\} \cap \overline{L_1}$$

$$L_1 = \{a^3 b^3\}$$

مثال

آیا زبان $L = \{w \in \{a, b, c\}^* : n_a(w) = n_b(w) = n_c(w)\}$ ، مستقل از متن است؟

خیر - اشتراک این زبان با زبان منظم $L(a^*b^*c^*)$ برابر $\{a^n b^n c^n : n \geq 0\}$ می باشد که

می دانیم مستقل از متن نمی باشد.

بنابراین L مستقل از متن نیست. (با توجه به قضیه اشتراک منظم)

مثال

آیا مکمل زبان $L = \{w \in \{a,b,c\}^* : n_a(w) = n_b(w) = n_c(w)\}$ ، مستقل از متن است؟

حل: بله - مکمل این زبان از اجتماع چهار حالت زیر تشکیل شده :

$$n_a(w) < n_b(w) \quad -1$$

$$n_a(w) > n_b(w) \quad -2$$

$$n_a(w) < n_c(w) \quad -3$$

$$n_a(w) > n_c(w) \quad -4$$

تمامی این چهار حالت مستقل از متن هستند. زبان های مستقل از متن تحت اجتماع بسته هستند.

تمرین

نشان دهید زبان های مستقل از متن تحت متمم بسته نیستند.

فرا دارس
FaraDars.org

ابهام در گرامر و زبان

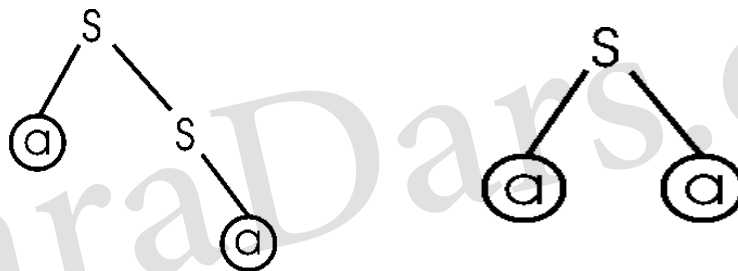
گرامر مستقل از متن G در صورتی مبهم می شود که یک رشته $w \in L(G)$ وجود داشته باشد که حداقل دو درخت اشتقاق مجزا داشته باشد.

به بیان دیگر، ابهام به طور ضمنی به معنای وجود دو یا چند اشتقاق چپ ترین یا راست ترین، نیز می باشد.

مثال

گرامر $S \rightarrow aS \mid aa \mid a$ مبهم است:

چون به طور نمونه برای تولید رشته aa ، دو درخت اشتقاق وجود دارد:



مثال

$$S \rightarrow AB \mid aaB$$
$$A \rightarrow a \mid Aa$$
$$B \rightarrow b$$

گرامر زیر مبهم است:

چون به طور نمونه برای تولید رشته aab ، دو اشتقاق چپ وجود دارد:
$$1 : S \Rightarrow aaB \Rightarrow aab$$
$$2 : S \Rightarrow AB \Rightarrow AaB \Rightarrow aaB \Rightarrow aab$$

مثال

گرامر $S \rightarrow aSbS \mid bSaS \mid \lambda$ مبهم است:

چون به طور نمونه برای تولید رشته $abab$ ، دو اشتقاق چپ وجود دارد:

$$1 : S \Rightarrow aSbS \Rightarrow abSaSbS \Rightarrow abaSbS \Rightarrow ababS \Rightarrow abab$$

$$2 : S \Rightarrow aSbS \Rightarrow abS \Rightarrow abaSbS \Rightarrow ababS \Rightarrow abab$$

مثال

گرامر $S \rightarrow aSb \mid SS \mid \lambda$ مبهم است:

چون به طور نمونه برای تولید رشته ab ، دو اشتقاق چپ وجود دارد:

$$1 : S \Rightarrow aSb \Rightarrow ab$$

$$2 : S \Rightarrow SS \Rightarrow aSbS \Rightarrow abS \Rightarrow ab$$

مثال

$$S \rightarrow aB \mid A$$
$$A \rightarrow aA \mid \lambda$$
$$B \rightarrow bB \mid a$$

گرامر زیر مبهم است:

چون به طور نمونه برای تولید رشته aa ، دو اشتقاق چپ وجود دارد:

1: $S \Rightarrow A \Rightarrow aA \Rightarrow aaA \Rightarrow aa$

2: $S \Rightarrow aB \Rightarrow aa$

بنابراین گرامر منظم نیز می تواند مبهم باشد.

مثال

زبان $L = \{a^i b^j c^k : i = j \text{ or } j = k\}$ ، یک زبان مستقل از متن **ذاتا مبهم** است.

$$S \rightarrow S_1 \mid S_2$$

$$S_1 \rightarrow S_1 c \mid A$$

$$A \rightarrow aAb \mid \lambda$$

$$S_2 \rightarrow aS_2 \mid B$$

$$B \rightarrow bBc \mid \lambda$$

این گرامر مبهم است.

اگر L مستقل از متن بوده و هر گرامر تولید کننده L ، مبهم باشد، آنگاه این زبان ذاتا مبهم است.

زبان های منظم نمی توانند ذاتا مبهم باشند.

ساده سازی گرامرهای مستقل از متن

در تعریف گرامرهای مستقل از متن، هیچ محدودیتی برای سمت راست قانون در نظر گرفته نشده است که این آزادی در برخی استدلال ها، مشکل ایجاد می کند. در بسیاری موارد بهتر است محدودیت شدیدی قائل شویم.

FaraDars.org

حذف متغیرها و قوانین بی فایده

یک متغیر مفید است اگر و تنها اگر در حداقل یک اشتقاق حضور داشته باشد.

عوامل غیرمفید بودن یک متغیر عبارتند از:

۱- قابل دسترس نبودن از طریق متغیر شروع گرامر

۲- ناتوانی در اشتقاق یک رشته پایانی

مثال

در گرامر زیر متغیر **B** بی فایده می باشد:
(از طریق متغیر شروع یعنی **S**، قابل دسترس نیست)

$$S \rightarrow A$$

$$A \rightarrow aA \mid \lambda$$

$$B \rightarrow bA$$

پس قانون $B \rightarrow bA$ بی فایده است و می توان آن را حذف کرد، بدون اینکه تغییری در زبان ایجاد کند.

مثال

در گرامر زیر، متغیر A بی فایده است:
(چون نمی تواند رشته ای از پایانی ها را تولید کند.)

$$S \rightarrow aSb \mid A \mid \lambda$$

$$A \rightarrow aA$$

بنابراین قانون $S \rightarrow A$ و $A \rightarrow aA$ را می توان حذف کرد، بدون اینکه تغییری در زبان ایجاد شود.

سؤال

$$S \rightarrow aS \mid A \mid C$$

$$A \rightarrow a$$

$$B \rightarrow aa$$

$$C \rightarrow aCb$$

در گرامر زیر کدام متغیرها بی فایده هستند؟

پاسخ:

متغیر C بی فایده است، چون یک رشته پایانی را تولید نمی کند.

متغیر B بی فایده است، چون از متغیر شروع قابل دستیابی نمی باشد.

گرامر نهایی :

$$S \rightarrow aS \mid A$$

$$A \rightarrow a$$

مثال

ساده سازی:

$S \rightarrow aAb \mid bBa \mid bCa$
 $A \rightarrow aaAb \mid ab$
 $B \rightarrow bBa \mid a$
 $C \rightarrow aC \mid bC$



$S \rightarrow aAb \mid bBa$
 $A \rightarrow aaAb \mid ab$
 $B \rightarrow bBa \mid a$

متغیر C ، غیر مفید است .

مثال

در گرامر زیر متغیرهای بی فایده را حذف کنید.

$S \rightarrow AC \mid BS \mid B$
 $A \rightarrow aA \mid aF$
 $B \rightarrow CF \mid b$
 $C \rightarrow cC \mid D$
 $D \rightarrow aD \mid BD \mid C$
 $E \rightarrow aA \mid BSA$
 $F \rightarrow bB \mid b$



$S \rightarrow BS \mid B$
 $A \rightarrow aA \mid aF$
 $B \rightarrow b$
 $E \rightarrow aA \mid BSA$
 $F \rightarrow bB \mid b$



$S \rightarrow BS \mid B$
 $B \rightarrow b$

ابتدا متغیرهایی را که به رشته‌ای از الفبا نمی‌رسند یعنی C و D و سپس متغیرهایی را که نمی‌توان از S به آنها رسید، یعنی A و E و F را حذف می‌کنیم.

حذف قوانین λ

هر قانونی از یک گرامر مستقل از متن به فرم $A \rightarrow \lambda$ را قانون λ می گویند. این قوانین در بعضی مواقع نامطلوب می باشند. هر متغیر A که اشتقاق $\lambda \Rightarrow^* A$ برای آن امکان پذیر باشد را متغیر میرا می نامند.

برخی گرامرها زبانهایی را تولید می کنند که هر چند فاقد λ هستند، تعدادی متغیر میرا یا قانون λ در آنها وجود دارند. در این موارد، می توان قوانین λ را حذف کرد.

مثال

گرامر زیر، زبان $\{a^n b^n : n \geq 1\}$ را تولید می کند. این زبان فاقد λ می باشد.

$$S \rightarrow aAb$$

$$A \rightarrow aAb \mid \lambda$$

برای حذف قانون $A \rightarrow \lambda$ ، دو قانون جدید که با جایگزینی λ در A های سمت راست بدست آمده اند را به گرامر اضافه می کنیم:

$$S \rightarrow aAb \mid ab$$

$$A \rightarrow aAb \mid ab$$

مثال

قوانین λ را در گرامر زیر حذف کنید.

$$S \rightarrow AB$$

$$A \rightarrow aA \mid \lambda$$

$$B \rightarrow bB \mid \lambda$$



$$S \rightarrow AB \mid A \mid B \mid \lambda$$

$$A \rightarrow aA \mid a$$

$$B \rightarrow bB \mid b$$

نیازی به حذف قانون $S \rightarrow \lambda$ نمی باشد، چون S ، متغیر شروع است.

مثال

$S \rightarrow ABaC$
 $A \rightarrow BC$
 $B \rightarrow b | \lambda$
 $C \rightarrow D | \lambda$
 $D \rightarrow d$



$S \rightarrow ABaC | AaC$
 $A \rightarrow BC | C$
 $B \rightarrow b$
 $C \rightarrow D | \lambda$
 $D \rightarrow d$



$S \rightarrow ABaC | AaC | ABa | Aa$
 $A \rightarrow BC | C | B | \lambda$
 $B \rightarrow b$
 $C \rightarrow D$
 $D \rightarrow d$

حذف قوانين λ :

$S \rightarrow ABaC | AaC | ABa | Aa | BaC | aC | Ba | a$
 $A \rightarrow BC | B | C$
 $B \rightarrow b$
 $C \rightarrow D$
 $D \rightarrow d$

حذف قوانین واحد

هر قانونی از یک گرامر مستقل از متن به فرم $A \rightarrow B$ که در آن $A, B \in V$ ، قانون واحد یا یک نامیده می شود.

این قوانین گاهی اوقات نامطلوب هستند و باید حذف شوند.

FaraDars.org

حذف قوانین واحد :

$S \rightarrow Aa \mid B$

$B \rightarrow A \mid bb$

$A \rightarrow bc \mid a \mid B$



$S \rightarrow Aa \mid bc \mid a \mid bb$

$B \rightarrow bc \mid a \mid bb$

$A \rightarrow bc \mid a \mid bb$

در اثر حذف قوانین واحد، متغیر B و قوانین مربوط به آن بی فایده شده اند.

ممکن است حذف قوانین λ ، باعث تولید قوانین واحد شود که قبلا وجود نداشته اند.

زبان L را یک زبان مستقل از متن فاقد λ فرض کنید. آنگاه یک گرامر مستقل از متن وجود خواهد داشت که L را تولید کرده و فاقد هرگونه قانون بی فایده، قانون و قانون واحد باشد.

مثال

در گرامر زیر قوانین λ را حذف کنید.

$$\begin{aligned} S &\rightarrow aA \\ A &\rightarrow BB \\ B &\rightarrow aBb \mid \lambda \end{aligned}$$

حل: با حذف $B \rightarrow \lambda$ گرامر زیر بدست می آید:

$$\begin{aligned} S &\rightarrow aA \mid a \\ A &\rightarrow BB \mid B \\ B &\rightarrow aBb \mid ab \end{aligned}$$

در این گرامر قانون واحد $A \rightarrow B$ تولید شده که قبلا وجود نداشت.

مثال

در گرامر زیر تمامی قوانین λ واحد و بی فایده را حذف کنید.

$$S \rightarrow aA \mid aBB$$

$$A \rightarrow aaA \mid \lambda$$

$$B \rightarrow bC \mid bbC$$

$$C \rightarrow B$$

حل: ابتدا قانون $A \rightarrow \lambda$ را حذف می کنیم:

$$S \rightarrow aA \mid aBB \mid a$$

$$A \rightarrow aaA \mid aa$$

$$B \rightarrow bC \mid bbC$$

$$C \rightarrow B$$

سپس قانون واحد $C \rightarrow B$ را حذف می کنیم:

$$S \rightarrow aA \mid aBB \mid a$$

$$A \rightarrow aaA \mid aa$$

$$B \rightarrow bC \mid bbC$$

$$C \rightarrow bC \mid bbC$$

نهایتا اینکه B و C بی فایده هستند، بنابراین داریم:

$$S \rightarrow aA \mid a$$

$$A \rightarrow aaA \mid aa$$

زبان تولید شده بوسیله این گرامر $L((aa)^*a)$ می باشد.

فرم‌های نرمال گرامر مستقل از متن

برای گرامر مستقل از متن، دو فرم نرمال چامسکی و گریباخ وجود دارد.

فرم نرمال چامسکی

یک گرامر مستقل از متن در صورتی در فرم نرمال چامسکی قرار دارد که تمام قوانین آن به یکی از دو

فرم $A \rightarrow BC$ یا $A \rightarrow a$ باشد. که در آن A, B, C عضو V بوده و a عضو T است.

مثال

گرامر زیر در فرم نرمال چامسکی قرار دارد:

$$S \rightarrow AS \mid BS \mid a$$
$$A \rightarrow SA \mid a$$
$$B \rightarrow SB \mid b$$

مثال

گرامر زیر را به فرم نرمال چامسکی تبدیل کنید.

$$S \rightarrow ABa$$

$$A \rightarrow aab$$

$$B \rightarrow Ac$$

حل:

قدم اول : معرفی متغیرهای جدید X,Y,Z	قدم دوم : نرمال کردن قانون اول و سوم
$S \rightarrow ABX$ $X \rightarrow a$ $A \rightarrow XXY$ $Y \rightarrow b$ $B \rightarrow AZ$ $Z \rightarrow c$	$S \rightarrow AT$ $T \rightarrow BX$ $A \rightarrow XF$ $F \rightarrow XY$ $B \rightarrow AZ$ $X \rightarrow a, Y \rightarrow b, Z \rightarrow c$

مثال

گرامر زیر را به فرم نرمال چامسکی تبدیل کنید.

$$S \rightarrow aAB$$

$$A \rightarrow aA \mid b$$

$$B \rightarrow bB \mid b$$

حل:

قدم اول	قدم دوم
$S \rightarrow TAB$ $T \rightarrow a$ $A \rightarrow TA \mid b$ $B \rightarrow FB \mid b$ $F \rightarrow b$	$S \rightarrow TK$ $K \rightarrow AB$ $T \rightarrow a$ $A \rightarrow TA \mid b$ $B \rightarrow FB \mid b$ $F \rightarrow b$

مثال

گرامر زیر را به فرم نرمال چامسکی تبدیل کنید.

$$S \rightarrow aAbB \mid ab$$

$$A \rightarrow ABS \mid a$$

$$B \rightarrow bb$$

حل:

قدم اول	قدم دوم	قدم سوم
$S \rightarrow TAFB \mid TF$ $A \rightarrow ABS \mid a$ $B \rightarrow FF$ $T \rightarrow a$ $F \rightarrow b$	$S \rightarrow TK \mid TF$ $K \rightarrow AFB$ $A \rightarrow AU \mid a$ $U \rightarrow BS$ $B \rightarrow FF$ $T \rightarrow a$ $F \rightarrow b$	$S \rightarrow TK \mid TF$ $K \rightarrow AX$ $X \rightarrow FB$ $A \rightarrow AU \mid a$ $U \rightarrow BS$ $B \rightarrow FF$ $T \rightarrow a$ $F \rightarrow b$

مثال

گرامر زیر را به فرم نرمال چامسکی تبدیل کنید.

$$S \rightarrow abAB$$

$$A \rightarrow bAB \mid \lambda$$

$$B \rightarrow BAa \mid A \mid \lambda$$

حل: ابتدا قواعد λ را حذف می کنیم:

حذف $A \rightarrow \lambda$	حذف $B \rightarrow \lambda$
$S \rightarrow abAB \mid abB$ $A \rightarrow bAB \mid bB$ $B \rightarrow BAa \mid A \mid Ba \mid \lambda$	$S \rightarrow abAB \mid abB \mid abA \mid ab$ $A \rightarrow bAB \mid bB \mid bA \mid b$ $B \rightarrow BAa \mid A \mid Ba \mid Aa \mid a$

سپس قاعده یکه $B \rightarrow A$ را حذف می کنیم:

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

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

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

در نهایت گرامر را به فرم نرمال چامسکی در می آوریم:

قدم اول : معرفی متغیرهای جدید Y, X	قدم دوم : معرفی متغیرهای جدید P, T, K, N, M
$S \rightarrow XYAB \mid XYB \mid XYA \mid XY$ $A \rightarrow YAB \mid YB \mid YA \mid b$ $B \rightarrow BAX \mid BX \mid AX \mid a \mid YAB \mid YB \mid YA \mid b$ $X \rightarrow a$ $Y \rightarrow b$	$S \rightarrow XM \mid XN \mid XK \mid XY$ $M \rightarrow YT$ $T \rightarrow AB$ $N \rightarrow YB$ $A \rightarrow YT \mid YB \mid YA \mid b$ $B \rightarrow BP \mid BX \mid AX \mid a \mid YT \mid YB \mid YA \mid b$ $P \rightarrow AX$ $X \rightarrow a$ $Y \rightarrow b$

در فرم نرمال چامسکی، تولید رشته‌ای به طول n دارای اشتقاقی با طول $2n - 1$ می‌باشد.
(n اشتقاق از $A \rightarrow a$ و $n - 1$ اشتقاق از $A \rightarrow BC$)

فرض کنید G یک گرامر مستقل از متن (CFG) در فرم نرمال چامسکی با b متغیر باشد. در این صورت اگر G بتواند رشته‌ای را با تعداد گام‌های اشتقاق بیشتر از b تولید کند، آنگاه $L(G)$ نامحدود است.

فرم نرمال گریباخ

یک گرامر مستقل از متن در فرم نرمال گریباخ است هرگاه تمام قوانین آن به فرم $A \rightarrow aX$ باشند، که در آن، $x \in V^*$, $a \in T$

گرامر $S \rightarrow abSb \mid aa$ را به فرم نرمال گریباخ تبدیل کنید.

حل: متغیرهای جدید A و B را معرفی می کنیم که مترادف با a و b هستند:

$$S \rightarrow aBSB \mid aA$$

$$A \rightarrow a$$

$$B \rightarrow b$$

مثال

گرامر $S \rightarrow ab \mid aS \mid aaS$ را به فرم نرمال گریباخ تبدیل کنید.

$$S \rightarrow aX \mid aS \mid aYS$$

$$X \rightarrow b$$

$$Y \rightarrow a$$

مثال

گرامر زیر را به فرم نرمال گریباخ تبدیل کنید.

$$S \rightarrow AB$$

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

$$B \rightarrow b$$

حل: قاعده $S \rightarrow AB$ با تعریف فرم نرمال گریباخ مغایر است. بنابراین به جای A در این

قانون از قوانین جایگزین آن استفاده می کنیم:

$$S \rightarrow aAB \mid bBB \mid bB$$

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

$$B \rightarrow b$$

به ازای هر گرامر مستقل از متن در صورتیکه شامل λ نباشد، یک گرامر معادل به فرم نرمال گریباخ وجود دارد.

در فرم نرمال گریباخ، برای تولید رشته‌ای به طول n ، به اشتقاقی با طول n نیاز است. چون در هر مرحله یکی از نمادهای رشته ایجاد می‌شود.

مثال

گرامر زیر را به فرم نرمال گریباخ تبدیل کنید.

$$S \rightarrow ABb \mid a$$

$$A \rightarrow aaA \mid B$$

$$B \rightarrow bAb$$

$$S \rightarrow ABb \mid a$$

$$A \rightarrow aaA \mid bAb$$

$$B \rightarrow bAb$$



$$S \rightarrow aaABb \mid bAbBb \mid a$$

$$A \rightarrow aaA \mid bAb$$

$$B \rightarrow bAb$$



$$S \rightarrow aXABY \mid bAYBY \mid a$$

$$A \rightarrow aXA \mid bAY$$

$$B \rightarrow bAY$$

$$X \rightarrow a$$

$$Y \rightarrow b$$

گرامر خطی

گرامر مستقل از متنی که در سمت راست تمام قواعد آن، حداکثر یک متغیر وجود داشته باشد.

FaraDars.org

مثال

آیا زبان $\{a^n b^m : m \leq n \leq 2m - 1\}$ خطی است؟

پاسخ: بله - چون می توان یک گرامر خطی برای آن نوشت :

$$S \rightarrow aAb \mid aaBb$$

$$A \rightarrow aAb \mid \lambda$$

$$B \rightarrow aaBb \mid aBb \mid ab \mid b$$

مثال

آیا زبان $\{a^n b^n c^m : n \geq 0, m \geq 0\}$ خطی است؟

حل: بله - چون می توان یک گرامر خطی برای آن نوشت :

$$S \rightarrow Sc \mid aAb \mid \lambda$$

$$A \rightarrow aAb \mid \lambda$$

۱- تمامی زبان های خطی، مستقل از متن هستند.

۲- خانواده زبان های خطی، تحت اجتماع و هم ریختی و معکوس بسته است.

۳- خانواده زبان های خطی، تحت اشتراک و الحاق بسته نیست.

این اسلاید ها بر مبنای نکات مطرح شده در فرادرس
«نظریه زبان ها و ماشین ها»
تهیه شده است.

برای کسب اطلاعات بیشتر در مورد این آموزش به لینک زیر مراجعه نمایید.

faradars.org/fvsft110