نظریهٔ زبان‌ها و خودکاره‌ها
ساده‌سازی دم‌ام
صورت‌های نرمال

محسن هوشمند
دانشکده تکنولوژی اطلاعات و علم رایانه
دانشگاه تحصیلات تکمیلی علوم پایه زنجان

ساده‌سازی

دستور مام

- بدون محدودیت روی سمت راست قوانین تولید (اشتقاق)

عدم‌ضرورت بی‌محدودیتی

- گاهی اوقات موجب بروز مشکل
  - تسهیل محاسبات با حذف شکل‌های $A \to \epsilon$ و $A \to B$
- پس گاهی اوقات بهتر است محدودیت‌هائی منظور شود

به دنبال رویه‌های روش‌مند و باقاعده جهت تبدیل دستورهای مام

- با رعایت بعضی محدودیت‌ها
- کمک در حل بعضی از مسائل
- حفظ قدرت زبان مام

ساده‌سازی

ابتدا بررسی تبدیل‌ها و جانشینی‌ها، سپس بررسی صورت‌های نرمال

صورت نرمال

- صورتی دستوری دارای محدودیت‌هائی روی قوانین
- هم‌زمان دارای گستردگی لازم جهت تبدیل هر دمام به آن

مثال

$$A \to a|aaA|aaBc$$

$$B \to abbA|b$$

$$A \to a|aaA|aaabbAc|aabc$$

منظور از ساده‌سازی

- حذف قوانین نامطلوب
- اما لزوما منجر به کاهش تعداد قوانین نمی‌شود

قضیه

دستور $G = (V,\Sigma,S,R)$ مستقل از متن است و قوانین تولید آن شامل $A \to x_1Bx_2$ و $B \to y_1|y_2|\dots|y_m$ باشد. حال اگر $\hat{G} = (V,\Sigma,S,\hat{R})$ دستوری باشد که در آن $\hat{R}$ با حذف دو قانون مذکور و افزودن $A \to x_1y_1x_2|x_1y_2x_2|\dots|x_1y_mx_2|$ به دست آید،آن‌گاه $\mathrm{L}(G) = L(\hat{G})$

اثبات

$S \overset{*}{\Rightarrow} w$ و $A \to x_1Bx_2$ در اشتقاق لازم است.

$$S \overset{*}{\Rightarrow}_G z_1Az_2 \Longrightarrow_G z_1x_1Bx_2z_2 \Longrightarrow_G z_1x_1(y_1|y_2|\dots|y_m)x_2z_2 \overset{*}{\Rightarrow}_G w$$

$$S \overset{*}{\Rightarrow}_{\hat{G}} z_1Az_2 \Longrightarrow_{\hat{G}} \underbrace{z_1\ (x_1y_1x_2|x_1y_2x_2|\dots|x_1y_mx_2)\ z_2}_{z_1x_1(y_1|y_2|\dots|y_m)x_2z_2} \overset{*}{\Rightarrow}_{\hat{G}} w$$

حذف قوانین و متغیرهای بی‌فایده

$S \rightarrow aSb|A|\epsilon$

$A \rightarrow aA$

اما قانون $S \rightarrow A$ نقشی ندارد

- چون هرگز به رشتهٔ پایانی (صرفا شامل حروف الفباء) ختم نمی‌شود

متغیری مفید است اگر و فقط اگر حداقل در یک اشتقاق استفاده شود

$S \rightarrow A$

$A \rightarrow aA|\epsilon$

$B \rightarrow aA$

دستور B بی‌فایده

- دسترس‌ناپذیر از متغیر آغاز
- امکان حذف آن، بدون تاثیر در زبان

حذف قوانین و متغیرهای بی‌فایده

عوامل نامفیدی یک متغیر

- دسترسی‌ناپذیری از متغیر آغاز دستور
- ناتوانی در اشتقاق رشتهٔ پایانی

مثال

$S \rightarrow aS|A|B$
$A \rightarrow a$
$C \rightarrow aa$ دسترس‌ناپذیر از آغاز
$B \rightarrow aBb$ دچار تسلسل و عدم خروجی

$\Rightarrow$

$S \rightarrow aS|A$
$A \rightarrow a$

متغیر مفید

دستور $G = (V,\Sigma,S,R)$ مستقل از متن است. متغیر $A \in V$ مفید است اگر و فقط اگر رشته‌ای $\mathrm{w} \in L(G)$ باشد که $S \overset{*}{\Rightarrow} xAy \overset{*}{\Rightarrow} w$

به دیگر سخن، متغیری مفید است اگر و فقط اگر حداقل در یک اشتقاق باشد

متغیری که استفاده‌ای نداشته باشد بی‌فایده است.

قانونی که دارای متغیر بی‌فایده باشد، قانون بی‌فایده است

مثال

$S \to aAB|bBa|bCa$
$A \to aaAb|ab$
$B \to aBb|a$
$C \to aC|bC$ تسلسل

$\Rightarrow$

$S \to aAB|bBa$
$A \to aaAb|ab$
$B \to aBb|a$

مثال

$$S \to AC|BS|B$$
$$A \to aA|aF$$
$$B \to CF|b$$
$$C \to cC|D$$ تسلسل
$$D \to aD|BD|C$$ تسلسل
$$E \to aA|BSA$$ دسترس‌ناپذیر از آغاز
$$F \to bB|b$$

$$S \to BS|B$$
$$A \to aA|aF$$ دسترس‌ناپذیر از آغاز
$$B \to b$$
$$F \to bB|b$$

$$S \to BS|B$$
$$B \to b$$
$$F \to bB|b$$ دسترس‌ناپذیر از آغاز

$$S \to BS|B$$
$$B \to b$$

$$S \to bS|b$$

قضیه

دستور $G = (V,\Sigma,S,R)$ مستقل از متن است. آن‌گاه $\hat{G} = (\hat{V},\hat{\Sigma},S,\hat{R})$ معادل آن وجود دارد که دارای هیچ متغیر و قانون بی‌فایده نیست.

اثبات

تولید دستور $G_1(V_1,\Sigma,S,R_1)$ با استفاده از روال

۱- $V_1 = \phi$

۲- $\forall A \in V$، اگر R شامل قانون $A \to x_1x_2 \dots x_n, x_i \in V_1 \cup \Sigma$، آن‌گاه $V_1 = V_1 \cup \{A\}$.

۳- تکرار مرحله دو تا زمان تغییر نکردن V_1

۴- R_1 شامل تمام قوانین R به طوری‌که همه علامت‌های آن عضو $V_1 \cup \Sigma$ باشد.

پایان‌پذیر؟

قضیه

تولید G_1 از $\hat{G}$

- رسم گراف وابستگی G_1
- یافتن متغیرهای دسترس‌ناپذیر از S
- حذف متغیرهای مذکور همراه با قوانین مرتبط با آنها
- حذف هر حرفی که در هیچ قانون مفیدی استفاده نشود

$$\hat{G} = (\hat{V}, \hat{\Sigma}, S, \hat{R})$$

حذف قوانین ϵ

قانون ϵ

- هر قانون از دستور مام به صورت $A \rightarrow \epsilon$
- گاهی اوقات مغضوب!

حذف قوانین ϵ

قانون ϵ

- هر قانون از دستور مام به صورت $A \rightarrow \epsilon$
- گاهی اوقات مغضوب!

متغیر میرا: هر متغیر که دارای اشتقاق $A \overset{*}{\Rightarrow} \epsilon$

حذف قوانین ϵ

قانون ϵ

- هر قانون از دستور مام به صورت $A \to \epsilon$
- گاهی اوقات مغضوب!

متغیر میرا: هر متغیر که دارای اشتقاق $A \overset{*}{\Rightarrow} \epsilon$

دستوری می‌تواند حاوی جمله تهی نباشد و در عین حال دارای قانون تهی و متغیرهای میرا باشد.

حذف قوانین ϵ

قانون ϵ

- هر قانون از دستور مام به صورت $A \to \epsilon$
- گاهی اوقات مغضوب!

متغیر میرا: هر متغیر که دارای اشتقاق $A \overset{*}{\Rightarrow} \epsilon$

دستوری می‌تواند حاوی جمله تهی نباشد و در عین حال دارای قانون تهی و متغیرهای میرا باشد.

$$S \to aS_1b$$

$$S_1 \to aS_1b|\epsilon$$

حذف قوانین ϵ

قانون ϵ

- هر قانون از دستور مام به صورت $A \to \epsilon$
- گاهی اوقات مغضوب!

متغیر میرا: هر متغیر که دارای اشتقاق $A \overset{*}{\Rightarrow} \epsilon$

دستوری می‌تواند حاوی جمله تهی نباشد و در عین حال دارای قانون تهی و متغیرهای میرا باشد.

$$S \to aS_1b$$

$$S_1 \to aS_1b|\epsilon$$

در حالی که دارای ϵ است ولی دارای جمله ϵ نیست

حذف قوانین ϵ

قانون ϵ

- هر قانون از دستور مام به صورت $A \to \epsilon$
- گاهی اوقات مغضوب!

متغیر میرا: هر متغیر که دارای اشتقاق $A \overset{*}{\Rightarrow} \epsilon$

دستوری می‌تواند حاوی جمله تهی نباشد و در عین حال دارای قانون تهی و متغیرهای میرا باشد.

$$S \to aS_1b$$

$$S_1 \to aS_1b|\epsilon$$

در حالی که دارای ϵ است ولی دارای جمله ϵ نیست

با حذف $S_1 \to \epsilon$ قوانین جدید جانشینی تعریف می‌کنیم

حذف قوانین ϵ

قانون ϵ

- هر قانون از دستور مام به صورت $A \to \epsilon$
- گاهی اوقات مغضوب!

متغیر میرا: هر متغیر که دارای اشتقاق $A \overset{*}{\Rightarrow} \epsilon$

دستوری می‌تواند حاوی جمله تهی نباشد و در عین حال دارای قانون تهی و متغیرهای میرا باشد.

$$S \to aS_1b$$

$$S_1 \to aS_1b|\epsilon$$

در حالی که دارای ϵ است ولی دارای جمله ϵ نیست

با حذف $S_1 \to \epsilon$ قوانین جدید جانشینی تعریف می‌کنیم

$$S \to aS_1b|ab$$

$$S_1 \to aS_1b|ab$$

مثال

$$
\begin{aligned}
S &\to ABaC \\
A &\to BC \\
B &\to \epsilon | b \\
C &\to \epsilon | D \\
D &\to d
\end{aligned}
$$

$$
\begin{aligned}
S &\to ABaC | AaC \\
A &\to BC | C \\
B &\to b \\
C &\to \epsilon | D \\
D &\to d
\end{aligned}
$$

$$
\begin{aligned}
S &\to ABaC | AaC | ABa | Aa \\
A &\to BC | C | B | \epsilon \\
B &\to b \\
C &\to D \\
D &\to d
\end{aligned}
$$

$$
\begin{aligned}
S &\to ABaC | AaC | ABa | Aa | BaC | aC | Ba | a \\
A &\to BC | C | B \\
B &\to b \\
C &\to D \\
D &\to d
\end{aligned}
$$

مثال

$$S \to AB$$
$$A \to aAA|\epsilon$$
$$B \to bBB|\epsilon$$

$$S \to AB|B$$
$$A \to aAA|aA|a$$
$$B \to bBB|\epsilon$$

$$S \to AB|B|A|\epsilon$$
$$A \to aAA|aA|a$$
$$B \to bBB|\ bB|b$$

مثال

$$S \to ASA|aB$$
$$A \to B|S$$
$$B \to b|\epsilon$$

$$S \to ASA|aB|a$$
$$A \to B|S|\epsilon$$
$$B \to b$$

$$S \to ASA|aB|a|SA|AS|S$$
$$A \to B|S$$
$$B \to b$$

قضیه

دستور $G = (V,\Sigma,S,R)$ مستقل از متن و فاقد جملهٔ تهی است. آن‌گاه $\hat{G} = (V,\Sigma,S,\hat{R})$ معادل آن وجود دارد که دارای قانون تهی نیست.

اثبات- ایجاد مجموعه متغیرهای میرا

$$v_0 = \{A \in V: (A \to \epsilon) \in R\}$$

$$v_{i+1} = v_i \cup \{B \in V: \exists B \to A_1 A_2 \ldots A_k) \in R, \forall j: 1 \leq j \leq k: A_j \in v_i\}$$

v_N پایان‌پذیر است و شامل تمامی متغیرهای میرا

$G_1(V,\Sigma,S,R_1)$

فرض $A \to X_1 X_2 \ldots X_k, k \geq 1$

- m مورد از X_i-ها متغیر میرا دستور G_1 دارای $2^m - 1$ نوع از این قانون
- اگر همهٔ k مورد میرا، آن‌گاه افزودن $A \to \epsilon$

ادامه اثبات

$A \overset{*}{\Rightarrow}_{G_1} w \Leftrightarrow A \overset{*}{\Rightarrow}_G w, w \neq \epsilon$

۱- فرض استقراء $A \overset{*}{\Rightarrow}_{G_1} w$ و حتما $w \neq \epsilon$

$A \to w \in G_1$ ساخت G_1 نشان‌دهندهٔ $A \to \alpha \in G$

- به‌طوری‌که α با صفر یا چند متغیر میرا برابر w است
  - پس $A \overset{*}{\Rightarrow}_G \alpha \overset{*}{\Rightarrow}_G w$

حکم استقراء

- فرض وجود $A \overset{*}{\Rightarrow}_{G_1} X_1X_2 \dots X_k \overset{*}{\Rightarrow}_{G_1} w$ این قانون از $A \to Y_1Y_2 \dots Y_l$ که Y_jها با حفظ ترتیب X_iها هستند با صفر یا چند متغیر میر
- هم‌چنین شکستن w به $w_1w_2 \dots w_k$ به طوری که $X_i \overset{*}{\Rightarrow}_{G_1} w_i$
  - اگر X_i پایانه $X_i = w_i$ و اگر متغیر باشد آن‌گاه اشتقاق $X_i \overset{*}{\Rightarrow}_{G_1} w_i$ کمتر از nمرحله $Y_j \overset{*}{\Rightarrow}_G w_i$
- در نتیجه می‌توان اشتقاق متناظر را در G ساخت
- $A \Rightarrow_G Y_1Y_2 \dots Y_l \overset{*}{\Rightarrow}_G X_1X_2 \dots X_k \overset{*}{\Rightarrow}_G w_1w_2 \dots w_k = w$
  - حکم برقرار است $\left(A \overset{*}{\Rightarrow}_{G_1} w\right) \Longrightarrow A \overset{*}{\Rightarrow}_G w$

ادامه اثبات

$A \overset{*}{\Rightarrow}_{G_1} w \Leftrightarrow A \overset{*}{\Rightarrow}_G w, w \neq \epsilon$

۲- فرض استقراء $A \overset{*}{\Rightarrow}_G w$ و حتما $w \neq \epsilon$

$A \to w \in G$ ساخت G نشان‌دهندهٔ وجود این قانون در G_1

- پس $A \overset{*}{\Rightarrow}_{G_1} w$

حکم استقراء-

- فرض برای n>1 اشتقاق به صورت $A \Rightarrow_G Y_1Y_2 \dots Y_m \overset{*}{\Rightarrow}_G w$ و $w = w_1w_2 \dots w_k$ به طوری‌که $Y_i \overset{*}{\Rightarrow}_G w_i$
- فرض $X_1X_2 \dots X_k$ به ترتیب Y_jهائی باشند که $w_j \neq \epsilon$ و چون $w \neq \epsilon$ باید $k \geq 1$
  - پس $A \to X_1X_2 \dots X_k$ قانونی در G_1
- ادعا $X_1X_2 \dots X_k \overset{*}{\Rightarrow}_{G_1} w$
  - زیرا Y_jهائی که در X حاضر نیستند جهت اشتقاق تهی استفاده می‌شوند و در اشتقاق w نقشی ندارند.
  - چون $Y_j \overset{*}{\Rightarrow}_G w_j$ کمتر از n مرحله نیاز دارد، از فرض استقراء استفاده و اگر رشته تهی نباشد: $Y_j \overset{*}{\Rightarrow}_{G_1} w_j$. پس $A \Rightarrow_{G_1} X_1X_2 \dots X_k \overset{*}{\Rightarrow}_{G_1} w$
- حکم برقرار است $\left(A \overset{*}{\Rightarrow}_G w\right) \Longrightarrow A \overset{*}{\Rightarrow}_{G_1} w$

دو طرف حکم ثابت است.

حذف m قانون تهی منجر به پیدایش $۲^m - 1$ قانون جدید می‌شود

- کارآمد نیست

حذف قوانین یکه

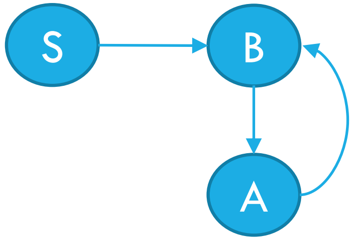
تعریف

- هر قانون به صورت $A \rightarrow B$ و $A,B \in V$ در زبان مام قانون-یکه است.
  - این‌ها هم مغضوب هستند
  - اهی اوقات نیاز به حذف آن‌ها

مثال

- $A \to A$ حذف بدون جانشینی
- $A \to B$ حذف و افزودن $A \to u$ به جای $B \to u$

$$S \to Aa|B$$
$$A \to bc|a|B$$
$$B \to A|bb$$



$$S \to Aa|bb|a|bc$$
$$A \to a|bc|bb$$
$$B \to bb|a|bc$$

- در اثر حذف قانون یکه، متغیر B دسترسی‌ناپذیر و به تبع قوانین آن بی‌فایده‌اند

$$S \to Aa|bb|a|bc$$
$$A \to a|bc|bb$$

مثال

- حذف $S \to S$
- حذف $A \to B$ و $A \to S$

$$S \to ASA|aB|a|SA|AS|S$$
$$A \to S|B$$
$$B \to b$$

$$S \to ASA|aB|a|SA|AS$$
$$A \to ASA|aB|a|SA|AS|b$$
$$B \to b$$

مثال

$$S \to XY \quad X \to A \quad A \to B|a \quad B \to b \quad Y \to T \quad T \to Y|cc$$

$$S \to XY \quad X \to B|a \quad A \to B|a \quad B \to b \quad Y \to Y|cc \quad T \to T|cc$$

$$S \to XY \quad X \to B|a \quad A \to B|a \quad B \to b \quad Y \to cc \quad T \to cc$$

$S \to XY$
$X \to b|a$
$A \to b|a$ دسترس‌ناپذیر
$B \to b$ دسترس‌ناپذیر
$Y \to cc$
$T \to cc$ دسترس‌ناپذیر

$$S \to XY \quad X \to b|a \quad Y \to cc$$

قانون یکه

غالبا روش گسترش قانون واحد تا حذف آن جواب می‌دهد. اما در صورت وجود دور ممکن است شکست بخورد

$A \to B$ و $B \to C$ و $C \to A$

- ابتدا جفت متغیرهای قوانین $A \to B$ *را می‌یابیم*

- *الگویتم کتاب هاپکرافت و المن،ویرایش سوم،صفحه ۲۷۰ جهت حذف قانون یکه کارا و جالب‌توجه*

قضیه

دستور $G = (V,\Sigma,S,R)$ مستقل از متن و فاقد جملهٔ تهی است. آن‌گاه $\hat{G} = (V,\Sigma,S,\hat{R})$ معادل آن وجود دارد که دارای قانون یکه نیست.

اثبات- ابتدا با قضایای قبل می‌توان فرض کرد که G دارای قانون تهی نیست.

دستور جانشینی اثبات می‌کند؟ نه همیشه $A \to B$ و $B \to A$

با استفاده از گراف وابستگی تمامی قوانین $A \overset{*}{\Rightarrow} B$ را پیدا می‌کنیم

دستور $\hat{R}$ با قرار دادن تمامی دستورهای غیر یکه R شروع می‌کنیم

سپس به ازای هر $A \to B$ به جای B قرار می‌دهیم $A \to y_1|y_2| \dots |y_m$

- به طوری که y_i متغیر یکه نیست و $B \to y_1|y_2| \dots |y_m$

در ادامه مانند اثبات قضیهٔ جانشینی

قضیه

دستور $G = (V, \Sigma, S, R)$ مستقل از متن و فاقد جملهٔ تهی است. آن‌گاه $\hat{G} = (V, \Sigma, S, \hat{R})$ معادل آن وجود دارد که دارای بی‌فایده و قانون تهی و قانون یکه نیست.

ترتیب

- ۱- حذف قوانین تهی در ابتدا
- ۲- سپس حذف قوانین یکه
- ۳- در انتها حذف قوانین بی‌فایده
  - چرا؟
    - حذف قانون تهی ممکن است موجب تولید قوانین یکه شود
    - جانشینی قوانین یکه ممکن است موجب شود بعضی متغیرها بی‌فایده (دسترس‌ناپذیر یا پایان‌ناپذیر) شوند

مثال

$$S \to aA \qquad\qquad S \to aA \qquad\qquad S \to aA|a \qquad\qquad S \to aA|a$$
$$A \to BB \quad \Rightarrow \quad A \to BB|B|\epsilon \quad \Rightarrow \quad A \to BB|B \quad \Rightarrow \quad A \to BB|aBb|ab$$
$$B \to aBb|\epsilon \qquad B \to aBb|ab \qquad B \to aBb|ab \qquad B \to aBb|ab$$

مثال

$$S \to ASA|aB$$
$$A \to B|S$$
$$B \to b|\epsilon$$

$$S \to ASA|aB|AS|SA|S|a$$
$$A \to B|S$$
$$B \to b$$

$$S \to ASA|aB|AS|SA|a$$
$$A \to ASA|aB|AS|SA|a|b$$
$$B \to b$$

مثال

$$
\begin{aligned}
S &\to aA|aBB \\
A &\to aaA|\epsilon \\
B &\to bC|bbC \\
C &\to B
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
S &\to aA|aBB|a \\
A &\to aaA|aa \\
B &\to bC|bbC \\
C &\to B
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
S &\to aA|aBB|a \\
A &\to aaA|aa \\
B &\to bC|bbC \quad \text{تسلسل} \\
C &\to bC|bbC \quad \text{تسلسل}
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
S &\to aA|a \\
A &\to aaA|aa
\end{aligned}
$$

$$(aa)^*a$$

صورت‌های نرمال دستور مام

چامسکی

گریباک(خ)

دستور صورت چامسکی

دستور مام صورت نرمال چامسکی (صچ) است اگر قوانین آن به شکل زیر باشد

$A \to BC$

$A \to b$

$S \to \epsilon$

$A,B,C \in V$ و $b \in \Sigma$

مثال ۱

$$S \rightarrow AS|BS|a$$

$$A \rightarrow SA|a$$

$$B \rightarrow SB|b$$

مثال ۲

$S \rightarrow AS|AAS$

$A \rightarrow SA|aa$

مثال ۳

تبدیل به صچ

نرمال‌سازی قوانین اول و سوم

$S \rightarrow AT$

$T \rightarrow BX$

$X \rightarrow a$

$A \rightarrow XF$

$F \rightarrow XY$

$Y \rightarrow b$

$B \rightarrow AZ$

$Z \rightarrow c$

تعریف X و Y و Z

$S \rightarrow ABX$

$X \rightarrow a$

$A \rightarrow XXY$

$Y \rightarrow b$

$B \rightarrow AZ$

$Z \rightarrow c$

$S \rightarrow ABa$

$A \rightarrow aab$

$B \rightarrow Ac$

مثال۴

تبدیل به صچ

$$S \rightarrow aAB$$
$$A \rightarrow aA|b$$
$$B \rightarrow bB|b$$

$$S \rightarrow TAB$$
$$T \rightarrow a$$
$$A \rightarrow TA|b$$
$$B \rightarrow FB|b$$
$$F \rightarrow b$$

$$S \rightarrow TK$$
$$K \rightarrow AB$$
$$T \rightarrow a$$
$$A \rightarrow TA|b$$
$$B \rightarrow FB|b$$
$$F \rightarrow b$$

قضیه

دستور $G = (V, \Sigma, S, R)$ مستقل از متن و فاقد جملهٔ تهی است. آن‌گاه $\hat{G} = (\hat{V}, \hat{\Sigma}, S, \hat{R})$ معادل با صورت چامسکی است.

اثبات-

$G_1(V_1, \Sigma_1, S, R_1)$ دارای قوانین تهی و یکه نیست.

$A \rightarrow x_1 x_2 \dots x_n$:

اگر n=1:

- $A \rightarrow x_1 \Longrightarrow x_1 \in \Sigma$
  - به R_1 اضافه شود

اگر $n \geq 2$:

- $A \rightarrow c_1 c_2 \dots c_n$ به R_1 اضافه شود.
  - $c_i = x_i, x_i \in \mathrm{V}$
- $c_i = B_i, x_i \in \Sigma$ و $B_i \rightarrow a$ به R_1 اضافه شود.

قضیه

مرحله دو

همه قوانین $n \leq 2$ را در $\hat{R}$ قرار می‌دهیم

قوانین $n > 2$

- قوانین جدید
  - $A \rightarrow C_1D_1$
  - $D_1 \rightarrow C_2D_2$
  - ...
  - $D_{n-2} \rightarrow C_{n-1}C_n$
  - همگی به $\hat{R}$ اضافه شوند
- آن‌گاه با استفاده از قضیهٔ جانشینی خواهیم داشت $\mathrm{L}(\hat{G}) = L(G)$.

مثال

$$S \rightarrow ASB$$
$$A \rightarrow aAS|a|\epsilon$$
$$B \rightarrow SbS|A|bb$$

$$S \rightarrow ASB|SB$$
$$A \rightarrow aAS|a|aS$$
$$B \rightarrow SbS|A|bb|\epsilon$$

$$S \rightarrow ASB|SB|AS$$
$$A \rightarrow aAS|a|aS$$
$$B \rightarrow SbS|A|bb$$

$$S \rightarrow AT|SB|AS$$
$$T \rightarrow SB$$
$$A \rightarrow UAS|a|US$$
$$U \rightarrow a$$
$$B \rightarrow SVS|A|VV$$
$$V \rightarrow b$$

$$S \rightarrow AT|SB|AS$$
$$T \rightarrow SB$$
$$A \rightarrow UX|a|US$$
$$X \rightarrow AS$$
$$U \rightarrow a$$
$$B \rightarrow SZ|A|VV$$
$$Z \rightarrow VS$$
$$V \rightarrow b$$

$$S \rightarrow AT|SB|AS$$
$$T \rightarrow SB$$
$$A \rightarrow UX|a|US$$
$$X \rightarrow AS$$
$$U \rightarrow a$$
$$B \rightarrow SZ|UX|a|US|VV$$
$$Z \rightarrow VS$$
$$V \rightarrow b$$

قضیه

تولید رشته با صورت نرمال چامسکی دارای اشتقاقی به طول 2n-1 است.

- n اشتقاق از نوع $A \to a$
- n-1 اشتقاق از نوع $A \to BC$

$$S \xRightarrow{1} AB \xRightarrow{2} AYZ \xRightarrow{3} WXYZ \xRightarrow{4} UVXYZ \xRightarrow{*} abcde$$

$$S \to AB$$
$$A \to WX$$
$$B \to YZ$$
$$W \to UV$$
$$U \to a$$
$$V \to b$$
$$X \to c$$
$$Y \to d$$
$$Z \to e$$

صورت چامسکی

مناسب جهت تجزیه parsing و اثبات قضایا

یافتن صورت چامسکی از دمام ساده است

صورت نرمال گریباک

محدودیت رو تعداد علامت‌ها وجود ندارد

ترتیب قرارگیری متغیرها و حرف الفباء

هم‌ارز دستورهای مام دشوار است

دستور صورت گریباک (ص‌گ)

دستور مام $G = (V, \Sigma, S, R)$ دارای صورت نرمال گریباک است اگر به $a \in \Sigma$ و $x \in V^*$ دستورها به یکی از صورت‌های زیر باشند: گریباک (ص‌گ) است اگر قوانین آن به شکل زیر باشد

$A \to ax$

$S \to \epsilon$

تعمیم دستورهای ساده

مثال ۱

$S \rightarrow AB$

$A \rightarrow aA|bB|b$

$B \rightarrow b$

قانون جانشینی

$S \rightarrow aAB|bBB|bB$

$A \rightarrow aA|bB|b$

$B \rightarrow b$

مثال ۲

قانون جانشینی

$S \rightarrow abSb|aa$

$$S \rightarrow aBSB|aA$$
$$A \rightarrow a$$
$$B \rightarrow b$$

مثال ۳

قانون جانشینی

$S \rightarrow ab|aS|aaS$

$S \rightarrow aB|aS|aAS|aA$

$A \rightarrow a$

$B \rightarrow b$

قضیه

تولید رشته با طول **n** در دمام با صورت نرمال گریباک اشتقاقی به طول **n** دارد.

- چون در هر دور یک حرف از الفباء رشته تولید می‌شود
- مناسب جهت تجزیه ولی یافتن ان‌ها معمولا سخت

حذف قاعده بازگشت چپ

$S \rightarrow b|Sa$

$$S \rightarrow bB$$
$$B \rightarrow aB|\epsilon$$

$S \rightarrow S\alpha_1| \dots |S\alpha_m|\beta_1| \dots |\beta_n$

$$S \rightarrow \beta_1 B| \dots |\beta_n B$$
$$B \rightarrow \alpha_1 B| \dots |\alpha_m B|\epsilon$$

مثال ۴

$S \to ASB|SB|AS$

$A \to aAS|a|aS$

$B \to SbS|aAS|a|aS|bb$

الگوریتم عضویت دستورهای مام کوک-یانگر-کاسامی CYK

صرفا برای صچ

استفاده از برنامه‌ریزی پویا

طراحی جداگانه

الگوریتم عضویت دستورهای مام کوک-یانگر-کاسامی CYK

$S \rightarrow AB|BC$

$A \rightarrow BA|a$

$B \rightarrow CC|b$

$C \rightarrow AB|a$

$w = baaba$

S,A,C				
	S,A,C			
	B	B		
S,A	B	S,C	A,S	
B	A,C	A,C	B	A,C
b	a	a	b	a

منابع

[سیپسر]

[لینز]

“