



نظریهٔ زبان‌ها و خودکاره‌ها دستور مستقل از متن

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

قضیه کلین (نحوه تشخیص منظم بودن یک زبان)

زبانی منظم است اگر آن را بتوان با خم نشان داد.

زبانی منظم است اگر آن را بتوان با خم نشان داد.

زبانی منظم است اگر آن را بتوان با عبارت منظم نشان داد.

زبانی منظم است اگر آن را بتوان با دستور منظم نشان داد.

انواع دستور

- ۱- منظم: قوانینی به صورت $A \rightarrow Bx|x$ یا $A \rightarrow xB|x$ دارد.
- ۲- مستقل از متن: در سمت چپ تمامی قوانین آن، صرفاً یک متغیر باشد.
- ۳- حساس به متن: قوانین آن $x \rightarrow y$ است که $x, y \in (V \cup \Sigma)^+$ و $|x| \leq |y|$.
- ۴- بدون محدودیت: هیچ محدودیتی روی قوانین تولید اعمال نشود مگر اینکه نباید ϵ در سمت چپ قوانین باشد.

دستور مستقل از متن

Context free language

زبان متن آزاد

زبان بافتار آزاد

زبان مستقل از متن

$$a^n b^n, n \geq 0$$

مهم‌ترین موضوع در نظریه زبان‌ها

مورد استفاده در زبان‌های برنامه‌سازی

دستور مستقل از متن

دستور منظم

$$A \rightarrow Bx|x$$

یا

$$A \rightarrow xB|x$$

دستور قوی تر

▪ کاهش محدودیت و قیود

دستور مستقل از متن

تعریف- دستور $G = (V, \Sigma, S, R)$ مستقل از متن است اگر قوانین آن به صورت $A \rightarrow x$ باشد،
به طوری که $A \in V$ و $x \in (V \cup \Sigma)^*$

زبان مستقل از متن است اگر و فقط اگر دستور مستقل از متن G موجود باشد به طوری که $L=L(G)$

هر دستور منظم، دستور مستقل از متن است.

مثال

$$S \rightarrow aSb|ab$$

$$S \Rightarrow ab$$

$$S \Rightarrow aSbS \Rightarrow aaSbbS \Rightarrow aaabbb$$

$$L = \{a^n b^n, n \geq 1\}$$

مثال

$a^n b^n; n \geq 0?$

$S \rightarrow aSb | \epsilon$

مثال

$\{a^n b^{n+1}; n \geq 0\}$?

1- $S \rightarrow aSb|b$

2- $a^n b^n b$

$S \rightarrow Ab$

$A \rightarrow aAb|\epsilon$

مثال

$\{a^{n+1}b^n; n \geq 0\}$?

1- $S \rightarrow aSb|a$

2- $aa^n b^n$

$S \rightarrow aA$

$A \rightarrow aAb|\epsilon$

مثال

$\{a^n b^n a^k b^k; n, k \geq 0\}$?

$S \rightarrow AB$

$A \rightarrow aAb | \epsilon$

$B \rightarrow aBb | \epsilon$

مثال

$$a^n b^m; n \neq m$$

دو حالت

$$n > m: a^{n-m} a^m b^m$$

$$S \rightarrow AS_1$$

$$A \rightarrow aA|a$$

$$S_1 \rightarrow aS_1 b | \epsilon$$

$$n < m:$$

$$S \rightarrow S_1 B$$

$$B \rightarrow bB|b$$

مثال

$\{a^n b^{2n}; n \geq 0\}$?

$S \rightarrow aSbb | \epsilon$

مثال

$\{a^n b^n; n \geq 0\} \cup \{b^n a^n; n \geq 0\}$?

$S \rightarrow S_1 | S_2$

$S_1 \rightarrow aS_1b | \epsilon$

$S_2 \rightarrow bS_2a | \epsilon$

مثال

$$\{a^n b^m c^k; m = n + k; n, k \geq 0\} = \{a^n b^n b^k c^k\}$$

$$S \rightarrow AB$$

$$A \rightarrow aAb | \epsilon$$

$$B \rightarrow bBc | \epsilon$$

مثال

$$\{a^n b^m c^k; k > n + m; n, m \geq 0\} = \{a^n b^m c^l c^m c^k\}$$

$$S \rightarrow aSc|A$$

$$A \rightarrow bAc|cY$$

$$Y \rightarrow cY|\epsilon$$

مثال

$$\{a^n b^m c^k; k = n \cdot m; n, m \geq 0\}?$$

مثال

$\{aa(bc)^n be(dde)^n; n \geq 0\}$?

$S \rightarrow aaA$

$A \rightarrow bBe$

$B \rightarrow cAdd|\epsilon$

$S \rightarrow aaA$

$A \rightarrow bcAdde|be$

مثال

$$S \rightarrow aSa|bSb|\epsilon$$

$$\begin{aligned} S &\rightarrow aA|bB|\epsilon \\ A &\rightarrow Sa \\ B &\rightarrow Sb \end{aligned}$$

دستور مولد $\{ww^r : w \in \{a,b\}^*\}$

مثال

دستور مولد $\{w: \#_a(w) = \#_b(w)\}$

$$S \rightarrow SS|aSb|bSa|\epsilon$$

مثال

دستور روبرو چه زبانی تولید می کند

$$S \rightarrow AB|\epsilon$$

$$A \rightarrow 1A|\epsilon$$

$$B \rightarrow 0B|\epsilon$$

دستور خطی و غیر خطی

دستور مام که در سمت راست قواعد تولید آن، حداکثر یک متغیر موجود باشد.
دستور مام $\{a^n b^n c^m; n \geq 0, m \geq 0\}$ خطی است؟

$$S \rightarrow Sc|aAb|\epsilon$$

$$A \rightarrow aAb|\epsilon$$

دستورهای خطی، مام هستند

▪ دستور مام $A \rightarrow x$ به طوری که $A \in V$ و $x \in (V \cup \Sigma)^*$

▪ دستور خطی $x \in \Sigma^* V \Sigma^*$

اشتقاق چپ و راست

دستور مستقل از متنی که خطی نباشد (در سمت راست بیش از یک متغیر باشد).
جانشینی متغیرها در این وضعیت اختیاری است.

$$1 - S \rightarrow AB$$

$$2 - A \rightarrow aaA|\epsilon$$

$$3 - B \rightarrow Bb|\epsilon$$

$$S \Rightarrow AB \Rightarrow aaAB \Rightarrow aaB \Rightarrow aab$$

$$S \Rightarrow AB \Rightarrow ABb \Rightarrow Ab \Rightarrow aaAb \Rightarrow aab$$

بهتر است ترتیب خاصی در نظر بگیریم.

تعریف

- اشتقاق چپ: اگر در هر مرحله سمت چپ‌ترین متغیر جانشین بپذیرد
- اشتقاق راست: اگر در هر مرحله سمت راست‌ترین متغیر جانشین بپذیرد

درخت اشتقاق

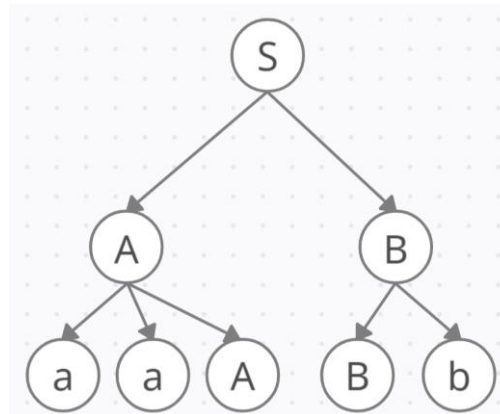
روش دیگر نمایش اشتقاق

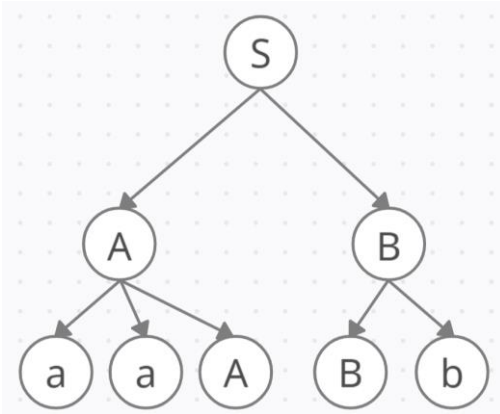
هر رأس $x \rightarrow y$ سمت چپ قانون تولیدی است و فرزندان آن رأس سمت راست آن قانون هستند.

$$1 - S \rightarrow AB$$

$$2 - A \rightarrow aaA|\epsilon$$

$$3 - B \rightarrow Bb|\epsilon$$





درخت اشتقاق

دستور $G = (V, \Sigma, S, R)$ مستقل از متن است. درخت اشتقاق درخت مرتبی است با ویژگی‌های

۱- S ریشه است

۲- هر برگ دارای مقداری از $\Sigma \cup \{\epsilon\}$ است.

۳- رأس‌های غیربرگ مقداری از V دارند

۴- اگر رأس $A \in V$ و فرزندان آن از چپ به راست $\{a_1, a_2, \dots, a_n\}$ باشند، آن‌گاه R دارای قانون

$A \rightarrow a_1 a_2, \dots, a_n$ است.

۵- رأس ϵ هم‌زاد ندارد.

درخت اشتقاق جزئی

دارای ویژگی‌های ۳ و ۴ و ۵ اما ویژگی ۱ لزوماً برای آن صدق نکند و ویژگی ۲ به صورت زیر
۲- هر برگ دارای مقداری از $V \cup \Sigma \cup \{\epsilon\}$ است.

$$1 - S \rightarrow aAB$$

$$2 - A \rightarrow bBb$$

$$3 - B \rightarrow A|\epsilon$$

قضیه

دستور $G = (V, \Sigma, S, R)$ مستقل از متن است. آن گاه

۱- هر $w \in L(G)$ درخت اشتقاق دارد

۲- محصول هر درخت اشتقاق در $L(G)$ است.

۳- اگر درختی اشتقاق جزئی G با ریشه S باشد، آن گاه جمله‌ای از G است.

قضیه

$v \rightarrow A$ و $|v| = k > 1$ آن گاه درخت اشتقاق $w \in L(G)$ ارتفاع h دارد به طوری که

$$\log_k |w| \leq h \leq \frac{|w| - 1}{k - 1}$$

ابهام و تجزیه

Ambiguity and parsing

وقتی بخواهیم بدانیم رشته‌ای را $L(G)$ می‌پذیرد یا نه.

الگوریتم عضویت: بررسی می‌کند رشته‌ای در $L(G)$ است یا نه.

تجزیه: یافتن دنباله‌ای از قوانین که w از آن‌ها مشتق شده است.

تجزیه و عضویت به طور مرتب و مرحله به مرحله تمامی اشتقاق‌های ممکن (چپ) را بدست آوریم و بررسی شود که w با یکی از آن‌ها منطبق است یا خیر (با شروع از ریشه S و افزایش طول)

ابهام و تجزیه

$$S \rightarrow SS|aSb|bSa|\epsilon$$

$$w = aabb$$

دور اول

$$S \Rightarrow SS$$

$$S \Rightarrow aSb$$

$$S \Rightarrow bSa \quad \times$$

$$S \Rightarrow \epsilon \quad \times$$

دور دوم

$$S \Rightarrow SS \Rightarrow SSS$$

$$S \Rightarrow SS \Rightarrow aSbS$$

$$S \Rightarrow SS \Rightarrow bSaS$$

$$S \Rightarrow SS \Rightarrow S$$

$$S \Rightarrow aSb \Rightarrow aSSb$$

$$S \Rightarrow aSb \Rightarrow aaSbb$$

$$S \Rightarrow aSb \Rightarrow abSab \quad \times$$

$$S \Rightarrow aSb \Rightarrow ab \quad \times$$

دور بعد

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aabb$$

ابهام و تجزیه

مشکلات

طولانی و اتلاف هزینه (ناکارا) اگر w عضو $L(G)$ نباشد ممکن است متوقف نشود

حل: دستکاری قوانین تهی $A \rightarrow \epsilon$

اگر شکل دو قانون $A \rightarrow B$ و $A \rightarrow \epsilon$ را نداشته باشیم، مشکلات کمتری داریم.

ابهام و تجزیه

$$S \rightarrow SS|aSb|bSa|\epsilon$$

ختم تجزیه جستجو کامل پس از حداکثر $|w|$ دور
▪ یا رسیدن به رشته یا $w \notin L(G)$

قضیه

دستور $G = (V, \Sigma, S, R)$ مستقل از متن است و دو قانون $A \rightarrow B$ و $A \rightarrow \epsilon$ را ندارد، آن گاه روش تجزیه جستجو کامل را می توان به گونه ای نوشت که به ازای هر $w \in \Sigma^*$ ، تجزیه یا w را تولید می کند و یا می گوید تجزیه ممکن نیست و $w \notin L(G)$.

اثبات

در هر قدم در اشتقاق طول و پایانه های هر جمله یک واحد افزایش پیدا می کند. نه طول و نه تعداد نشانه های الفبا نمی توانند بیش از طول w باشند. پس اشتقاق نمی تواند بیشتر از $2|w|$ دور باشد، یا در این تعداد دورها تجزیه w را می یابد یا $w \notin L(G)$.

← نظری و نه کاربردی: تولید بسیار زیادی از جمله

دستور ساده

S-Grammar

دستور $G = (V, \Sigma, S, R)$ مستقل از متن با قواعد $A \rightarrow ax$ و $a \in \Sigma$ و $A \in V$ و $x \in V^*$ و (A, a) حداکثر یکبار در R ظاهر شود

مثال $S \rightarrow aS|bSS|c$

مثال $S \rightarrow aS|bSS|aSS|c$ ❌

دستور ساده

اگر دستور ساده

▪ هر رشته با $O(|w|)$ تجزیه پذیر

$$w = a_1 a_2 \dots a_n$$

▪ به دلیل وجود یک قاعده که S در سمت چپ و شروع با a_1 در سمت راست

$$S \Rightarrow a_1 A_1 \dots A_m$$

▪ جانشینی A_1

▪ صرفاً یک انتخاب

$$S \Rightarrow a_1 a_2 B_1 \dots B_k \dots A_m$$

▪ پس تولید یک حرف در هر مرحله

▪ کل مراحل پردازش برابر $|w|$

ابهام

چند درخت اشتقاق w را تولید می کند.

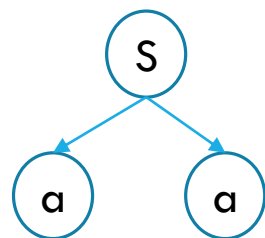
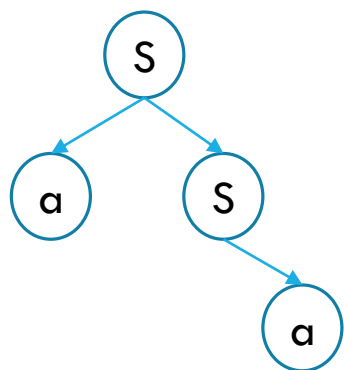
تعریف

دستور مام مبهم است اگر رشته‌ای دارای حداقل دو درخت اشتقاق متفاوت باشد.

به طور ضمنی به معنای وجود دو یا چند اشتقاق چپ و یا اشتقاق راست.

ابهام

$$S \rightarrow aS|aa|a$$



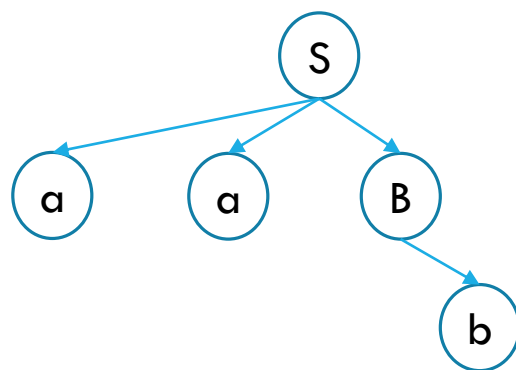
رشته aa

ابهام

$S \rightarrow AB|aaB$

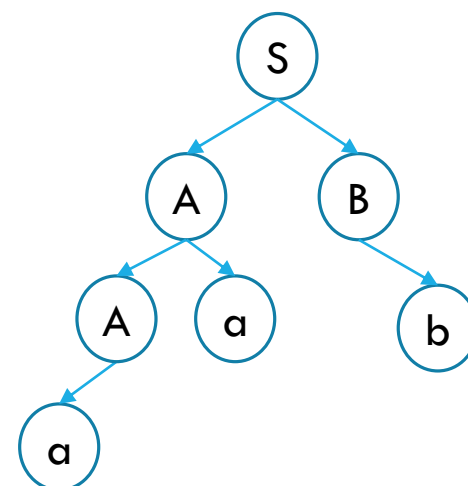
$A \rightarrow a|Aa$

$B \rightarrow b$



$S \Rightarrow aaB \Rightarrow aab$

$S \Rightarrow AB \Rightarrow AaB \Rightarrow aaB \Rightarrow aab$



ابهام

دستور منظم چگونه؟

$$S \rightarrow aB|A$$

$$A \rightarrow \epsilon|Aa$$

$$B \rightarrow bB|a$$

رشته aa

$$S \Rightarrow aB \Rightarrow aa$$

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

پس دستور منظم نیز می تواند مبهم باشد.

ولی ذاتا مبهم نیست.

دستورهای مبهم

امکان اشتقاق جمله‌ای با دو درخت تجزیه متفاوت

مثال - جمله $id := id + id + id$

$E \rightarrow id$

$E \rightarrow num$

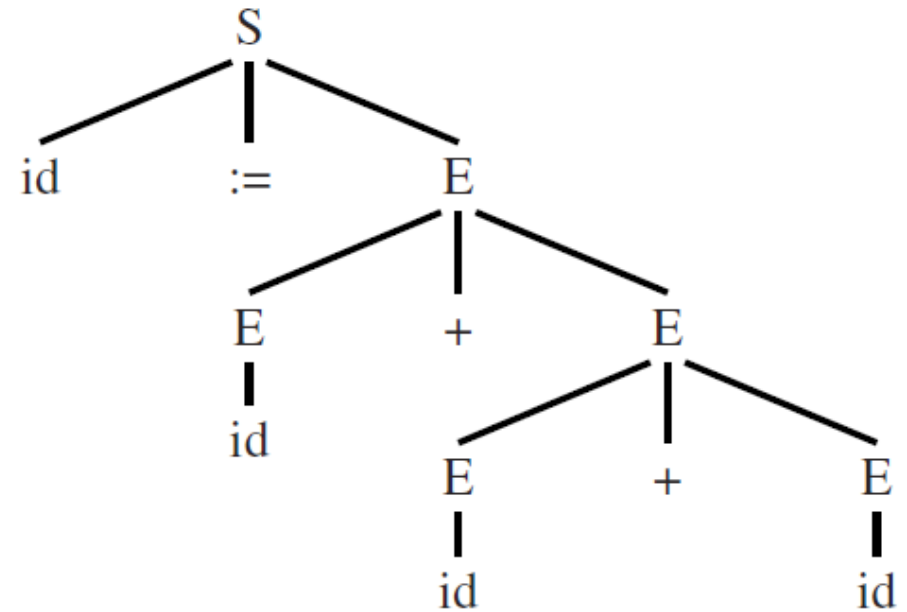
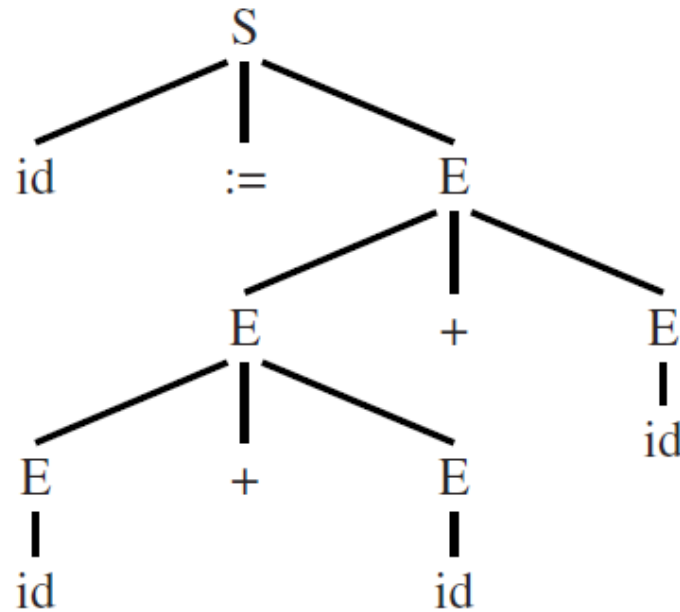
$E \rightarrow E * E$

$E \rightarrow E / E$

$E \rightarrow E + E$

$E \rightarrow E - E$

$E \rightarrow (E)$



دستورهای مبهم - مثال ۲

دستوری مبهم با دو درخت تجزیه متفاوت برای جمله ۱-۲-۳

$$E \rightarrow id$$

$$E \rightarrow num$$

$$E \rightarrow E * E$$

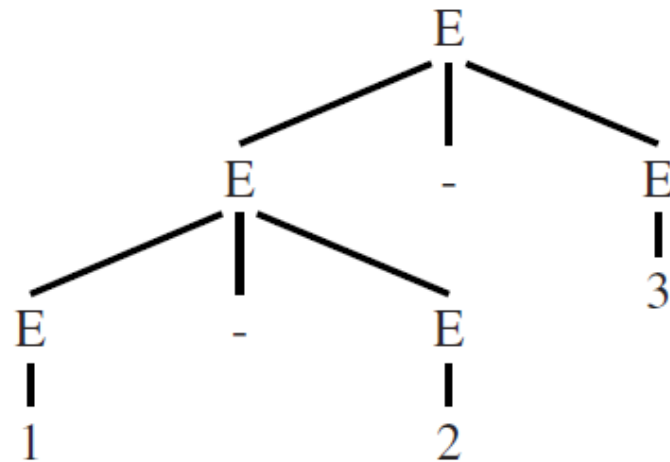
$$E \rightarrow E / E$$

$$E \rightarrow E + E$$

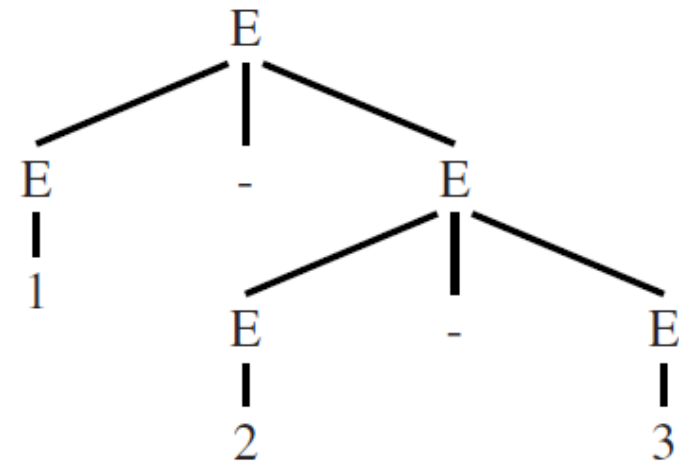
$$E \rightarrow E - E$$

$$E \rightarrow (E)$$

1-2-3



$$(1-2)-3 = -4$$



$$1-(2-3) = 2$$

دستورهای مبهم - مثال ۳

دستوری مبهم با دو درخت تجزیه متفاوت برای جمله

$1+2*3$

$E \rightarrow id$

$E \rightarrow num$

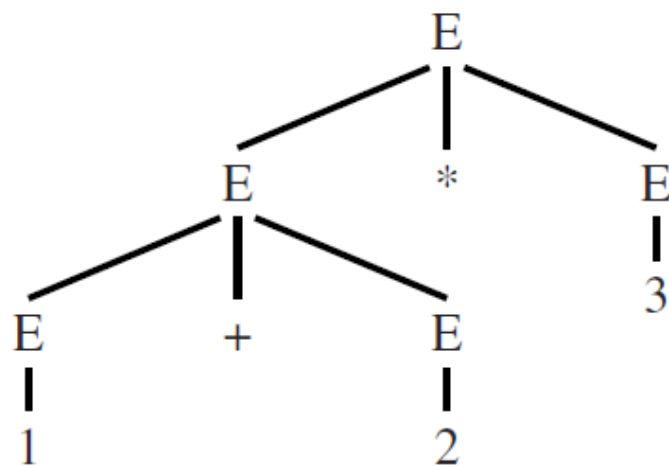
$E \rightarrow E * E$

$E \rightarrow E / E$

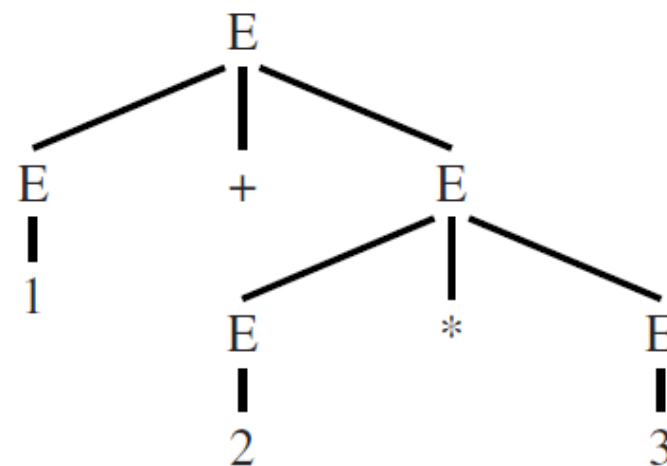
$E \rightarrow E + E$

$E \rightarrow E - E$

$E \rightarrow (E)$



$(1+2)*3=9$



$1+(2*3)=7$

ابهام

راه حل

▪ اعطای اولویت

▪ خارج حیطه دستور

▪ پس دستور بهتر است طوری نوشته شود که هر رشته یک اشتقاق داشته باشد

تعریف زبان ذاتا مبهم

▪ اگر زبانی مام باشد و هر دستور مولد آن مبهم باشد.

$$S \rightarrow aB|A$$

$$A \rightarrow \epsilon|Aa$$

$$B \rightarrow bB|a$$

مثال

زبان ذاتا مبهم است $\{a^n b^m c^k; m = n \text{ یا } m = k\}$ ؟

$$S \rightarrow S_1 | S_2 a B | A$$

$$S_1 \rightarrow S_1 c | A$$

$$A \rightarrow \epsilon | a A b$$

$$S_1 \rightarrow a S_2 | B$$

$$B \rightarrow b B | a$$

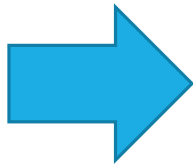
رشته $a^2 b^2 c^2$

$$S \Rightarrow S_1 \Rightarrow S_1 c \Rightarrow S_1 c c \Rightarrow A c c \Rightarrow a A b c c \Rightarrow a a A b b c c \Rightarrow a a b b c c$$

$$S \Rightarrow S_2 \Rightarrow a S_2 \Rightarrow a a S_2 \Rightarrow a a B \Rightarrow a a b B c \Rightarrow a a b b B c c \Rightarrow a a b b c c$$

رفع ابهام دستور

$E \rightarrow id$
 $E \rightarrow num$
 $E \rightarrow E * E$
 $E \rightarrow E / E$
 $E \rightarrow E + E$
 $E \rightarrow E - E$
 $E \rightarrow (E)$



$E \rightarrow E + T$
 $E \rightarrow E - T$
 $E \rightarrow T$
 $T \rightarrow T * F$
 $T \rightarrow T / F$
 $T \rightarrow F$
 $F \rightarrow id$
 $F \rightarrow num$
 $F \rightarrow (E)$

تغییرات

- جهت
- اعمال اولویت‌ها
- اعمال شرکت‌پذیری چپ عملگرها
- شامل
- افزودن متغیرهای T و F
- E مخفف Expression
- T مخفف Term
- F مخفف Factor
- قرارداد
- T-ها نمایشگر ضرب‌پذیرها (تقسیم)
- F-ها نمایشگر جمع‌پذیرها (تفریق)

رفع ابهام دستور - ادامه

$E \rightarrow id$

$E \rightarrow num$

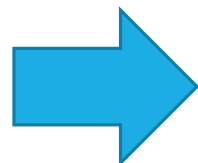
$E \rightarrow E * E$

$E \rightarrow E / E$

$E \rightarrow E + E$

$E \rightarrow E - E$

$E \rightarrow (E)$



$E \rightarrow E + T$

$E \rightarrow E - T$

$E \rightarrow T$

$T \rightarrow T * F$

$T \rightarrow T / F$

$T \rightarrow F$

$F \rightarrow id$

$F \rightarrow num$

$F \rightarrow (E)$

دو دستور متناظر

▪ تولید یکسانی از جملات

دستور جدید بر عکس دستور قبلی

▪ هر جمله دارای دقیقاً «یک» درخت تجزیه

رفع ابهام دستور - ادامه

$$E \rightarrow E + T$$

$$E \rightarrow E - T$$

$$E \rightarrow T$$

$$T \rightarrow T * F$$

$$T \rightarrow T / F$$

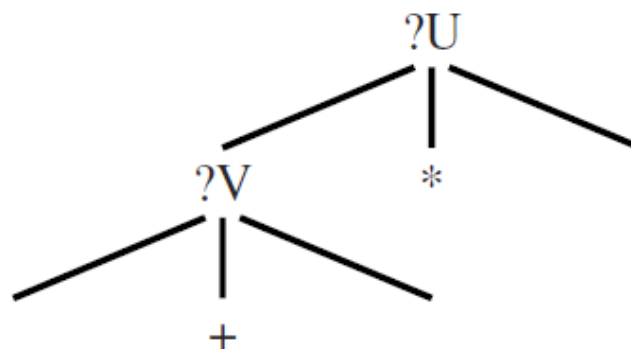
$$T \rightarrow F$$

$$F \rightarrow id$$

$$F \rightarrow num$$

$$F \rightarrow (E)$$

عدم امکان تولید درخت تجزیه زیر از دستور مقابل



دو عملگر با اولویت متفاوت

▪ عملگر با اولویت بالاتر در سطح پائین تر درخت

به چه معنی؟

▪ «اعمال و کد کردن قانون» در «ساختار»

رفع ابهام دستور - ادامه

$$E \rightarrow E + T$$

$$E \rightarrow E - T$$

$$E \rightarrow T$$

$$T \rightarrow T * F$$

$$T \rightarrow T / F$$

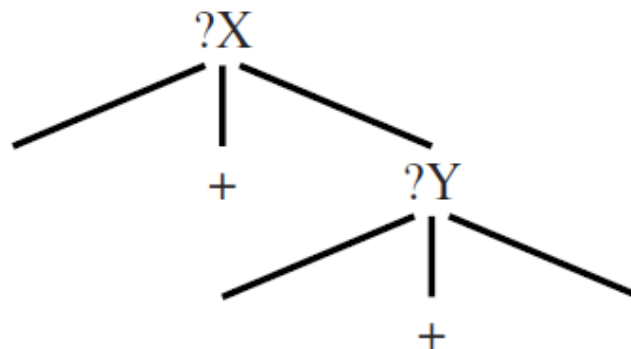
$$T \rightarrow F$$

$$F \rightarrow id$$

$$F \rightarrow num$$

$$F \rightarrow (E)$$

عدم امکان تولید درخت تجزیه زیر از دستور مقابل



دو عملگر با اولویت یکسان
▪ شرکت پذیری چپ عملگرهای با اولویت یکسان ذخیره شده در سطح پایین تر درخت

به چه معنی؟
▪ باز هم «اعمال و کد کردن قانون» در «ساختار»

در صورت ایجاد شرکت پذیری راست
▪ تغییر به $T \rightarrow F * T$

منابع

[سیپسرا]

[لینزا]

[بیرسيز]

[اژدرها]