

Bài 4.

PHÂN TÍCH CÚ PHÁP

Hoàng Anh Việt
Viện CNTT&TT - ĐHBKHN

Nội dung

1. Vai trò của bộ phân tích cú pháp (PTCP)
2. Văn phạm của ngôn ngữ lập trình
3. Phân tích cú pháp từ trên xuống
- 4. Phân tích cú pháp từ dưới lên**
5. Bộ sinh bộ PTCP

4. Phương pháp phân tích từ dưới lên

- Thí dụ 4.6. Cho văn phạm G.

$S \rightarrow aABe$

$A \rightarrow Abc|b$

$B \rightarrow d$

Phân tích câu $w = \mathbf{abbcde}$.

4. Phương pháp phân tích từ dưới lên

(1) $\begin{matrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a & b & b & c & d & e \end{matrix}$

Các lá của cây cú pháp

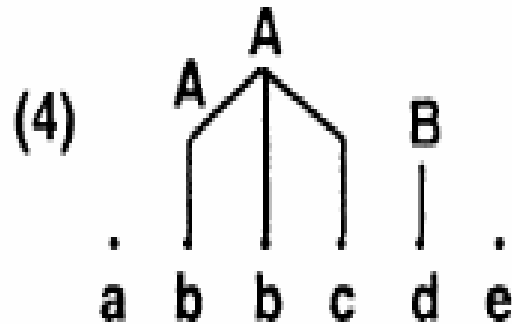
(2) $\begin{matrix} & A & & & & \\ & | & & & & \\ \cdot & & \cdot & \cdot & \cdot & \cdot \\ a & b & b & c & d & e \end{matrix}$

Thu giảm bằng $A \rightarrow b$

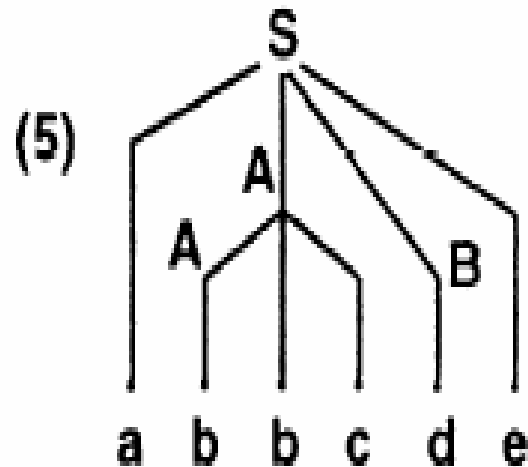
(3) $\begin{matrix} & & A & & & \\ & \swarrow & | & \searrow & & \\ A & & & & & \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a & b & b & c & d & e \end{matrix}$

Thu giảm bằng $A \rightarrow Abc$

4. Phương pháp phân tích từ dưới lên



Thu giảm bằng $B \rightarrow d$



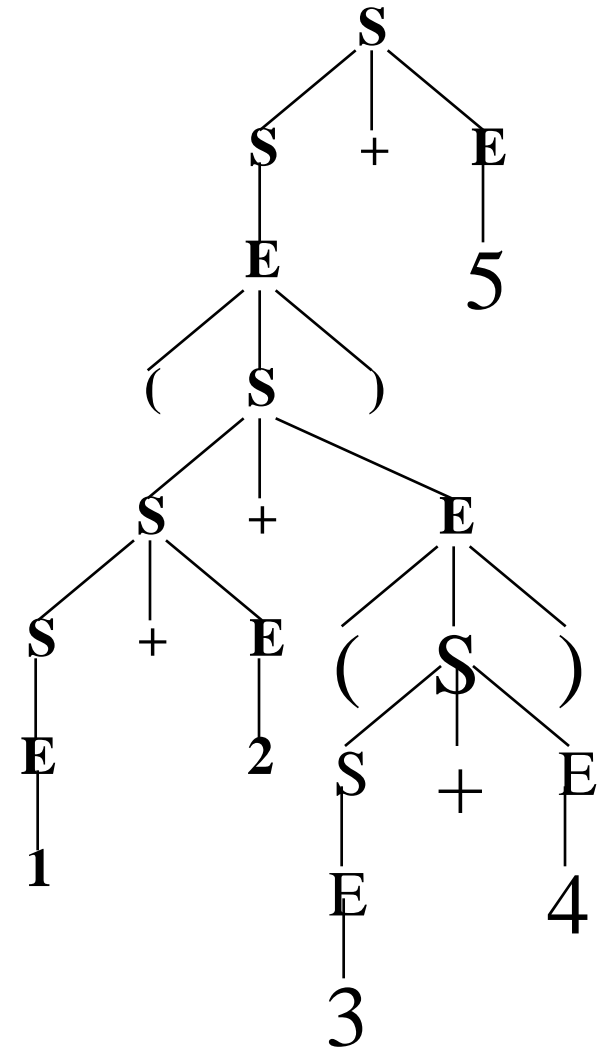
Thu giảm bằng $S \rightarrow aABe$

Phân tích từ dưới lên (bottom-up parsing)

- Kỹ thuật phân tích mạnh hơn
- Văn phạm lớp LR có khả năng mô tả mạnh hơn văn phạm lớp LL, có thể mô tả văn phạm đệ quy trái (có trong hầu hết các ngôn ngữ lập trình)
- Dễ dàng mô tả các ngôn ngữ lập trình thông thường
- Bộ phân tích cú pháp **gạt – thu gọn** (Shift-Reduce parsing)
 - Xây dựng cây suy diễn phải
 - Tự động xây dựng bộ phân tích cú pháp
VD: yacc, CUP
 - Phát hiện lỗi ngay khi xuất hiện
 - Cho phép phục hồi khi lỗi xảy ra

Phân tích trên xuống

- Suy dẫn trái
- Toàn bộ cây phía trên một kí hiệu được sinh ra
- Phải có khả năng đoán trước được sản xuất



Phân tích dưới lên (1)

$S \rightarrow S+E \mid E$ $E \rightarrow \text{số} \mid (S)$
--

- Suy dẫn phải
- Cây suy dẫn được xây dựng ngược lại
 - Bắt đầu từ kí hiệu kết thúc
 - Kết thúc tại kí hiệu bắt đầu

- Ví dụ

$(1+2+(3+4))+5 \Leftarrow (E+2+(3+4))+5 \Leftarrow$
 $(S+2+(3+4))+5 \Leftarrow (S+E+(3+4))+5 \Leftarrow$
 $(S+(3+4))+5 \Leftarrow (S+(E+4))+5 \Leftarrow (S+(S+4))+5 \Leftarrow$
 $(S+(S+E))+5 \Leftarrow (S+(S))+5 \Leftarrow (S+E)+5 \Leftarrow$
 $(S)+5 \Leftarrow E+5 \Leftarrow S+5 \Leftarrow S+E \Leftarrow S$

Phân tích dưới lên (2)

Suy dẫn phải

$$(1+2+(3+4))+5 \quad \Leftarrow$$
$$(\mathbf{E}+2+(3+4)) + 5 \Leftarrow$$
$$(\mathbf{s}+2+(3+4)) + 5 \Leftarrow$$
$$(S + \mathbf{E} + (3 + 4)) + 5 \Leftarrow$$
$$(\mathbf{s} + (3+4)) + 5 \Leftarrow$$
$$(S + (\mathbf{E} + 4)) + 5 \Leftarrow$$
$$2 + (s+4) + 5 \Leftarrow$$
$$(S + (S + \mathbf{E})) + 5 \Leftarrow$$
$$(\overline{S+}(\mathbf{S})) + 5 \quad \Leftarrow$$
$$(S+\mathbf{E})+5 \Leftarrow$$
$$(\mathbf{S}) + 5 \Leftarrow$$

E+5 \Leftarrow

$$S + \mathbf{E} \rightleftharpoons$$

S

$$(1+2+(3+4))+5$$
$$(1 + 2 + (3 + 4)) + 5$$
$$(1 + 2 + (3 + 4)) + 5$$
$$(1+2 + (3+4)) + 5$$
$$(1+2+(3+4))+5$$
$$(1+2+(3+4))+5$$
$$(1+2+(3+4))+5$$
$$(1+2+(3+4)) + 5$$
$$(1+2+(3+4)) + 5$$
$$(1+2+(3+4)) + 5$$
$$(1+2+(3+4)) + 5$$
$$(1+2+(3+4)) \quad +5$$
$$(1+2+(3+4)) + 5$$
$$(1+2+(3+4)) + 5$$

Phân tích dưới lên (3)

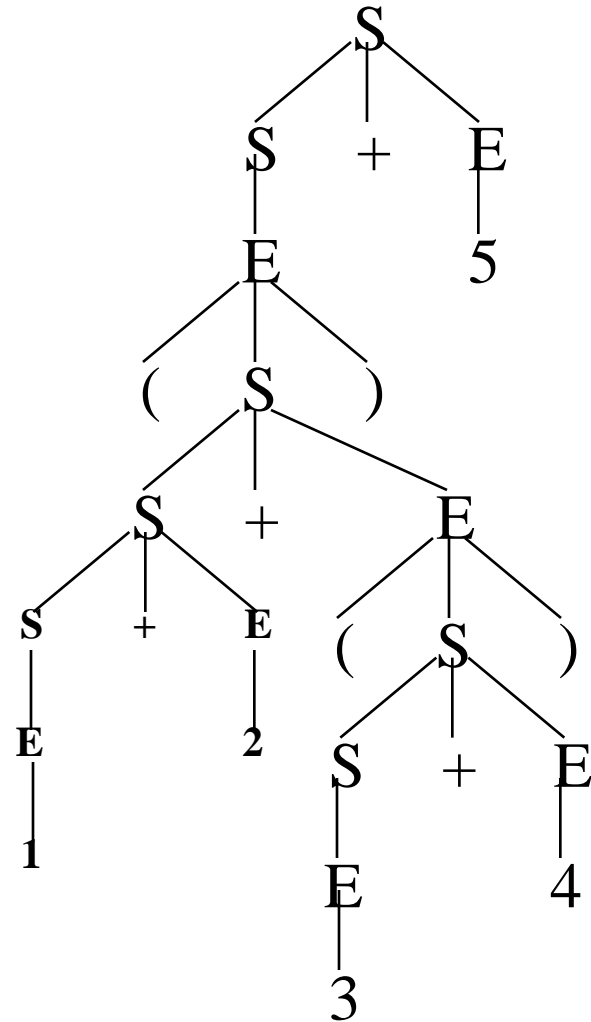
$(1+2+(3+4))+5$

$\Leftarrow (E+2+(3+4))+5$

$\Leftarrow (S+2+(3+4))+5$

$\Leftarrow (S+E+(3+4))+5 \dots$

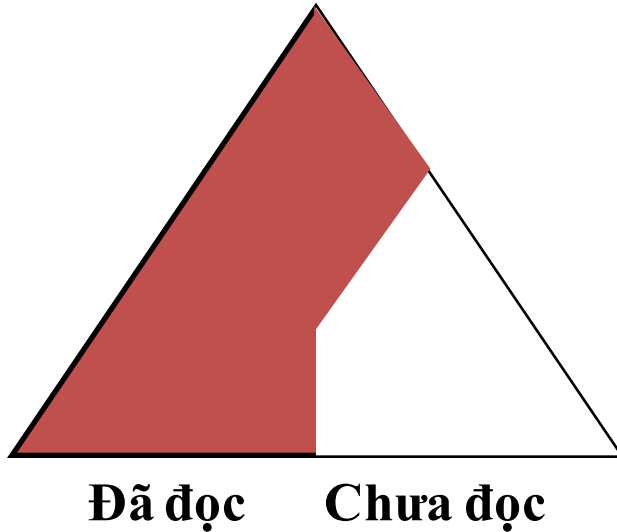
- Phân tích dưới lên có nhiều thông tin hơn khi phân tích



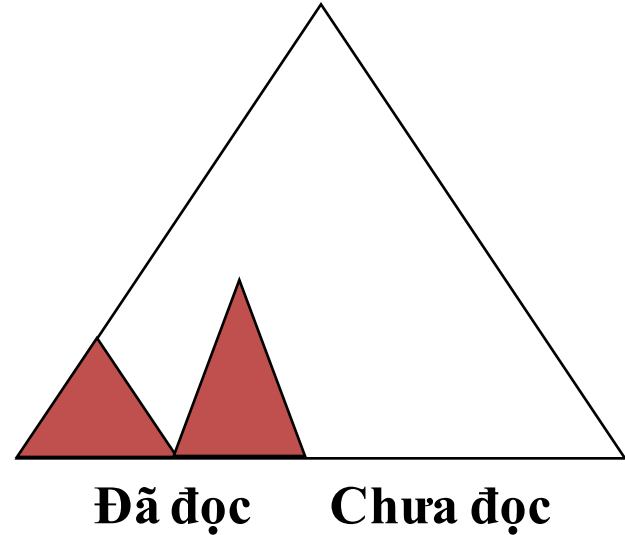
Phân tích dưới lên và phân tích trên xuống

- Phân tích dưới lên không cần sinh ra toàn bộ cây suy diễn trong quá trình phân tích

Phân tích trên xuống



Phân tích dưới lên



4.1 Phân tích gọt – thu gọn (1)

- Phân tích bằng một dãy thao tác: *gọt* và *thu gọn*
- Mỗi thời điểm, trạng thái của bộ phân tích là ngăn xếp các kí hiệu kết thúc và không kết thúc
- Cấu hình tại mỗi thời điểm gồm:
ngăn xếp + xâu các kí hiệu chưa đọc

Suy dẫn	Ngăn xếp	Chưa đọc
$(1+2+(3+4)) + 5 \Leftarrow$		$(1+2+(3+4)) + 5$
$(E+2+(3+4)) + 5 \Leftarrow$	(E	$+2+(3+4)) + 5$
$(S+2+(3+4)) + 5 \Leftarrow$	(S	$+2+(3+4)) + 5$
$(S+E+(3+4)) + 5 \Leftarrow$	(S+E	$+ (3+4)) + 5$

4.1 Phân tích gọt – thu gọn (2)

- **Gọt:** Đọc và đưa một kí hiệu kết thúc của xâu vào stack

Ngăn xếp	Chưa đọc	Thao tác
((1	$1+2+(3+4)+5$ $+2+(3+4)+5$	Gọt 1

- **Thu gọn:** Thay thế một xâu γ ở đỉnh của ngăn xếp bằng kí hiệu không kết thúc X với $X \rightarrow \gamma$ (pop γ , push X)

Ngăn xếp	Chưa đọc	Thao tác
(S+E (S	$+(3+4)+5$ $+(3+4)+5$	Thu gọn: $S \rightarrow S+E$

4.1 Phân tích gọt – thu gọn (3)

Suy dẫn	Ngăn xếp	Chưa đọc	Thao tác
$(1+2+(3+4)) + 5 \Leftarrow$		$(1+2+(3+4)) + 5$	<i>gọt</i> (
$(1+2+(3+4)) + 5 \Leftarrow$	($1+2+(3+4)) + 5$	<i>gọt</i> 1
$(1+2+(3+4)) + 5 \Leftarrow$	(1	$+2+(3+4)) + 5$	<i>thu gọn</i> $E \rightarrow 1$
$(E+2+(3+4)) + 5 \Leftarrow$	(E	$+2+(3+4)) + 5$	<i>thu gọn</i> $S \rightarrow E$
$(S+2+(3+4)) + 5 \Leftarrow$	(S	$+2+(3+4)) + 5$	<i>gọt</i> +
$(S+2+(3+4)) + 5 \Leftarrow$	(S+	$2+(3+4)) + 5$	<i>gọt</i> 2
$(S+2+(3+4)) + 5 \Leftarrow$	(S+2	$+(3+4)) + 5$	<i>thu gọn</i> $E \rightarrow 2$
$(S+E+(3+4)) + 5 \Leftarrow$	(S+E	$+(3+4)) + 5$	<i>thu gọn</i> $S \rightarrow S+E$
$(S+(3+4)) + 5 \Leftarrow$	(S	$+(3+4)) + 5$	<i>gọt</i> +
$(S+(3+4)) + 5 \Leftarrow$	(S+	$(3+4)) + 5$	<i>gọt</i> (
$(S+(3+4)) + 5 \Leftarrow$	(S+($3+4)) + 5$	<i>gọt</i> 3
$(S+(3+4)) + 5 \Leftarrow$	(S+(3	$+4)) + 5$	<i>thu gọn</i> $E \rightarrow 3$
$(S+(E+4)) + 5 \Leftarrow$	(S+(E	$+4)) + 5$	<i>thu gọn</i> $S \rightarrow E$
$(S+(S+4)) + 5 \Leftarrow$	(S+(S	$+4)) + 5$	<i>gọt</i> +
$(S+(S+4)) + 5 \Leftarrow$	(S+(S+	$4)) + 5$	<i>gọt</i> 4
...

Các vấn đề nảy sinh

- Cần xác định khi nào gọt hoặc thu gọn hoặc thu gọn với sản xuất nào?
- Thu gọn sản xuất rỗng

$$X \rightarrow \varepsilon$$

- Có nhiều cách thu gọn

$$S \rightarrow E \text{ hay } S \rightarrow S+E$$

Lựa chọn thao tác

- Tại mỗi thời điểm, từ cấu hình
 $\langle S - \text{ngăn xếp}, a - \text{từ tổ nhìn trước} \rangle$
- Xác định
 - *Gạt* a , ngăn xếp trở thành $\langle Sa \rangle$
 - *Thu gọn* $X \rightarrow \gamma$, nếu $S = \alpha\gamma$,
ngăn xếp trở thành $\langle \alpha X \rangle$
- Nếu $S = \alpha\gamma$, cần lựa chọn *gạt* a hoặc *thu gọn* $X \rightarrow \gamma$ dựa vào tiền tố α
 - Với mỗi khả năng thu gọn $X \rightarrow \gamma$ có một α
 - Cần tìm cách đánh dấu các khả năng thu gọn

Trạng thái của bộ phân tích gát – thu gọn

- **Mục tiêu:** Xác định khả năng thu gọn hợp lệ tại từng thời điểm
- **Ý tưởng:** gộp các khả năng có thể có của tiền tố α thành trạng thái của bộ phân tích
- Các vấn đề nảy sinh:
 - Tính toán các trạng thái của bộ phân tích
 - Tính toán các trạng thái kết thúc
 - Phân tích tất định (loại văn phạm nào)
 - Kích cỡ của bộ phân tích (số lượng trạng thái)

4.2 Bộ phân tích cú pháp LR

Phân tích cú pháp LR(k):

- **L** (left - to - right): Duyệt chuỗi nhập từ trái sang phải.
- **R** (rightmost derivation): Xây dựng chuỗi dẫn xuất phải nhất đảo ngược.
- **k** : Số lượng ký hiệu nhập được xét tại mỗi thời điểm dùng để đưa ra quyết định phân tích. Khi không đề cập đến k, hiểu ngầm là $k = 1$.

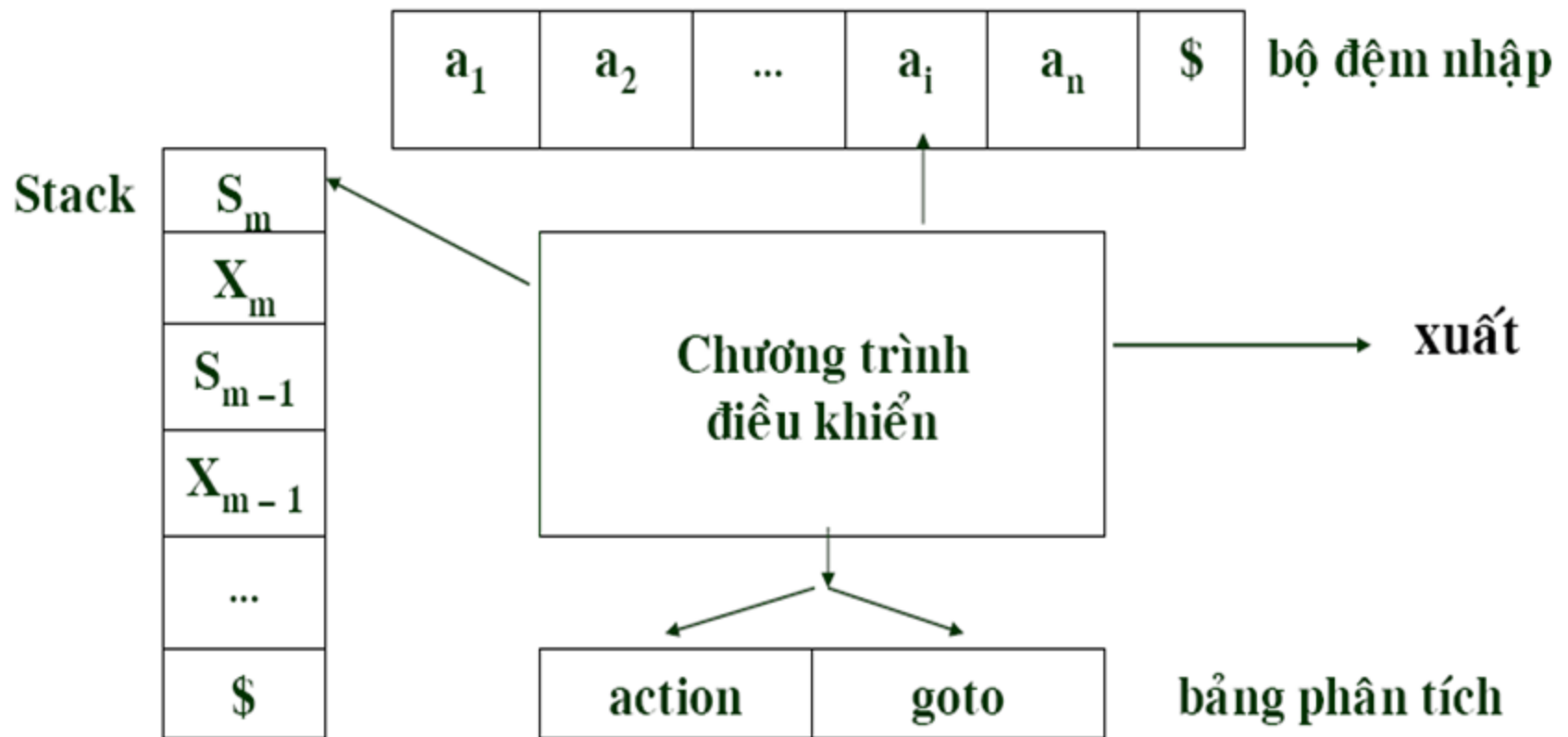
4.2 Bộ phân tích cú pháp LR

Các tính chất của phương pháp phân tích LR(k):

- Bộ phân tích LR có thể nhận dạng được cấu trúc cú pháp của **các** ngôn ngữ lập trình do văn phạm phi ngữ cảnh tạo ra.
- Phương pháp LR là phương pháp tổng quát nhất của phương pháp phân tích gọt và thu gọn, không bị quay lui.
- Lớp văn phạm có thể dùng phương pháp LR là một lớp rộng lớn hơn lớp văn phạm có thể sử dụng **phương pháp dự đoán**.
- Bộ phân tích cú pháp LR cũng có thể xác định lỗi cú pháp nhanh ngay trong khi duyệt dòng nhập từ trái sang phải.

Nhược điểm?

Cấu tạo bộ phân tích LR



Mô hình bộ phân tích LR

Cấu tạo bộ phân tích LR

- **Stack** được dùng để chứa chuỗi ký hiệu có dạng $s_0X_1s_1X_2\ldots X_ms_m$, với s_m nằm trên đỉnh stack, X_i được gọi là ký hiệu văn phạm, s_i là trạng thái tóm tắt thông tin bên dưới stack. Cặp (s_i, X_i) sẽ xác định một trị được lưu chứa trong bảng phân tích.
- **Cấu hình (configuration)** của một bộ phân tích cú pháp LR là một cặp, trong đó thành phần đầu là nội dung của Stack, phần sau là chuỗi nhập chưa phân tích:

$$(s_0X_1s_1X_2s_2 \ldots X_ms_m, a_ia_{i+1}\ldots a_n\$)$$

Cấu tạo bộ phân tích LR

- **Bảng phân tích** bao gồm 2 phần: hàm action và hàm goto:
 - **action**[s_m, a_i] có thể có một trong 4 giá trị :
 1. **shift** s : đẩy s , trong đó s là một trạng thái.
 2. **reduce** $A \rightarrow \beta$: thu gọn bằng luật sinh $A \rightarrow \beta$.
 3. **accept**: Chấp nhận
 4. **error**: Báo lỗi
 - **Goto** lấy 2 tham số là một trạng thái và một ký hiệu văn phạm, nó sinh ra một trạng thái.

Cấu hình

- Với s_m là ký hiệu nằm trên đỉnh Stack, a_i là ký hiệu nhập hiện tại thì cấu hình có được tại mỗi bước:
 - Nếu **action** $[s_m, a_i] = \text{Shift } s$: Thực hiện phép đẩy để được cấu hình mới:

$$(s_0 X_1 s_1 X_2 s_2 \dots X_m s_m a_i s, a_{i+1} \dots a_n \$)$$

- Nếu **action** $[s_m, a_i] = \text{Reduce}(A \rightarrow \beta)$ thì thực hiện phép thu gọn để được cấu hình:

$$(s_0 X_1 s_1 X_2 s_2 \dots X_{m-i} s_{m-i} A s, a_i a_{i+1} \dots a_n \$)$$

Trong đó: $s = \text{goto}[s_{m-i}, A]$

- Nếu **action** $[s_m, a_i] = \text{accept}$: quá trình phân tích kết thúc.
- Nếu **action** $[s_m, a_i] = \text{error}$: gọi thủ tục phục hồi lỗi.

Giải thuật LR

- **Nhập:** chuỗi nhập w , bảng phân tích action goto của văn phạm G (*giả sử đã có*).
- **Xuất:** nếu w thuộc $L(G)$, nó tạo ra sự phân tích từ dưới lên. Ngược lại, bộ phân tích sẽ báo lỗi.
- **Phương pháp:**
 - Thời điểm ban đầu stack có trạng thái s_0 .
 - Chuỗi $w\$$ nằm trên bộ đệm nhập.
 - Bộ phân tích đặt đầu đọc (con trỏ ip) vào ký hiệu nhập đầu tiên của w .

Giải thuật LR

`c:=false; /*c là biến luận lý, báo cho biết quá trình phân tích kết thúc*/`

repeat

- Đặt s là trạng thái trên đỉnh stack a là ký hiệu nhập được ip chỉ đến

if *action* [s, a] = *shift*(s') **then begin**

(a) đẩy a lên stack (b) sau đó đẩy s' lên đỉnh stack

(c) chuyển ip sang ký hiệu nhập kế tiếp; **end**

else if *action* [s, a] = *reduce*($A \rightarrow \beta$) **then**

begin

(a) đẩy ($2 * |\beta|$) ký hiệu ra khỏi stack, s' là trạng thái trên đỉnh stack

(b) Tìm $j = \text{goto } [s', A]$; (c) đẩy A và sau đó là j lên đỉnh stack; (d) xuất luật $A \rightarrow \beta$

end

else if *action* [s, a] = *accept* **then** $c := \text{true}$

else *error* ();

until c ;

Ví dụ

Cho văn phạm G

$$(1) E \rightarrow E + T$$

$$(2) E \rightarrow T$$

$$(3) T \rightarrow T * F$$

$$(4) T \rightarrow F$$

$$(5) F \rightarrow (E)$$

$$(6) F \rightarrow id$$

Phân tích câu $w = id * id + id$

Bảng phân tích cho văn phạm ví dụ

Trong đó:

si : chuyển trạng thái i

ri : thu gọn bởi luật sinh i

acc: accept (chấp nhận)

error : khoảng trống

<i>Trạng thái</i>	<i>action</i>						<i>goto</i>		
	id	+	*	()	\$	E	T	F
0	s ₅			s ₄			1	2	3
1		s ₆				acc			
2		r ₂	s ₇		r ₂	r ₂			
3		r ₄	r ₄		r ₄	r ₄			
4	s ₅			s ₄			8	2	3
5		r ₆	r ₆		r ₆	r ₆			
6	s ₅			s ₄				9	3
7	s ₅			s ₄					10
8		s ₆			s ₁₁	<u>s₁₁</u>			
9		r ₁	s ₇		r ₁	r ₁			
10		r ₃	r ₃		r ₃	r ₃			
11		r ₅	r ₅		r ₅	r ₅			

Các bước chuyển trạng thái trên stack và nội dung bộ đệm nhập

$$w = id * id + id$$

STACK	INPUT	ACTION
(1) 0	id * id + id \$	Shift
(2) 0 id 5	* id + id \$	Reduce by $F \rightarrow id$
(3) 0 F 3	* id + id \$	Reduce by $T \rightarrow F$
(4) 0 T 2	* id + id \$	Shift
(5) 0 T 2 * 7	id + id \$	Shift
(6) 0 T 2 * 7 id 5	+ id \$	Reduce by $F \rightarrow id$
(7) 0 T 2 * 7 F 10	+ id \$	Reduce by $T \rightarrow T * F$
(8) 0 T 2	+ id \$	Reduce by $E \rightarrow T$
(9) 0 E 1	+ id \$	Shift
(10) 0 E 1 + 6	id \$	Shift
(11) 0 E 1 + 6 id 5	\$	Reduce by $F \rightarrow id$
(12) 0 E 1 + 6 F 3	\$	Reduce by $T \rightarrow F$
(13) 0 E 1 + 6 T 9	\$	Reduce by $E \rightarrow E + T$
(14) 0 E 1	\$	Thành công

Bài tập

- Xây dựng bước chuyển trạng thái trên stack và bộ đệm cho chuỗi nhập (với cùng văn phạm ở ví dụ trên):

$$w = (id + id) * id + id$$

4.3 Xây dựng bảng phân tích SLR

- **Định nghĩa:** thực thể LR (0) gọi tắt là thực thể của văn phạm G là luật sinh của G với các điểm chấm ở các vị trí nào đó của vế phải.
- **Thí dụ:** G có luật sinh $A \rightarrow XYZ$, sẽ cho bốn thực thể:

$A \rightarrow \bullet XYZ$

$A \rightarrow X \bullet YZ$

$A \rightarrow XY \bullet Z$

$A \rightarrow XYZ \bullet$

Nếu $A \rightarrow \in$ sẽ cho ta thực thể $A \rightarrow \bullet$

Giải thuật tính bao đóng–Closure.

```
Function closure (I : item) : item;  
begin J := I;  
  repeat  
    for với mỗi thực thể  $A \rightarrow a \bullet B \beta$  trong J và với mỗi luật  
    sinh  
     $B \rightarrow \gamma$  trong G sao cho  
    thực thể  $B \rightarrow \bullet \gamma$  chưa có trong J do  
    thêm  $B \rightarrow \bullet \gamma$  vào J;  
  until không thể thêm thực thể mới vào J;  
  closure := J;  
end;
```

Ví dụ

- Xét văn phạm:

$$E' \rightarrow E$$

$$E \rightarrow E + T \mid T$$

$$T \rightarrow T * F \mid F$$

$$F \rightarrow (E) \mid \mathbf{id}$$

Ví dụ

Nếu I là tập hợp chỉ gồm văn phạm $\{ E' \rightarrow \bullet E \}$ thì $\text{closure}(I)$ bao gồm:

$$E' \rightarrow \bullet E$$

$$E \rightarrow \bullet E + T$$

$$E \rightarrow \bullet T$$

$$T \rightarrow \bullet T * F$$

$$T \rightarrow \bullet F$$

$$F \rightarrow \bullet (E)$$

$$F \rightarrow \bullet \text{id}$$

Giải thuật tính goto

- $\text{Goto}(I, X)$, trong đó I là một tập các mục và X là một ký hiệu văn phạm, là bao đóng của tập hợp các mục $A \rightarrow \alpha X \bullet \beta$ sao cho $A \rightarrow \alpha \bullet X \beta \in I$.
- Cách tính $\text{goto}(I, X)$:
 1. Tạo một tập $I' = \emptyset$.
 2. Nếu $A \rightarrow \alpha \bullet X \beta \in I$ thì đưa $A \rightarrow \alpha X \bullet \beta$ vào I' , tiếp tục quá trình này cho đến khi xét hết tập I .
 3. $\text{Goto}(I, X) = \text{closure}(I')$

Ví dụ

- Giả sử $I = \{ E' \rightarrow E \bullet, E \rightarrow E \bullet + T \}$.

Tính goto ($I, +$) ?

- Ta có $I' = \{ E \rightarrow E + \bullet T \}$

(goto ($I, +$) = closure(I') bao gồm các mục :

$E \rightarrow E + \bullet T$ (Luật 1)

$T \rightarrow \bullet T * F$ (Luật 2)

$T \rightarrow \bullet F$ (Luật 2)

$F \rightarrow \bullet (E)$ (Luật 2)

$F \rightarrow \bullet id$ (Luật 2)

Giải thuật tính tập tuyến các tập thực thể

Procedure items (G');

begin

$C := \{\text{closure}(\{S' \rightarrow \bullet S\})\}$

repeat

for với mỗi tập thực thể I trong C và với
mỗi ký hiệu văn phạm X sao cho phép goto(I ,
 X) không rỗng và không có trong C do

thêm goto(I , X) vào C ;

until không thể thêm tập thực thể mới vào C ;

end;

Ví dụ

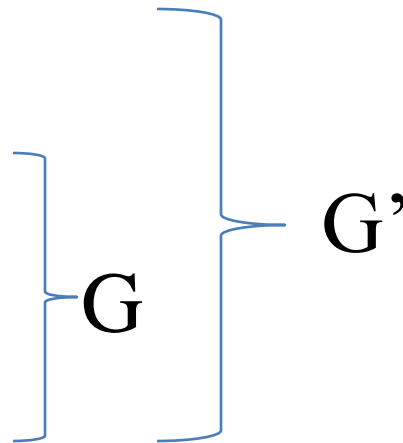
- Xét văn phạm:

$$E' \rightarrow E$$

$$E \rightarrow E + T \mid T$$

$$T \rightarrow T * F \mid F$$

$$F \rightarrow (E) \mid \mathbf{id}$$



- $C := \text{Closure}(\{E' \rightarrow \bullet E\})$: 

$$I_0: E' \rightarrow \bullet E$$

$$E \rightarrow \bullet E + T$$

$$E \rightarrow \bullet T$$

$$T \rightarrow \bullet T * F$$

$$T \rightarrow \bullet F$$

$$F \rightarrow \bullet (E)$$

$$F \rightarrow \bullet \mathbf{id}$$

Ví dụ(2)

Goto (I_0 , E) I_1 : $E' \rightarrow E \bullet$
 $E \rightarrow E \bullet + T$

Goto (I_0 , T) I_2 : $E \rightarrow T \bullet$
 $T \rightarrow T \bullet * F$

Goto (I_0 , F) I_3 : $T \rightarrow F \bullet$

Goto (I_0 , () I_4 : $F \rightarrow (\bullet E)$
 $E \rightarrow \bullet E + T$
 $E \rightarrow \bullet T$
 $T \rightarrow \bullet T * F$
 $T \rightarrow \bullet F$
 $F \rightarrow \bullet (E)$
 $F \rightarrow \bullet id$

Goto (I_0 , id) I_5 : $F \rightarrow id \bullet$

Ví dụ (3)

Goto (I_1 , +)

I_6 : $E \rightarrow E + \bullet T$

$T \rightarrow \bullet T * F$

$T \rightarrow \bullet F$

$F \rightarrow \bullet (E)$

$F \rightarrow \bullet id$

Goto (I_2 , *)

I_7 : $T \rightarrow T * \bullet F$

$F \rightarrow \bullet (E)$

$F \rightarrow \bullet id$

Goto (I_4 , E)

I_8 : $T \rightarrow (E \bullet)$

$E \rightarrow E \bullet + T$

Goto (I_6 , T)

I_9 : $E \rightarrow E + T \bullet$

$T \rightarrow T \bullet * F$

Goto (I_7 , F)

I_{10} : $T \rightarrow T * F \bullet$

Goto (I_8 ,)

I_{11} : $F \rightarrow (E) \bullet$

Xây dựng bảng phân tích

- Nhập: văn phạm gia tố G'
- Xuất: bảng phân tích SLR với hàm action và goto cho văn phạm G'
- Phương pháp:

1. Xây dựng $C = \{I_0, I_1, \dots, I_n\}$.

2. i là trạng thái đại diện cho tập thực thể I_i .

2.1. Nếu $A \rightarrow \alpha \bullet a \beta$ là thực thể ở trong I_i và $\text{goto}(I_i, a) = I_j$ thì phần tử $\text{action}[i, a] = \text{shift}(j)$, với a phải là ký hiệu kết thúc.

2.2. Nếu $A \rightarrow \alpha \bullet$ ở trong I_i thì $\text{action}[i, a] = \text{reduce}(A \rightarrow \alpha)$ với a là tất cả các ký hiệu nằm trong $\text{follow}(A)$. A không phải là S' (ký hiệu mục tiêu mới).

2.3. Nếu $S' \rightarrow S \bullet$ ở trong I_i thì $\text{action}[i, \$] = \text{accept}$.

Xây dựng bảng phân tích (2)

3. Cho tất cả các ký hiệu không kết thúc A. Nếu $\text{goto}[\mathbf{Ii}, \mathbf{A}] = \mathbf{Ij}$ thì hàm $\text{goto}[\mathbf{i}, \mathbf{A}] = \mathbf{j}$.
4. Tất cả các phần tử của bảng phân tích không được xác định bằng quy tắc 2 và 3, chúng ta coi là lỗi.
5. Trạng thái bắt đầu của bộ phân tích là tập thực thể có chứa thực thể $S' \rightarrow \bullet S$.

Ví dụ xây dựng bảng phân tích

- Xét văn phạm:

$$E' \rightarrow E$$

$$E \rightarrow E + T \mid T$$

$$T \rightarrow T * F \mid F$$

$$F \rightarrow (E) \mid \mathbf{id}$$

$$(0) \quad E' \rightarrow E$$

$$(1) \quad E \rightarrow E + T$$

$$(2) \quad E \rightarrow T$$

$$(3) \quad T \rightarrow T * F$$

$$(4) \quad T \rightarrow F$$

$$(5) \quad F \rightarrow (E)$$

$$(6) \quad F \rightarrow \mathbf{id}$$

$$1. C = \{ I_0, I_1, \dots, I_{11} \}$$

$$2. \text{FOLLOW}(E) = \{ +,), \$ \}$$

$$\text{FOLLOW}(T) = \{ *, +,), \$ \}$$

$$\text{FOLLOW}(F) = \{ *, +,), \$ \}$$

Ví dụ xây dựng bảng phân tích(1)

- Trước tiên xét tập mục I_0 : Mục $F \rightarrow \bullet (E)$ cho ra $\text{action}[0, (] = \text{"shift 4"}$, và mục $F \rightarrow \bullet \mathbf{id}$ cho $\text{action}[0, \mathbf{id}] = \text{"shift 5"}$. Các mục khác trong **I_0** không sinh được hành động nào.
- Bây giờ xét **I_1** : Mục $E' \rightarrow E \bullet$ cho $\text{action}[1, \$] = \text{"accept"}$, mục $E \rightarrow E \bullet + T$ cho $\text{action}[1, +] = \text{"shift 6"}$.

Ví dụ xây dựng bảng phân tích(2)

- Kế đến xét **I2** : $E \rightarrow T \cdot$
 $T \rightarrow T \cdot * F$
- Vì $FOLLOW(E) = \{+,), \$\}$, làm cho $action[2, \$] = action[2,+] = action[2,)] = \text{"reduce 2"}$. Mục thứ hai làm cho $action[2,*] = \text{"shift 7"}$.
- Tiếp tục theo cách này, ta thu được bảng phân tích cú pháp SLR đã trình bày.

Tổng kết Bài 4

- Các kiến thức cần nhớ:
 - Phân tích từ trên xuống
 - Phân tích dự đoán
 - Phân tích từ dưới lên

Bài học phần sau

Bài 5: Phân tích ngữ nghĩa

Thảo luận

