

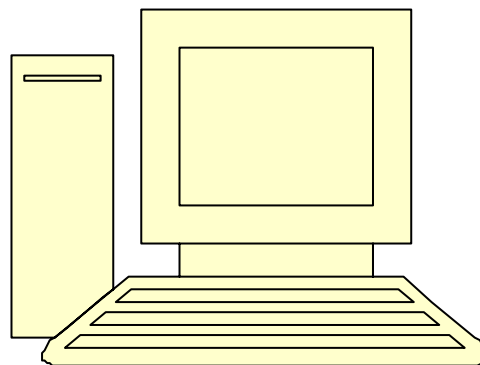


ĐẠI HỌC ĐÀ NẴNG

TRƯỜNG ĐẠI HỌC BÁCH KHOA

KHOA CÔNG NGHỆ THÔNG TIN

CHƯƠNG TRÌNH DỊCH



Giới thiệu

Mục tiêu giáo trình

1. Cung cấp những kiến thức cơ bản về chương trình dịch
2. Cung cấp các phương pháp phân tích từ vựng, phân tích cú pháp.
3. Cơ sở cho việc tìm hiểu các ngôn ngữ lập trình.
4. Rèn luyện kỹ năng lập trình cho sinh viên



Giới thiệu

Nội dung giáo trình

CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

CHƯƠNG 5. PHÂN TÍCH NGŨ NGHĨA

CHƯƠNG 6. XỬ LÝ LỖI VÀ SINH MÃ



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

1. Các khái niệm cơ bản
2. Đặc trưng của ngôn ngữ lập trình (NNLT) bậc cao
3. Các qui tắc từ vựng và cú pháp
4. Các chức năng của một trình biên dịch



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

1. Các khái niệm cơ bản

1.1. Sự phát triển của ngôn ngữ lập trình

1.2. Khái niệm chương trình dịch

1.3. Phân loại chương trình dịch

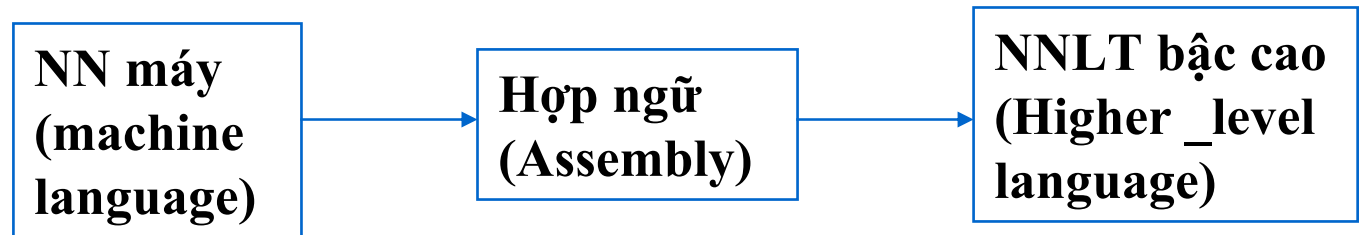
1.4. Các ứng dụng khác của kỹ thuật dịch



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

1. Các khái niệm cơ bản

1.1. Sự phát triển của ngôn ngữ lập trình



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

1. Các khái niệm cơ bản

1.2. Khái niệm chương trình dịch

Chương trình dịch là chương trình dùng để dịch một chương trình (CT nguồn) viết trên NNLT nào đó (NN nguồn) sang một chương trình tương đương (CT đích) trên một NN khác (NN đích)

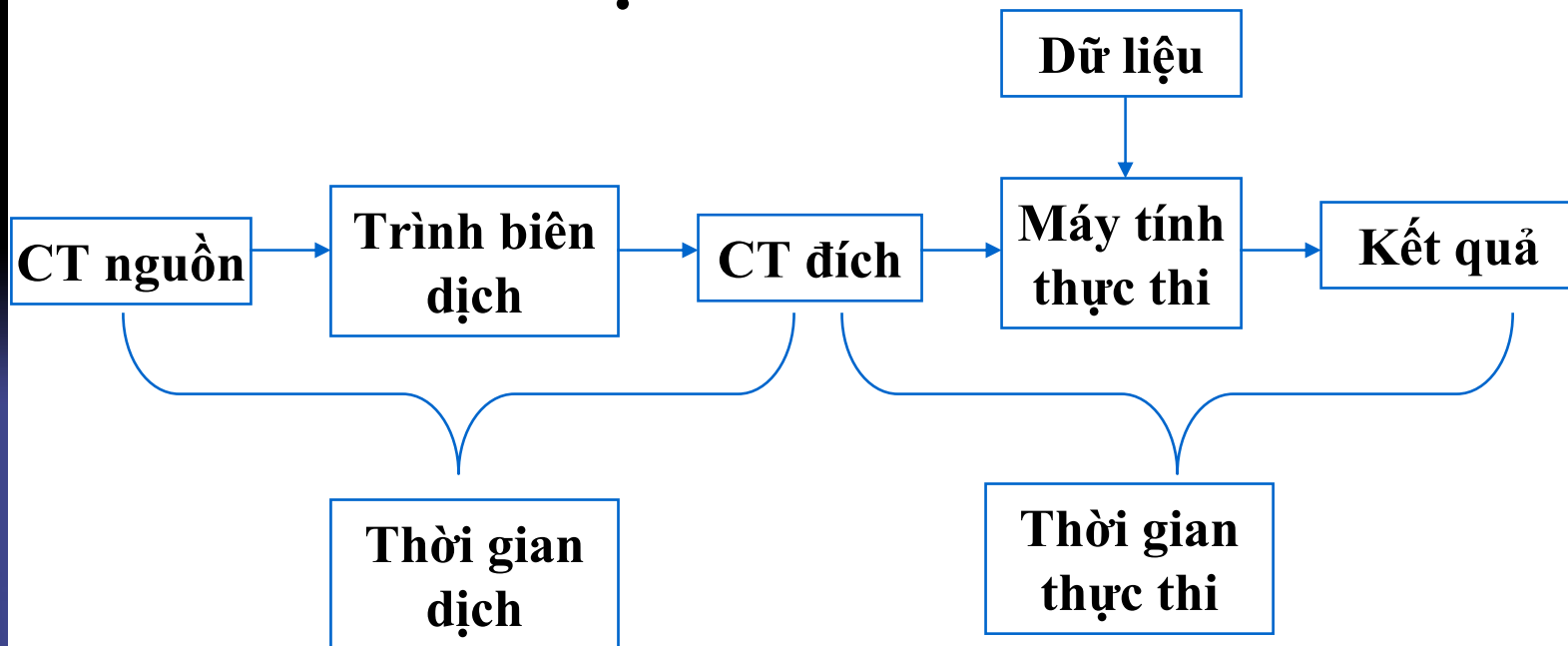


CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

1. Các khái niệm cơ bản

1.3. Phân loại chương trình dịch

❖ Trình biên dịch

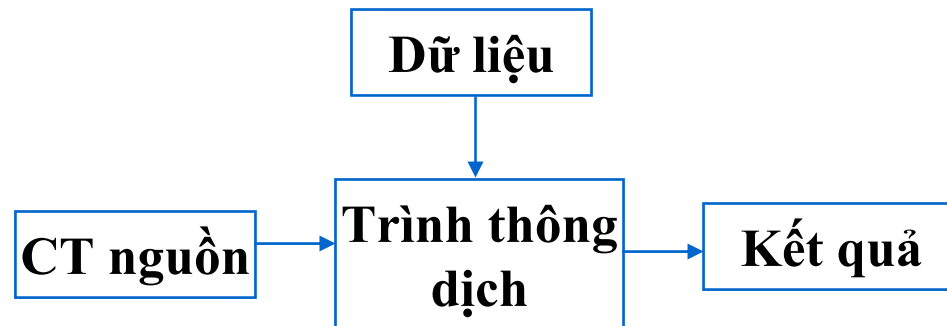


CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

1. Các khái niệm cơ bản

1.3. Phân loại chương trình dịch

❖ Trình thông dịch



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

1. Các khái niệm cơ bản

1.4. Các ứng dụng khác của kỹ thuật dịch

- Trong các hệ thống: phần giao tiếp giữa người và máy thông qua các câu lệnh.
- Hệ thống xử lý NN tự nhiên: dịch thuật, tóm tắt văn bản.



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

2. Đặc trưng của NNLT bậc cao

- Tính tự nhiên
- Tính thích nghi
- Tính hiệu quả
- Tính đa dạng



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các quy tắc từ vựng và cú pháp

3.1. Bản chữ cái

- Gồm những ký hiệu được phép sử dụng để viết chương trình
- Số lượng, ý nghĩa sử dụng của các ký tự trong bản chữ cái của các NN là khác nhau.
- Nhìn chung bản chữ cái của các NNLT:
 - + 52 chữ cái: $A \rightarrow Z, a \rightarrow z$
 - + 10 chữ số: $0 \rightarrow 9$
 - + Các ký hiệu khác: $*, /, +, -, \dots$



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các quy tắc từ vựng và cú pháp

3.2. Từ tổ (Token)

- Từ tổ là đơn vị nhỏ nhất có nghĩa
- Từ tổ được xây dựng từ bản chữ cái
- Ví dụ: hằng, biến, từ khoá, các phép toán,...



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các qui tắc từ vựng và cú pháp

3.3. Phạm trù cú pháp

- Phạm trù cú pháp là một dãy từ tổ kết hợp theo một qui luật nào đó
- Các cách biểu diễn cú pháp thông thường

+ BNF(Backus Naus Form):

$\langle \text{lệnhgán} \rangle ::= \langle \text{tên biến} \rangle := \langle \text{biểu thức} \rangle$



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các qui tắc từ vựng và cú pháp

3.3. Phạm trù cú pháp

+ Biểu đồ cú pháp:

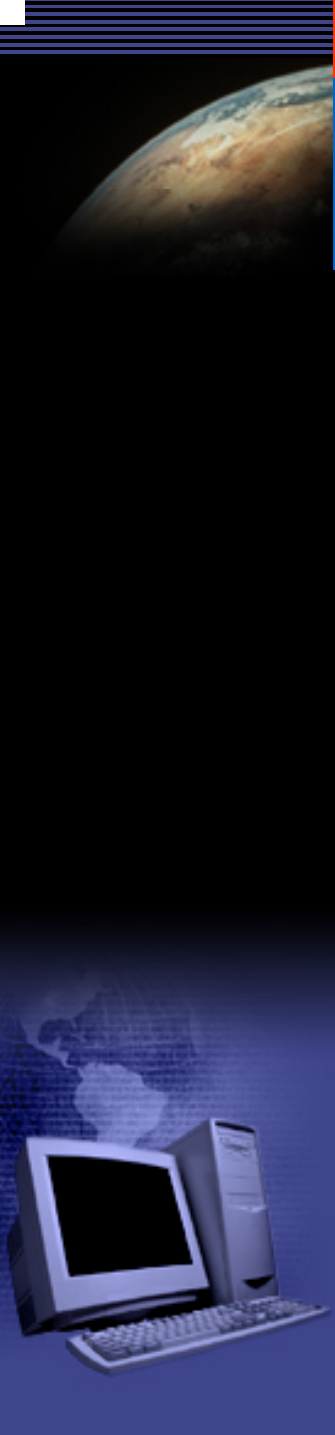
Chương trình → Program → Danh biểu → Khối

Khối → - var...

- procedure → Danh biểu → Khối

- begin → lệnh → end →.

- Mục tiêu của phạm trù cú pháp là việc định nghĩa được khái niệm chương trình đến mức độ tự có



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các qui tắc từ vựng và cú pháp

3.4. Các qui tắc từ vựng thông dụng

- Cách sử dụng khoảng trống(dấu trắng), dấu tab(‘\t’), dấu sang dòng(‘\n’)
- Đối với liên kết tự do, có thể sử dụng nhiều khoảng trống thay vì một khoảng trống.



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các qui tắc từ vựng và cú pháp

3.4. Các qui tắc từ vựng thông dụng

- Một khoảng trống là bắt buộc giữa các từ tố: từ khoá và tên,...

Ví dụ: `program tenct;`

- Khoảng trống không bắt buộc: số và các phép toán, tên biến và các phép toán

Ví dụ: `x:=x+3*3;`

- Cách sử dụng chú thích và xâu ký tự



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các qui tắc từ vựng và cú pháp

3.5. Modun hoá và chuyển giao dữ liệu

- Modun hoá là khả năng tách một công việc lớn thành hệ thống những công việc nhỏ phân cấp.
- Có 2 hình thức chuyển giao dữ liệu:

❖ Dùng chung:

- Dữ liệu được khai báo ở cấp cao hơn có thể được xử lý ở cấp thấp hơn (các biến cục bộ của hệ thống ctc Pascal)
- Khai báo những dữ liệu dùng chung cho các modun (các biến chung của C)



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

3. Các qui tắc từ vựng và cú pháp

3.5. Modun hoá và chuyển giao dữ liệu

- ❖ Truyền tham số giữa CT gọi và CT được gọi
 - Truyền theo tham biến
 - Truyền theo tham trị



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

- Phân tích từ vựng
- Phân tích cú pháp
- Phân tích ngữ nghĩa
- Xử lý lỗi
- Sinh mã trung gian
- Tối ưu mã trung gian
- Sinh mã đối tượng



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.1. Phân tích từ vựng

- CT nguồn là một dãy các ký tự.
- Phân tích từ vựng là phân tích CT nguồn thành các từ tố (Token).
- Các Token này sẽ là dữ liệu đầu vào của phân tích cú pháp.



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.2. Phân tích cú pháp

- Đầu vào sẽ là dãy các Token nối nhau bằng một qui tắc nào đó.
- Phân tích xem các Token có tuân theo qui tắc cú pháp của ngôn ngữ không



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.3. Phân tích ngữ nghĩa

- Kiểm tra tính hợp lệ của các phép toán và các phép xử lý

- Ví dụ:

- Biến phải khai báo trước khi sử dụng (Pascal)
- Kiểm tra tính tương thích kiểu dữ liệu của biến và biểu thức



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.4. Xử lý lỗi

- CT nguồn vẫn có thể xảy ra lỗi.
- Phần xử lý lỗi sẽ thông báo lỗi cho NSD
- Lỗi ở phần nào báo ở phần đó.



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.4. Xử lý lỗi

- Có các loại lỗi:

- Lỗi từ vựng (trong Pascal sử dụng biến mà chưa khai báo)
- Lỗi cú pháp ((a+5; lỗi thiếu dấu ‘)’)
- Lỗi ngữ nghĩa (x=3.5; nhưng khai báo int x)
- Lỗi thực hiện (phép chia 0)



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.5. sinh mã trung gian

- Sau giai đoạn phân tích ngữ nghĩa
- Mã trung gian là một dạng trung gian của CT nguồn có 2 đặc điểm:
 - Dễ được sinh ra
 - Dễ dịch sang ngôn ngữ đích



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.6. Tối ưu mã trung gian

- Bỏ bớt các lệnh thừa.
- Cải tiến lại mã trung gian để khi sinh mã đối tượng thì thời gian thực thi mã đối tượng sẽ ngắn hơn



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

4.7. Sinh mã đối tượng

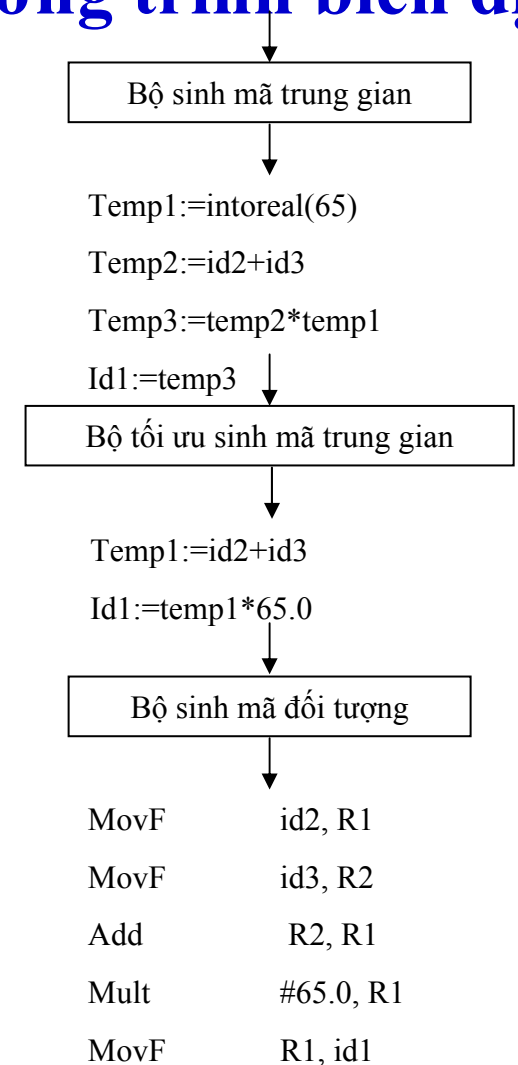
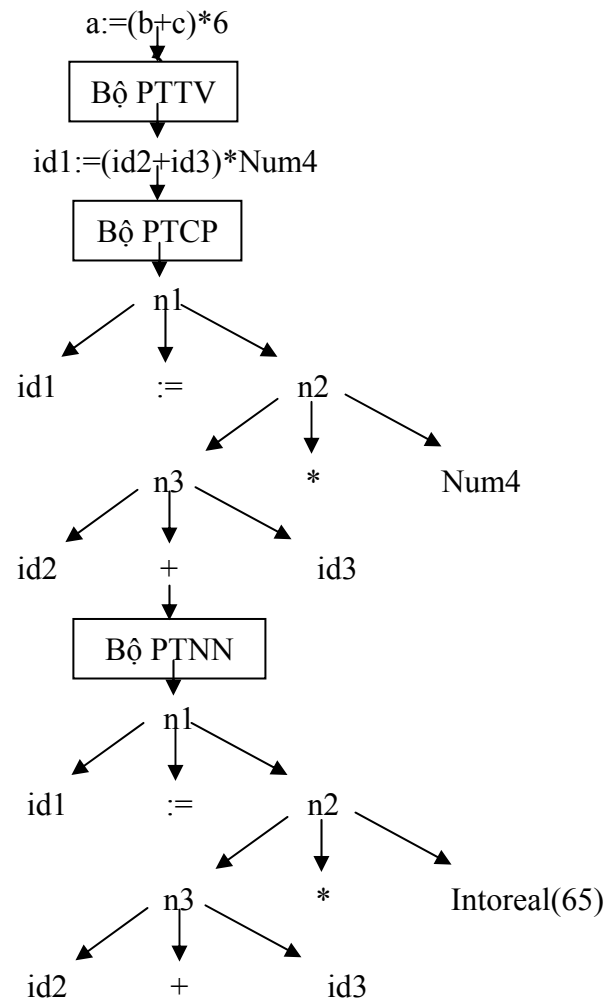
- Giai đoạn cuối của trình biên dịch.
- Mã đối tượng có thể là mã máy, hợp ngữ hay một ngôn ngữ khác ngôn ngữ nguồn.
- Các pha (giai đoạn) có thể thực hiện song hành
- Một vài pha có thể ghép lại thành lượt (chuyển
- Một lượt sẽ đọc toàn bộ CT nguồn hay một dạng trung gian của CT nguồn, sau đó ghi kết quả để lượt sau đọc và xử lý tiếp.



CHƯƠNG 1. NHẬP MÔN CHƯƠNG TRÌNH DỊCH

4. Các chức năng của một chương trình biên dịch

Ví dụ:



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

- **Mục đích**
- **Nội dung**
- **Otomat hữu hạn đơn định**
- **Bộ phân tích từ vựng**
- **Bảng danh biểu**



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

1. Mục đích

- Chia cắt xâu vào (CT nguồn) thành dãy các từ tố.
- Hai cách cài đặt
 - Sử dụng một lượt cho việc phân tích từ vựng → dãy các token → phân tích cú pháp.
 - Phân tích từ vựng dùng chung một lượt với phân tích cú pháp. Một lần chỉ phát hiện 1 token gọi là từ tố tiếp đến.



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

2. Nội dung

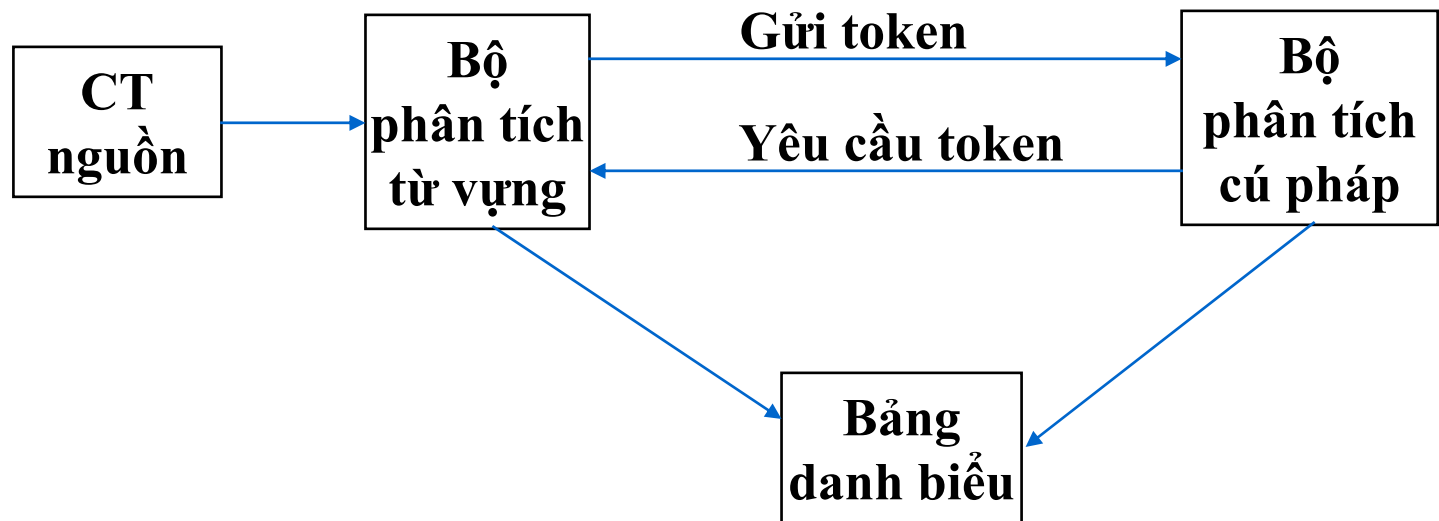
- Đọc sâu vào từng ký tự một → gom lại thành token đến khi gặp ký tự không thể kết hợp thành token.
- Luôn luôn đọc trước một ký tự.
- Loại bỏ các ký tự trống và chú thích.
- Chuyển những thông tin của những từ tố (văn bản, mã phân loại) vừa phát hiện cho bộ phân tích cú pháp.
- Phát hiện lỗi.



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

2. Nội dung

- Sự giao tiếp giữa bộ phân tích từ vựng và bộ phân tích cú pháp



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.1. Định nghĩa: $M(\Sigma, Q, \delta, q_0, F)$

Σ : bộ chữ vào

Q : tập hữu hạn các trạng thái

$q_0 \in Q$: trạng thái đầu

$F \subseteq Q$: tập các trạng thái kết thúc

δ : hàm chuyển trạng thái có dạng $\delta(q, a) = p$

Với $q, p \in Q, a \in \Sigma$

➤ $\delta(q, a) = p$: nghĩa là ở trạng thái q , đọc a , chuyển sang trạng thái p



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.2. Biểu diễn các hàm chuyển trạng thái

❖ **Dùng bảng: sử dụng ma trận δ có:**

- **Chỉ số hàng: trạng thái**
- **Chỉ số cột: ký hiệu vào**
- **Giá trị tại hàng q , cột a là trạng thái p , sao cho $\delta(q,a)=p$**



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.2. Biểu diễn các hàm chuyển trạng thái

❖ Dùng bảng:

Ví dụ: có hàm chuyển của một Otomat như sau: $\delta(1,a)=2$, $\delta(2,b)=2$, $\delta(2,c)=2$

δ	a	b	c
1	2		
2		2	2



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.2. Biểu diễn các hàm chuyển trạng thái

❖ Hình vẽ:

- mỗi trạng thái $q \in Q$ được đặt trong các vòng tròn.
- Trạng thái bắt đầu q_0 có thêm dấu ' $>$ ' ở đầu.
- Trạng thái kết thúc $q \in F$ được đặt trong vòng tròn kép.
- Các cung nối từ trạng thái q sang trạng thái p có mang các nhãn $a \in \Sigma$, có nghĩa $\delta(q, a) = p$



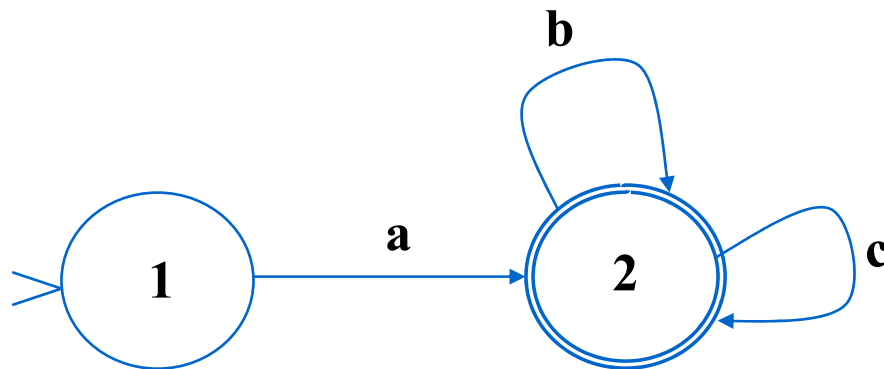
CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.2. Biểu diễn các hàm chuyển trạng thái

❖ Hình vẽ:

Ví dụ: có hàm chuyển của một Otomat như sau: $\delta(1,a)=2$, $\delta(2,b)=2$, $\delta(2,c)=2$



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.2. Biểu diễn các hàm chuyển trạng thái

❖ Nhận xét:

- Biểu diễn hàm chuyển trạng thái bằng hình vẽ có ưu điểm hơn. Trong hình vẽ ta xác định đầy đủ tất cả các thành phần của Otomat.
- Biểu diễn bằng bảng xác định hàm chuyển trạng thái, tập các trạng thái, bộ chữ vào nhưng không phân biệt được trạng thái bắt đầu và trạng thái kết thúc.



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.3. Hoạt động của Otomat

- Đọc các ký hiệu của xâu vào từ trái sang phải, bắt đầu từ trạng thái q_0 .
- Mỗi bước đọc một ký hiệu thì chuyển sang trạng thái theo δ . Có thể đọc xong hay không đọc xong xâu vào.



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.3. Hoạt động của Otomat

- Đọc xong xâu vào đến một trạng thái $p \in F$ thì xâu vào được đoán nhận (xâu đúng).
- Đọc xong xâu vào mà rơi vào trạng thái $p \notin F$ thì xâu vào không được đoán nhận.
- Không đọc xong xâu vào (do δ rơi vào điểm không xác định) thì xâu vào không được đoán nhận.



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định

3.4. Ví dụ: Xác định Otomat đoán nhận số nh phân. $M(\Sigma, Q, \delta, q_0, F)$

$\Sigma: \{0, 1, \text{trắng}\}$

$Q: \{0, 1, 2\}$

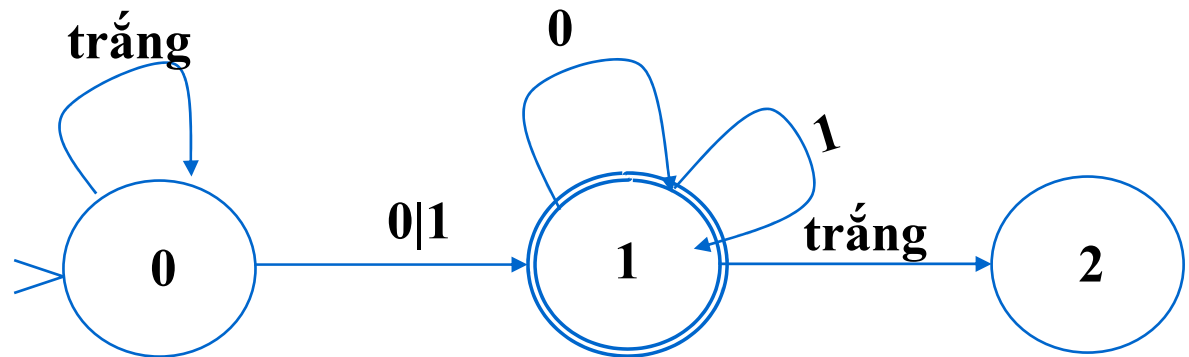
$q_0: 0$

$F : \{2\}$

$\delta: \delta(0, \text{trắng})=0, \delta(0,0)=1, \delta(0,1)=1,$
 $\delta(1,0)=1, \delta(1,1)=1, \delta(1, \text{trắng})=2$



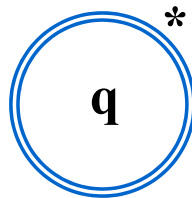
CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

3. Otomat hữu hạn đơn định**3.4. Ví dụ: Xác định Otomat đoán nhận số nhị phân**

CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

4. Lập bộ phân tích từ vựng

Ngoài các hình qui ước của Otomat thông thường lại có thêm:



Trạng thái kết thúc và
trả lui ký tự vừa đọc

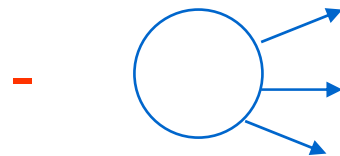


CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

4. Lập bộ phân tích từ vựng

4.1. Phương pháp mô phỏng

- **Mỗi trạng thái: tương ứng với một đoạn chương trình**
- **Nối tiếp các trạng thái: nối tiếp 2 đoạn chương trình tương ứng**



Lệnh rẽ nhánh



Lệnh lặp



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

4. Lập bộ phân tích từ vựng

4.1. Phương pháp mô phỏng

Max=10; {độ dài tối đa của 1 danh biểu}

Type

Loaikytu=(conso,cham, Ttu, trang, Ccai);

Loaituto=(nguyen,thuc,Toantu,
Danhbieu);

Xau=Array[1..max] of char ;

Var Kytutiep:char;

Procedure Dockytu(var c:char);

... {Đọc ký tự tiếp, ký tự này luôn luôn được
đọc trước}

Function LoaiKT(c:char):Loaikytu;

... {Cho biết loại của ký tự c}

Procedure Baoloi;

... {Cho một thông báo lỗi}

Procedure Tuvung(var ma:Loaituto;var x:xau);

Var i:0..max;

Begin

For i:=1 to max do x[i]:='';

I:=0;

While loaikytu(kytutiep)=trang do

 Dockytu(kytutiep);

Case loaikytu(kytutiep) of

 Conso: Begin

 Repeat I:=i+1;

 x[i]:=kytutiep;

 Dockytu(kytutiep);

 Until Loaikytu(kytutiep) <> conso;

 Ma:=nguyen;



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

4. Lập bộ phân tích từ vựng

4.1. Phương pháp mô phỏng

```

If loaikytu(kytutiep)=cham then
    Begin Repeat      I:=i+1;
                      x[i]:=kytutiep;
                      Dockytu(kytutiep);
    Until loaikytu(kytutiep)<>Conso
    Ma:=thuc;
End;
End;
Cham: Begin Baoloi;
      Repeat I:=i+1;
            x[i]:=kytutiep;
            Dockytu(kytutiep);
      Until loaikytu(kytutiep)<>conso;
      Ma:=thuc; End;

```

```

Ttu: begin I:=i+1; x[i]:=kytutiep;
        ma:=toantu; Dockytu(kytutiep);
      end;
Ccai: begin
      Repeat
        If i<max then
          Begin I:=i+1;
                x[i]:=kytutiep;
          end;
          Dockytu(kytutiep);
        Until (loaikytu(kytutiep)<>Ccai)
        and (loaikytu(kytutiep)<>conso);
        Ma:=danhbieu;
      End;
End: {case} End: {tuvung}

```

CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

4. Lập bộ phân tích từ vựng

4.1. Phương pháp điều khiển bằng bảng

Var bangchuyen: array[1..6,loaikytu] of 0..6;

Mảng này được nạp dữ liệu như sau:

	trang	Conso	Cham	Ttu	Ccai
1	1	2	4	5	6
2	0	2	3	0	0
3	0	3	0	0	0
4	0	3	0	0	0
5	0	0	0	0	0
6	0	0	0	0	6

CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

4. Lập bộ phân tích từ vựng

4.1. Phương pháp điều khiển bằng bảng

Procedure Tuvung(var ma:loaituto; var x:xau);

Begin

 trangthai:=1;

 trangthaitiep:=bangchuyen[trangthai,
 loaikytu(kytutiep)];

 i:=0;

Repeat

 i:=i+1;

 x[i]:=kytutiep;

 trangthai:=trangthaitiep;

 Dockytu(kytutiep);

 trangthaitiep:= bangchuyen[trangthai,
 loaikytu(kytutiep)];

Until trangthaitiep=0;

Case trangthai of

 2: ma:=nguyen;

 3: ma:=thuc;

 4: baoloi;

 5:ma:=toantu;

 6: ma:=danhbieu;

End; {case}

End; {Tuvung}



CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

5. Bảng danh biểu

Gồm các token và các thuộc tính của token

Chỉ số	Token	Trị từ vựng	Các thuộc tính khác
01			
02	Num	45	
03	Id	A	
04	Id	B	
05			
06	Relation	<	
07	Then	Then	
08	operator	+	

CHƯƠNG 2. PHÂN TÍCH TỪ VỰNG

- 6. Các cấu trúc dữ liệu cho bảng các danh biểu**
- Tổ chức tuần tự: mảng, danh sách liên kết, danh sách móc nối
 - Tổ chức cây tìm kiếm nhị phân



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

- Một số vấn đề về ngôn ngữ
- Văn phạm phi ngữ cảnh
- Đại cương về phân tích cú pháp
- Các phương pháp phân tích cú pháp



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.1. Xâu

- Bộ chữ (bảng chữ) là tập hợp hữu hạn các ký hiệu

Ví dụ: $\{0,1\}$ bộ chữ gồm 2 ký hiệu 0 và 1

$\{a,b,c,\dots,z\}$ bộ chữ gồm các ký hiệu $a \rightarrow z$

Tập các chữ cái tiếng việt



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.1. Xâu

- Xâu trên bộ chữ V là 1 dãy các ký hiệu của V

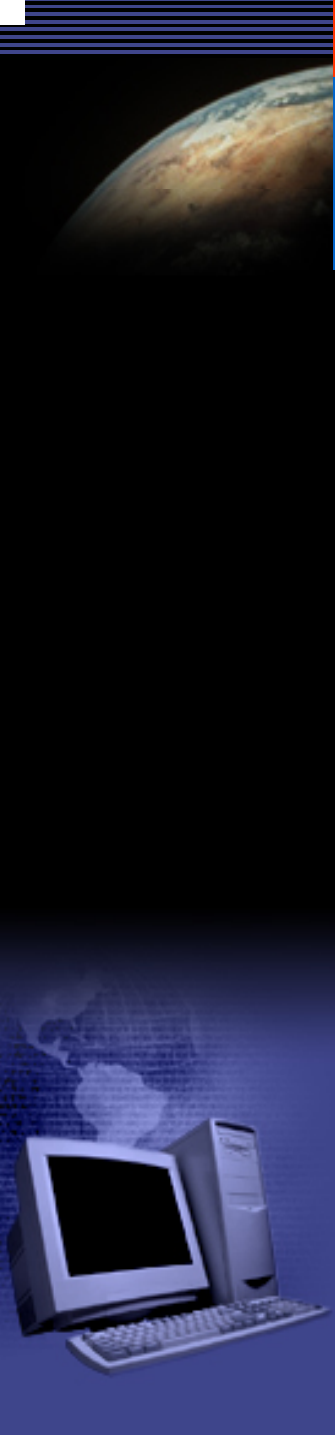
Ví dụ: 0110 là xâu trên bộ chữ $\{0,1\}$

a, ab, giathanh là xâu trên bộ chữ $\{a,b,\dots,z\}$

- Độ dài xâu là số các ký hiệu trong xâu

Ký hiệu: độ dài xâu x là $|x|$

Ví dụ: $|01110|=5$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.1. Xâu

- Xâu rỗng là xâu có độ dài bằng 0

Ký hiệu: $\epsilon, |\epsilon|=0$

- Tập tất cả các xâu trên V là V^* , $\{\epsilon\} \subseteq V^*$

$$V^+ = V^* - \{\epsilon\}$$

V^* : tập vô hạn đếm được

Ví dụ: $V=\{a,b\} \rightarrow V^*=\{\epsilon,a,b,aa,bb,ab,ba,\dots\}$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.1. Xâu

- Các phép toán trên xâu
- Ghép tiếp: cho 2 xâu x, y . Ghép tiếp của x, y là $x.y$ hay xy là 1 xâu viết x trước, rồi đến y sau chứ không có dấu cách.

Ví dụ: $x=01$

$y=0110$

$xy=010110$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.1. Xâu

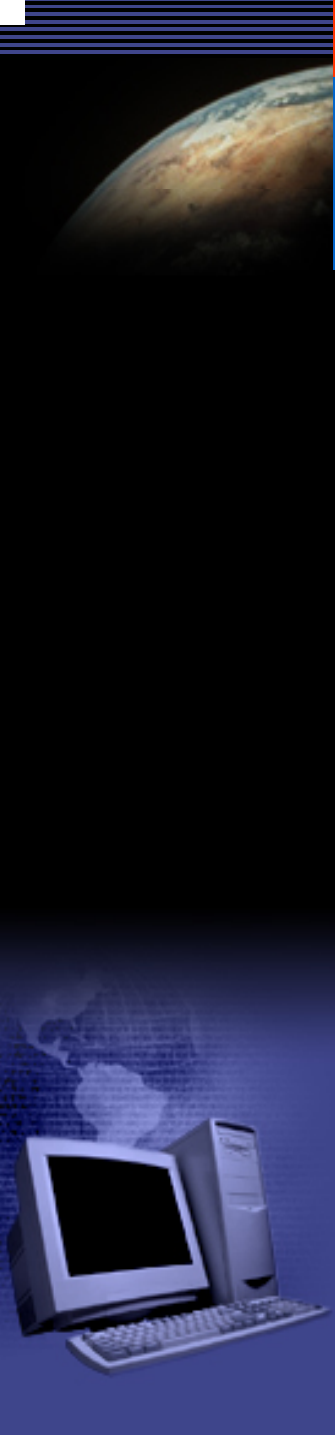
- Đảo ngược xâu x (x^r): xâu được viết theo thứ tự ngược lại của xâu x

Ví dụ: $x=0101 \rightarrow x^r=1010$

Chú ý: $\varepsilon^r = \varepsilon, 1^r = 1$

- Xâu x mà $x=x^r$ thì x là xâu hình tháp (xâu đối xứng)

Ví dụ: $x=0110 \rightarrow x^r=0110$, x : xâu hình tháp



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.2. Ngôn ngữ

- Ngôn ngữ L trên bộ chữ V là tập hợp các xâu trên V , $L \subseteq V^*$
- Các phép toán trên ngôn ngữ
- Vì ngôn ngữ là tập hợp nên có các phép toán tập hợp: \cap (giao), \cup (hợp), $-$ (hiệu, bù)



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.2. Ngôn ngữ

- Ghép tiếp 2 ngôn ngữ

Cho 2 ngôn ngữ L_1, L_2 . Ta gọi ghép tiếp $L_1.L_2$ (L_1L_2) của L_1 và L_2 là một tập hợp $L_1L_2 = \{xy / (x \in L_1) \text{ và } (y \in L_2)\}$

$$x.x = x^2; x.x.x = x^3; x^0 = \varepsilon; x^i = x^{i-1}.x$$

$$L^0 = \{\varepsilon\}; L^i = L^{i-1}.L$$

- $L^* = L^0 \cup L^1 \cup L^2 \cup \dots \cup \varepsilon; L^+ = L^1 \cup L^2 \cup \dots \cup$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.3. Biểu diễn ngôn ngữ

❖ Ngôn ngữ đơn giản

- Phương pháp liệt kê: ngôn ngữ có số xâu là hữu hạn và có thể xác định được.

Ví dụ: ngôn ngữ là các số tự nhiên nhỏ hơn 20 và lớn hơn 12

$$L=\{13, 14, 15, 16, 17, 18, 19\}$$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

1. Một số vấn đề về ngôn ngữ

1.3. Biểu diễn ngôn ngữ

❖ Ngôn ngữ đơn giản

- Phương pháp sử dụng tân từ $P(x)$: ngôn ngữ mà các xâu có cùng các đặc điểm.

Ví dụ: ngôn ngữ là các số thực nhỏ hơn 5.

$$L = \{x / (x \in \mathbb{R}) \text{ và } (x < 5)\}$$

❖ Ngôn ngữ phức tạp

Văn phạm: cơ chế để sản sinh ra ngôn ngữ



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.1. Định nghĩa: $G=(\Sigma, \Delta, s, p)$ trong đó:

Σ : tập hữu hạn các ký hiệu kết thúc.

Δ : tập hữu hạn các ký hiệu chưa kết thúc.

s : ký hiệu bắt đầu; $s \in \Delta$

p : tập hữu hạn các sản xuất có dạng
 $A \rightarrow \alpha$ với $A \in \Delta$ và $\alpha \in (\Sigma \cup \Delta)^*$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.2. Ví dụ: $G=(\Sigma, \Delta, s, p)$ trong đó:

$\Sigma: \{0,1\}$

$\Delta: \{S\}$

$s: S$

$p: S \rightarrow 0S \mid 1S \mid 0 \mid 1$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

❖ Qui ước:

- Ký hiệu kết thúc được viết bằng chữ thường
- Ký hiệu chưa kết thúc được viết bằng chữ in
- Ký hiệu chưa kết thúc nằm bên trái của sản xuất đầu tiên là ký hiệu bắt đầu.



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ **Xâu (câu) và dạng câu:**

- α gọi là **xâu** khi $\alpha \in \Sigma^*$
- α gọi là **dạng câu** khi $\alpha \in (\Sigma \cup \Delta)^*$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Quan hệ suy dẫn:

- A có quan hệ suy dẫn ra α hay α được suy dẫn từ A, có nghĩa là từ A áp dụng các sản xuất sinh ra được α
- Quan hệ suy dẫn trực tiếp: từ A áp dụng một sản xuất sinh được α

Ký hiệu: $A \Rightarrow \alpha$ với $A \in \Delta$ và $\alpha \in (\Sigma \cup \Delta)^*$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Quan hệ suy dẫn:

- Quan hệ suy dẫn nhiều lần: từ A áp dụng nhiều sản xuất mới sinh được α

Ký hiệu: $A \Rightarrow^+ \alpha$ với $A \in \Delta$ và $\alpha \in (\Sigma \cup \Delta)^*$

- Độ dài suy dẫn: số lần áp dụng các sản xuất
- Độ dài của suy dẫn trực tiếp bằng 1



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Quan hệ suy dẫn:

- Nếu luôn luôn thay thế ký hiệu chưa kết thúc ở bên trái nhất gọi là suy dẫn trái. Tương tự ta có suy dẫn phải



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

- **Cây suy dẫn: cây thoả mãn các điều kiện:**
 - **Mỗi nút có 1 nhãn: ký hiệu kết thúc hoặc chưa kết thúc**
 - **Nhãn của nút gốc: ký hiệu bắt đầu**
 - **Nhãn của nút lá: ký hiệu kết thúc**
 - **Nếu một nút có nhãn A có các nút con của nó từ trái sang phải có nhãn $x_1, x_2, x_3, \dots, x_n$ thì $A \rightarrow x_1x_2x_3\dots x_n \in p$**



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Cây suy dẫn

- Suy dẫn trái tạo cây suy dẫn trái.
- Suy dẫn phải tạo cây suy dẫn phải.
- Ví dụ: cho văn phạm phi ngữ cảnh sau:

$$E \rightarrow E^{(1)}+E \mid E^{(2)}*E \mid (E)^{(3)} \mid a^{(4)}$$

Vẽ cây suy dẫn trái, phải sinh xâu: $a+a*a$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Văn phạm đơn nghĩa

Văn phạm $G=(\Sigma, \Delta, s, p)$ sản sinh ra ngôn ngữ $L(G)=\{w \in \Sigma^*\}$. Ta nói G là văn phạm đơn nghĩa (không nhập nhằng) nếu với mỗi chuỗi $w \in L(G)$ chỉ có một cây suy dẫn duy nhất, trái lại thì G là văn phạm nhập nhằng.



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Văn phạm tương đương

Văn phạm G_1 và G_2 được gọi là tương đương \Leftrightarrow bất kỳ chuỗi x được sinh ra từ G_1 thì G_2 cũng sinh ra được và ngược lại



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Văn phạm đệ qui

Cho văn phạm PNC G , với $A \in \Delta$ mà $\exists A \Rightarrow^+ \alpha A \beta$ thì A gọi là ký hiệu đệ qui, G gọi là văn phạm đệ qui. Với $\alpha, \beta \in (\Sigma \cup \Delta)^*$

- Nếu $\alpha = \varepsilon$: đệ qui trái
- Nếu $\beta = \varepsilon$: đệ qui phải



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

2.3. Các khái niệm

➤ Văn phạm đệ quy

Ví dụ: $S \xrightarrow{(1)} S0 \mid \xrightarrow{(2)} S1 \mid \xrightarrow{(3)} 0 \mid \xrightarrow{(4)} 1$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

Bài tập

(1) Xác định ngôn ngữ được sản sinh bởi Văn phạm:

a. $S \rightarrow S(S)S \mid \varepsilon$

b. $S \rightarrow aSb \mid bSa \mid \varepsilon$

c. $S \rightarrow + S S \mid * S S \mid a$

d. $S \rightarrow 0S1 \mid \varepsilon$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

2. Văn phạm phi ngữ cảnh

Bài tập

(2) Xây dựng văn phạm sản sinh ra ngôn ngữ:

a. Số nhị phân lẻ

b. Số nguyên không dấu

c. Số nguyên có dấu

d. Số thực, số nguyên không và có dấu



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.1. Mục đích

Cho $G=(\Sigma, \Delta, s, p)$

Một xâu vào $x \in \Sigma^*$

x có viết đúng cú pháp
của văn phạm G ?



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.2. Phương pháp giải quyết

- Bắt đầu từ S áp dụng các sản xuất để tìm x:
PTCP từ trên xuống
- Nếu tìm được x: x viết đúng cú pháp của văn phạm G
- Nếu k0 tìm được x: x viết không đúng cú pháp của văn phạm G



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.2. Phương pháp giải quyết

- Bắt đầu từ x áp dụng các suy dẫn ngược 1 sản xuất để thu S : PTCP từ dưới lên
- Nếu thu được S : x viết đúng cú pháp của văn phạm G
- Nếu không thu được S : x viết không đúng cú pháp của văn phạm G



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.2. Phương pháp giải quyết

Ví dụ: Cho văn phạm PNC G sau:

$$\overset{(1)}{S \rightarrow B}$$

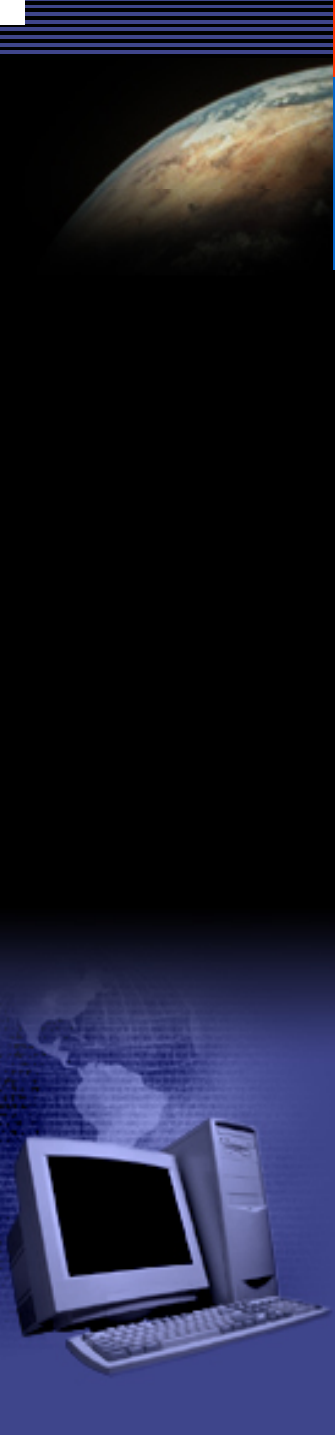
$$\overset{(2)}{B \rightarrow R} \mid \overset{(3)}{(B)}$$

$$\overset{(4)}{R \rightarrow E=E}$$

$$E \rightarrow \overset{(5)}{a} \mid \overset{(6)}{b} \mid \overset{(7)}{(E+E)}$$

Xâu x: $(a=(b+a))$

Hỏi xâu x có viết đúng cú pháp của G không?



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.2. Phương pháp giải quyết

Ví dụ:

➤ Phương pháp từ trên xuống

$$\begin{aligned}
 \underline{S} &\stackrel{(1)}{=>} \underline{B} \stackrel{(3)}{=>} (\underline{B}) \stackrel{(2)}{=>} (\underline{R}) \stackrel{(4)}{=>} (E=\underline{E}) \\
 &\stackrel{(7)}{=>} (E=(E+\underline{E})) \stackrel{(5)}{=>} (E=(\underline{E}+a)) \\
 &\stackrel{(6)}{=>} (\underline{E}=(b+a)) \stackrel{(5)}{=>} (a=(b+a)) : \text{xâu } x
 \end{aligned}$$

Vậy xâu x viết đúng cú pháp của G



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.2. Phương pháp giải quyết

Ví dụ:

➤ Phương pháp từ dưới lên

Stt	Dạng câu	Cán	Sx dùng
(0)	(<u>a</u> =(b+a))	a	$E \rightarrow a$
(1)	(E=(<u>b</u> +a))	b	$E \rightarrow b$
(2)	(E=(E+ <u>a</u>))	a	$E \rightarrow a$
(3)	(E=(E+ <u>E</u>))	(E+E)	$E \rightarrow (E+E)$

CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.2. Phương pháp giải quyết

Ví dụ:

➤ Phương pháp từ dưới lên

(4)	<u>(E=E)</u>	E=E	R→E=E
(5)	<u>(R)</u>	R	B→R
(6)	<u>(B)</u>	(B)	B→(B)
(7)	<u>B</u>	B	S→B
(8)	S		

Vậy xâu x viết đúng cú pháp của G



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.3. Sơ đồ chung giải thuật PTCP từ dưới lên

Biết α_i tìm α_{i-1}

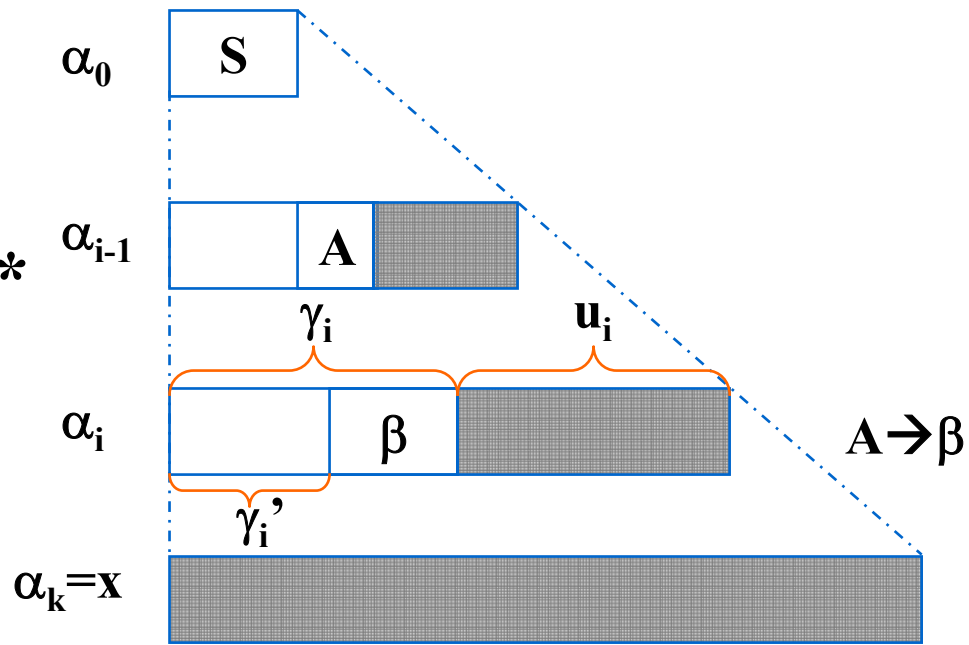
$$\alpha_i = \gamma_i u_i$$

$$\gamma_i \in (\Sigma \cup \Delta)^*; u_i \in \Sigma^*$$

$$\gamma_i = \gamma_i' \beta$$

$$\alpha_k = x = u_k; \gamma_k = \varepsilon$$

$$\alpha_0 = S = \gamma_0; u_0 = \varepsilon$$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.3. Sơ đồ chung giải thuật PTCP từ dưới lên

➤ Thuật toán:

Sử dụng: 1 stack và 1 Buffer

Khởi tạo: - stack: \$

- Buffer: x\$

Lặp: If (Stack là \$S) và (Buffer là \$) Then

- x đúng cú pháp của vp G

- Dừng vòng lặp



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.3. Sơ đồ chung giải thuật PTCP từ dưới lên

➤ Thuật toán:

Else

If (cán β xuất hiện ở đỉnh stack) Then

- Lấy cán β ra khỏi stack
- Đẩy A vào stack với $A \rightarrow \beta$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.3. Sơ đồ chung giải thuật PTCP từ dưới lên

➤ Thuật toán:

Else

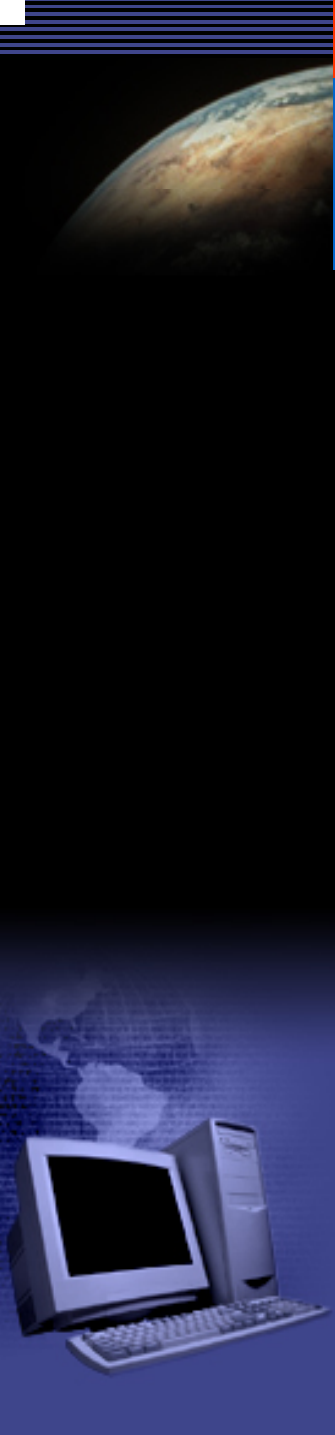
If (Buffer \neq \$) Then

D/c k/h ở đỉnh của Buffer \rightarrow Stack

Else

-Báo lỗi x không đúng cú pháp VP G

-Dừng vòng lặp



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.3. Sơ đồ chung giải thuật PTCP từ dưới lên

➤ Ví dụ: $S \rightarrow \text{if } DK \text{ then } L ;$

$DK \rightarrow \text{true} \mid \text{false}$

$L \rightarrow \text{write}(\text{ID}) \mid \text{read}(\text{ID})$

$\text{ID} \rightarrow a \mid b$

Xâu x: $\text{if true then read}(a)$; có đúng cú pháp
vp trên?



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.4. Sơ đồ chung giải thuật PTCP từ trên xuống

Biết α_i tìm α_{i+1}

$$\alpha_i = u_i \gamma_i$$

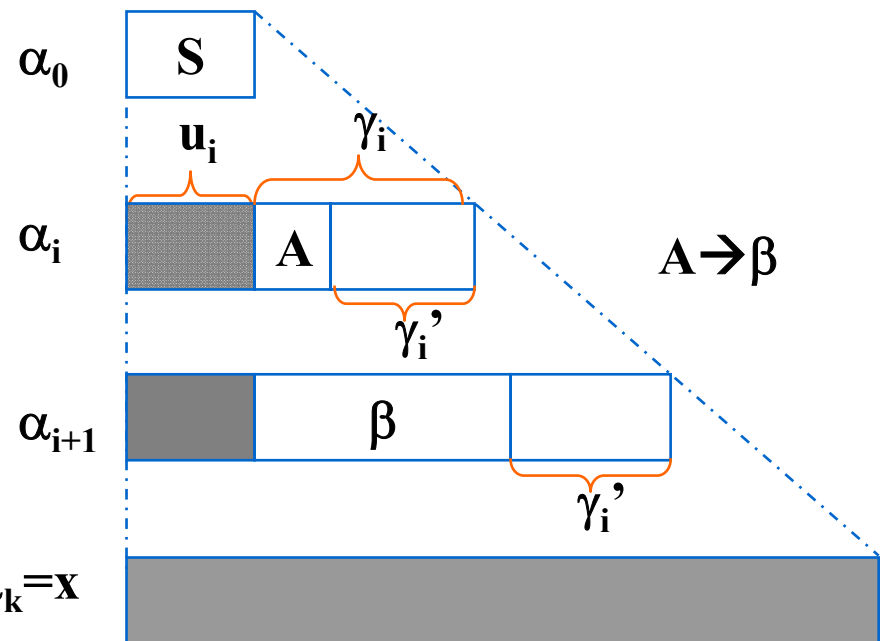
$$\gamma_i \in (\Sigma \cup \Delta)^*; u_i \in \Sigma^*$$

$$\gamma_i = A \gamma_i'$$

$$\alpha_k = x = u_k; \gamma_k = \varepsilon$$

$$\alpha_0 = S = A = \gamma_0;$$

$$\gamma_0' = u_0 = \varepsilon$$



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.4. Sơ đồ chung giải thuật PTCP từ trên xuống

➤ Thuật toán:

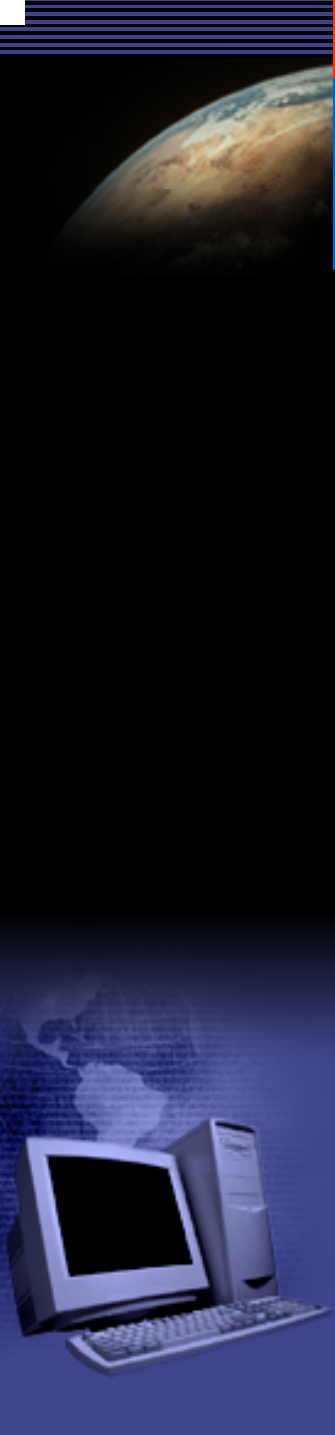
Sử dụng: 1 stack và 1 buffer

Khởi tạo: - stack: S\$

- Buffer: x\$

Lặp: If (Stack là \$) và (Buffer là \$) Then

- x đúng cú pháp của VP G



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.4. Sơ đồ chung giải thuật PTCP từ trên xuống

➤ Thuật toán:

- Dừng vòng lặp

Else

If $(A \in \Delta)$ xuất hiện ở đỉnh Stack Then

Chọn sx thích hợp $A \rightarrow \beta$

Triển khai A bằng β ở đỉnh Stack



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.4. Sơ đồ chung giải thuật PTCP từ trên xuống

➤ Thuật toán:

Else

If $(a \in \Sigma)$ xuất hiện ở đỉnh Stack và
Buffer Then

Lấy a ra khỏi Stack và Buffer {đối
sánh}



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.4. Sơ đồ chung giải thuật PTCP từ trên xuống

➤ Thuật toán:

Else

- Báo lỗi x không đúng cú pháp của G
- Dừng vòng lặp



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

3. Đại cương về phân tích cú pháp

3.4. Sơ đồ chung giải thuật PTCP từ trên xuống

➤ Ví dụ: $S \rightarrow aA$

$A \rightarrow bA \mid c$

Xâu x: abbc có đúng cú pháp của VP trên ?



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

4. Các phương pháp phân tích cú pháp

4.1. Từ trên xuống

- Phương pháp LL(k)
- Phương pháp đệ qui trên xuống



CHƯƠNG 3. CÁC VẤN ĐỀ CƠ BẢN VỀ PHÂN TÍCH CÚ PHÁP

4. Các phương pháp phân tích cú pháp

4.2. Từ dưới lên

- Phương pháp ưu tiên toán tử
- Phương pháp thứ tự yếu
- Phương pháp LR(k)



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CỤ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Văn phạm ưu tiên toán tử

Văn phạm phi ngữ cảnh thỏa mãn các ĐK:

- Không có 2 sản xuất có cùng vế phải
- Không có vế phải là ε
- Không có 2 ký hiệu chưa kết thúc đứng liền nhau



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ **Mỗi quan hệ ưu tiên giữa các ký hiệu**

Với $a, b \in \Sigma$ có:

- **$a \triangleleft b$: a kém ưu tiên hơn b**
- **$a \doteq b$: a ưu tiên bằng b**
- **$a \triangleright b$: a ưu tiên hơn b**



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Quy tắc xác định mối quan hệ

(1) $\exists Sx$ mà vế phải có dạng $\alpha ab\beta$
hay $\alpha aCb\beta$ $\Rightarrow a \doteq b$

(2) $\exists Sx$ mà vế phải có dạng $\alpha aB\beta$
mà $B \Rightarrow^+ b\gamma$ hay $B \Rightarrow^+ Cb\gamma$ $\Rightarrow a < b$

(3) $\exists Sx$ mà vế phải có dạng $\alpha Ab\beta$
mà $A \Rightarrow^+ \gamma a$ hay $A \Rightarrow^+ \gamma aC$ $\Rightarrow a > b$

CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Qui tắc xác định mối quan hệ

$$(4) \quad \left. \begin{array}{l} S \Rightarrow^+ \gamma a \text{ hay } S \Rightarrow^+ \gamma a C \\ \text{hay } S \Rightarrow^+ a \gamma \text{ hay } S \Rightarrow^+ C a \gamma \end{array} \right\} \Rightarrow a \succ \$$$

Với $a, b \in \Sigma$; $A, B, C \in \Delta$; $\alpha, \beta, \gamma \in (\Sigma \cup \Delta)^*$



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Thuật toán

Sử dụng: 1 stack và 1 Buffer

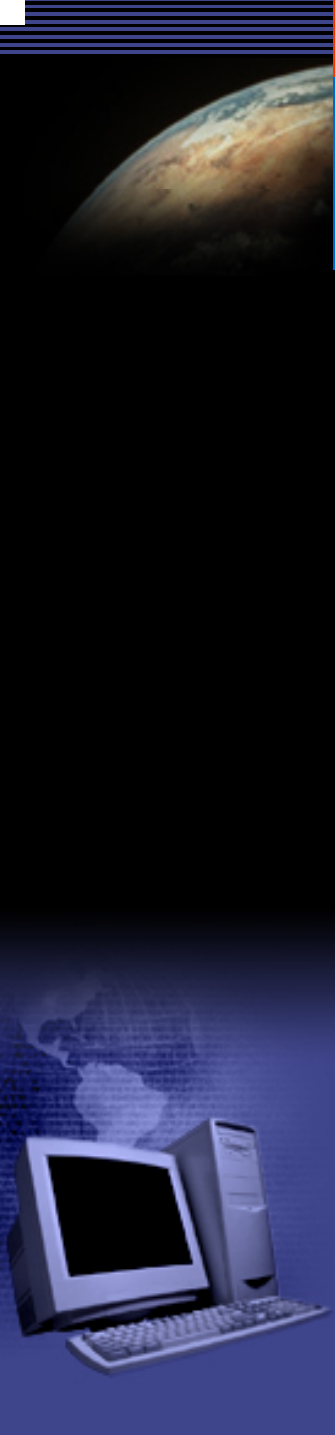
Khởi tạo: - stack: \$

- Buffer: x\$

Lặp: If (Stack là \$S) và (Buffer là \$) Then

- x đúng cú pháp của vp G

- Dừng vòng lặp



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Thuật toán

Else *{giả sử k/h kết thúc gần đỉnh stack
nhất là a và buffer là b}*

If (a>b) Then

- Tìm cán β ở đỉnh stack
- Lấy cán β ra khỏi stack
- Đẩy A vào stack với $A \rightarrow \beta$



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Thuật toán

Else

If $(a < b)$ or $(a = b)$ Thì

D/c b từ Buffer \rightarrow Stack

Else

- Báo lỗi x không đúng cú pháp G
- Dừng vòng lặp



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Ví dụ: $S \rightarrow \text{if } DK \text{ then } L ;$

$DK \rightarrow \text{true} \mid \text{false}$

$L \rightarrow \text{write}(\text{ID}) \mid \text{read}(\text{ID})$

$\text{ID} \rightarrow a \mid b$

Xâu x: $\text{if true then read}(a)$; có đúng cú pháp vp trên?



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.1. Phương pháp ưu tiên toán tử

➤ Ví dụ

- Xác định tất cả các mối quan hệ
- Phân tích



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.2. Phương pháp thứ tự yếu



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.3. Phương pháp SLR (LR(0))



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

1. Phương pháp phân tích cú pháp dưới lên

1.4. Phương pháp Canonical LR (LR(1))



CHƯƠNG 4. CÁC PHƯƠNG PHÁP PHÂN TÍCH CÚ PHÁP

2. Phương pháp phân tích cú pháp trên xuống

