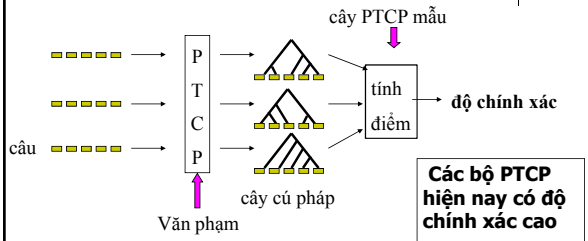
Phân tích cú pháp - SLIDE phân tích cú pháp trong xử lý ngôn ngữ tự nhiên

Công nghệ thông tin (Học viện Công nghệ Bưu chính Viễn thông)

Phân tích cú pháp

Lê Thanh Hương
Bộ môn Hệ thống Thông tin
Viện CNTT &TT – Trường ĐHBKHN
Email: huonglt-fit@mail.hut.edu.vn

Bài toán PTCP



Các bộ PTCP hiện nay có độ chính xác cao
(Eisner, Collins, Charniak, etc.)

Khái niệm về văn phạm

- Phân tích câu "Bò vàng gặm cỏ non"
- Cây cú pháp:
- Tập luật
  - C → CN VN
  - CN → DN
  - VN → ĐgN
  - ĐgN → ĐgT DN
  - DN → DT TT

Văn phạm

- Một văn phạm sản sinh là một hệ thống
- G = (T, N, S, R), trong đó
- T (terminal) – tập ký hiệu kết thúc
- N (non terminal) – tập ký hiệu không kết thúc
- S (start) – ký hiệu khởi đầu
- R (rule) – tập luật
- R = { $\alpha \rightarrow \beta \mid \alpha, \beta \in (T \cup N)$ }
- $\alpha \rightarrow \beta$ gọi là luật sản xuất

Dạng chuẩn Chomsky

- Mọi NNPNC không chứa ε đều có thể sinh từ một văn phạm tnđó mọi sản xuất đều có dạng A → BC hoặc A → a, với A,B,C∈N và a ∈T
- Ví dụ: Tìm dạng chuẩn Chomsky cho văn phạm G với T = {a,b}, N ={S,A,B}, R như sau:
  - S → bA|aB
  - A →bAA|aS|a
  - B → aBB|bS|b

Nhắc lại về văn phạm

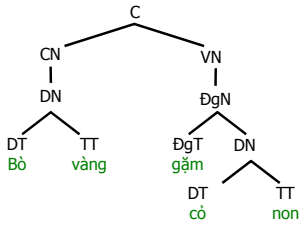
- Văn phạm: 1 tập luật viết lại
- Ký hiệu kết thúc: các ký hiệu không thể phân rã được nữa.
- Ký hiệu không kết thúc: các ký hiệu có thể phân rã được.
- Xét văn phạm G:
  S → NP VP
  NP → John, garbage
  VP → laughed, walks

G có thể sinh ra các câu sau:
John laughed. John walks.
Garbage laughed. Garbage walks.

Cấu trúc ngữ pháp

Cây cú pháp biểu diễn cấu trúc ngữ pháp của một câu.

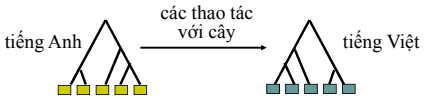
Bò vàng gặm cỏ non.



7

Các ứng dụng của PTCP

- Dịch máy (Alshawi 1996, Wu 1997, ...)



- Nhận dạng tiếng nói sử dụng PTCP (Chelba et al 1998)

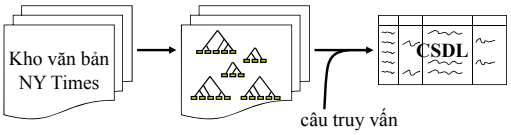
Put the file in the folder.

Put the file and the folder.

8

Các ứng dụng của PTCP

- Kiểm tra ngữ pháp (Microsoft)
- Trích rút thông tin (Hobbs 1996)



9

Văn phạm phi ngữ cảnh (Context-Free Grammar)

… còn gọi là văn phạm cấu trúc đoạn

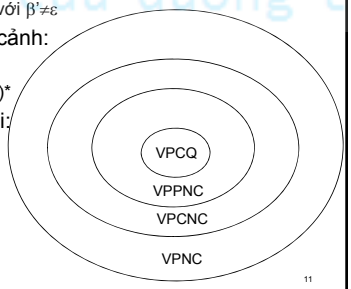
- G = <T,N,P,S,R>
  - T – tập các ký hiệu kết thúc (terminals)
  - N - tập các ký hiệu không kết thúc (non-terminals)
  - P – ký hiệu tiền kết thúc (preterminals), khi viết lại trở thành ký hiệu kết thúc. $P \subset N$
  - S – ký hiệu bắt đầu
  - R: $X \rightarrow \gamma$, X là ký hiệu không kết thúc; γ là chuỗi các ký hiệu kết thúc và không kết thúc (có thể rỗng)
  - Văn phạm G sinh ra ngôn ngữ L
- Bộ nhận dạng: trả về yes hoặc no
- Bộ PTCP: trả về tập các cây cú pháp

So với văn phạm cảm ngữ cảnh R: $\alpha A\gamma \Rightarrow \alpha\beta\gamma$

10

- Văn phạm ngữ cấu:
  - $\alpha \rightarrow \beta$, với $\alpha \in V+$, $\beta \in V^*$
- Văn phạm cảm ngữ cảnh:
  - r = $\alpha \rightarrow \beta$, với $\alpha \in V+$, $\beta \in V^*$, $|\alpha| \leq |\beta|$
  - và $\alpha 1 A \alpha 2 \rightarrow \alpha 1 \beta' \alpha 2$ với $\beta' \neq \varepsilon$
- Văn phạm phi ngữ cảnh:
  - $A \rightarrow \theta$, $A \in N$,
  - với $\theta \in V^* = (T \cup N)^*$
- Văn phạm chính qui:
  - $A \rightarrow aB$,
  - $A \rightarrow Ba$,
  - $A \rightarrow a$,

  với A, B ∈ N, a ∈ T.



11

Văn phạm phi ngữ cảnh

S	→	NP VP	DT	→	*the*
NP	→	DT NNS / DT NN / NP PP	NNS	→	*children* / *students* / *mountains*
VP	→	VP PP / VBD / VBD NP	VBD	→	*slept* / *ate* / *saw*
PP	→	IN NP	IN	→	*in* / *of*
			NN	→	*cake*

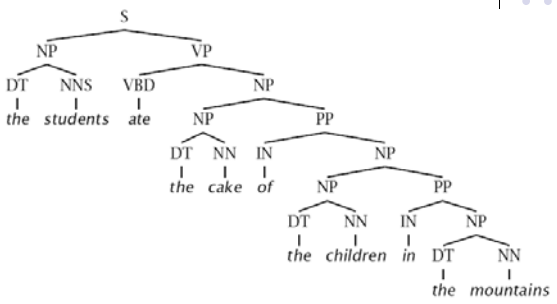
12

Áp dụng tập luật ngữ pháp

- S
  → NP VP
  → DT NNS VBD
  → *The children slept*
- S
  → NP VP
  → DT NNS VBD NP
  → DT NNS VBD DT NN
  → *The children ate the cake*

13

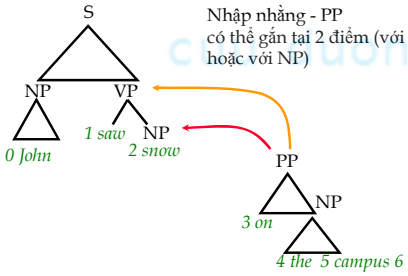
Cấu trúc đoạn đệ qui



14

Văn phạm cho ngôn ngữ tự nhiên có nhập nhằng

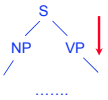
John saw snow on the campus



Nhập nhằng - PP có thể gắn tại 2 điểm (với VP hoặc với NP)

15

PTCP kiểu trên xuống

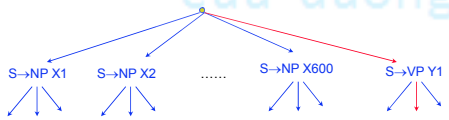


- Hướng đích
- Khởi đầu với 1 danh sách các ký hiệu cần triển khai (S, NP,VP,…)
- Viết lại các đích trong tập đích bằng cách:
  - tìm luật có vế trái trùng với đích cần triển khai
  - triển khai nó với vế phải luật, tìm cách khớp với câu đầu vào
- Nếu 1 đích có nhiều cách viết lại → chọn 1 luật để áp dụng (bài toán tìm kiếm)
- Có thể sử dụng tìm kiếm rộng (breadth-first search) hoặc tìm kiếm sâu (depth-first search)

16

Khó khăn với PTCP trên xuống

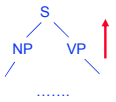
- Các luật đệ qui trái
- PTCP trên xuống rất bất lợi khi có nhiều luật có cùng vế trái



- Nhiều thao tác thừa: triển khai tất cả các nút có thể phân tích trên xuống
- PTCP trên xuống sẽ làm việc tốt khi có chiến lược điều khiển ngữ pháp phù hợp
- PTCP trên xuống không thể triển khai các ký hiệu tiền kết thúc thành các ký hiệu kết thúc. Trên thực tế, người ta thường sử dụng phương pháp dưới lên để làm việc này.
- Lặp lại công việc: bất cứ chỗ nào có cấu trúc giống nhau

17

PTCP dưới lên



- Hướng dữ liệu
- Khởi tạo với xâu cần phân tích
- Nếu chuỗi trong tập đích phù hợp với vế phải của 1 luật → thay nó bằng vế trái của luật.
- Kết thúc khi tập đích = {S}.
- Nếu vế phải của các luật khớp với nhiều luật trong tập đích, cần lựa chọn luật áp dụng (bài toán tìm kiếm)
- Có thể sử dụng tìm kiếm rộng (breadth-first search) hoặc tìm kiếm sâu (depth-first search)

18

Khó khăn với PTCP dưới lên

- Không hiệu quả khi có nhiều nhập nhằng mức từ vựng
- Lặp lại công việc: bất cứ khi nào có cấu trúc con chung
- Cả PTCP TD (LL) và BU (LR) đều có độ phức tạp là hàm mũ của độ dài câu.

Thuật toán CKY (bộ nhận dạng)

- Vào: xâu n từ
- Ra: yes/no
- Cấu trúc ngữ pháp: bảng n x n (chart table)
  - hàng đánh số 0 đến n-1
  - cột đánh số 1 đến n
  - cell [i,j] liệt kê tất cả các nhãn cú pháp giữa i và j

Thuật toán CKY (bottom-up)

- for i := 1 to n
  - Thêm tất cả từ loại của từ thứ i vào ô [i-1,i]
- for width := 2 to n
  - for start := 0 to n-width
    - end := start + width
    - for mid := start+1 to end-1
      - for mọi nhãn cú pháp X trong [start,mid]
        - for mọi nhãn cú pháp Y trong [mid,end]
          - for mọi cách kết hợp X và Y (nếu có)
            - Thêm nhãn kết quả vào [start,end] nếu chưa có nhãn này

Ví dụ

Bò vàng gặm cỏ non

	1	2	3	4	5
0	DT	DN, CN			C
1		TT			
2			ĐgT		VN, ĐgN
3				DT	DN
4					TT

Văn phạm phi ngữ cảnh

1. Start→ S
2. S → NP VP
3. NP → Det Noun
4. NP → Name
5. NP → Name PP
6. PP → Prep NP
7. VP → V NP
8. VP → V NP PP
9. V → *ate*
10. Name → *John*
11. Name → *ice-cream, snow*
12. Noun → *ice-cream, pizza*
13. Noun → *table, guy, campus*
14. Det → *the*
15. Prep → *on*

Luật kết hợp

- Ô Cell[i,j] chứa nhãn X nếu
  - Có luật X→YZ;
  - Cell[i,k] chứa nhãn Y và ô Cell[k,j] chứa nhãn Z, với k nằm giữa i và j;
- VD: NP → DT [0,1] NN[1,2]

CKY phải sử dụng luật nhị phân

- Chuyển VP→V NP PP thành:
  8.a. VP→V Arguments
  8.b. Arguments → NP PP

25

CKY chart

"The guy ate the ice-cream on the table"

	1	2	3	4	5	6	7	8
0	DT							
1		NN						
2			VBD					
3				DT				
4					NN			
5						IN		
6							DT	
7								NN

26

Áp dụng thao tác 'dán'

	1	2	3	4	5	6	7	8
0	DT	NP						
1		NN						
2			VBD					
3				DT				
4					NN			
5						IN		
6							DT	
7								NN

27

Nhập nhằng!

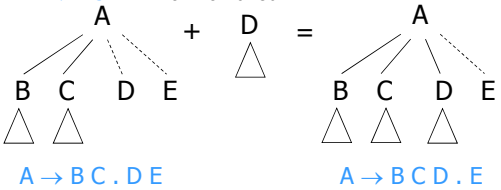
5. NP → NN PP
8.a. VP→V Arguments
8.b. Arguments → NP PP

	1	2	3	4	5	6	7	8
0	DT	NP						S
1		NN						
2			VBD					VP
3				DT	NP			NP, Args
4					NN			
5						IN		PP
6							DT	NP
7								NN

28

Thuật toán Earley (top-down)

- Tìm các nhãn và các nhãn thiếu (partial constituents) từ đầu vào
  - A → B C . D E là nhãn thiếu:



29

  - Tiến hành dần từ trái sang phải

Ví dụ

ROOT	→ S	NP	→ Papa
S	→ NP VP	N	→ caviar
NP	→ Det N	N	→ spoon
NP	→ NP PP	V	→ ate
VP	→ VP PP	P	→ with
VP	→ V NP	Det	→ the
PP	→ P NP	Det	→ a

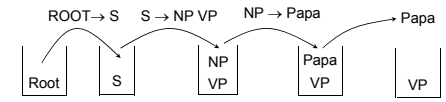
30

0 Papa 1 ate 2 the 3 caviar 4 with 5 a 6 spoon 7

Recursive Descent (Đệ quy)

ROOT	→ S	VP	→ VP PP	NP	→ Papa	V	→ ate
S	→ NP VP	VP	→ V NP	N	→ caviar	P	→ with
NP	→ Det N	PP	→ P NP	N	→ spoon	Det	→ the
NP	→ NP PP					Det	→ a

- 0 ROOT → . S 0
  - 0 S → . NP VP 0
    - 0 NP → . Papa 0
    - 0 NP → Papa . 1
  - 0 S → NP . VP 1



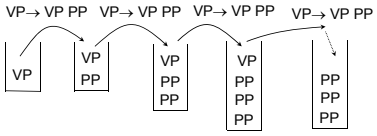
Goal stack

0 Papa 1 ate 2 the 3 caviar 4 with 5 a 6 spoon 7

Recursive Descent

ROOT	→ S	VP	→ VP PP	NP	→ Papa	V	→ ate
S	→ NP VP	VP	→ V NP	N	→ caviar	P	→ with
NP	→ Det N	PP	→ P NP	N	→ spoon	Det	→ the
NP	→ NP PP					Det	→ a

- 0 S → NP . VP 1
  - 1 VP → . VP PP 1
    - 1 VP → . VP PP 1
      - 1 VP → . VP PP 1
        - 1 VP → . VP PP 1 stack overflowed



0 Papa 1 ate 2 the 3 caviar 4 with 5 a 6 spoon 7

Recursive Descent

ROOT	→ S	VP	→ V NP	NP	→ Papa	V	→ ate
S	→ NP VP	VP	→ VP PP	N	→ caviar	P	→ with
NP	→ Det N	PP	→ P NP	N	→ spoon	Det	→ the
NP	→ NP PP					Det	→ a

- 0 ROOT → . S 0
  - 0 S → . NP VP 0
    - 0 NP → . Papa 0
    - 0 NP → Papa . 1
  - 0 S → NP . VP 1
    - 1 VP → . V NP 1 sau . = nonterminal, lặp đi lặp lại việc tìm ký hiệu này ("predict")
      - 1 V → . ate 1 sau . = terminal, tìm nó ở đầu vào ("scan")
      - 1 V → ate . 2 sau . = rỗng, đích con của cha nó đã hoàn chỉnh ("attach")
    - 1 VP → V . NP 2 predict (đích con tiếp theo)
      - 2 NP → . ... 2 phân tích tiếp và cuối cùng …
      - 2 NP → ... . 7 we hoàn thành đích con NP của cha nó → attach
    - 1 VP → V NP . 7 attach
  - 0 S → NP VP . 7 attach

0 Papa 1 ate 2 the 3 caviar 4 with 5 a 6 spoon 7

Recursive Descent

ROOT	→ S	VP	→ V NP	NP	→ Papa	V	→ ate
S	→ NP VP	VP	→ VP PP	N	→ caviar	P	→ with
NP	→ Det N	PP	→ P NP	N	→ spoon	Det	→ the
NP	→ NP PP					Det	→ a

- 0 ROOT → . S 0
  - 0 S → . NP VP 0
    - 0 NP → . Papa 0
    - 0 NP → Papa . 1
  - 0 S → NP . VP 1
    - 1 VP → . V NP 1
      - 1 V → . ate 1
      - 1 V → ate . 2
    - 1 VP → V . NP 2
    - 1 VP → V NP . 7
  - 0 S → NP VP . 7

thực hiện bằng lời gọi hàm:
S() gọi NP() và VP(), VP được triển khai 1 cách đệ qui

cần quay lại để thử 1 luật VP khác

0 Papa 1 ate 2 the 3 caviar 4 with 5 a 6 spoon 7

Recursive Descent

ROOT	→ S	VP	→ V NP	NP	→ Papa	V	→ ate
S	→ NP VP	VP	→ VP PP	N	→ caviar	P	→ with
NP	→ Det N	PP	→ P NP	N	→ spoon	Det	→ the
NP	→ NP PP					Det	→ a

- 0 ROOT → . S 0
  - 0 S → . NP VP 0
    - 0 NP → . Papa 0
    - 0 NP → Papa . 1
  - 0 S → NP . VP 1
    - 1 VP → . VP PP 1
    - 1 VP → . V NP 1 chỗ này cũng cần quay lại
      - 1 V → . ate 1
      - 1 V → ate . 2
    - 1 VP → V . NP 2
      - 2 NP → . ... 2 phân tích tiếp và cuối cùng…
      - 2 NP → ... . 4 … đoạn NP đúng là từ 2 đến 4

0 Papa 1 ate 2 the 3 caviar 4 with 5 a 6 spoon 7

Recursive Descent

ROOT	→ S	VP	→ V NP	NP	→ Papa	V	→ ate
S	→ NP VP	VP	→ VP PP	N	→ caviar	P	→ with
NP	→ Det N	PP	→ P NP	N	→ spoon	Det	→ the
NP	→ NP PP					Det	→ a

- 0 ROOT → . S 0
  - 0 S → . NP VP 0
    - 0 NP → . Papa 0
    - 0 NP → Papa . 1
  - 0 S → NP . VP 1
    - 1 VP → . VP PP 1
      - 1 VP → . VP PP 1
        - 1 VP → . VP PP 1
          - 1 VP → . VP PP 1
            - 1 VP → . VP PP 1

stack overflowed
không giải quyết được gì
– cần thay đổi tập luật để loại trừ đệ qui trái

Thuật toán Earley

- Thuật toán Earley giống thuật toán đệ qui nói trên, nhưng giải quyết được vấn đề đệ qui trái.
- Sử dụng bảng phân tích giống thuật toán CKY, nhằm lưu lại các thông tin đã tìm thấy → lập trình động **"Dynamic programming."**

Các thao tác của thuật toán

- Xử lý phần đi sau dấu . theo kiểu đệ qui :
  - Nếu là từ, quét (**scan**) đầu vào để xem có phù hợp không
  - Nếu là ký hiệu không kết thúc, đoán (**predict**) các khả năng để khớp nó (giảm số phép tiên đoán bằng cách nhìn trước k ký hiệu từ đầu vào và chỉ sử dụng các luật phù hợp với k ký hiệu đó)
  - Nếu rỗng, ta đã hoàn thành một thành phần ngữ pháp, gắn (**attach**) nó vào những chỗ liên quan

37

0
0 ROOT . S

khởi tạo

tương đương với (0, ROOT → . S)

38

0
0 ROOT . S
0 S . NP VP

predict luật có vế trái là S

(0, S → . NP VP)

39

0
0 ROOT . S
0 S . NP VP
0 NP . Det N
0 NP . NP PP
0 NP . Papa

predict luật có VT = NP
(có 3 luật phù hợp)

40

0
0 ROOT . S
0 S . NP VP
0 NP . Det N
0 NP . NP PP
0 NP . Papa
0 Det . the
0 Det . a

predict luật có VT = Det (2 luật)

41

0
0 ROOT . S
0 S . NP VP
0 NP . Det N
0 NP . NP PP
0 NP . Papa
0 Det . the
0 Det . a

predict luật có VT = NP
ta đã làm việc này ở bước trước, vì vậy không làm lại
Chú ý: ta phải làm lại việc này với luật đệ qui trái

42

0	Papa	1
0 ROOT . S	0 NP Papa .	
0 S . NP VP		
0 NP . Det N		
0 NP . NP PP		
0 NP . Papa		
0 Det . the		
0 Det . a		

scan: từ phù hợp từ đầu vào

43

0	Papa	1
0 ROOT . S	0 NP Papa .	
0 S . NP VP		
0 NP . Det N		
0 NP . NP PP		
0 NP . Papa		
0 Det . the		
0 Det . a		

scan: không phù hợp

44

0	Papa	1
0 ROOT . S	0 NP Papa .	
0 S . NP VP		
0 NP . Det N		
0 NP . NP PP		
0 NP . Papa		
0 Det . the		
0 Det . a		

scan: không phù hợp

45

0	Papa	1
0 ROOT . S	0 NP Papa .	
0 S . NP VP		
0 NP . Det N		
0 NP . NP PP		
0 NP . Papa		
0 Det . the		
0 Det . a		

attach NP mới tạo (bắt đầu từ 0) với các phần liên quan (các phần chưa hoàn thành kết thúc tại 0 và có NP sau dấu .)

46

0	Papa	1
0 ROOT . S	0 NP Papa .	
0 S . NP VP	0 S NP . VP	
0 NP . Det N	0 NP NP . PP	
0 NP . NP PP	1 VP . V NP	
0 NP . Papa	1 VP . VP PP	
0 Det . the		
0 Det . a		

predict

47

0	Papa	1
0 ROOT . S	0 NP Papa .	
0 S . NP VP	0 S NP . VP	
0 NP . Det N	0 NP NP . PP	
0 NP . NP PP	1 VP . V NP	
0 NP . Papa	1 VP . VP PP	
0 Det . the	1 PP . P NP	
0 Det . a		

predict

48

0 Papa 1	
0 ROOT . S	0 NP Papa .
0 S . NP VP	0 S NP . VP
0 NP . Det N	0 NP NP . PP
0 NP . NP PP	1 VP . V NP
0 NP . Papa	1 VP . VP PP
0 Det . the	1 PP . P NP
0 Det . a	1 V . ate

predict

49

0 Papa 1	
0 ROOT . S	0 NP Papa .
0 S . NP VP	0 S NP . VP
0 NP . Det N	0 NP NP . PP
0 NP . NP PP	1 VP . V NP
0 NP . Papa	1 VP . VP PP
0 Det . the	1 PP . P NP
0 Det . a	1 V . ate

predict

50

0 Papa 1	
0 ROOT . S	0 NP Papa .
0 S . NP VP	0 S NP . VP
0 NP . Det N	0 NP NP . PP
0 NP . NP PP	1 VP . V NP
0 NP . Papa	1 VP . VP PP
0 Det . the	1 PP . P NP
0 Det . a	1 V . ate
	1 P . with

predict

51

0 Papa 1 ate 2		
0 ROOT . S	0 NP Papa .	1 V ate .
0 S . NP VP	0 S NP . VP	
0 NP . Det N	0 NP NP . PP	
0 NP . NP PP	1 VP . V NP	
0 NP . Papa	1 VP . VP PP	
0 Det . the	1 PP . P NP	
0 Det . a	1 V . ate	
	1 P . with	

scan: thành công!

52

0 Papa 1 ate 2		
0 ROOT . S	0 NP Papa .	1 V ate .
0 S . NP VP	0 S NP . VP	
0 NP . Det N	0 NP NP . PP	
0 NP . NP PP	1 VP . V NP	
0 NP . Papa	1 VP . VP PP	
0 Det . the	1 PP . P NP	
0 Det . a	1 V . ate	
	1 P . with	

scan: không hợp

53

0 Papa 1 ate 2		
0 ROOT . S	0 NP Papa .	1 V ate .
0 S . NP VP	0 S NP . VP	1 VP V . NP
0 NP . Det N	0 NP NP . PP	
0 NP . NP PP	1 VP . V NP	
0 NP . Papa	1 VP . VP PP	
0 Det . the	1 PP . P NP	
0 Det . a	1 V . ate	
	1 P . with	

attach

54

0 Papa 1 ate 2		
0 ROOT . S	0 NP Papa .	1 V ate .
0 S . NP VP	0 S NP . VP	1 VP V . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa
0 Det . the	1 PP . P NP	
0 Det . a	1 V . ate	
	1 P . with	

predict

55

0 Papa 1 ate 2		
0 ROOT . S	0 NP Papa .	1 V ate .
0 S . NP VP	0 S NP . VP	1 VP V . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the
0 Det . a	1 V . ate	2 Det . a
	1 P . with	

predict *(các bước sau tương tự)*

56

0 Papa 1 ate 2		
0 ROOT . S	0 NP Papa .	1 V ate .
0 S . NP VP	0 S NP . VP	1 VP V . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the
0 Det . a	1 V . ate	2 Det . a
	1 P . with	

predict

57

0 Papa 1 ate 2		
0 ROOT . S	0 NP Papa .	1 V ate .
0 S . NP VP	0 S NP . VP	1 VP V . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the
0 Det . a	1 V . ate	2 Det . a
	1 P . with	

scan *(lúc này thất bại vì Papa không phải là từ tiếp theo)*

58

0 Papa 1 ate 2 the 3			
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .
0 S . NP VP	0 S NP . VP	1 VP V . NP	
0 NP . Det N	0 NP NP . PP	2 NP . Det N	
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	
0 NP . Papa	1 VP . VP PP	2 NP . Papa	
0 Det . the	1 PP . P NP	2 Det . the	
0 Det . a	1 V . ate	2 Det . a	
	1 P . with		

scan: thành công!

59

0 Papa 1 ate 2 the 3			
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .
0 S . NP VP	0 S NP . VP	1 VP V . NP	
0 NP . Det N	0 NP NP . PP	2 NP . Det N	
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	
0 NP . Papa	1 VP . VP PP	2 NP . Papa	
0 Det . the	1 PP . P NP	2 Det . the	
0 Det . a	1 V . ate	2 Det . a	
	1 P . with		

60

0 Papa	1 ate	2 the	3
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N
0 NP . Det N	0 NP NP . PP	2 NP . Det N	
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	
0 NP . Papa	1 VP . VP PP	2 NP . Papa	
0 Det . the	1 PP . P NP	2 Det . the	
0 Det . a	1 V . ate	2 Det . a	
	1 P . with		

61

0 Papa	1 ate	2 the	3
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon
0 NP . Papa	1 VP . VP PP	2 NP . Papa	
0 Det . the	1 PP . P NP	2 Det . the	
0 Det . a	1 V . ate	2 Det . a	
	1 P . with		

62

0 Papa	1 ate	2 the	3	caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .		3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N		
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar		
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon		
0 NP . Papa	1 VP . VP PP	2 NP . Papa			
0 Det . the	1 PP . P NP	2 Det . the			
0 Det . a	1 V . ate	2 Det . a			
	1 P . with				

63

0 Papa	1 ate	2 the	3	caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .		3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N		
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar		
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon		
0 NP . Papa	1 VP . VP PP	2 NP . Papa			
0 Det . the	1 PP . P NP	2 Det . the			
0 Det . a	1 V . ate	2 Det . a			
	1 P . with				

64

0 Papa	1 ate	2 the	3	caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N		
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar		
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon		
0 NP . Papa	1 VP . VP PP	2 NP . Papa			
0 Det . the	1 PP . P NP	2 Det . the			
0 Det . a	1 V . ate	2 Det . a			
	1 P . with				

65

attach

0 Papa	1 ate	2 the	3	caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar		
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon		
0 NP . Papa	1 VP . VP PP	2 NP . Papa			
0 Det . the	1 PP . P NP	2 Det . the			
0 Det . a	1 V . ate	2 Det . a			
	1 P . with				

66

attach

0 Papa	1 ate	2 the	3 caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP
0 Det . a	1 V . ate	2 Det . a		
	1 P . with			

attach

67

0 Papa	1 ate	2 the	3 caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP
	1 P . with			

68

0 Papa	1 ate	2 the	3 caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP
	1 P . with			0 ROOT S .

attach

69

0 Papa	1 ate	2 the	3 caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP
	1 P . with			0 ROOT S .

70

0 Papa	1 ate	2 the	3 caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP
	1 P . with			4 P . with

71

0 Papa	1 ate	2 the	3 caviar	4
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP
	1 P . with			0 ROOT S .

72

0 Papa	1 ate	2 the	3 caviar	4 with	5
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	4 P with .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	
	1 P . with			0 ROOT S .	
				4 P . with	

73

0 Papa	1 ate	2 the	3 caviar	4 with	5
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	4 P with .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	
	1 P . with			0 ROOT S .	
				4 P . with	

74

0 Papa	1 ate	2 the	3 caviar	4 with	5
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	4 P with .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	5 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	
	1 P . with			0 ROOT S .	
				4 P . with	

75

0 Papa	1 ate	2 the	3 caviar	4 with	5
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	4 P with .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	5 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	5 Det . the
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	5 Det . a
	1 P . with			0 ROOT S .	
				4 P . with	

76

0 Papa	1 ate	2 the	3 caviar	4 with	5
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	4 P with .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	5 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	5 Det . the
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	5 Det . a
	1 P . with			0 ROOT S .	
				4 P . with	

77

0 Papa	1 ate	2 the	3 caviar	4 with	5
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	4 P with .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	5 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	5 Det . the
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	5 Det . a
	1 P . with			0 ROOT S .	
				4 P . with	

78

0 Papa	1 ate	2 the	3 caviar	4 with	5
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	4 P with .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	5 NP . Papa
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	5 Det . the
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	5 Det . a
	1 P . with			0 ROOT S .	
				4 P . with	

79

ate	2 the	3 caviar	4 with	5 a	6
.	1 V ate .	2 Det the .	3 N caviar .	4 P with .	5 Det a .
)	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP	
PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N	
)	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP	
P	2 NP . Papa		0 S NP VP .	5 NP . Papa	
)	2 Det . the		1 VP VP . PP	5 Det . the	
	2 Det . a		4 PP . P NP	5 Det . a	
			0 ROOT S .		
			4 P . with		

80

ate	2 the	3 caviar	4 with	5 a	6
.	1 V ate .	2 Det the .	3 N caviar .	4 P with .	5 Det a .
)	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP	5 NP Det . N
PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N	5 NP Det N
)	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP	
P	2 NP . Papa		0 S NP VP .	5 NP . Papa	
)	2 Det . the		1 VP VP . PP	5 Det . the	
	2 Det . a		4 PP . P NP	5 Det . a	
			0 ROOT S .		
			4 P . with		

81

ate	2 the	3 caviar	4 with	5 a	6
.	1 V ate .	2 Det the .	3 N caviar .	4 P with .	5 Det a .
)	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP	5 NP Det . N
PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N	6 N . caviar
)	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP	6 N . spoon
P	2 NP . Papa		0 S NP VP .	5 NP . Papa	
)	2 Det . the		1 VP VP . PP	5 Det . the	
	2 Det . a		4 PP . P NP	5 Det . a	
			0 ROOT S .		
			4 P . with		

82

ate	2 the	3 caviar	4 with	5 a	6
.	1 V ate .	2 Det the .	3 N caviar .	4 P with .	5 Det a .
)	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP	5 NP Det . N
PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N	6 N . caviar
)	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP	6 N . spoon
P	2 NP . Papa		0 S NP VP .	5 NP . Papa	
)	2 Det . the		1 VP VP . PP	5 Det . the	
	2 Det . a		4 PP . P NP	5 Det . a	
			0 ROOT S .		
			4 P . with		

83

ate	2 the	3 caviar	4 with	5 a	6 spoon	7
.	1 V ate .	2 Det the .	3 N caviar .	4 P with .	5 Det a .	6 N spoon .
)	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP	5 NP Det . N	
PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N	6 N . caviar	
)	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP	6 N . spoon	
P	2 NP . Papa		0 S NP VP .	5 NP . Papa		
)	2 Det . the		1 VP VP . PP	5 Det . the		
	2 Det . a		4 PP . P NP	5 Det . a		
			0 ROOT S .			
			4 P . with			

84

ate	2	the	3	caviar	4	with	5	a	6	spoon
.	1 V ate .	2 Det the .	3 N caviar .	4 P with .	5 Det a .	6 N spoon .
?	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP	5 NP Det . N	5 NP Det N .
PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N	6 N . caviar	
?	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP	6 N . spoon	
P	2 NP . Papa		0 S NP VP .	5 NP . Papa		
?	2 Det . the		1 VP VP . PP	5 Det . the		
	2 Det . a		4 PP . P NP	5 Det . a		
			0 ROOT S .			
			4 P . with			

ate	2	the	3	caviar	4	with	5	a	6	spoon
.	1 V ate .	2 Det the .	3 N caviar .	4 P with .	5 Det a .	6 N spoon .
?	1 VP V . NP	2 NP Det . N	2 NP Det N .	4 PP P . NP	5 NP Det . N	5 NP Det N .
PP	2 NP . Det N	3 N . caviar	1 VP V NP .	5 NP . Det N	6 N . caviar	4 PP P NP .
?	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP . NP PP	6 N . spoon	5 NP NP . PP
P	2 NP . Papa		0 S NP VP .	5 NP . Papa		
?	2 Det . the		1 VP VP . PP	5 Det . the		
	2 Det . a		4 PP . P NP	5 Det . a		
			0 ROOT S .			
			4 P . with			

0	Papa	1	ate	2	the	3	caviar	4	with a spoon
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		
	1 P . with			0 ROOT S .		
				4 P . with		

0	Papa	1	ate	2	the	3	caviar	4	with a spoon
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		
				4 P . with		

0	Papa	1	ate	2	the	3	caviar	4	with a spoon
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP

0	Papa	1	ate	2	the	3	caviar	4	with a spoon
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP
						0 S NP VP .
						1 VP VP . PP

0	Papa	1	ate	2	the	3	caviar	4	with a spoon

0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP
						0 S NP VP .
						1 VP VP . PP
						7 P . with

0	Papa	1	ate	2	the	3	caviar	4	with a spoon

0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP
						0 S NP VP .
						1 VP VP . PP
						7 P . with

0	Papa	1	ate	2	the	3	caviar	4	with a spoon

0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP
						0 S NP VP .
						1 VP VP . PP
						7 P . with

0	Papa	1	ate	2	the	3	caviar	4	with a spoon

0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP
						0 S NP VP .
						1 VP VP . PP
						7 P . with

0	Papa	1	ate	2	the	3	caviar	4	with a spoon

0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP
						0 S NP VP .
						1 VP VP . PP
						7 P . with
						0 ROOT S .

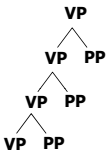
0	Papa	1	ate	2	the	3	caviar	4	with a spoon

0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	...	6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .		5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .		4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP		5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .		2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP		1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP		7 PP . P NP
	1 P . with			0 ROOT S .		1 VP V NP .
				4 P . with		2 NP NP . PP
						0 S NP VP .
						1 VP VP . PP
						7 P . with
						0 ROOT S .

0 Papa 1 ate 2 the 3 caviar 4 with a spoon 7					
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	... 6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	7 PP . P NP
	1 P . with			0 ROOT S .	1 VP V NP .
				4 P . with	2 NP NP . PP
					0 S NP VP .
					1 VP VP . PP
					7 P . with
					0 ROOT S .

0 Papa 1 ate 2 the 3 caviar 4 with a spoon 7					
0 ROOT . S	0 NP Papa .	1 V ate .	2 Det the .	3 N caviar .	... 6 N spoon .
0 S . NP VP	0 S NP . VP	1 VP V . NP	2 NP Det . N	2 NP Det N .	5 NP Det N .
0 NP . Det N	0 NP NP . PP	2 NP . Det N	3 N . caviar	1 VP V NP .	4 PP P NP .
0 NP . NP PP	1 VP . V NP	2 NP . NP PP	3 N . spoon	2 NP NP . PP	5 NP NP . PP
0 NP . Papa	1 VP . VP PP	2 NP . Papa		0 S NP VP .	2 NP NP PP .
0 Det . the	1 PP . P NP	2 Det . the		1 VP VP . PP	1 VP VP PP .
0 Det . a	1 V . ate	2 Det . a		4 PP . P NP	7 PP . P NP
	1 P . with			0 ROOT S .	1 VP V NP .
				4 P . with	2 NP NP . PP
					0 S NP VP .
					1 VP VP . PP
					7 P . with
					0 ROOT S .

Vấn đề với PTCP trên xuống: đệ qui trái



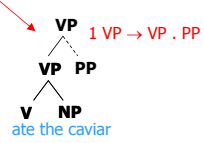
gắn liên tục các luật mới vào cây trước khi thấy PPs
→ cần đoán trước số PP cần ở đầu vào

… nhưng thuật toán Earley Ok!

1 VP → . VP PP

(cột 1)

attach



ate the caviar

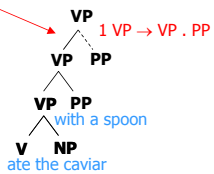
(cột 4)

… nhưng thuật toán Earley Ok!

1 VP → . VP PP có thể dùng lại

(cột1)

attach



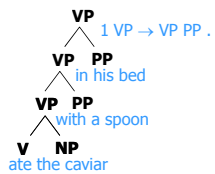
ate the caviar

(cột 7)

… nhưng thuật toán Earley Ok!

1 VP → . VP PP có thể dùng lại

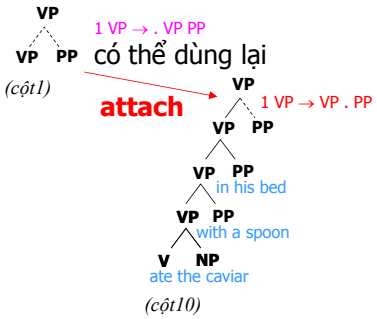
(cột1)



ate the caviar

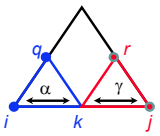
(cột 10)

… nhưng thuật toán Earley Ok!



Phục hồi cây cú pháp

[s,i]trong tập trạng thái *j* Sử dụng thuật toán dùng queue đơn giản, dựa trên các thành phần có ích



- 1 thành phần ở trạng thái kết thúc là *có ích*
- *If* $s=[A \rightarrow \alpha \bullet B, i]$ trong tập đích *k* & *có ích*
- *then* $q=[A \rightarrow \alpha B \bullet, k]$ & item $r=[B \rightarrow \gamma \bullet, j]$ là *có ích*

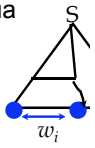
[s,i] : một thành phần với luật *s* & trả về con trỏ *i*.

Đánh dấu tất cả các thành phần trong tập trạng thái *Sn* ở dạng $Start \rightarrow \alpha S \bullet, 0$

for j=n downto 0 do
  for i=0 to *j do*
    for mọi bộ đã đánh dấu [s,i] trong tập trạng thái *j do*
      for k=i to j do
        if $[q,i] \in Sk$ & $[r,k] \in Sj$ & $s = q \otimes r$ *then*
          đánh dấu [q,i] và [r,k]

Ưu điểm

- Thuật toán Earley thực hiện một vài phép lọc *top-down*: bất cứ thành phần nào (state, or triple) được đưa vào tập trạng thái cần tương thích với phần đã được sinh ra ở bên trái. Ví dụ: $S \overset{*}{\Rightarrow} w_i$ trong đó w_i là phần của câu đã được duyệt qua



Nhược điểm

- Biểu diễn luật: Explicit representation of rules: wastes time building them.
- Thực hiện phép lọc bên trái nhưng không lọc bên phải

Phép lọc nhìn trước cho ký hiệu không kết thúc *A:*

$FIRST(A) = \{x | A \Rightarrow x\delta\}$, *x= 1 token*

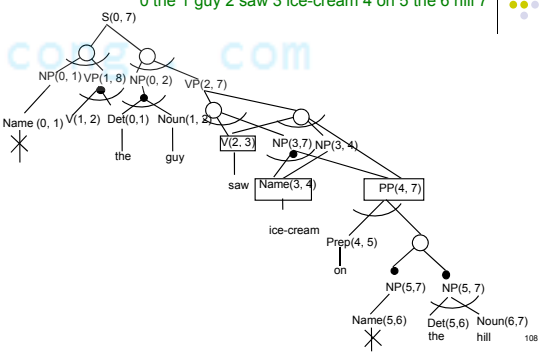
v.d., FIRST(S)= who, did, the, etc.

Các phương pháp khác

- Các phương pháp khác ứng với các cách khác nhau để tìm các đoạn
- Đoạn X[*i, j*] là đoạn có nhãn X phủ đầu vào từ I đến j Example:
  $_0$ *John* $_1$ *ate* $_2$ *ice-cream* $_3$ *on* $_4$ *the* $_5$ *table* $_6$
  PP[3,6]; S[0,6]; …
- Biểu diễn không gian tìm kiếm như cây and-or
  - Disjuncts (or) = các đường phân tích khác nhau
  - Conjuncts (and) = vế phải của luật, ví dụ vế phải của S là NP VP

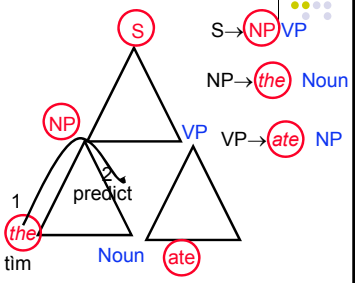
PTCP là việc tìm kiếm

0 the 1 guy 2 saw 3 ice-cream 4 on 5 the 6 hill 7



PTCP góc trái (Left-corner parsing)

- Nhìn từ dưới lên để tìm ký hiệu đầu tiên (left-corner) của đoạn, sau đó phân tích phần còn lại theo kiểu trên xuống
- Tìm cách kết hợp các đặc trưng tốt nhất của tìm phân tích trên xuống và dưới lên



Phương pháp này làm việc tốt với ngôn ngữ với thành phần quan trọng đặt ở đầu như tiếng Anh. Các tiếng Đức, Hà Lan, Nhật là ngôn ngữ có phần quan trọng đặt cuối.