

[言語情報科学論A] Chart Parsing

2007年06月18日

言語情報科学講座
林 良彦 教授

Text: Courtesy of Dr. Jurafsky, D. and Dr. Martin, J.H: “Speech and Language Processing, 1st edition (Prentice Hall, 2000) & 2nd edition”,
<http://www.cs.colorado.edu/~martin/slp2.html>

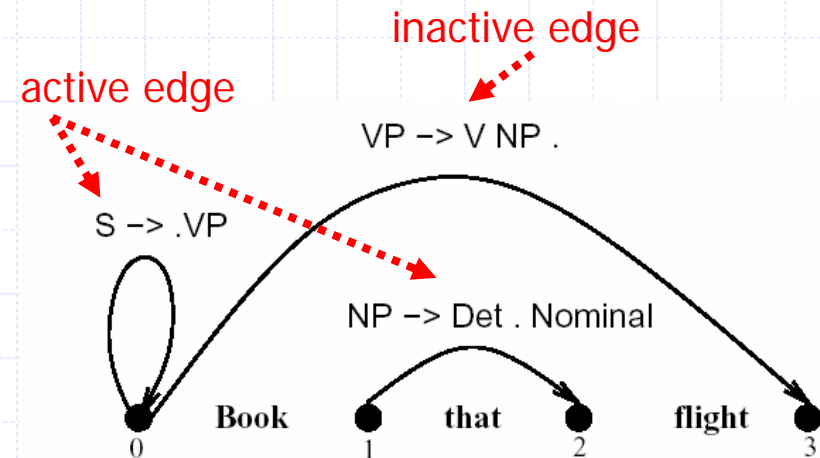
Review: The Early Algorithm

◆ DP (Dynamic Programming) に基づく

- 部分問題 (sub-problem) に対する解を記録しておく
- parsingでは, 入力のある部分に対する部分木を記録しておく
 - ◆ 処理の重複を避ける
 - ◆ 記憶されている部分木の組み合わせ方の違いとして, 曖昧性を取り扱う
 - ◆ left-recursionの問題も解消

◆ 基本的な要素

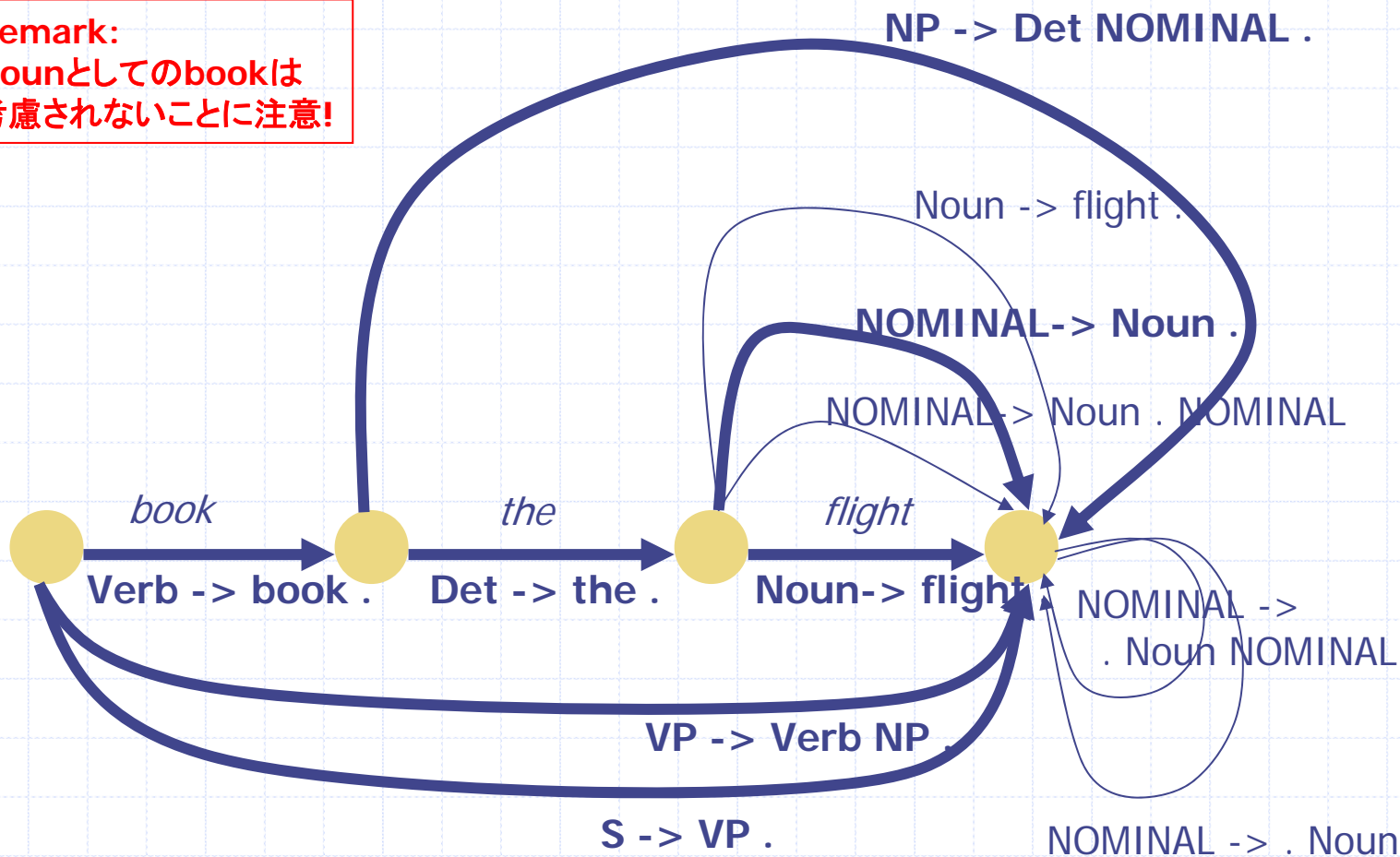
- Chart
 - ◆ DAG (Directed Acyclic Graph)
- dotted rule
 - ◆ $S \rightarrow \cdot VP, [0,0]$ 状態:state
 - ◆ $NP \rightarrow Det \cdot Nominal, [1,2]$
 - ◆ $VP \rightarrow V NP \cdot, [0,3]$



An Example

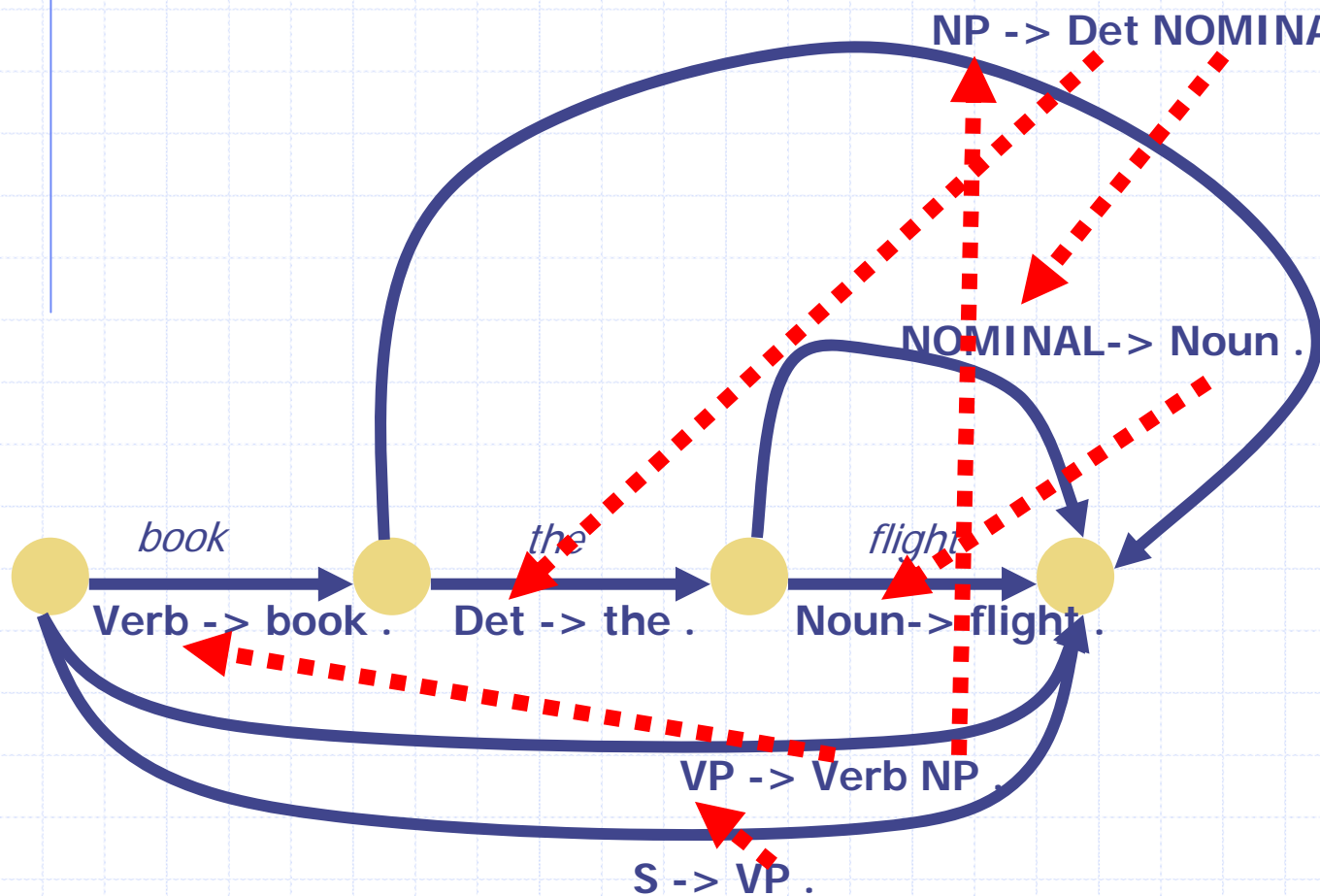
この図は最終ステップ Chart[3]の状況を示す

Remark:
Nounとしてのbookは
考慮されないことに注意!



Chartから構文木をretrieveする

どの不活性エッジを用いたかのポインタを張るようにCOMPLEMENTERを改良する



Courtesy of Dr. Jurafsky, D. and Dr. Martin, J.H: "Speech and Language Processing, 2nd edition", CH. 14

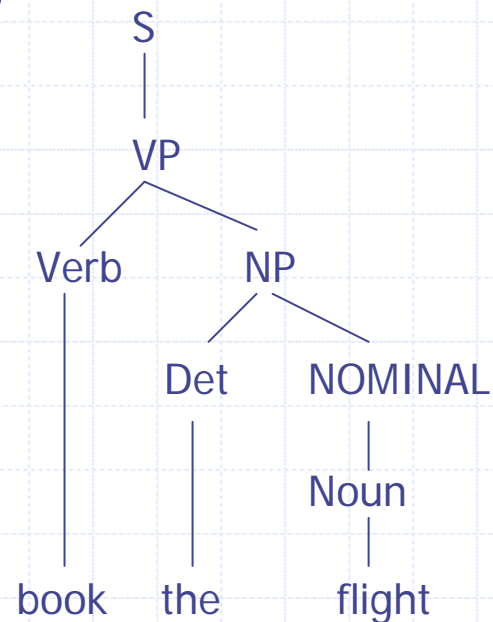


Chart Parser

◆ Remark: Early Algorithmはtop-down型

◆ Chart Parser

- 同じくchartを基本的なデータ構造として用いる
- ただし, 新しいedgeをすぐにchartに追加するのではなく, agendaを用いる. これにより, 動作を制御する
 - ◆ active edgeの提案の仕方によって, T-DかB-Uかを制御
 - ◆ agendaへのedgeの仕方によって, DFSかBFSかを制御
 - 先頭へ追加する (stack; LIFO): depth-first
 - 最後尾に追加する (queue; FIFO): breadth-first

◆ terminology

- ラベル: 生成規則の左辺の記号のこと
- 空所リスト: dotの右側にある記号のリスト

Chart Parserのアルゴリズム

- ◆ chartの初期化: terminal symbolsに対応するinactive edgesをagendaに登録する
 - さらに, T-Dのときには, Sに関するdummyのactive edgeをagendaの先頭に登録する
- ◆ ループ: 解が見つかるか, agendaが空になるまで繰り返す
 - edgeの選択: agendaからedgeを取り出してchartへ追加
 - edgeの結合:
 - ◆ もしそれがactive edgeなら, その右側のinactive edgeと結合を試みる
 - ◆ もしそれがinactive edgeならば, その左側のactive edgeとの結合を試みる
 - ◆ ここで, 結合OKの条件: active edgeの空所リストの第一要素とinactive edgeのラベルが一致すること
 - edgeの登録: 結合が成功し, そのedgeと同じものがchart, agendaに存在しなければ, agendaに登録する
 - active edgeの提案:
 - ◆ T-Dの場合: 現在のedgeがactiveなら, この空所リストの第一要素を左辺とするような規則を用いて, このactive edgeの終了位置に新しいactive edgeを作成し, agendaに登録する
 - ◆ B-Uの場合: 現在のedgeがinactiveなら, このinactive edgeを右辺の第一要素とするような規則を用いて, このinactive edgeの開始位置に新しいactive edgeを作成し, agendaに登録する

Bottom-Up/BFSの動作例

さらに簡略化
したGrammar

S -> VP

VP -> Verb NP

NP -> Det NOM

NOM -> Noun

Verb -> *book*

Det -> *the*

Noun -> *flight*

Agenda

~~Verb -> book . [0,1]~~

~~Det -> the . [1,2]~~

~~Noun -> flight. [2,3]~~

~~VP -> . Verb NP [0,0]~~

~~NP -> . Det NOM [1,1]~~

~~NOM -> . Noun [2,2]~~

~~VP -> Verb . NP [0,1]~~

~~NP -> Det . NOM [1,2]~~

~~NOM -> Noun . [2,3]~~

~~NP -> Det NOM . [1,3]~~

~~VP -> Verb NP . [0,3]~~

~~S -> . VP [0,0]~~

~~S -> VP . [0,3]~~

Chart

