

言語解析論

講師 竹内孔一

本日の内容

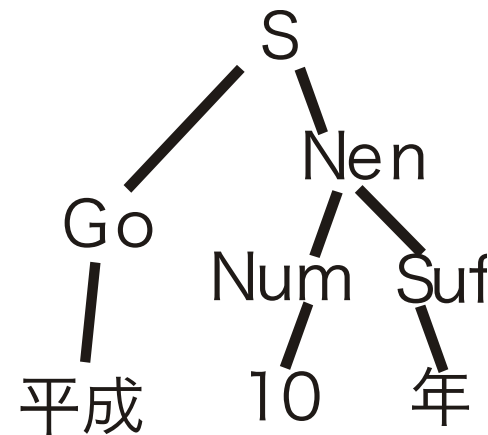
- 文脈自由文法
- pda (プッシュダウンスタックオートマトン)

BNF

- CFGの文法を記述する便利な表記
 - Backus Naur Form (バックス・ノーアフォーム)
 - 同じ非終端記号から派生する生成規則をまとめて書く
 - ものの定義を記述する

例) 「平成□□年」を受け付ける文脈自由文法をBNFで書く

```
S ::= Go Nen
Nen ::= Num Suf
Go ::= 平成
Num ::= 1 | 2 | 3 | ... | 21
Suf ::= 年
```



練習6

- BNFを利用して次の文を受理するCFGを記述せよ

–「彼の姉は犯人は弟であることを知っていた」
ただし

$P ::= \text{は} \mid \text{の} \mid \text{を} \mid \text{で}$ (助詞)

$V ::= \text{ある} \mid \text{知っていた}$ (動詞)

$N ::= \text{彼} \mid \text{姉} \mid \text{犯人} \mid \text{弟} \mid \text{こと}$ (名詞)

を利用すること

文脈自由文法

- 正規文法との違い
 - $\{a^n b^n \mid n = 0, 1, 2, \dots\}$ という離れた関係が捉えられる
- 練習
 - 上記の表現を捉えられる簡単なCFGを記述せよ

解答

$$V_N = \{\sigma\}$$

$$V_T = \{a, b\}$$

$$P = \{\sigma \rightarrow a\sigma b, \sigma \rightarrow ab\}$$

$$G = \langle V_N, V_T, P, \sigma \rangle$$

非終端記号に代入して繰り返すと

$$\sigma \Rightarrow a\sigma b$$

$$\Rightarrow a(a\sigma b)b$$

$$\Rightarrow a^2(a\sigma b)b^2$$

$$\Rightarrow a^n b^n$$

捉えられる現象

- CFGの特徴生かした言語

例 “That ...that he is tall is false ... is false. ”

パターン (that)ⁱ he is tall (is false)ⁱ

→ 最初の i 回を覚えておいて次の i 回を繰り返す

例 彼は 彼は 姉は犯人でない ことを知っている ことを知っている

パターンを記述してみよう

Push down automaton (PDA)

- 定義

- 状態集合 Q
- 入力文字集合 Σ
- スタックの文字集合 Γ
- 状態遷移集合 σ
- 初期状態 q_0
- 終了状態集合 F

$$P = \langle Q, \Sigma, \Gamma, \sigma, q_0, F \rangle$$

- 事例

- $\{a^n b^n \mid n = 0, 1, 2, \dots\}$ を受理する場合

- 定性的な説明

- なぜ離れた関係を捉えられるのか
 - スタックに非終端を積むから
- $$P = \left\{ \sigma \rightarrow a\sigma b, \sigma \rightarrow ab \right\}$$

入力文字列 aaabbb



有限制御部

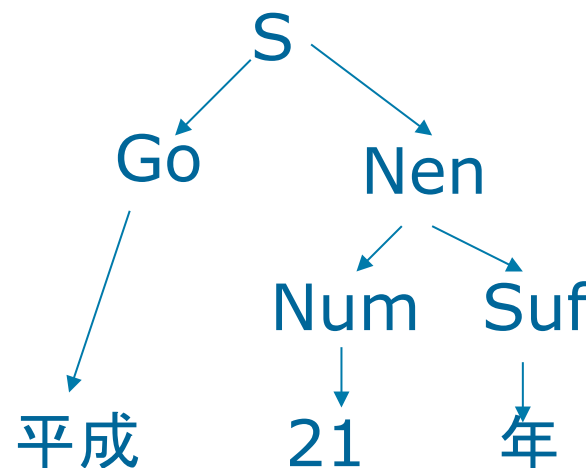


スタック
(補助記憶装置)

練習7

- 下記CFGと等価なPDAで「平成21年」を受けよ

$S ::= \text{Go Nen}$
 $\text{Nen} ::= \text{Num Suf}$
 $\text{Go} ::= \text{平成}$
 $\text{Num} ::= 1 \mid 2 \mid 3 \mid \dots \mid 21$
 $\text{Suf} ::= \text{年}$



ヒント: スタックの先頭はSが初期値

● 定義

- 状態集合 Q
- 入力文字集合 Σ
- スタックの文字集合 Γ
- 状態遷移集合 σ
- 初期状態 q_0
- 終了状態集合 F

$S ::= \text{Go Nen}$
 $\text{Nen} ::= \text{Num Suf}$
 $\text{Go} ::= \text{平成}$
 $\text{Num} ::= 1 \mid 2 \mid 3 \mid \dots \mid 21$
 $\text{Suf} ::= \text{年}$

$P = \langle Q, \Sigma, \Gamma, \sigma, q_0, F \rangle$

初期状態 q_0 は最初の非終端S

終了: 最後の文字が読み
終わったときにスタックも空

入力文字列 Σ



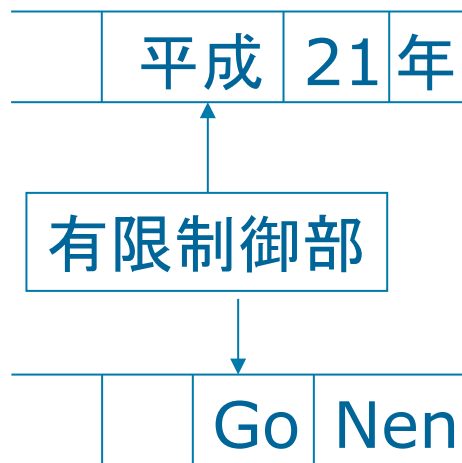
このSはCFGの
非終端記号の最初

スタック
(補助記憶装置)

S ::= Go Nen
Nen ::= Num Suf
Go ::= 平成
Num ::= 1 | 2 | 3 | ... | 21
Suf ::= 年

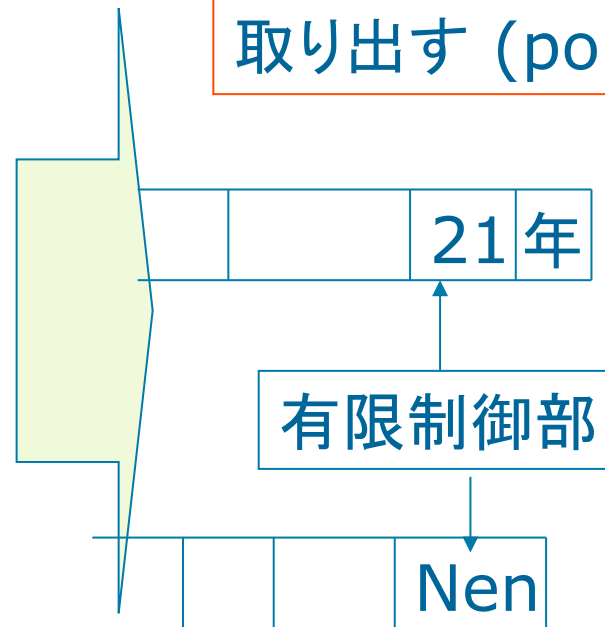
Step 1

平成を受け付ける前に
S を展開してGo
まで持っていく



Step2

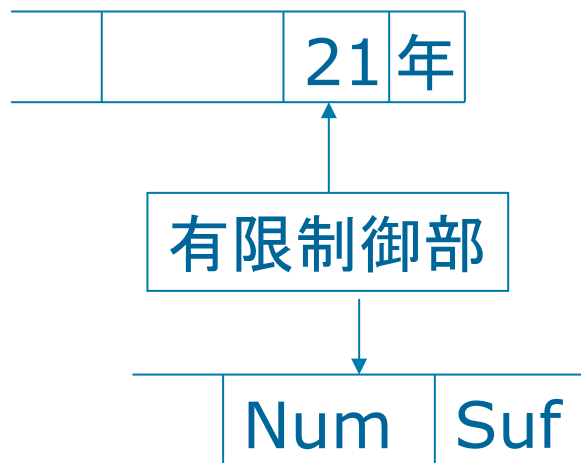
平成を読み込んで
Go をスタックから
取り出す (pop)



S ::= Go Nen
Nen ::= Num Suf
Go ::= 平成
Num ::= 1 | 2 | 3 | ... | 21
Suf ::= 年

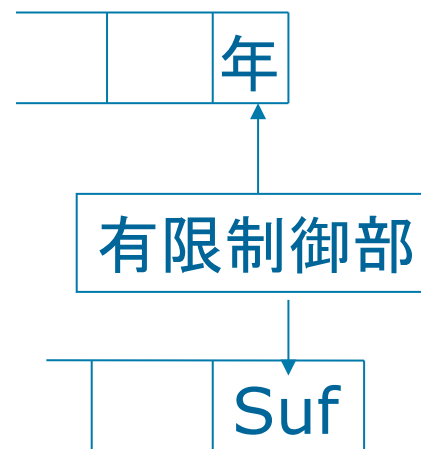
Step 3

21を受け付ける前に
Nenを展開してNum
までもっていく



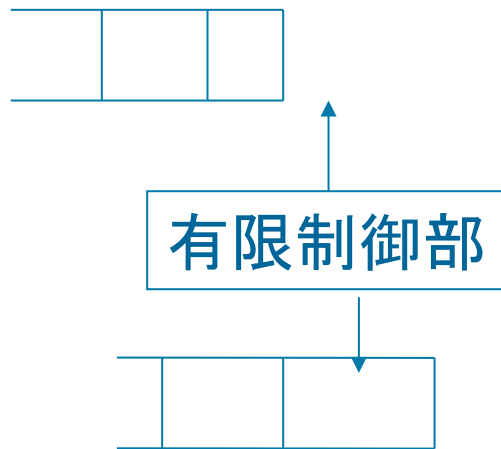
Step4

「21」を読み込んで
Num をスタックから
取り出す (pop)



S ::= Go Nen
Nen ::= Num Suf
Go ::= 平成
Num ::= 1 | 2 | 3 | ... | 21
Suf ::= 年

Step5
「年」を読み込んで
Suf をスタックから
取り出す (pop)



受理

pda の状態遷移関数とcfgの生成規則

S ::= Go Nen
Nen ::= Num Suf
Go ::= 平成
Num ::= 1 | 2 | 3 | ... | 21
Suf ::= 年

状態遷移関数

$$\begin{aligned}\delta(q, \varepsilon, S) &= \{(q, \text{Go Nen})\} \\ \delta(q, \varepsilon, \text{Nen}) &= \{(q, \text{Num Suf})\} \\ \delta(q, \text{平成}, \text{Go}) &= \{(q, \varepsilon)\} \\ \delta(q, 1, \text{Num}) &= \{(q, \varepsilon)\} \\ &\dots \\ \delta(q, \text{年}, \text{Suf}) &= \{(q, \varepsilon)\}\end{aligned}$$

様相間のステップ

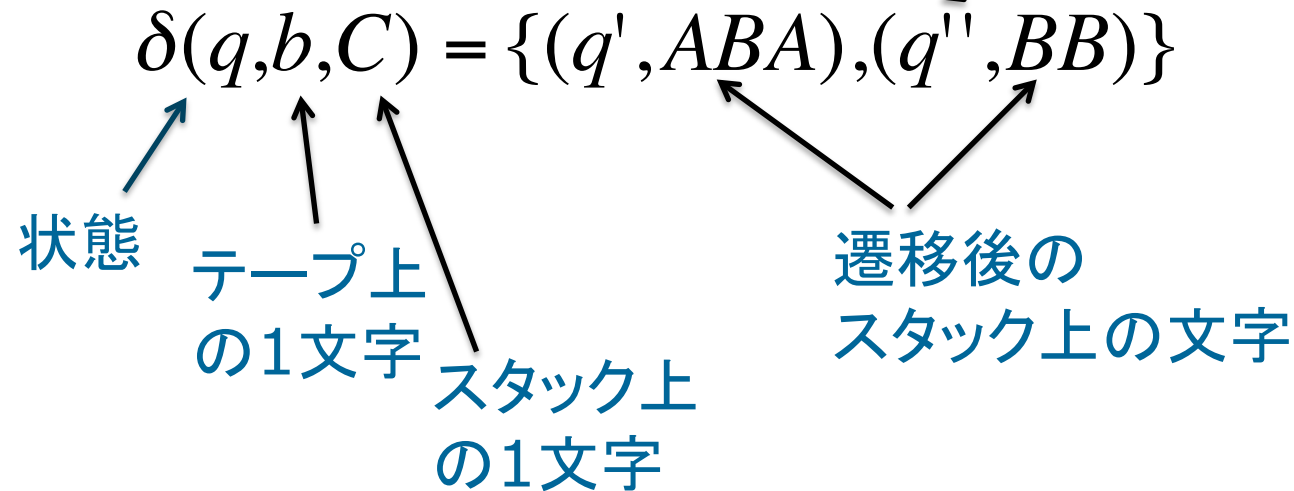
$$\begin{aligned}(q, \text{平成21年}, \text{GoNen}) \\ \Rightarrow (q', 21\text{年}, \text{Nen})\end{aligned}$$

状態遷移関数

状態遷移関数の説明

(例) 状態 q から q' または q'' に遷移

遷移後の
状態 q' 遷移後の
状態 q''



pda は非決定性なので、遷移先は複数あって良い

様相間のステップ

- ある状態から1ステップで次の状態に移った場合の記述

qからq'に1ステップで様相が変わった

$$\underline{(q, \text{平成21年}, GoNen)} \Rightarrow (q', \text{21年}, Nen)$$

様相 状態遷移関数との違いに注意

テープ上の1文字(単語)「平成」とスタックの`Go`を取り除き、テープには「21年」が残り、スタック上には`Nen`のみが残っている状態となった。この変化が1ステップで起こったことを上記の式は示している。

$$(q, y, S) \stackrel{*}{\Rightarrow} (q', \varepsilon, \varepsilon)$$

受理とは有限回遷移でテープ上の文字列もスタックも空になること

考えてみよう

下記の状態遷移はどんなことを表してるか説明せよ

$$\delta(q, \varepsilon, S) = \{(q', NV)\}$$

$$\delta(q, b, A) = \{(q', AB), (q'', \varepsilon)\}$$

練習

- 下記の生成規則に対応するpdaの状態遷移関数を記述せよ

S ::= Go Nen
Nen ::= Num Suf
Go ::= 平成
Num ::= 1 | 2 | 3 | ... | 21
Suf ::= 年