



# アルゴリズムとデータ構造III

## 2回目: 10月18日

---

### 文脈自由文法

授業資料 <http://ir.cs.yamanashi.ac.jp/~ysuzuki/algorithm3/index.html>

# 授業の予定(中間試験まで)

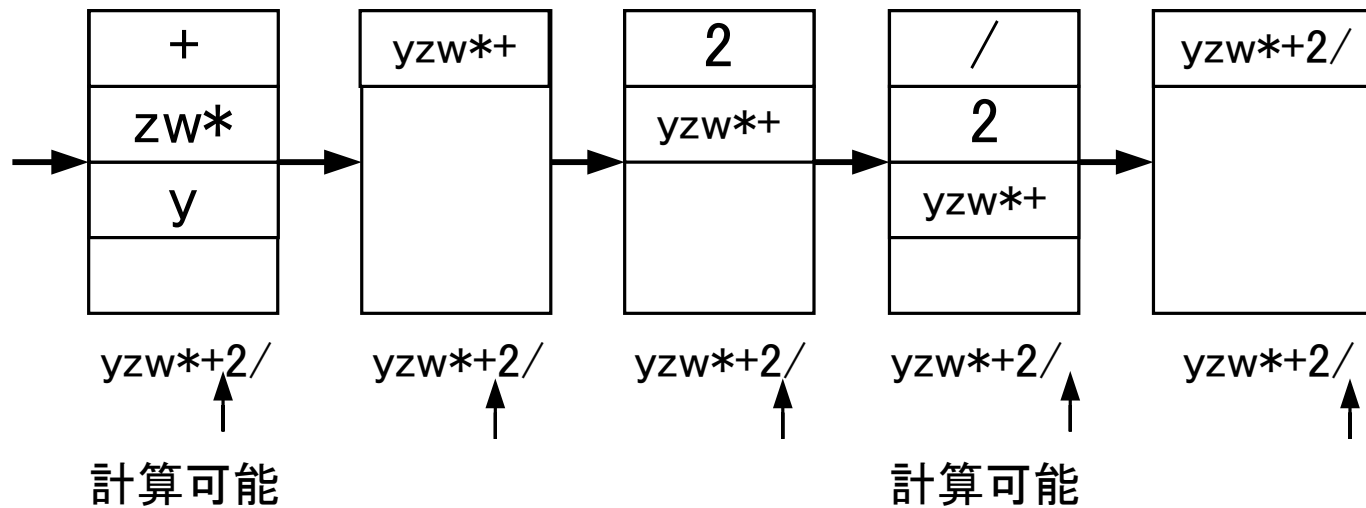
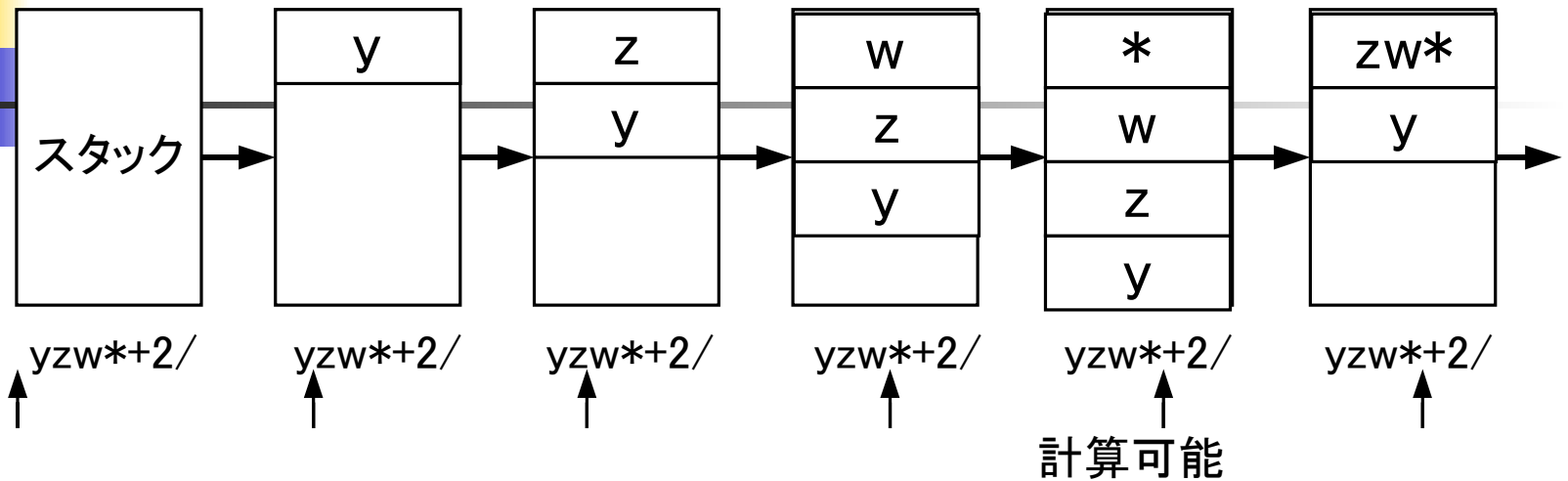
- スタック (後置記法で書かれた式の計算)
- 文脈自由文法
- 構文解析 CKY法
- 構文解析 チャート法, LR法
- グラフ(ダイクストラ法, 動的計画法, DPマッチング)
- グラフ(ビームサーチ, A\*アルゴリズム)
- グラフ(トライ構造, トライサーチ)
- 中間試験

# 授業の予定(中間試験以降)

1. 全文検索アルゴリズム (simple search, KMP, BM)
2. 全文検索アルゴリズム (Aho-Corasick)
3. テキスト圧縮 暗号 (例: モールス信号, 黄金虫, 踊る人形, ハフマン符号, Zipfの法則)
4. テキスト圧縮 zip
5. 音声圧縮 ADPCM, MP3
6. 音声圧縮 (CELP), 画像圧縮 (JPEG)
7. 期末試験

# 練習問題2の解答

## yzw\*+2/の計算方法(スタックの変化)



# 7 2 3 + - を計算してみよう

## (アセンブリ言語でプログラミング)

数式(7 2 3 + -)をメモリ(データ領域)に書き込まれている

### 3. データ領域から1文字読み込む

1. 数字(アスキーコード: 30H~39H)なら
  - 数値に変換し, 数値をスタックにプッシュ
2. 演算子なら
  1. 一旦スタックにプッシュし, ポップする.
  2. スタックからポップし, 数値をBレジスタに取り込む
  3. スタックからポップし, 数値をAレジスタ(アキュムレータ)に取り込む
  4. 演算子が+なら
    - A + B を計算し, Aレジスタに計算結果を格納
  5. 演算子が-なら
    - A - B を計算し, Aレジスタに計算結果を格納
  6. Aレジスタの内容をスタックにプッシュ

### 4. データ領域すべてを読み終えるまで続ける.

# 簡単な計算の例 7 2 3 + -

```
;後置記法 7 2 3 + - の計算
ORG 8000H ;
LD HL, DATA ; 数式の先頭番地を指定
LOOP: LD A, (HL)
      CP 00H
      JP Z, OWARI ; 数式を全部読み込んだら終わり

      LD E, (HL)
      LD D, 0H
      LD A, (HL)
      INC HL
      CP 2BH
      JP Z, LOOPA ; +なら加算処理へ
      CP 2DH
      JP Z, LOOPS ; -なら減算処理へ
      LD A, E
      SUB 30H ; 数字なら数値に変換
; Aレジスタの内容をスタックへプッシュ
STPUSH: LD E, A
        LD D, 0H
        PUSH DE ; 読み込んだ数値をスタックへプッシュ

        JP LOOP ; つぎの文字読み込みへ
```

```
;加算
LOOPA: PUSH DE ; 演算子をスタックへプッシュ
       POP DE ; 演算子をスタックからポップ
       POP DE ; 数値をスタックからポップ
       LD B, E ; スタックトップの値をBレジスタへ
       POP DE ; 数値をスタックからポップ
       LD A, E ; スタックトップの値をAレジスタへ
       ADD A, B ; 加算( A <= A + B )
       JP STPUSH

;減算
LOOPS: PUSH DE ; 演算子をスタックへプッシュ
       POP DE ; 演算子をスタックからポップ
       POP DE ; 数値をスタックからポップ
       LD B, E ; スタックトップの値をBレジスタへ
       POP DE ; 数値をスタックからポップ
       LD A, E ; スタックトップの値をAレジスタへ
       SUB B ; 減算( A <= A - B )
       JP STPUSH

;
OWARI: HALT
;入力データ
DATA:  DEFB 37H ;7
       DEFB 32H ;2
       DEFB 33H ;3
       DEFB 2BH ;+
       DEFB 2DH ;-
       DEFB 00H ;END
```

END



# 文脈自由文法

---

- 「オートマトンと言語」の最後の授業の続き



# 形式文法Gの定義

---

- $G=(N,T,P,S)$ 
  - N: 非終端記号の集合
  - T: 終端記号の集合
  - P: プロダクション
  - S: 開始記号



# 文脈自由文法と文脈依存文法

## ■ 文脈自由文法 (CFG)

- 文脈自由プロダクションのみから構成される
- 文脈自由プロダクション
  - $a \rightarrow \beta$
  - ただし,  $a \in N$ ,  $\beta \in V^*$
  - $N$ : 非終端記号の集合,  $T$ : 終端記号の集合,  $V$ :  $N$ と $T$ の直和
  - 左辺は変数1つ

# 文脈自由文法と文脈依存文法

## ■ 文脈依存文法(CSG)

- 文脈依存プロダクションを含むプロダクションから構成される
- 文脈依存プロダクション
  - $uav \rightarrow u\beta v$  ただし,  $a \in N$ ,  $u, v \in V^*$ ,  $\beta \in V^+$
  - $N$ : 非終端記号の集合,  $T$ : 終端記号の集合,  $V$ :  $N$ と $T$ の直和
  - $u=v=\varepsilon$ のとき  $(a \rightarrow \beta)$  文脈自由プロダクションとなる



# 文脈自由文法の例(例題5.9)

---

- CFG  $G=(N,T,P,S)$ 
  - $N$ (非終端記号) =  $\{B,S\}$
  - $T$ (終端記号) =  $\{a,b\}$
  - $P: S \rightarrow aSB \mid ab$   
 $B \rightarrow b$
- 語  $aaabbb$  の導出過程
- $L(G)$ はどのような言語か

- CFG  $G=(N,T,P,S)$

- $N=\{B,S\}$

- $T=\{a,b\}$

- $P: S \rightarrow aSB \mid ab$

## 例題5.9の解答例

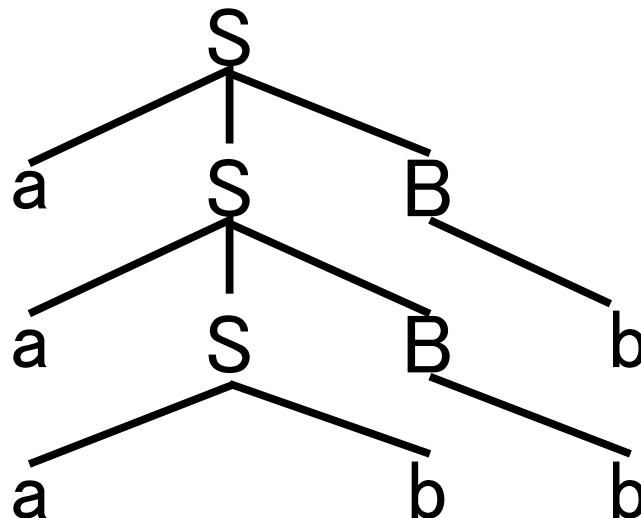
- $S \Rightarrow aSB \Rightarrow aaSB \Rightarrow aaabBB \Rightarrow aaabbB \Rightarrow aaabbb$

- $L(G): a^n b^n$

- 正規表現では表せない

- プッシュダウンオートマトンでは表現可能

- 構文木



# 練習問題1

## 例題5.10 文脈依存文法の例

---

- CSG  $G=(N,T,P,S)$
- $N=\{A,B,S\}$
- $T=\{a,b\}$
- $P: S \rightarrow aSBA \mid abA, AB \rightarrow BA, bB \rightarrow bb,$   
 $bA \rightarrow ba, aA \rightarrow aa$
- 語  $aabbaa$  の導出過程
- $L(G)$  はどのような言語か

# 練習問題1 解答

## 例題5.10 aabbaa

---

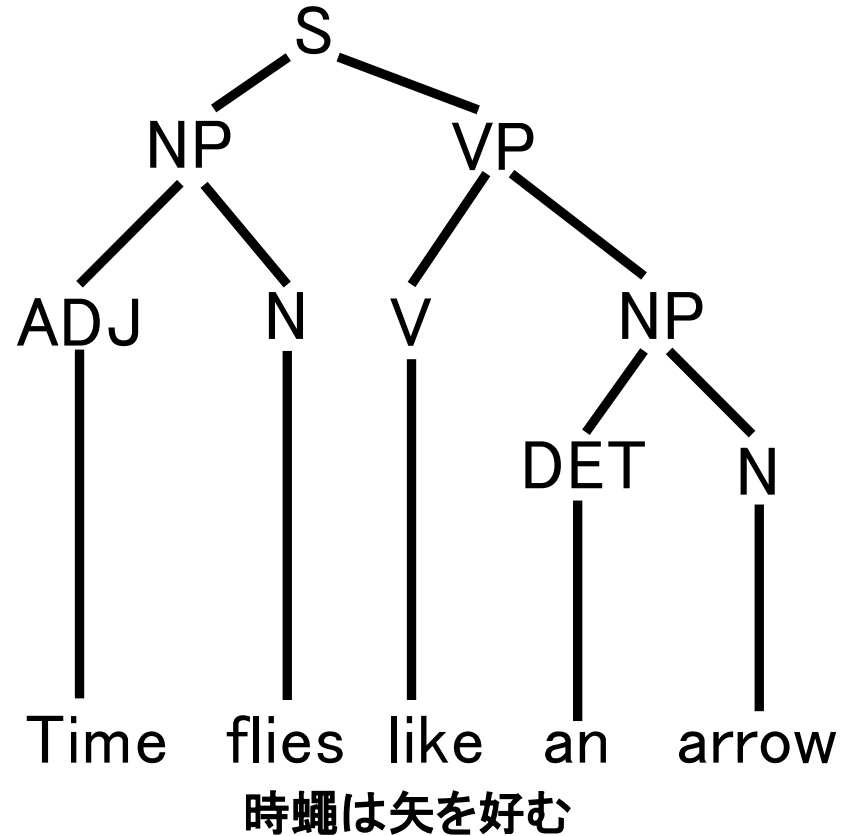
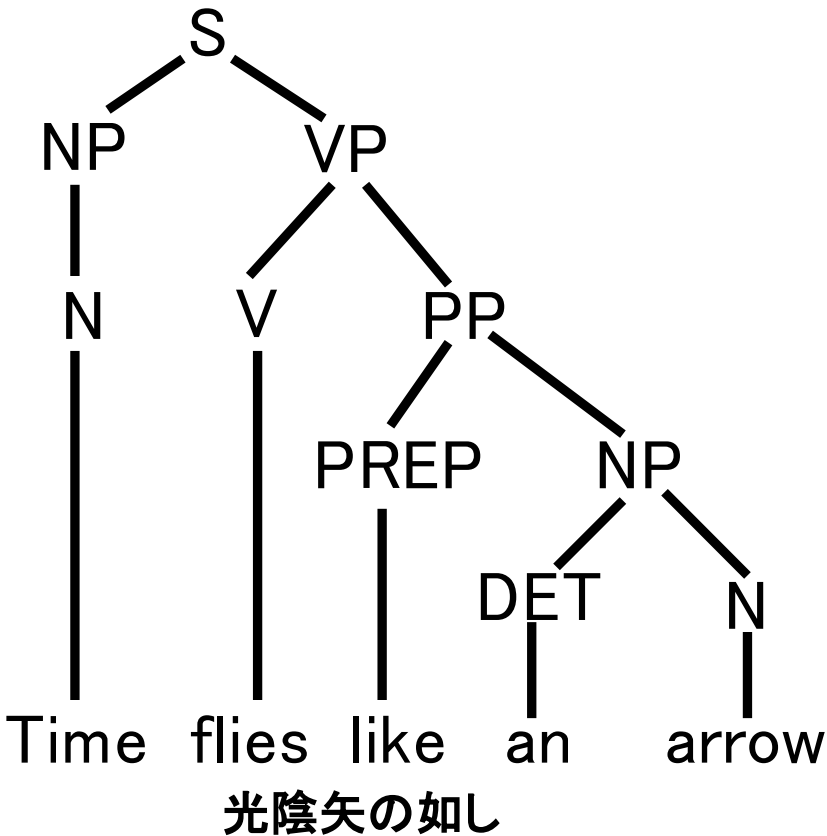
- $S \Rightarrow aSBA \Rightarrow aabABA \Rightarrow aabBAA \Rightarrow aabbAA$
- $\Rightarrow aabbaA \Rightarrow aabbaa$
- $L(G): a^n b^n a^n$

# 構文木(導出木)

- Time flies like an arrow.

S → NP VP  
NP → N | DET N | ADJ N  
VP → V PP | V NP  
PP → PREP NP  
N → Time | arrow | flies  
V → flies | like  
PREP → like  
DET → an

2種類の導出木

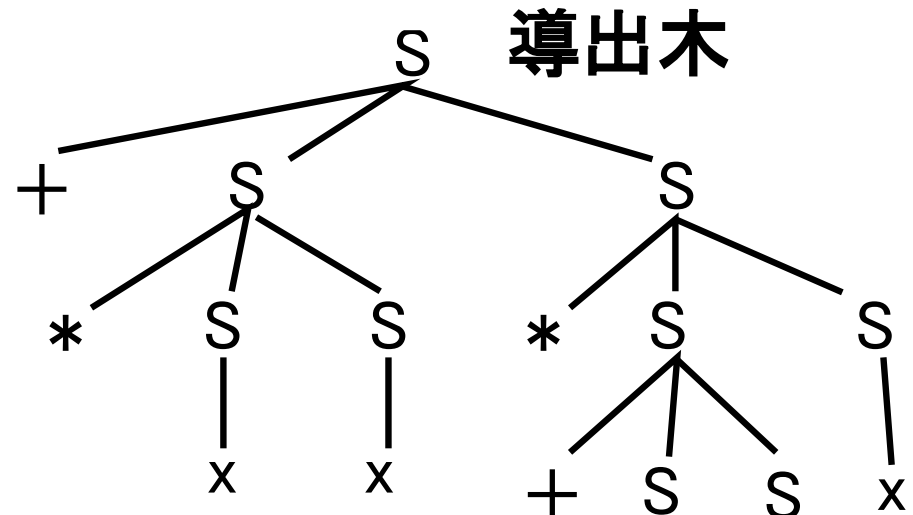


# ■ 問題: 例題5.11

■ 文法  $N=\{S\}, T=\{x, +, *\}, P=\{S \rightarrow +SS | *SS | x\}$

■ 語  $w = +*xx^*+xxx$  を導出せよ

■ 語  $w$  の導出木



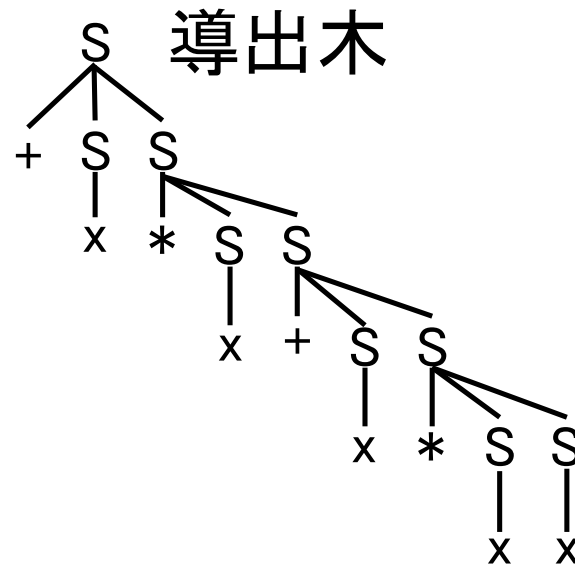
■ 解答

■ 導出:  $S \Rightarrow +SS \Rightarrow +*SSS \Rightarrow +*xSS \Rightarrow +*xxS \Rightarrow +*xx^*SS \Rightarrow +*xx^*+SSS \Rightarrow +*xx^*+xxx$



# 例題5.12 ①

- 問題
- 文法  $N=\{S\}, T=\{x, +, *\}, P=\{S \rightarrow +SS \mid *SS \mid x\}$
- 中置記法  $x+x*(x+x*x)$



- 解答例
- 前置記法  $+x*x+x*xx$
- $S \Rightarrow +SS \Rightarrow +xS \Rightarrow +x*SS \Rightarrow +x*xS \Rightarrow +x*x+SS$   
 $\Rightarrow +x*x+xS \Rightarrow +x*x+x*SS \Rightarrow +x*x+x*xS \Rightarrow +x*x+x*xx$

## 練習問題2 例題5.12 ②

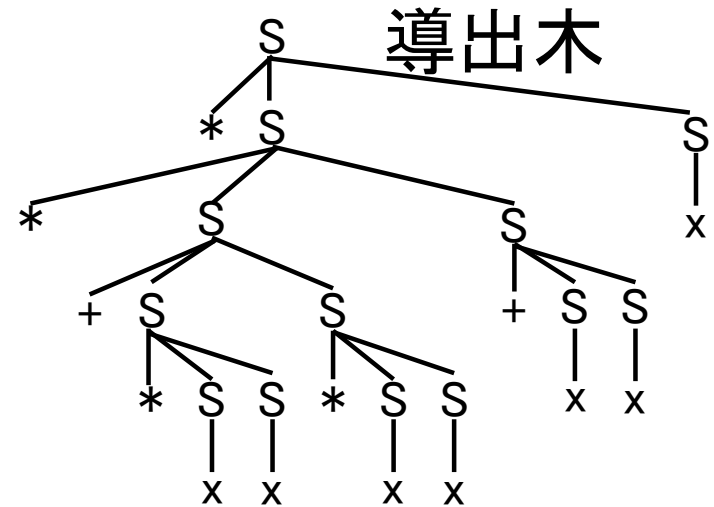
- 問題
- 文法  $N=\{S\}, T=\{x, +, *\}, P=\{S \rightarrow +SS \mid *SS \mid x\}$
- 中置記法  $(x*x+x*x)*(x+x)*x$
- 前置記法？
- 最左導出？
- 構文木？

# 練習問題2 例題5.12 ②の解答例

■ 問題

■ 文法  $N=\{S\}, T=\{x, +, *\}, P=\{S \rightarrow +SS \mid *SS \mid x\}$

■ 中置記法  $(x*x+x*x)*(x+x)*x$



■ 解答例

■ 前置記法  $**+*xx*xx+xxx$

■  $S \Rightarrow *SS \Rightarrow **SSS \Rightarrow **+SSSS \Rightarrow **+*SSSSS \Rightarrow **+*xSSSSS$   
 $\Rightarrow **+*xxSSS \Rightarrow **+*xx*SSSS \Rightarrow **+*xx*xSSS \Rightarrow **$   
 $+*xx*xxSS \Rightarrow **+*xx*xx+SSS$

■  $\Rightarrow **+*xx*xx+xxx$



# 文脈自由文法の曖昧性

---

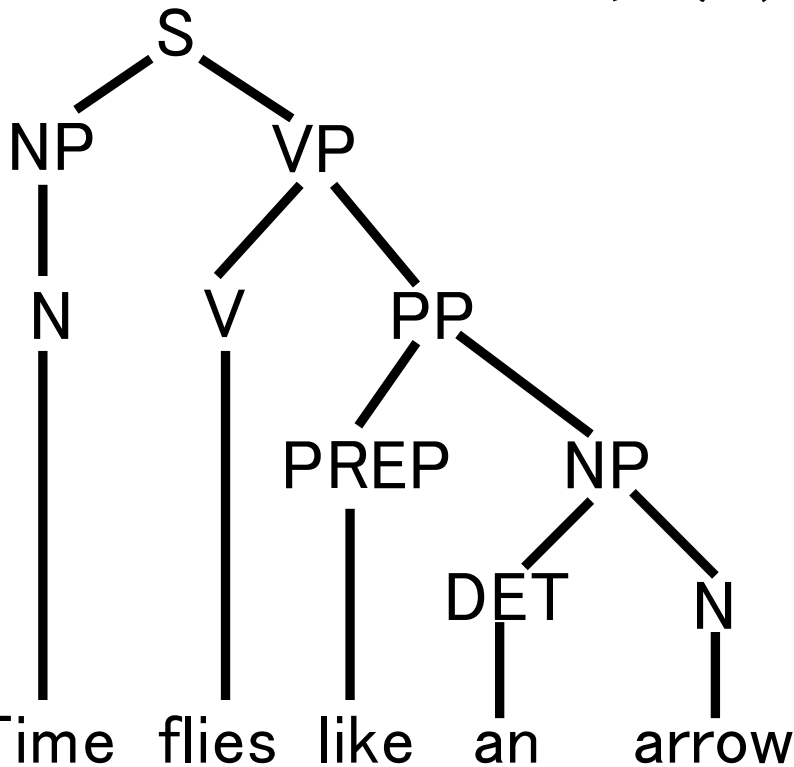
- どのような導出を行っても同じ導出木が得られる
- ⇒ 文法Gは曖昧でない
  
- 複数の異なった導出木が構成できるような語を含む
- ⇒ 文法Gは曖昧である

# 構文木(導出木)

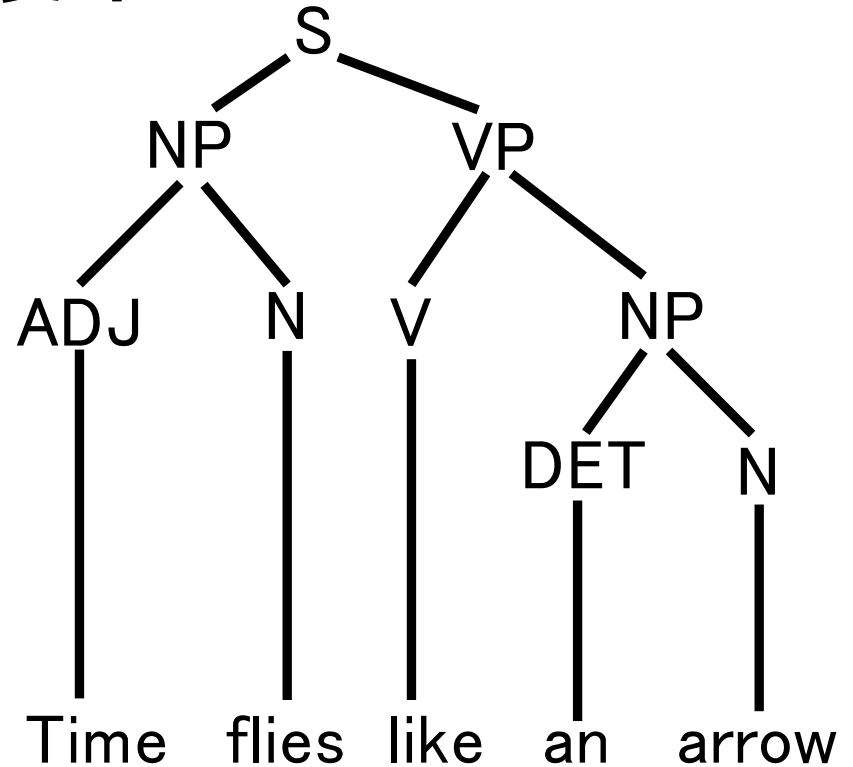
- Time flies like an arrow.

S → NP VP  
NP → N | DET N | ADJ N  
VP → V PP | V NP  
PP → PREP NP  
N → Time | arrow | flies  
V → flies | like  
PREP → like  
DET → an

2種類の導出木  
→ 文法が曖昧



光陰矢の如し



時蠅は矢を好む



## 例題5.26

---

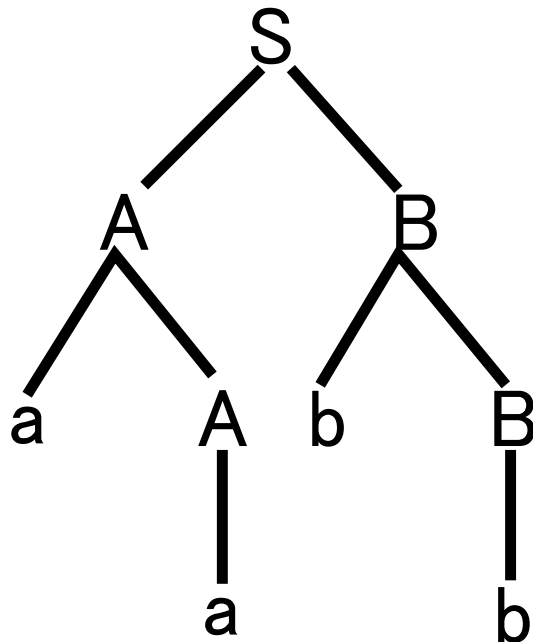
- 文法  $G=(N,T,P,S)$  において,  $N=\{S,A,B\}, T=\{a,b\}$ ,
- $P: S \rightarrow AB \mid aAB, A \rightarrow aA \mid a, B \rightarrow bB \mid b$
- この文法が曖昧であることを示せ

# 例題5.26 答え

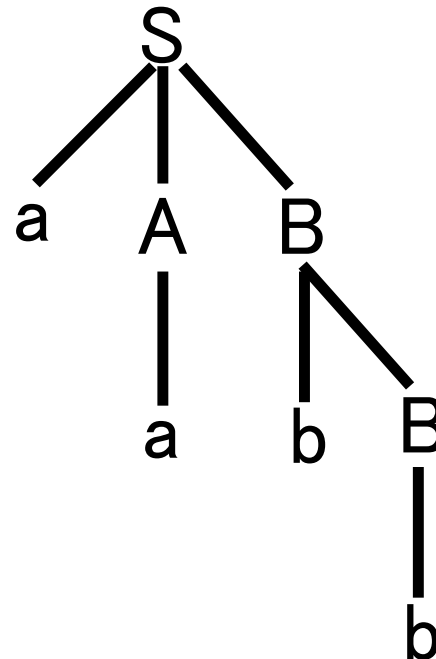
同一文字列に対して2種類の導出  
木が構成可能→曖昧である

- 1.  $S \rightarrow AB \rightarrow aAB \rightarrow aAbB \rightarrow aabB \rightarrow aabb$
- 2.  $S \rightarrow aAB \rightarrow aaB \rightarrow aabB \rightarrow aabb$

1.



2.



# 練習問題3

## 例題5.27

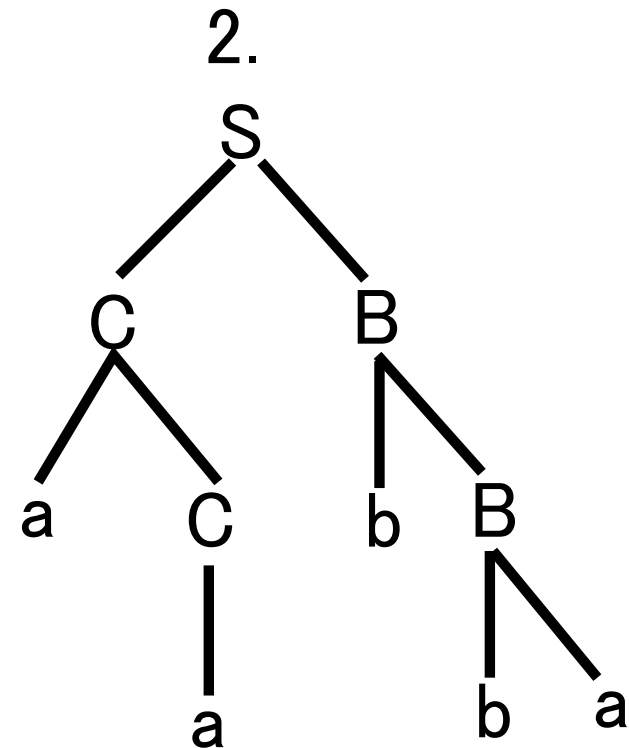
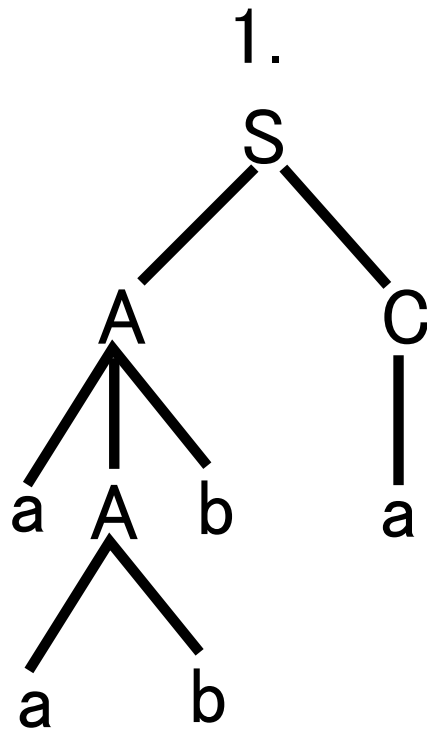
- 文法 $G=(N,T,P,S)$ において,
  - $N=\{S,A,B,C\}, T=\{a,b\}$ ,
  - $P: S \rightarrow AC|CB, A \rightarrow aA|a, A \rightarrow aAb|ab, B \rightarrow bB|ba$
  - $C \rightarrow aC|a$
- 
- この文法が曖昧であることを,  $aabba$ の導出木を構成して示せ



# 練習問題3 例題5.27 答え

同一文字列に対して2種類の導出木が構成可能 → 曖昧である

- 1.  $S \rightarrow AC \rightarrow aAbC \rightarrow aAba \rightarrow aabba$
- 2.  $S \rightarrow CB \rightarrow aCB \rightarrow aCbB \rightarrow aabB \rightarrow aabba$



# CFGの構文図式

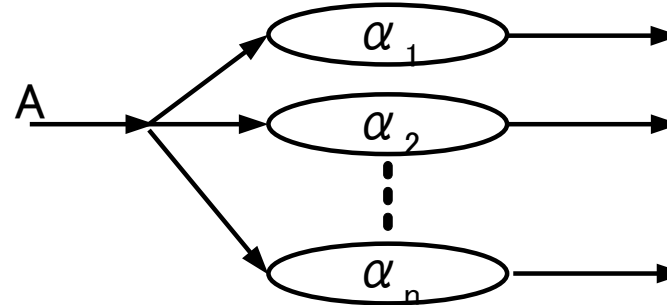
文脈自由プロダクション

$$A \rightarrow \alpha$$

構文図式



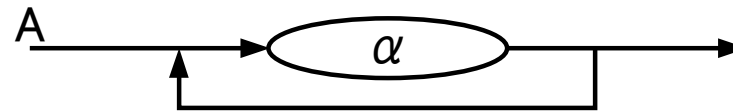
$$A \rightarrow \alpha_1 | \alpha_2 | \dots | \alpha_n$$



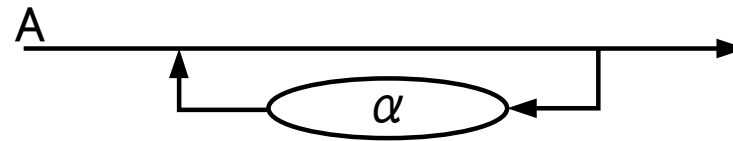
$$A \rightarrow \alpha_1 \alpha_2 \dots \alpha_n$$



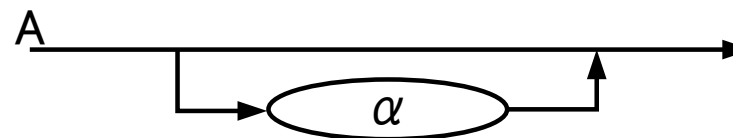
$$A \rightarrow \alpha | \alpha A |$$



$$A \rightarrow \varepsilon | \alpha | \alpha A |$$



$$A \rightarrow \varepsilon | \alpha$$





# 構文解析アルゴリズム

---

1. CKY (Cocke-Kasami-Younger)法
2. チャート法, LR法