前回の復習

Σ: アルファベット。言語を構成する記号の集合
Σ^*: 有限の長さの記号全体の集合
言語: Σ^* の部分集合

句構造文法
$G = (N, T, P, S)$:
P: $\alpha \to \beta$, $\alpha \in (N \cup T)^+$, $\beta \in (N \cup T)^*$ という形の書き換え規則の集合
N: 非終端記号の集合
T: 終端記号の集合
S: 開始記号の集合

$L(G)$: G の開始記号を書き換えて得られる終端記号列の集合。

演習問題解答

演習問題1 ある言語 L について L^* が空列 ϵ を含まないことはあるか？
答え：ない。

$$L^* = L^0 \cup L^1 \cup \cdots, \quad L^0 = \{\epsilon\}$$

なので。

演習問題2 L^+ が空列 ϵ を含むのは、言語 L がどのような条件を満たすときか？
答え： L が空列 ϵ を含むとき (答えは十分条件、例えば $L = \{\epsilon\}$ などで正解)

演習問題3 $N = \{E, T', F\}, T = \{a, b, c, +, -, \times, /, (,)\}$ としたときに $\epsilon \to a$ という生成規則を持つ句構造文法は有り得るか？理由を付けて答えよ。
答え：有り得ない。なぜならば、生成規則の左辺は $(N \cup T)^+$ の要素で、$(N \cup T)^+$ は空列を含まないから。

演習問題4 $N=\{E,T',F\}, T=\{a,b,c,+,-,\times,/,(,)\}$ としたときに $aa \to a$ という生成規則を持つ句構造文法は有り得るか？理由を付けて答えよ。
答え：有り得る。左辺が終端記号だけからなる列であるのが奇妙な感じがするが、定義上ありえる。しかし実際には左辺が終端記号だけからなる生成規則を含む文法を扱うことは稀である。

演習問題5 式の文法 $G=(N,T,P,E), N=\{E,T',F\}$, $T=\{a,b,c,+,-,\times,/,(,)\}, P=\{E\to T'|E+T'|E-T', T'\to F|T'\times F|T'/F, F\to a|b|c|(E)\}$ に対して、G の文を一つ挙げ、その文の導出を $\alpha_1 \Longrightarrow \alpha_2 \Longrightarrow \cdots \Longrightarrow \alpha_n$ という形式で書け。なるべく周りの人とは異なる文を挙げること。
答え：講義で例を一つ挙げたので省略

演習問題6 問題5の文法の文型を一つ挙げよ。なるべく周りの人とは異なる文型を挙げること。
答え：$E+T'$ など。非終端記号を含む列を書かないと文になってしまうので不正解。

演習問題7 問題5で解答した導出に対応する導出木を書け
授業でやったので解答略

演習問題 8 アルファベット $\Sigma = \{0, 1\}$ 上の言語 $L = \{1, 01\}^*$ を生成する句構造文法を書け
答え：正解はたくさん有る。例えば

$$G = (\{S\}, \{0, 1\}, \{S \to \epsilon | SS | 1 | 01\}, S)$$

など

コメントへの回答

コメント１：OHP をめくるのが速いので、ノートを取ることに気を取られて話を聞けず、内容が良く分からなかった
回答１：OHP のコピーを配布するようにした

コメント２：もっと例を出して欲しい
回答２：そうしたつもり

コメント３：演習の時間を長くしてほしい
回答３：今日はそうするつもり

コメント４：授業内容をすくなくしてまとめて欲しい
回答４：すでに内容をかなり削っている。これ以上削ると「オートマトンと言語」ではなくなってしまう

あまり内容を削ると、この授業を理解しても大学院入試問題を解けなくなる

正規文法

プログラムの中の符号無し二進整数：0 または 1 の後 $\{0,1\}$ が 0 個以上続く続く文字列

これを表す句構造文法

$G = (\{D,U\},\{0,1\},P,U)$,

$P = \{U \to 0|1D, \quad D \to \epsilon|0|1|0D|1D\}$

例: $U \Longrightarrow 0$

$U \Longrightarrow 1D \Longrightarrow 1$

$U \Longrightarrow 1D \Longrightarrow 10$

上の文法は生成規則が $A \to aB$ または $A \to a$ または $A \to \epsilon$ (但し A, B は非終端記号で、a は終端記号) という特徴がある。

C 言語で関数名や変数名になり得る文字列も上のような形の言語で記述できる。このような言語を正規言語と呼び、この講義の前半で学ぶ。

正規文法の定義

授業で使う定義: 正規文法とは、句構造文法 $G = (N, T, P, S)$ で生成規則がすべて

$$\begin{aligned} A &\rightarrow \epsilon, \\ A &\rightarrow a, \\ A &\rightarrow aB \end{aligned}$$

のいずれかであるもの。但し、大文字は非終端記号で、小文字は終端記号。

教科書 42 ページの定義: 正規文法とは、句構造文法 $G = (N, T, P, S)$ で右辺に ϵ を含まない生成規則はすべて

$$\begin{aligned} A &\rightarrow a, \\ A &\rightarrow aB \end{aligned}$$

のいずれかであり、生成規則が右辺に ϵ を含む規則は有るとすれば $S \rightarrow \epsilon$ だけであるもの。

正規言語：正規文法により生成される言語

正規文法の定義2

教科書の定義の方が制限が強いので、授業の定義で表せるが、教科書の定義で表せない言語がありそうに見える。しかし、教科書の定義による文法から授業の定義を満たす等価な文法を作ることができる。作り方は後に時間があれば説明する。

正規文法はどのように使われるのか？

ある記号列がある正規文法により定義される言語に属するかどうかは、後で述べるオートマトンにより容易に判定できる。オートマトンは１文字ずつ記号列を読み、そこまで読んだ記号列が言語に属するかどうか判定する。そこで、プログラムを１文字ずつオートマトンに入力して言語に属するかどうか判定し、言語に属する記号列ごとに区切る。

$L_1 \cup L_2$ も正規言語

L_1, L_2 が正規言語であるとき $L_1 \cup L_2$ も正規言語である。

正規文法の記述を簡潔に行う正規表現をあとで学ぶが、上の事実から、正規表現が正規文法を表せることがわかる。

$L_1 = L(G_1), L_2 = L(G_2), G_1 = (N_1, T_1, P_1, S_1), G_2 = (N_1, T_1, P_1, S_1)$ とする。$L_1 \cup L_2$ を生成する文法を作る

1. 非終端記号を置き換えて $N_1 \cap N_2 = \emptyset$ にする
2. $G = (N_1 \cup N_2, T_1 \cup T_2, P, S)$ で生成規則の集合を
$$P = P_1 \cup P_2 \cup \{S \to \alpha_1 | \cdots | \alpha_m | \beta_1 | \cdots | \beta_n\}$$
とする。但し $S_1 \to \alpha_1 | \cdots | \alpha_m \in P_1$, $S_2 \to \beta_1 | \cdots | \beta_n \in P_2$.

例: $P_1 = \{S \to \epsilon | a | aS\}$, $P_2 = \{S \to \epsilon | b | bS\}$

1. 非終端記号を置き換えて $P_1' = \{S_1 \to \epsilon | a | aS_1\}$, $P_2' = \{S_2 \to \epsilon | b | bS_2\}$ とする
2. $P = P_1' \cup P_2' \cup \{S \to \epsilon | a | aS_1 | b | bS_2\}$
$= \{S_1 \to \epsilon | a | aS_1,\ S_2 \to \epsilon | b | bS_2,$
$S \to \epsilon | a | aS_1 | b | bS_2\}$.

$L_1 L_2$ も正規言語

$L_1,\ L_2$ が正規言語であるとき $L_1 L_2$ も正規言語である。

正規文法の記述を簡潔に行う正規表現をあとで学ぶが、上の事実から、正規表現が正規文法を表せることがわかる。

$L_1 = L(G_1), L_2 = L(G_2), G_1 = (N_1, T_1, P_1, S_1), G_2 = (N_1, T_1, P_1, S_1)$ とする。$L_1 L_2$ を生成する文法を作る

1. 非終端記号を置き換えて $N_1 \cap N_2 = \emptyset$ にする
2. P_1 の生成規則を以下のように置き換え、置き換えた生成規則の集合を P_1'' とする

置き換え前	置き換え後
$A \to a$	$A \to aS_2$
$A \to \epsilon$	$A \to$ (S_2 の右辺)

3. 求める文法は $(N_1 \cup N_2, T_1 \cup T_2, P_1'' \cup P_2, S_1)$

$L_1 L_2$ も正規言語2

例: $P_1 = \{S \to \epsilon | a | aS\}$, $P_2 = \{S \to \epsilon | b | bS\}$

1. 非終端記号を置き換えて $P_1' = \{S_1 \to \epsilon | a | aS_1\}$, $P_2' = \{S_2 \to \epsilon | b | bS_2\}$ とする

2. P_1'' のための置き換え

置き換え前	置き換え後
$S_1 \to aS_1$	$S_1 \to aS_1$
$S_1 \to a$	$S_1 \to aS_2$
$S_1 \to \epsilon$	$S_1 \to \epsilon$ $S_1 \to b$ $S_1 \to bS_1$

3. 求める文法は $(\{S_1, S_2\}, \{a, b\}, P_1'' \cup P_2', S_1)$.

L^* も正規言語

L が正規言語であるとき L^* も正規言語である。

$L = L(G)$, $G = (N, T, P, S)$ とする。L^* を生成する文法を作るには、$S \to \epsilon$ および 以下の表の右の規則を P に追加する

P の規則	新たに追加する規則
$A \to a$	$A \to aS$
$A \to \epsilon$	$A \to$ (S の右辺)

例: $P = \{S \to \epsilon|a|b\}$。このとき $L(G) = \{\epsilon, a, b\}$.

P の規則	新たに追加する規則
$S \to a$	$S \to aS$
$S \to b$	$S \to bS$
$S \to \epsilon$	$S \to a$
$S \to \epsilon$	$S \to b$

生成規則の重複を除いて、できる生成規則の集合は

$$P' = \{S \to \epsilon|a|b|aS|bS\}$$

授業の定義による正規文法を教科書の定義に変換する方法

ϵ 生成規則：右辺が ϵ である生成規則

授業の定義による正規文法を $G=(N,T,P,S)$ とし、教科書の定義による文法 $G'=(N',T',P',S')$ を以下のように作る。

1. $P'=\emptyset$ とする
2. P の ϵ 生成規則でない規則をそのまま P' に追加
3. ϵ 生成規則の左辺の非終端記号をすべて列挙 (これを ϵ 生成記号と呼ぶ)
4. $P \ni A \to aB$ について B が ϵ 生成記号ならば $A \to a$ を P' に追加
5. $S \to \epsilon \in P$ なら、$S' \to \epsilon|$ (P' の S の規則の右辺) を P' に追加し、$N' = N \cup \{S'\}$ とする。$S \to \epsilon \notin P$ なら、$S'=S$, $N'=N$.

例: $P=\{S \to \epsilon|aS|bB, B \to \epsilon|bB\}$
2. $P'=\{S \to aS|bB, B \to bB\}$
3. S, B
4. $P'=\{S \to a|b|aS|bB, B \to b|bB\}$
5. $P'=\{S' \to \epsilon|a|b|aS|bB, S \to a|b|aS|bB,$
$B \to b|bB\}$

演習問題一覧

答案に学籍番号、名前、ふりがなを書いて下さい。

演習問題10 アルファベット $\Sigma = \{0, 1\}$ 上の言語 $L_1 = \{1, 01\}^*$ を生成する正規文法を書け (言語自体は問8と同じ)

演習問題11 アルファベット $\Sigma = \{0, 1\}$ 上の言語 $L_2 = \{00, 11\}$ を生成する正規文法を書け (注意: $S \to 00|11$ は正規文法ではない)

演習問題12 L_2^* を生成する正規文法を問11の解答を用いて授業で説明したやり方に沿って作れ

演習問題13 $L_1 L_2^*$ (L_1 と L_2^* の連接) を生成する正規文法を問10,12の解答を用いて授業で説明したやり方に沿って作れ

演習問題14 $L_1 \cup L_2^*$ (L_1 と L_2^* の和集合) を生成する正規文法を問10,12の解答を用いて授業で説明したやり方に沿って作れ

演習問題15 今日の授業でわかりにくい所や要望を書いて下さい

来週は休講です