平成 28 年度 言語・オートマトン 演習：生成文法と文脈自由言語

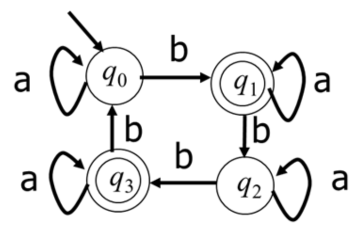
1. 講義中に与えた句構造文法 G_1 について次の各問に答えよ.
   (a) aaaabbbb $\in L(G_1)$ であることを示せ.
   (b) aaabbaa$\notin L(G_1)$, aaaabb$\notin L(G_1)$ である理由を説明せよ.
2. 講義中に与えた句構造文法 G_4 について次の各問に答えよ.
   (a) aaabbbccc$\in L(G_4)$ であることを示せ.
   (b) aabbccc$\notin L(G_4)$, aaabbccc$\notin L(G_4)$ である理由を説明せよ.
3. 講義中に与えた句構造文法 G_2 について次の各問に答えよ.
   (a) $L(G_2) = \{$ ababbabbabb $\}$ であることを示せ.
   (b) ababbabbabb の構文木を描け.
4. * 次の生成規則の集合 P を用いた句構造文法 $G = (\{S\}, \{\mathrm{a}, \mathrm{b}\}, P, S)$ について, (1)aaaaabbb$\in L(G)$ であることを示せ. (2) bbbaaaaa$\in L(G)$ であるか? 理由をつけて答えよ.

$$\{S \to \mathrm{a}\ S,\ S \to S\ \mathrm{b},\ S \to \mathrm{a},\ S \to \mathrm{b}\}$$

5. 次の生成規則の集合 P を用いた句構造文法 $G = (\{S\}, \{\mathrm{i,\ e,\ a}\}, P, S)$ があいまいであることを示す文字列 $w \in L(G)$ を構成せよ.

$$\{S \to \mathrm{i}\ S,\ S \to \mathrm{i}\ S\ \mathrm{e}\ S,\ S \to \mathrm{a}\}$$

6. 2 項演算子を引数の前に置いて表記する方法を**前置記法**あるいは**ポーランド記法** (Polish notation) という. たとえば, $a + b$, $c * (a + b)$ をポーランド記法で表すとそれぞれ $+\,a\,b$, $*\,c\ +a\,b$ となる. 2 項演算子についてポーランド記法を用いると括弧を用いる必要がないことに注意しよう.
   (a) $((a + c) + b) * (a * (b + c))$ をポーランド記法で表せ.
   (b) $\Sigma = \{a, b, c, +*\}$ とするとき, ポーランド記法で表された式を生成する文脈自由文法 G を与えよ.
   (c) 前小問で与えた G はあいまいな文法であるか? (上述の注意はあいまいではない文法が構成できることを示唆している)
7. 図に示す有限状態オートマトン M について, $L(M)$ を生成する正則文法を与えよ.



8. ** 句構造文法 $G = (N, \Sigma, P, S)$ において, P の中のすべての生成規則が

$$A \to c, \qquad A \to Bc, \qquad A \to \varepsilon \qquad (A, B \in N,\ c \in \Sigma)$$

のいずれかの形をしているとき, **左正則文法**であるという [1]. ある言語 L が左正則文法で生成されることと正則言語であることは同値であることを証明せよ.

[1] 講義中で定義した正則文法を左正則文法と**右正則文法**ともよばれる.

9. アルファベットを Σ とするとき, 文字列 $w \in \Sigma^*$ を逆順に並べて得られる文字列を w^R と表すことにする．形式言語 $L_{Pal} = \{w \in \Sigma^* \mid w = w^R\}$ の要素を**回文 (palindrome)** という．以下では $\Sigma = \{a, b\}$ とする．
   (a) $L(G_{Pal}) = L_{Pal}$ であるような文脈自由文法 G_{Pal} を与えよ．
   (b) 前問で与えた G_{Pal} を (Chomsky 標準型でない場合) は Chomsky 標準型に変換せよ．
   (c) CYK 法により aababaa $\in L_{Pal}$ であることを示せ．また, aababba $\notin L_{Pal}$ であることを示せ．
   (d) G_{Pal} を Greibach 標準型に変換せよ．

10. 正しい数式から括弧 (‘(’, ‘)’, ‘[’, ‘]’ など) だけを取り出して得られる言語を **Dyck 言語**とよぶ．括弧では分かり辛いので, $\Sigma_{Dyck} = \{a, \overline{a}, b, \overline{b}\}$ としておく [2].Dyck 言語は文脈自由文法 $G_{Dyck} = (\{S\}, \Sigma_{Dyck}, P_{Dyck}, S\}$ で定義される．ここで,

$$P_{Dyck} \quad = \quad \{S \to SS,\ S \to aS\overline{a},\ S \to bS\overline{b},\ S \to \varepsilon\}$$

である．
   (a) G_{Dyck} を Chomsky 標準型に変換せよ．
   (b) CYK 法により $ab\overline{b}\overline{a}b\overline{b} \in L_{Dyck}$ であることを示せ．また, $ba\overline{b}\overline{a}b\overline{b} \notin L_{Dyck}$ であることを示せ．
   (c) G_{Dyck} を Greibach 標準型に変換せよ．
   (d) 通常の数式では括弧に階層が定義されている．例えば, (‘(’, ‘)’ は ‘[’, ‘]’ よりも階層が低いとされ, [()()] は許されるが, ([]) は許されない．そこで Σ_{Dyck} において, a と $\overline{a}$ の階層が b と $\overline{b}$ より低いと仮定して, G_{Dyck} を修正し, その Chomsky 標準型と Greibach 標準型を与えよ．

11. 文脈自由文法 $G = (N, \Sigma, P, S)$ において, $L(G)$ の要素の生成に用いられない非終端記号 X は**無効記号 (useless symbol)** とよばれる．数学的には $S \Rightarrow \cdots \Rightarrow \alpha X \beta \Rightarrow \cdots \Rightarrow w \in \Sigma^*$ であるような導出が**構成できない**ような非終端記号 X が無効記号である．

   文脈自由文法 $G_5 = (\{S, A, B, C\}, \{a, b\}, P_6, S)$ において無効記号を求め, $L(G_6) = L(G_6')$ なる G_1' で無効記号のないものを求めよ．ここで

$$P_6 \quad = \quad \left\{ \begin{array}{llllllll} S \to AB & , & S \to CA & , & A \to a & , & & \\ B \to BC & , & B \to AB & , & C \to aB & , & C \to b \end{array} \right\}.$$

12. 生成規則が $X_1 \to X_2$ $(X_1, X_2 \in N)$ という形をしているとき, **単位生成規則 (unit production)** とよぶ．文脈自由文法 $G_2 = (\{S, A, B\}, \{a, b\}, P_7, S)$ において $L(G_7) = L(G_7')$ なるで単位生成規則がないものを求めよ．もし無効記号がもし存在すればそれも除去すること．さらに得られた文法を Greibach 標準型に変換せよ．ここで

$$P_7 \quad = \quad \left\{ \begin{array}{llllllll} S \to ASB & , & S \to \varepsilon & , & A \to aAS & , & A \to a & , \\ B \to SbS & , & B \to A & , & B \to bb & & & \end{array} \right\}.$$

13. 左正則文法 $G = (\{S, A\}, \{a, b\}, \{S \to Ab,\ A \to Sa, A \to a\}, S)$ を右正則文法に変換せよ．

14. これまでの演習問題で用いた文法とその結果を利用して以下の言語を受理する NPDA を構成せよ．
   (a) L_{Pal}
   (b) $L_{Dyck} = L(G_{Dyck})$

[2] マークアップ言語である HTML や XML における**タグ**に相当する．

(c) $L(G_6)$

(d) $L(G_7)$

15. アルファベット Σ が与えられたとき, $\sharp \notin \Sigma$ なる新しい文字 $\sharp$ を用意して, $\{\Sigma \cup \{\sharp\}$ 上の言語 $L^{\sharp}_{Pal}$ を $L^{\sharp}_{Pal} = \{w\sharp(w^R) \mid w \in \Sigma^*\}$ と定義する. 以下では $\Sigma = \{a, b\}$ とする.

   (a) 語 $w \in (\Sigma \cup \{\sharp\})^*$ で $w \in L^{\sharp}_{Pal}$ かつ $|w| = 5$ であるようなものを列挙せよ.

   (b) $L^{\sharp}_{Pal}$ を空スタックで受理する DPDA を与えよ.

   (c) 前問の L_{Pal} を受理するような DPDA は (空スタックによる受理であっても終了状態による受理であっても) 構成することはできない. その理由を述べよ.