句構造文法（正規文法）

句構造文法の定義（再掲）

定義2：句構造文法 $G = (V, \Sigma, P, S)$

- V:記号の集合
- Σ(⊂V)：アルファベット

  要素は終端記号(terminal symbol)という。
- $V_N = V - \Sigma$

  要素は非終端記号(nonterminal symbol)

  $V = \sum \cup V_N \quad \sum \cap V_N = \phi$
- P:ルール（構文規則）の集合

  要素はα→β,　ただし　$\alpha \in V^+, \beta \in V^*$
- S:開始記号(initial symbol)

  $S \in V_N$

文法に４つの型(ルールに制限)

- タイプ0文法G0: 制限なし
- タイプ1文法G1: 文脈依存文法(csg)
- タイプ2文法G2:文脈自由文法（cfg）
- タイプ3文法G3:正規文法(rg)

$$A \to a \in P, A \to aB \in P \quad a \in \sum, \; A, B \in V_N$$

- タイプ3文法から生成される言語:

⇒　正規言語（rl）

例

$G_h=(\{S,S_l,S_p,S_{pl},S_r\}\cup\Sigma,\Sigma,P_h,S)$

$\Sigma=\{a,+,\times,(,)\}$

$P_h=\{S\rightarrow a,\quad S\rightarrow aS_p,\quad S\rightarrow (S_l,\quad S\rightarrow +S,$

$S_p\rightarrow +S,\quad S_p\rightarrow \times S,$

$S_l\rightarrow aS_r,\quad S_l\rightarrow aS_{pl},\ S_l\rightarrow +S_l,$

$S_{pl}\rightarrow +S_l,\ S_{pl}\rightarrow \times S_l,$

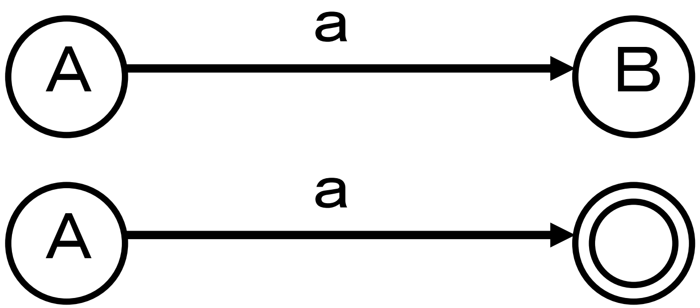
$S_r\rightarrow),\quad S_r\rightarrow)S_p\}$

$S \Rightarrow aS_p \Rightarrow a\times S \Rightarrow a\times (S_1 \Rightarrow a\times (aS_{p1} \Rightarrow a\times (a +S_1 \Rightarrow a\times (a + aS_r$
$\Rightarrow a\times (a + a)$

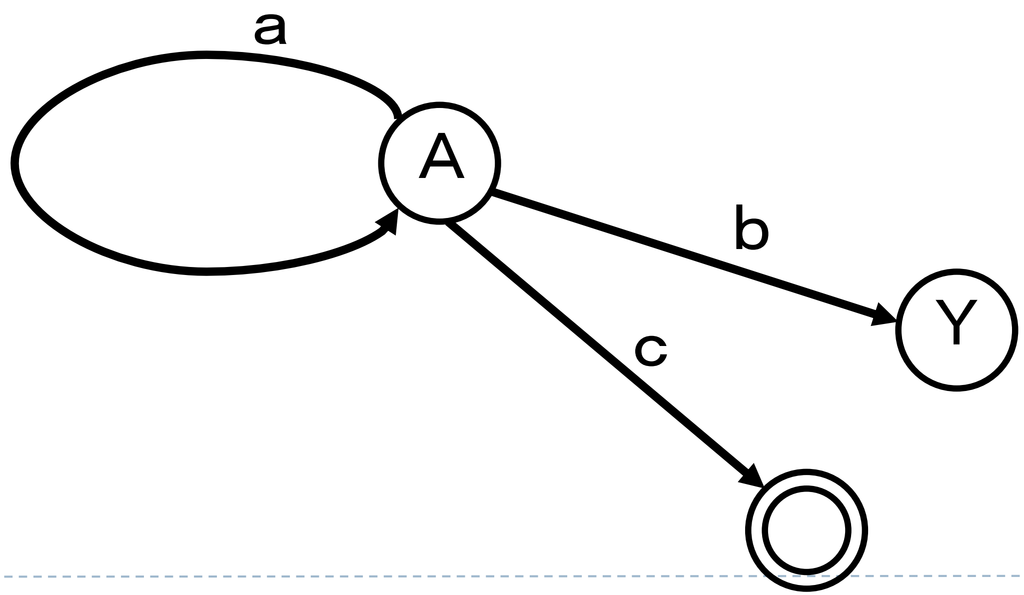
タイプ３文法から有限オートマトンの生成

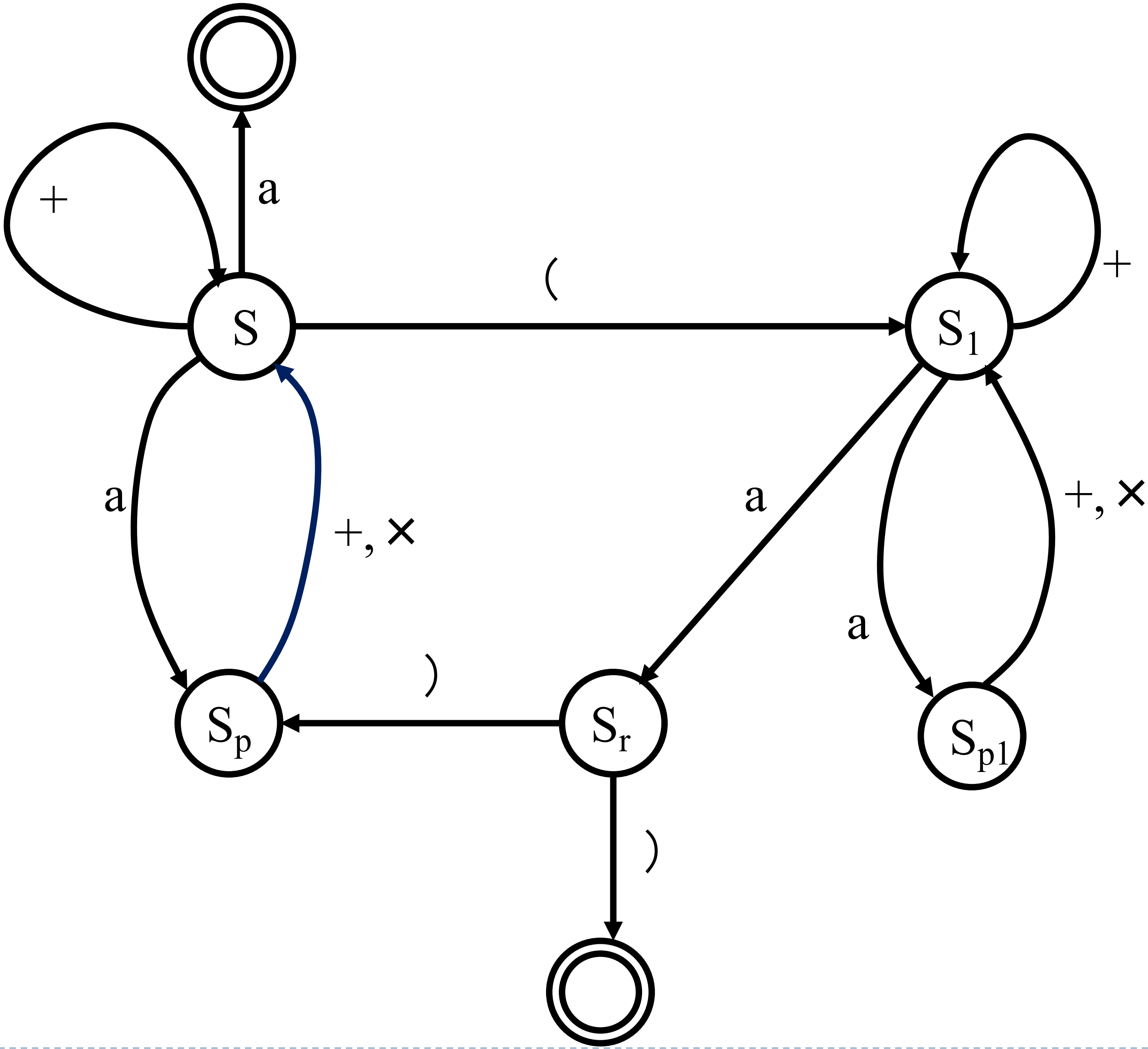
A→aB

A→a

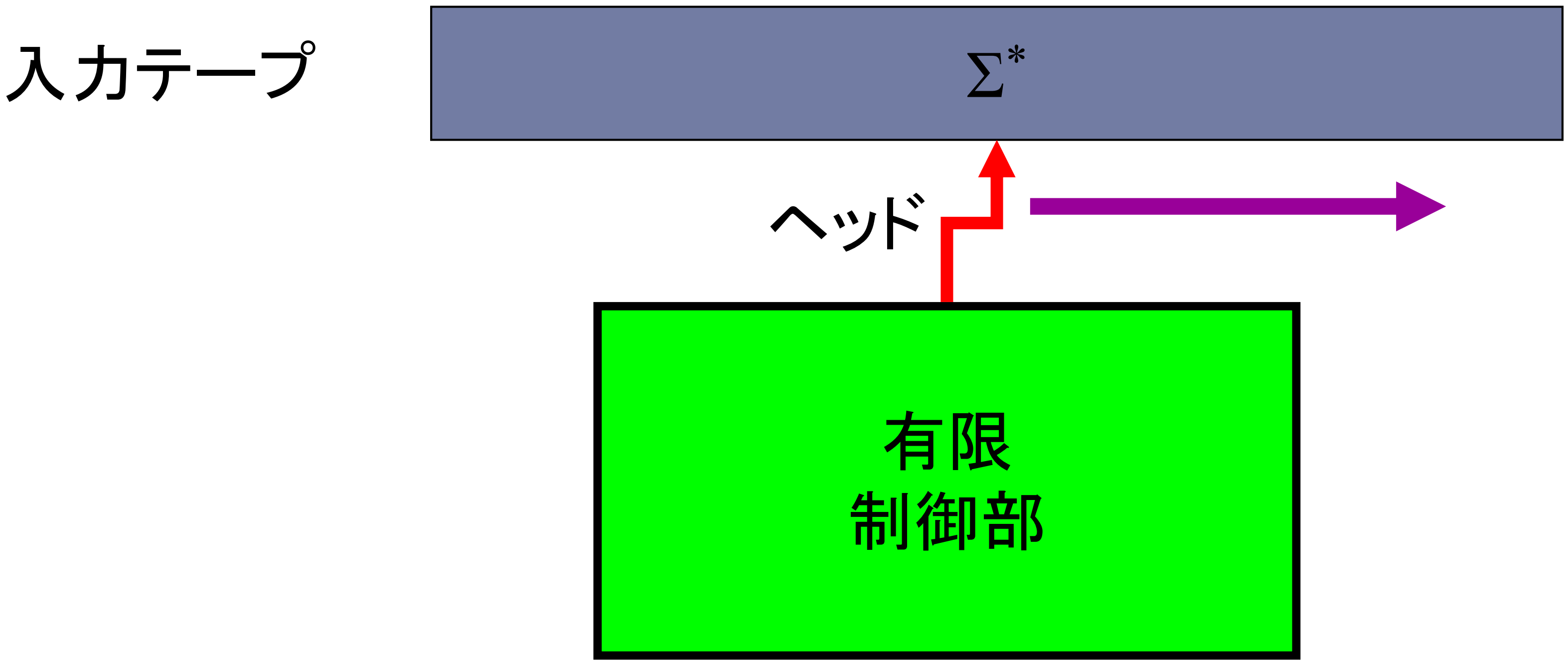


A→aA
A→bY
A→c




+
a
(
+
S
S1
a
+, ×
a
+, ×
a
)
Sp
Sr
Sp1
)

有限オートマトン



有限オートマトンの定義

定義 $A=(K,\Sigma,\delta,q_0,F)$

- K:状態の集合
- Σ:入力テープ上の記号の集合
- δ:状態遷移関数
- $q_0 \in K$: 初期状態
- $F \subset K$: 最終状態の集合

決定性FA(DFA)

状態遷移関数

$\delta(q, a) = p$

次の状態

現在の状態

(一意に決まる)

入力

$\delta(q, xa) = \delta(\delta(q, x), a)$

つまり、$\delta(q, x_1 x_2 \ldots x_n) =$

$\delta(\ldots\delta(\delta(q, x_1), x_2), \ldots, x_n)$

DFA Aで受理される言語

定義

dfa A= (K,Σ, δ,q_0,F)で受理される言語

$T(A) = \{w|w\in\Sigma^*,\delta(q_0,w)\in F\}$

例 L_m={連続する２つのａを含む語}

$A_{m1}=(\{q_0,q_1,q_e\},\{a,b\},\delta_{m1},q_0,\{q_e\})$

$\delta_{m1}(q_0,a)=q_1$, $\delta_{m1}(q_0,b)=q_0$,

$\delta_{m1}(q_1,a)=q_e$, $\delta_{m1}(q_1,b)=q_0$,

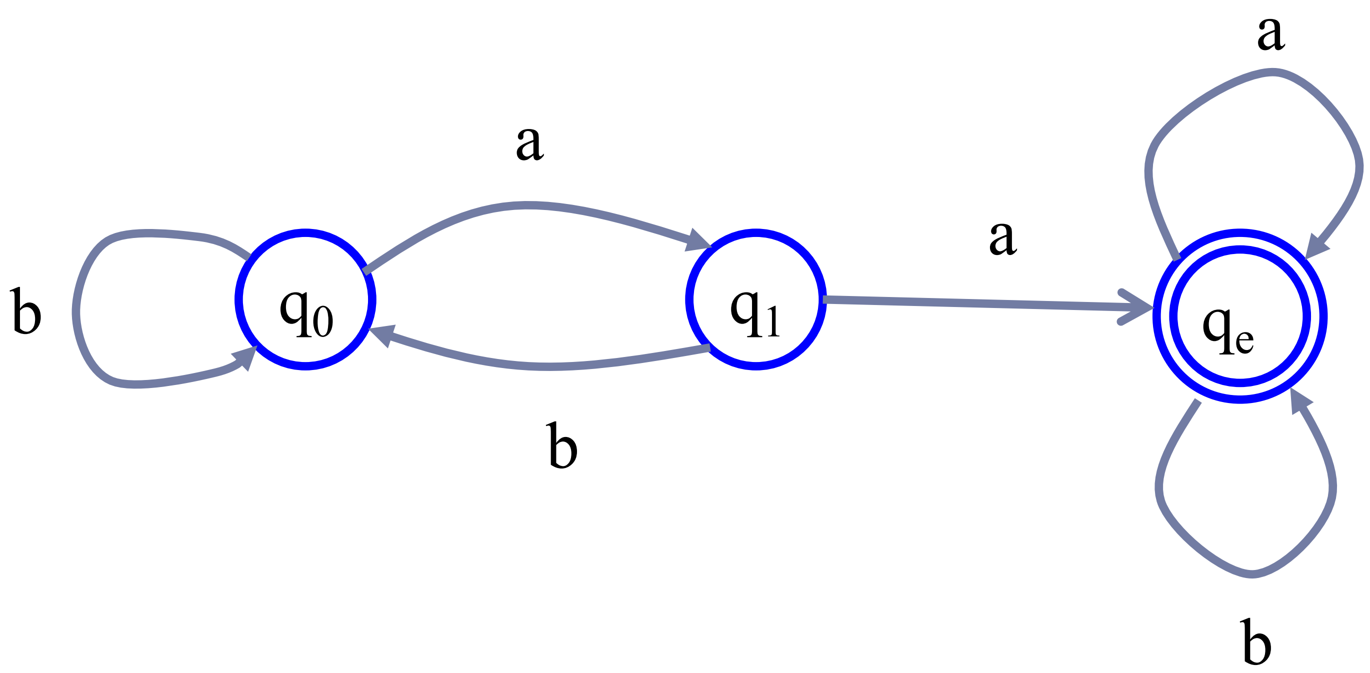
$\delta_{m1}(q_e,a)=q_e$, $\delta_{m1}(q_e,b)=q_e$,

関数表現

δ_{m1}	q_0	q_1	q_e
a	q_1	q_e	q_e
b	q_0	q_0	q_e

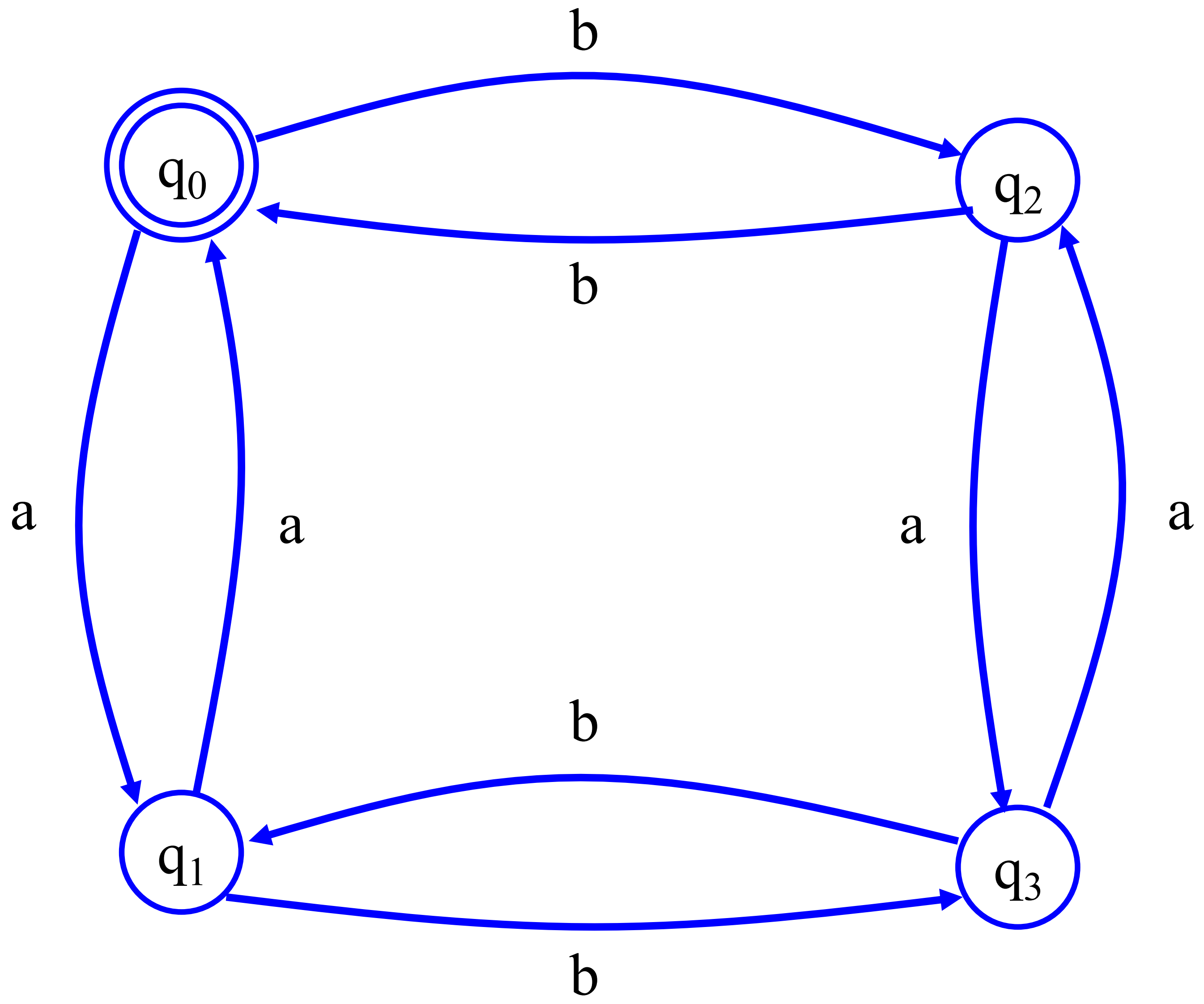
表表現

例 L_m={連続する2つのaを含む語}



グラフ図表現

例 Ln={偶数個のａと偶数個のbよりなる語}



非決定性FA（NFA)

状態遷移関数

$\delta(q, a) = \{p_1,\ldots,p_k\}$

次の状態

現在の状態

(複数あってよい）

入力

$\delta(q, \varepsilon) = \{q\}$

$\delta(q, xa) = \cup_{p\in\delta(q,x)}\delta(p,a)$

つまり

$$\delta(q, x_1x_2x_3\ldots x_n)$$

$$= \left\{\begin{array}{l} S_1=\delta(q, x_1) \\ S_2=\cup_{p\in S_1}\delta(p, x_2) \\ S_3=\cup_{p\in S_2}\delta(p, x_3) \\ \ldots \\ S_n= \cup_{p\in S_{n-1}}\delta(p, x_n) \\ \text{としたときの}S_n \end{array}\right.$$

NFA Aで受理される言語

定義

nfa A= (K,Σ, δ,q_0,F)で受理される言語

$T(A) = \{w|w\in\Sigma^*,p\in\delta(q_0,w),p\in F\}$

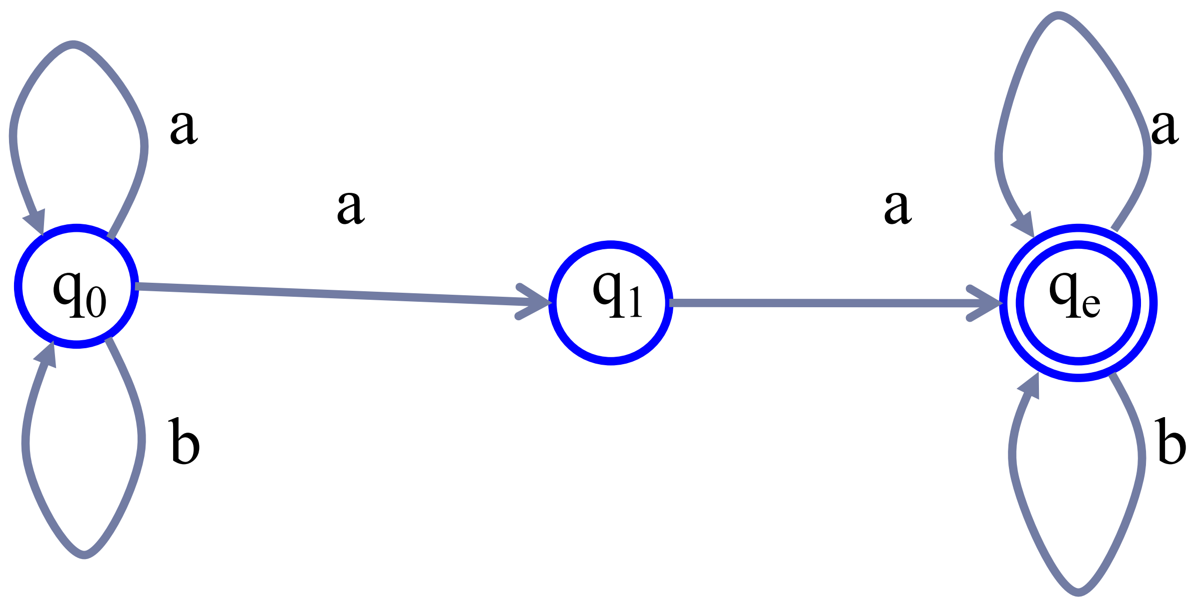
例

$A_{m2}=(\{q_0,q_1,q_e\},\{a,b\},\delta_{m2},q_0,\{q_e\})$

$\delta_{m2}(q_0,a)=\{q_0,q_1\}$, $\delta_{m2}(q_0,b)=\{q_0\}$,

$\delta_{m2}(q_1,a)=\{q_e\}$,

$\delta_{m2}(q_e,a)=\{q_e\}$, $\delta_{m2}(q_e,b)=\{q_e\}$,



定理

L:あるnfaによって受理

iff

Lを受理するdfaが存在

証明

nfa $A=(K,\Sigma,\delta,q_0,F)$, $L=T(A)$から$L=T(A')$となる

dfa $A'=(K',\Sigma,\delta',q_0',F')$を作る。

① K'＝2^K

Kの要素……… $q_1,q_2,\ldots$

K'の要素……… $\{q_1\}, \{q_2\}, \ldots, \{q_1,q_2\},\ldots$

② $q_0' = \{q_o\}$ ……q_0のみからなる集合

A'における状態は、Aの状態の集合

③ $F' = \{Q | \exists q \in Q \in K', q \in F\}$

Fの要素を少なくとも1つ含むK'の要素を集めたもの

④ $\delta'(\{q_1,\ldots,q_i\}, a) = \cup_{q \in \{q1,\ldots,qi\}} \delta(q,a)$

すべてのq_j （j＝１～i)について$\delta(q_j,a)$をあつめたもの。

最後に、w=$a_0a_1\ldots a_n\in T(A)$に対して、得られたdfa A'がw
を受理することを確認する

まず、

$$q_0 \xrightarrow{a_0} q_1 \xrightarrow{a_1} q_2 \rightarrow \cdots \xrightarrow{a_n} q_e$$
が存在する。

ただし、$\delta(q_0, a_0) \ni q_1$

$\delta(q_1, a_1) \ni q_2$

$\cdots$

$\delta(q_n, a_n) \ni q_e$

よって、$\delta'(\{\ldots q_0\ldots\}, a_0)=\{\ldots q_1\ldots\}$

$\delta'(\{\ldots q_1\ldots\}, a_1)=\{\ldots q_2\ldots\}$

$\cdots$

$\delta'(\{\ldots q_n\ldots\}, a_n)=\{\ldots q_e\ldots\}$

よって

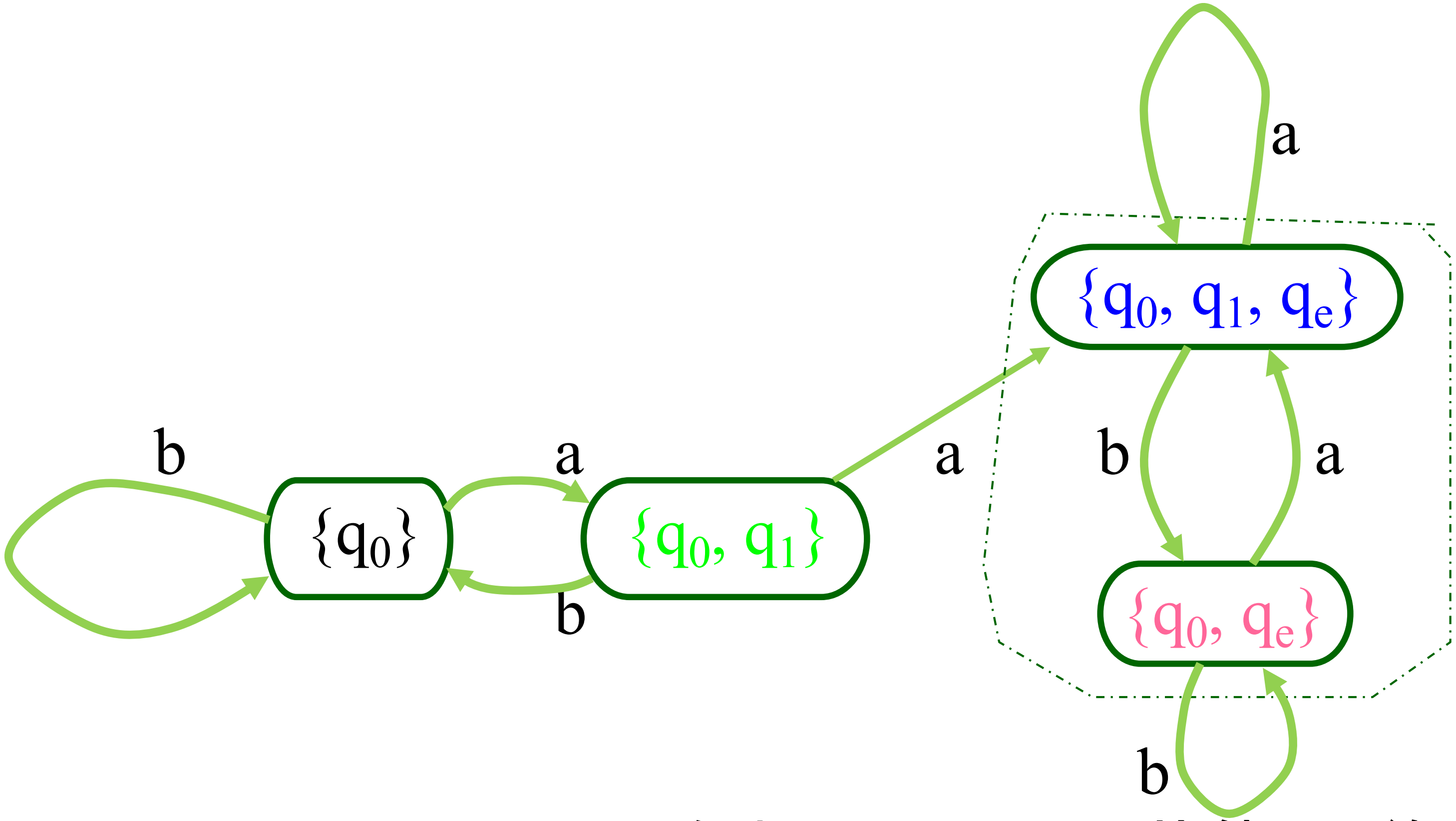
$$\{\ldots q_0\ldots\} \xrightarrow{a_0} \{\ldots q_1\ldots\} \xrightarrow{a_1} \{\ldots q_2\ldots\} \rightarrow \overset{\ldots}{\cdots} \xrightarrow{a_n} \{\ldots q_e\ldots\}$$

が存在する。

例

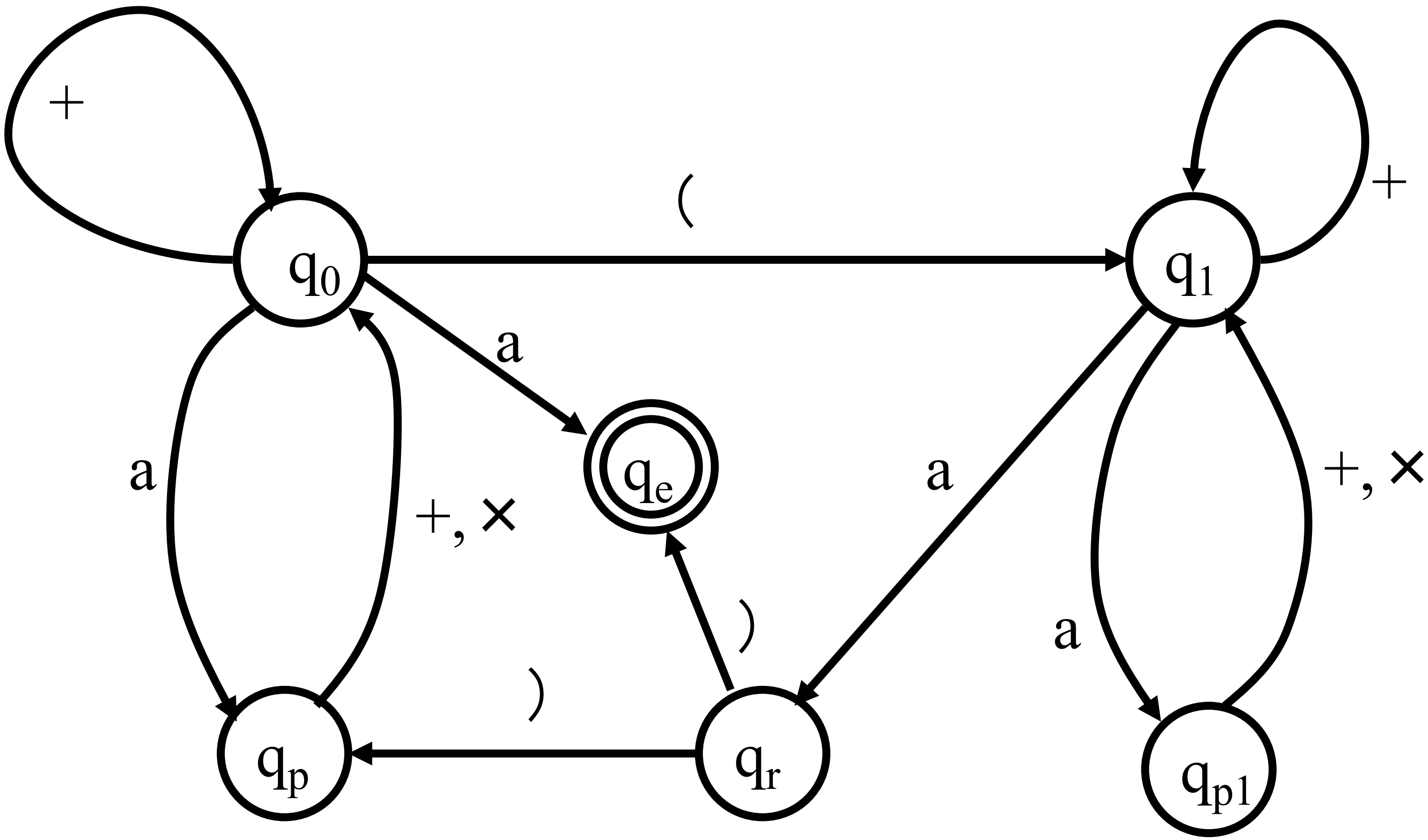
1. $\delta'(\{q_0\}, a) = \{q_0, q_1\}$
2. $\delta'(\{q_0\}, b) = \{q_0\}$
3. $\delta'(\{q_0, q_1\}, a) = \delta(q_0, a) \cup \delta(q_1, a)$
   $= \{q_0, q_1\} \cup \{q_e\}$
   $= \{q_0, q_1, q_e\}$
4. $\delta'(\{q_0, q_1\}, b) = \delta(q_0, b) \cup \delta(q_1, b)$
   $= \{q_0\} \cup \varphi$
   $= \{q_0\}$
5. $\delta'(\{q_0, q_1, q_e\}, a) = \delta(q_0, a) \cup \delta(q_1, a) \cup \delta(q_e, a)$
   $= \{q_0, q_1\} \cup \{q_e\} \cup \{q_e\}$
   $= \{q_0, q_1, q_e\}$
6. $\delta'(\{q_0, q_1, q_e\}, b) = \delta(q_0, b) \cup \delta(q_1, b) \cup \delta(q_e, b)$
   $= \{q_0\} \cup \varphi \cup \{q_e\}$
   $= \{q_0, q_e\}$

7. $\delta'(\{q_0, q_e\}, a) = \{q_0, q_1\} \cup \{q_e\} = \{q_0, q_1, q_e\}$
8. $\delta'(\{q_0, q_e\}, b) = \{q_0\} \cup \{q_e\} = \{q_0, q_e\}$



$\{q_0, q_1, q_e\}$と$\{q_0, q_e\}$をまとめて1つの状態（最終）
qe'と考えることができる。

例



前の例の決定性FAへの変換

	$\{q_0\}$	$\{q_1\}$	$\{q_p,q_e\}$	$\{q_r,q_{p1}\}$
a	$\{q_p,q_e\}$	$\{q_r,q_{p1}\}$	～	～
+	$\{q_0\}$	$\{q_1\}$	$\{q_0\}$	$\{q_1\}$
×	～	～	$\{q_0\}$	$\{q_1\}$
(	$\{q_1\}$	～	～	～
)	～	～	～	$\{q_p,q_e\}$

定理

タイプ３文法G=(V,Σ, P, S)

→ fa A=(K,Σ, δ, S,F), T(A)=L(G)

証明

① $K = V_N \cup \{q_e\}$

② $F = \{q_e\}$

③ $B \to a \in P$なら $q_e \in \delta(B,a)$

$B \to aC \in P$なら$C \in \delta(B,a)$

$w=a_1a_2\ldots a_k \in$ L(G)について

S$\Rightarrow a_1Q_1 \Rightarrow a_1a_2Q_2 \Rightarrow a_1a_2a_3Q_3 \Rightarrow \ldots \Rightarrow a_1\ldots a_k$

なる$Q_1,\ldots,Q_{k-1}$が存在する。

$Q_1\in\delta(S, a_1), Q_2\in\delta(Q_1, a_2), \ldots, q_e\in\delta(Q_{k-1}, a_k)$

のように状態遷移関数をつくれば、

$\delta(S,w)\ni q_e$

つまり、$w\in T(A)$

定理　（前の定理の逆）

fa A = (K,Σ,δ,q_0,F)

iff

タイプ3文法 G=(K,Σ,P,q_0), L(G)=T(A)

(証明は省略)

文法G $\Longrightarrow$ nfa A $\Longrightarrow$ dfa A'

まとめ

1. 正規文法の定義(再掲)
2. 有限オートマトン
    1. 決定性有限オートマトン(DFA)の定義
    2. 非決定性有限オートマトン(NFA)の定義
3. 変換:文法⇒NFA⇒DFA

課題

次の文法から、

- 非決定性faを書け（図でよい）。
- 上のnfaをdfaに変換せよ（関数をつくれ）。

G=({A,B,C,D,E,F,a,+,(,)},{a,+,(,)},P,A)

P={A→a, A→aB, A→(C, B→+A,

C→aD, C→aE, C→aF,

D→+C, E→)B, F→)}