Kyushu University 100th 2011
知の新世紀を拓く

形式言語理論

Formal Language Theory

6．正規言語の閉性

九州大学
KYUSHU UNIVERSITY

本日の内容

- 準備
  - 閉性
- 和集合，連接，Kleene閉包に関する閉性
- 補集合，積集合に関する閉性
- 準同型写像に関する閉性

準備

閉性(closure property)

- 正整数全体の集合 $\mathbb{N}$ 上で定義された加法演算 $+$ について，次の性質が成り立つ.
  - 任意の $a, b \in \mathbb{N}$ に対して $a+b$ は $\mathbb{N}$ の要素である.
- このことを
  集合 $\mathbb{N}$ は演算+に関して閉じている
  という.
- 集合 $\mathbb{N}$ は乗法演算・についても閉じている.
- 集合 $\mathbb{N}$ は減法演算 $-$ と除法演算 $/$ については閉じていない.

例

- $r = 0,1$ に対して $\mathbb{N}_r = \{ x \in \mathbb{N} \mid x \bmod 2 = r \}$とおく.
  - $\mathbb{N}_0$ は正の偶数の集合.
  - $\mathbb{N}_1$ は正の奇数の集合.
- $\mathbb{N}_0$は加法演算 + に関して閉じている.
  - 任意の $a, b \in \mathbb{N}_0$ に対して $a+b \in \mathbb{N}_0$
- $\mathbb{N}_1$は加法演算 + に関して閉じていない.
  - 任意の $a, b \in \mathbb{N}_1$ に対して $a+b \notin \mathbb{N}_1$ （$a+b \in \mathbb{N}_0$）

より一般的に：2項演算

- $\mathcal{F}$を集合とし, μ を$\mathcal{F}$上の2項演算とする.
  - すなわち, $\mu : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$

- $\mathcal{F}$の部分集合$\mathcal{A}$が演算 μ について閉じているとは,
  任意の$x, y \in \mathcal{A}$に対して, $\mu(x, y) \in \mathcal{A}$
  が成り立つときをいう.

（例）
言語の集合 $\mathcal{A}$ が, 和集合演算$\cup$に関して閉じているとは,
任意の$L_1, L_2 \in \mathcal{A}$に対して, $L_1 \cup L_2 \in \mathcal{A}$
が成り立つときをいう.

より一般的に: n項演算

- $\mathcal{F}$を集合とし, μ を$\mathcal{F}$上のn項演算とする.
  - すなわち, $\mu : \underbrace{\mathcal{F} \times \cdots \times \mathcal{F}}_{n} \to \mathcal{F}$
- $\mathcal{F}$の部分集合$\mathcal{A}$が演算 μ について閉じているとは, 任意の$x_1, \ldots, x_n \in \mathcal{A}$に対して, $\mu(x_1, \ldots, x_n) \in \mathcal{A}$ が成り立つときをいう.

和集合，連接，Kleene閉包に関する閉性

和集合，連接，Kleene閉包

【定理3.2】正規言語の族は和集合，連接，Kleene閉包に関して閉じている．

定理3.2の証明

- 言語L_1, L_2が正規ならば言語$L_1 \cup L_2$, L_1L_2, $L_1{}^*$は正規言語であることを示せばよい.
- L_1, L_2は正規であるから定理3.1より, $L_1 = \|r_1\|, L_2 = \|r_2\|$となる正規表現$r_1, r_2$が存在する.
- このとき， $r_1+r_2, r_1r_2, r_1{}^*$は正規表現であり, 対応する言語

$$\|r_1+r_2\| = L_1 \cup L_2,\ \|r_1r_2\| = L_1L_2,\ \|r_1{}^*\| = L_1{}^*$$

は正規言語である.

QED

補集合，積集合に関する閉性

補集合に関する閉性

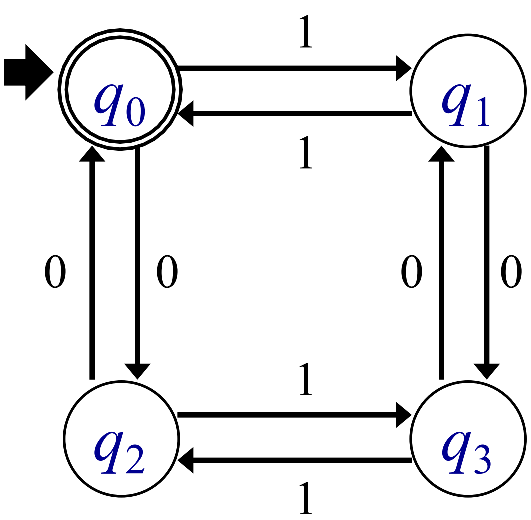
【定理3.3】正規言語族は補集合に関して閉じている.

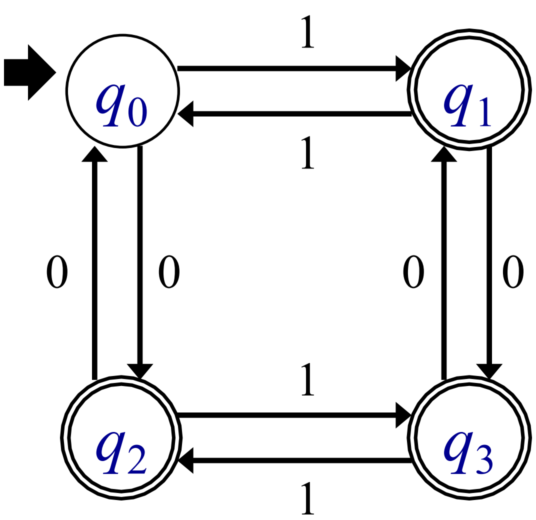
定理3.3の証明

- $L \subseteq \Sigma^*$ を任意の正規言語とするとき $\Sigma^* - L$ が正規であることを示せばよい.
- L を受理するDFAを $M = (Q, \Sigma, \delta, q_0, F)$ とする.
- このとき，DFA $M' = (Q, \Sigma, \delta, q_0, Q - F)$ は，言語 $\Sigma^* - L$ を受理する.
- よって，$\Sigma^* - L$ は正規である.

QED

例





積集合に関する閉性

【定理3.4】正規言語族は積集合に関して閉じている.

定理3.4の証明

- 言語$L_1, L_2 \subseteq \Sigma^*$が正規ならば $L_1 \cap L_2$が正規であることを示せばよい.
- $L_1 \cap L_2 = \Sigma^* - ((\Sigma^* - L_1) \cup (\Sigma^* - L_2))$と書ける.
- 定理3.2, 3.3より正規言語族は和集合と補集合について閉じている．よって，$L_1 \cap L_2$は正規である.

QED

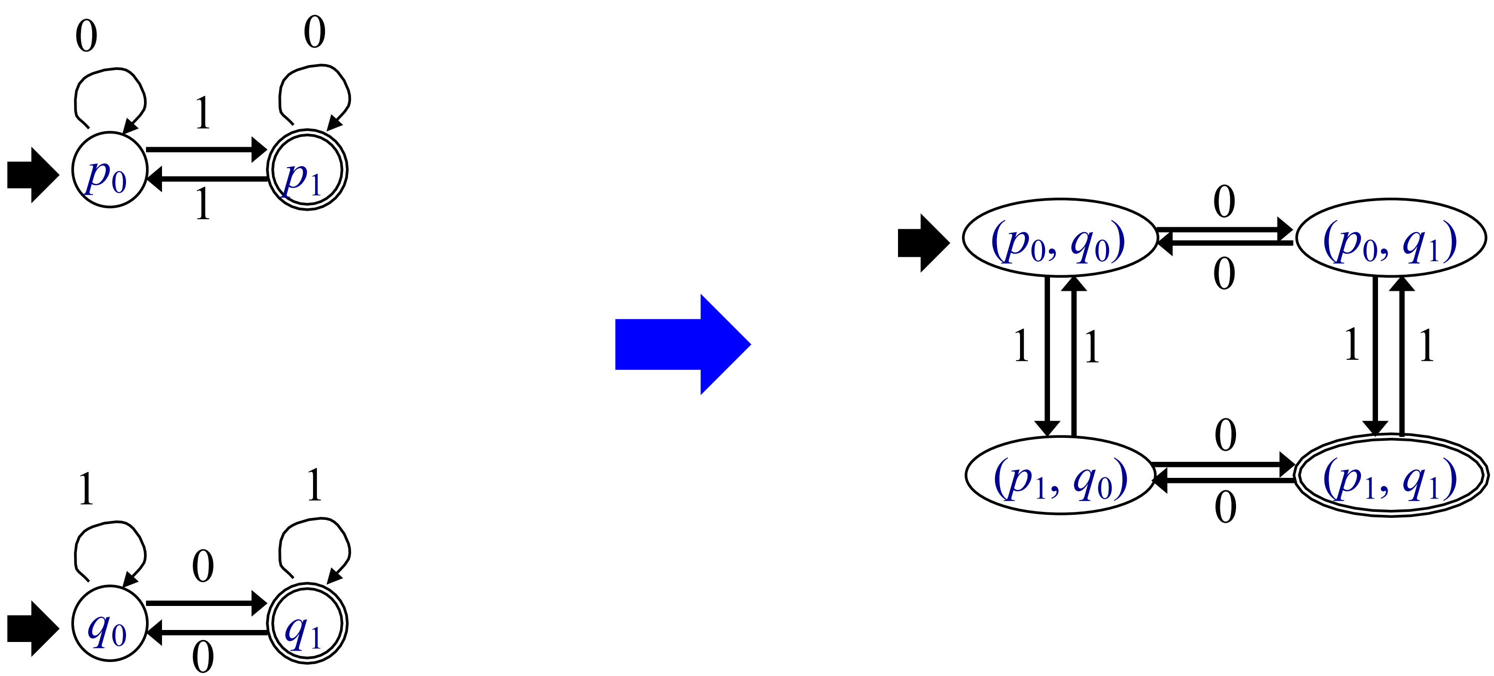
定理3.4の証明(別証明)

- 言語$L_1, L_2 \subseteq \Sigma^*$が正規ならば，$L_1 \cap L_2$が正規であることを示せばよい．
- L_iを受理するDFAを $M^{(i)} = (Q^{(i)}, \Sigma, \delta^{(i)}, q_0{}^{(i)}, F^{(i)})$ とし DFA $M' = (Q, \Sigma, \delta, q_0, F)$を次のように構成する．
  - $Q = Q^{(1)} \times Q^{(2)}$.
  - $q_0 = (q_0{}^{(1)}, q_0{}^{(2)})$.
  - 写像$\delta : Q \times \Sigma \to Q$は次で定まる．
    - $\delta((p,q), a) = (\delta^{(1)}(p,a), \delta^{(2)}(q,a)) \quad ((p,q) \in Q^{(1)} \times Q^{(2)}, a \in \Sigma)$
  - $F = F^{(1)} \times F^{(2)}$とする．

直積オートマトン

- このとき，M'は言語$L_1 \cap L_2$を受理する．
- よって，$L_1 \cap L_2$は正規である．

例



準同型写像に関する閉性

準同型写像

【定義】 Σ, Δ をアルファベットとする．写像 $h: \Sigma^* \to \Delta^*$ が準同型写像であるとは，次を満たすときをいう．

任意の $x, y \in \Sigma^*$ に対して $h(xy) = h(x)h(y)$

- 上式で $x{=}y{=}\varepsilon$ とおくと $h(\varepsilon){=}h(\varepsilon)h(\varepsilon)$ となり，$h(\varepsilon) = \varepsilon$ を得る．
- 準同型写像 hは，各$a \in \Sigma$ に対する値 $h(a){\in}\Delta^*$ が定まれば一意に定まる．

- 以下のように，言語 $L{\subseteq}\Sigma^*$ を言語 $h(L){\subseteq}\Delta^*$に割り当てる写像に拡張する．
  - $h(L) = \bigcup_{x \in L} h(x) \quad (\forall L \subseteq \Sigma^*)$

例

- $\Sigma=\{0,1\}, \Delta=\{a, b\}$, $h(0)=aa, h(1)=b$とするとき:
  - $h(010)=aabaa$.
  - $L_1=\{01, 11\}$のとき, $h(L_1)= \{aab, bb\}$.
- $\Sigma=\{0,1\}, \Delta=\{a, b\}$, $h(0)=aa, h(1)=aba$とするとき:
  - $h(010)=aaabaaa$.
  - $L_1=\{01\}^*$のとき, $h(L_1)= \{aaaba\}^*$.

準同型写像に関する閉性

【定理3.5】正規言語族は準同型写像に関して閉じている.

すなわち, Σ上の任意の正規言語 L は, Σ^* から Δ^* への準同型写像 h によって Δ 上の正規言語 $h(L)$ に写される.

定理3.5の利用例（その1）

- 言語$L_1=\{\mathrm{a}^n\mathrm{b}^n | n\geq 1\}$が正規でないことの証明.
  - ただし，言語 $L=\{0^n1^n | n\geq 1\}$が正規でないことは既知とする.
- 準同型写像 h: $\{0, 1\}^* \rightarrow \{\mathrm{a}, \mathrm{b}\}^*$を

  $h(0)=\mathrm{a}$, $h(1)=\mathrm{b}$

  で定めると，$h(L) = L_1$となる.
- L は正規ではないことから，定理3.5より，L_1は正規ではない.

定理3.5の利用例(その2)

- 言語$L_2=\{a^{2n}b^{2n} | n \geq 1\}$が正規でないことの証明.
  - ただし，言語 $L=\{0^n 1^n | n \geq 1\}$が正規でないことは既知とする.
- 準同型写像 h: $\{0, 1\}^* \rightarrow \{a, b\}^*$を

  $h(0)=aa$, $h(1)=bb$

  で定めると，$h(L) = L_2$となる.
- L は正規ではないことから，定理3.5より，L_2は正規ではない.

定理3.5の利用例(その3)

- 言語$L_3=\{a^{2n}b^{3n} | n \geq 1\}$が正規でないことの証明.
  - ただし，言語 $L=\{0^n1^n | n \geq 1\}$が正規でないことは既知とする.

- 準同型写像 h: $\{0, 1\}^* \rightarrow \{a, b\}^*$を
  $$h(0)=aa,\ h(1)=bbb$$
  で定めると，$h(L) = L_3$となる.
- L は正規ではないことから，定理3.5より，
  L_3は正規ではない.

定理3.5の証明

- 次の補題から明らか.

【補題3.1】 h を Σ^* から Δ^* への準同型写像とする.
Σ 上の任意の正規表現 r に対し, r における文字 $a \in \Sigma$ の出現を $h(a) \in \Delta^*$ で置換して得られる記号列 $h(r)$ は Δ 上の正規表現であり,
$\|h(r)\| = h(\|r\|) = \bigcup_{x \in \|r\|} h(x)$ が成り立つ.

例

$\Sigma=\{0, 1\}, \Delta=\{$あ, い$\}$ $0 \rightarrow$ ああ
$1 \rightarrow$ あいあ

$0+10$ → ああ+あいああ

$01(0+1)^*$ → ああああいあ(ああ+あいあ)*

補題3.1の証明（1／3）

- 正規表現 r に含まれる演算子の個数 m に関する帰納法で証明する.
- $m=0$のとき.
  - $r=a\in\Sigma$ ならば $h(a)=w\in\Delta^*$ はΔ上の正規表現であり,
    $\|h(r)\| = \|w\|=\{w\},\quad h(\|r\|)= h(\{a\})=\{w\}$.
  - $r=\varepsilon$ならば $h(r)=\varepsilon$ はΔ上の正規表現であり,
    $\|h(r)\| = \|\varepsilon\|=\{\varepsilon\}, h(\|r\|) = h(\{\varepsilon\}) =\{\varepsilon\}$.
  - $r=\varnothing$ならば $h(r)=\varnothing$ はΔ上の正規表現であり,
    $\|h(r)\| = \|\varnothing\|=\varnothing,\ h(\|r\|) = h(\varnothing)=\varnothing$.

補題3.1の証明（２／３）

- $m>0$のとき.
  - $r = r_1 \cdot r_2$ (r_1, r_2は正規表現)ならば，帰納法の仮定より
    $h(r_1), h(r_2)$ は Δ上の正規表現であり，
    $\|h(r_1)\|= h(\|r_1\|), \|h(r_2)\|= h(\|r_2\|)$.
    よって，$h(r)= h(r_1) \cdot h(r_2)$も$\Delta$ 上の正規表現であり，
    $\|h(r)\| = \|h(r_1) \cdot h(r_2)\|= \|h(r_1)\| \cdot \|h(r_2)\| = h(\|r_1\|) \cdot h(\|r_2\|)$,
    $h(\|r\|)= h(\|r_1 \cdot r_2\|)= h(\|r_1\| \cdot \|r_2\|)= h(\|r_1\|) \cdot h(\|r_2\|)$.
  - $r = r_1 + r_2$ (r_1, r_2は正規表現)ならば，帰納法の仮定より
    $h(r_1), h(r_2)$ は Δ上の正規表現であり，
    $\|h(r_1)\|= h(\|r_1\|), \|h(r_2)\|= h(\|r_2\|)$.
    よって，$h(r)= h(r_1)+h(r_2)$もΔ 上の正規表現であり，
    $\|h(r)\| = \|h(r_1)+h(r_2)\|=\|h(r_1)\| \cup \|h(r_2)\| = h(\|r_1\|) \cup h(\|r_2\|)$,
    $h(\|r\|)= h(\|r_1+r_2\|)= h(\|r_1\| \cup \|r_2\|)= h(\|r_1\|) \cup h(\|r_2\|)$.

補題3.1の証明（3／3）

- $m>0$のとき.
  - $r=r_1^*$ $(r_1$ は正規表現)ならば，帰納法の仮定より $h(r_1)$ は Δ 上の正規表現であり，$\|h(r_1)\|=h(\|r_1\|)$.
    よって，$h(r)=h(r_1)^*$ も Δ 上の正規表現であり，
    $\|h(r)\|=\|h(r_1)^*\|=\|h(r_1)\|^*=h(\|r_1\|)^*$,
    $h(\|r\|)=h(\|r_1{}^*\|)=h(\|r_1\|^*)=h(\|r_1\|)^*$.
- よって，帰納法により，補題が成立する.

QED

逆準同型写像に関する閉性

逆準同型写像

- 写像 $h: \Sigma^* \to \Delta^*$ に対しその逆写像 h^{-1}を以下のように定める.
  - $h^{-1}(L) = \{x \in \Sigma^* \mid h(x) \in L\} \quad (\forall L \subseteq \Delta^*)$.
- 準同型写像 h に対しその逆写像 h^{-1} を逆準同型写像と呼ぶ.

例

- $\Sigma=\{0,1\}$, $\Delta=\{a, b\}$, $h(0)=ab$, $h(1)=a$とするとき:
  - $L_2=\{a, aa, ab\}$のとき, $h^{-1}(L_2)=\{0, 1, 11\}$.
  - $L_2=\{a, ab\}^*$のとき, $h^{-1}(L_2)=\{0, 1\}^*$.
  - $L_2=\{a, b\}^*$のとき, $h^{-1}(L_2)=\{0, 1\}^*$.
  - $L_2=\{a\}\{ab\}^*$のとき, $h^{-1}(L_2)=\{1\}\{0\}^*$.
  - $L_2=\{ab\}^*\{aa\}$のとき, $h^{-1}(L_2)=\{0\}^*\{11\}$.

例

- $\Sigma=\{0,1\}$, $\Delta=\{a, b\}$, $h(0)=aa$, $h(1)=aba$とするとき:
  - $L_2=(\{ab\}\cup\{ba\})^*\{a\}$のとき，$h^{-1}(L_2)=\{1\}$.
    - $h(\varepsilon)=\varepsilon\notin L_2$.
    - $h(1)\in L_2$.
    - $h(0)=aa$より，すべての$x\in 0\{0,1\}^*$に対して$h(x)\notin L_2$.
    - $h(10)=abaaa$より，すべての$x\in 10\{0,1\}^*$に対して$h(x)\notin L_2$.
    - $h(11)=abaaba$より，すべての$x\in 11\{0,1\}^*$に対して$h(x)\notin L_2$.
    - よって，「$h(x)\in L_2 \Rightarrow x=1$」が成り立つ.

準同型写像の逆写像

【定理3.6】正規言語族は準同型写像の逆写像に関して閉じている.

すなわち，Δ上の任意の正規言語 L は，Σ^* から Δ^* への準同型写像 h の逆写像 h^{-1} によって Σ 上の正規言語 $h^{-1}(L)$ に写される.

定理3.6の利用例

言語$L_1=\{0^n1^n|n\geq 1\}$が正規でないことは既知とし，言語$L_2=\{a^nba^n|n\geq 1\}$が正規でないことを示す．

- $L_2=\{a^nba^n|n\geq 1\}$が正規であると仮定する．
- 準同型写像 $h_1:\{a,b,c\}^* \to \{a,b\}^*$, $h_2:\{a,b,c\}^* \to \{0,1\}^*$を次で定める．
  - $h_1(a)=a$, $h_1(b)=ba$, $h_1(c)=a$
  - $h_2(a)=0$, $h_2(b)=1$, $h_2(c)=1$

要証明

- このとき，$h_2(h_1^{-1}(L_2)\cap\|a^*b^*c^*\|)=L_1$が成り立つ．
- 定理3.4, 3.5, 3.6よりL_1も正規でなければならないがこれは矛盾．よってL_2は正規でない．

QED

定理3.6の証明

- LをΔ上の任意の正規言語とし，$M=(Q, \Delta, \delta, q_0, F)$を$L$を受理するDFAとする．また，$h:\Sigma^* \to \Delta^*$を任意の準同型写像とする．
- Mが文字列$h(a) \in \Delta^*$を読んで行う動作を，文字$a \in \Sigma$に対して行うDFA $M' = (Q, \Delta, \delta', q_0, F)$を次で定める．
  - $\delta'(q,a)=\delta(q, h(a)) \quad (q \in Q, a \in \Sigma)$
- このとき，$\delta'(q_0, x) = \delta(q_0, h(x))\ (x \in \Sigma^*)$が成立し，$L(M')=h^{-1}(L)$を得る．よって，$h^{-1}(L)$は正規言語である．

QED