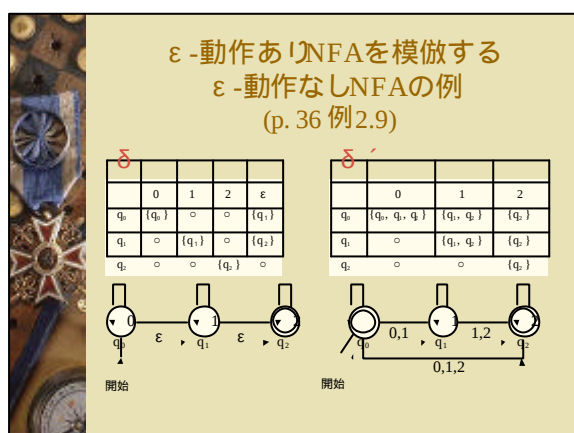
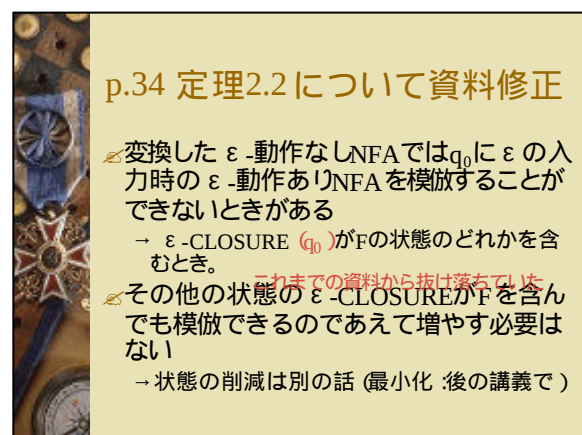
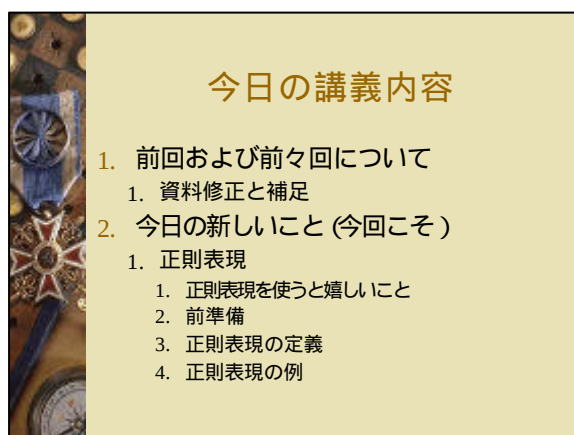


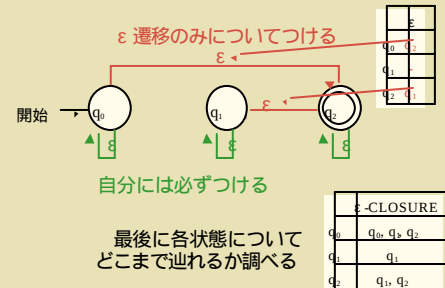


ε-CLOSUREの定義

ある状態 q からε-動作のみで移れる先の状態の集合

- ε-CLOSURE(q)
= 遷移図からラベルがεでない有向辺を取り去った上で、 q から到達可能な頂点の集合
- ε-CLOSURE(P): P は状態の集合
= $\bigcup_{q \in P} \text{ε-CLOSURE}(q)$

例1 ε-CLOSURE



ε-動作を含むNFA → NFA 例2

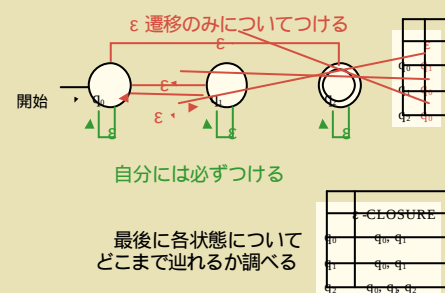
ε-動作を含むNFA

($\{q_0, q_1, q_2\}, \{0, 1\}, \delta, q_0, \{q_2\}$)

δは下表

	0	1	ε
q_0	q_0	-	q_1
q_1	-	q_1, q_2	q_2
q_2	q_1	q_2	q_2

例2 ε-CLOSURE



今日の新しいこと

正則表現

1. 正則表現を使うと嬉しいこと
2. 前準備
3. 正則表現の定義
4. 正則表現の例

正則表現を使えると嬉しいこと

パターンマッチにしょっちゅう使う

- UNIXのシェル
 - sh, csh, ksh, tcsh, bash
 - ファイルの名前
- UNIXのコマンド、プログラミング言語
 - egrep, awk, sed, perl
 - ファイルの中の文字列の処理

Bash のファイル名展開

```
% ls
aaa ,aab ,aac ,aba ,aca ,ada ,abc ,bbb ,bcb

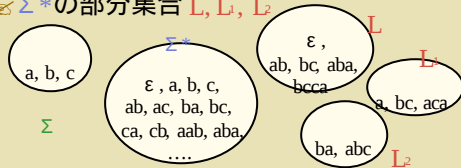
% echo a[abc]a
aaa aba aca
```

Perl での処理

```
Sun May 27 21:51:40 JST 2001
↓
s/¥w+ (¥w+) (¥d¥d) ¥d¥d:¥d¥d:¥d¥d ¥w+
(¥d+)/Today is ¥1 ¥2, ¥3./;
↓
Today is May 27, 2001.
```

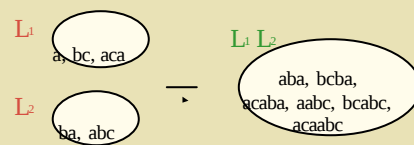
前準備 その1 (記号列の集合)

- アルファベット Σ
- Σ 上の記号列 Σ^*
- Σ^* の部分集合 L, L_1, L_2

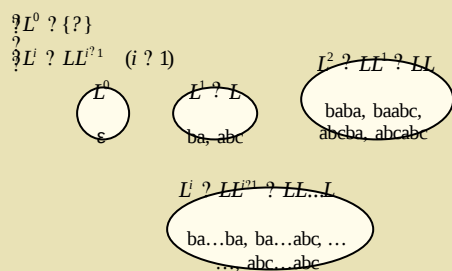


前準備 その2 (記号列の集合の演算)

$L_1 L_2 = \{xy \mid x \in L_1, y \in L_2\}$ – 接続

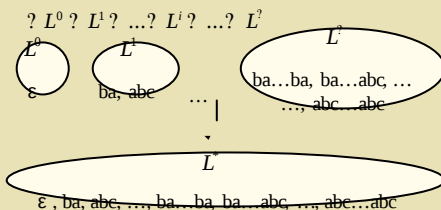


前準備 その3 (記号列の集合の演算)



前準備 その4 (Kleene閉包)

$L^* = \bigcup_{i \geq 0} L^i$ L の Kleene 閉包 (または単に閉包)



前準備 その5

(正閉包)

$L^? \supset \square L^i$ L の正閉包

L^1

ba, abc

...

$L^?$

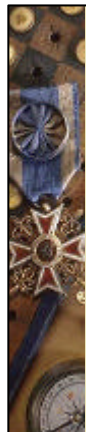
ba...ba, ba...abc, ...
..., abc...abc

↓

$L^?$ が?を含むのは L が?を含むときのみ

$L^?$

ba, abc, ..., ba...ba, ba...abc, ..., abc...abc

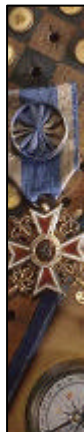


教科書の例

(p. 37 例2.10)

$L_1 ? \{10, 1\}$
 $L_2 ? \{011, 11\}$
 $L_1 L_2 ? \{10011, 1011, 111\}$

$\{10, 11\}^* ? \{?, 10, 11, 1010, 1011, 1110, 1111, \dots\}$
 $\{10, 11\}^? ? \{10, 11, 1010, 1011, 1110, 1111, \dots\}$



正則表現の定義

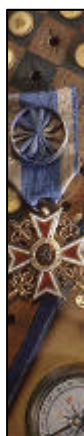
(p. 37)

1. $\emptyset$ は正則表現で、その表す集合は空集合である。
2. ε は正則表現で、その表す集合は $\{\varepsilon\}$ である。
3. Σ の各元 a に対して a は正則表現で、その表す集合は $\{a\}$ である。
4. r と s がそれぞれ言語 R と S を表す正則表現のとき、 $(r+s)$ 、 (rs) 、および (r^*) は正則表現で、その表す集合はそれぞれ、 $R \cup S$ 、 RS 、 R^* である。

正則表現の例

$\textcircled{\text{ } \cup \text{ } } \varepsilon, a$
 $\textcircled{\text{ } \cup \text{ } } \{ \varepsilon \} = \{ \varepsilon \}$
 $\textcircled{\text{ } \cup \text{ } } \{ a \} = \{ a \}$
 $\textcircled{\text{ } \cup \text{ } } \{ \varepsilon \} \quad \{ a \} = \{ \varepsilon, a \}$
 $\textcircled{a+b} = \{ a \} \quad \{ b \} = \{ a, b \}$
 $\textcircled{\varepsilon} = \{ \} \quad \{ \varepsilon \} = \{ \}$
 $\textcircled{a} = \{ \} \quad \{ a \} = \{ \}$
 $\varepsilon a = \{ \varepsilon \} \quad \{ a \} = \{ a \}$
 $ab = \{ a \} \{ b \} = \{ ab \}$

$\textcircled{\text{ }^*} = \{ \}^* = \{ \varepsilon \}$
 $\varepsilon^* = \{ \varepsilon \}^* = \{ \varepsilon \}$
 $a^* = \{ a \}^* = \{ \varepsilon, a, aa, aaa, \dots \}$
 $((\textcircled{\text{ } \cup \text{ } } + \textcircled{\text{ } \cup \text{ } }) + \textcircled{\text{ } \cup \text{ } }) = \{ \varepsilon \}$
 $\{ a \} = \{ \varepsilon, a \}$
 $((\textcircled{\text{ } \cup \text{ } } + \textcircled{\text{ } \cup \text{ } }) = \{ \{ a, b \} = \{ \}$
 $((ab)^+ \varepsilon) = \{ ab \} \quad \{ \varepsilon \} = \{ \varepsilon, ab \}$
 $((ab)^*) = \{ ab \}^* = \{ \varepsilon, ab, abab, ababab, \dots \}$



正規表現の演算の強さ

* > 連接 > +

- $((0(1^*)) + 0) \rightarrow 01^* + 0$
- $(1 + (10))^* \rightarrow (1 + 10)^*$
- $((1(1(1^*))) + (01)) \rightarrow (111^* + 01)$

教科書の例 その1

(p. 38, 例2.11)

$\varnothing 00 = \{00\}$

$\varnothing (0+1)^* = \{0, 1\}^*$
 $= \{\varepsilon, 0, 1, 00, 01, 10, 11, \dots\}$

$\varnothing (1+10)^* = \{1, 10\}^*$
 $= \{\varepsilon, 1, 10, 11, 110, 101, 1010, \dots\}$

$\varnothing (0+\varepsilon)(1+10)^* = \{0, \varepsilon\} \{1, 10\}^*$
 $= \{0, \varepsilon\} \{\varepsilon, 1, 10, 11, 110, 101, 1010, \dots\}$
 $= \{\varepsilon, 0, 1, 01, 10, 010, 11, 011, 110, 0110, 101, 0101, 1010, 01010, \dots\}$

教科書の例 その2 (p. 38, 例2.11)

- $(0+1)^*011 = \{0, 1\}^*011$
 $= \{ \varepsilon, 0, 1, 00, 01, 10, 11, \dots \} \{011\}$
 $= \{011, 0011, 1011, 00011, 01011, 10011, 11011, \dots\}$
 $0^*1^*2^* = \{0\}^* \{1\}^* \{2\}^*$ ← 図2.8のNFA
 $= \{ \varepsilon, 0, 00, \dots \} \{ \varepsilon, 1, 11, \dots \} \{ \varepsilon, 2, 22, \dots \}$
 $= \{ \varepsilon, 0, 1, 01, 012, 00, 001, 0011, 0012, 00112, 001122, 000, 0001, 00011, 00012, 000111, 000112, 0001112, 00011122, 000111222, \dots \}$

$(1+10)^*$ の性質

- $\hookrightarrow$ 1で始まり、連続した0を含まない列か空列から成る集合
 $\rightarrow$ 帰納法で示す。

$$(1?10)^*? \bigcup_{i \geq 0} \{1,10\}^i$$

$$i?n \text{ のとき } \bigcup_{i \geq 0} \{1,10\}^i \text{ が}$$

1 で始まり連続した0を含まない列か空列のみから成るためには
 $\{1,10\}^0, \dots, \{1,10\}^n$ がそれぞれ
 連続した0を含まない列か空列のみからなっていればよい。

帰納法での証明つづき

- $i?0$ のとき、 $\{1,10\}^0 = \{\varepsilon\}$
- $i?1$ のとき、 $\{1,10\}^1 = \{1,10\}$
 1 で始まり連続した0を含まない列のみを含む。
- $i?n$ のとき、
 $\{1,10\}^n$ が $\{1x_1, \dots, 1x_n\}$ のように
 1 で始まり連続した0を含まない列のみを含む
 とすれば $(x_1, \dots, x_n)$ は連続した0を含まない列、
 $i?n+1$ のとき、
 $\{1,10\}^{n+1} = \{1,10\}^n? \{1,10\} = \{11x_1, 101x_1, \dots, 11x_n, 101x_n\}$
 となり、 1 で始まり連続した0を含まない列のみを含むようになる。
 よって1,2,3より、
 $(1?10)^*$ は 1 で始まり連続した0を含まない列か空列のみを含む。

演習問題 2.16に注意

$$\begin{array}{ll}
 r?s?s?r & \text{✓} \\
 (r?s)?t?r?(s?t) & (r^*)^*?r^* \\
 (rs)t?r(st) & (?r)^*?r^* \\
 r(s?t)?rs?rt & (r^*s^*)^*?(r?s)^* \\
 (r?s)t?rt?st &
 \end{array}$$

証明できるかな?

おまけ

- $\hookrightarrow$ 正閉包に対応する正則表現

$$\begin{aligned}
 rr^*?r^? \bigcup_{i \geq 1} r^i \\
 00^*1^*22^*?0^?1^?2^? \\
 \quad ? \{0, 00, \dots\} \{1, 11, \dots\} \{2, 22, \dots\} \\
 \quad ? \{012, 0122, 0112, 01122, 0012, \\
 \quad \quad 00122, 00112, 001122, \dots\} \\
 \quad ? 0^*1^*2^*
 \end{aligned}$$

今日のミニテスト

- $\hookrightarrow$ ミニテスト
 - 教科書 資料を見ても良い
 $\hookrightarrow$ 資料、ミニテストがない人は前へ
 $\hookrightarrow$ 提出したら帰ってよし
 $\hookrightarrow$ 次回 (こそ)
 - 有限オートマトンと正則表現の等価性