

オートマトンと言語 7回目 5月23日(水)

正規表現

授業資料

<http://www.ccn.yamanashi.ac.jp/~ysuzuki/automaton/>

授業の予定(中間試験まで)

回数	月日	内容
1	4月11日	オートマトンとは, オリエンテーション
2	4月18日	2章(数式の記法, スタック, BNF)
3	4月25日	2章(BNF), 3章(グラフ)
4	5月02日	3章(グラフ)
5	5月09日	4章 有限オートマトン1
6	5月16日	有限オートマトン2 2・3章の小テスト
7	5月23日	正規表現
8	5月30日	正規表現, 非決定性有限オートマトン
9	6月06日	中間試験, 前半のまとめ

出張などにより, 授業日が変更になる場合があります。

授業の予定

回数	月日	内容
10	6月13日	NFA→DFA
11	6月20日	DFAの最小化
12	6月27日	DFAの最小化, 有限オートマトンの応用
13	7月04日	プッシュダウンオートマトン, チューリング機械
14	7月11日	形式言語理論, 文脈自由文法
15	7月18日	期末試験, まとめ

出張などにより, 授業日が変更になる場合があります。

前回のまとめ

- 状態遷移図 → 状態遷移関数, 出力関数
- 状態遷移図, 入力系列 → 状態系列, 出力系列
- 動作の説明 → 状態遷移図
- 有限オートマトンのモデル
- 入力語の受理, 不受理
- 有限オートマトンの定義
- 言語の受理
- 正規表現

今日のメニュー

- 小テストの解説
- 正規表現
 - 正規表現の性質
 - 状態 q_i で受理する言語
 - 有限オートマトン(状態遷移図) → 正規表現
 - 漸化式 → 連立方程式 → 正規表現
- 非決定性有限オートマトン

小テスト

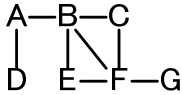
5月16日 10:15~10:30

- 試験範囲 2章, 3章
- 解答し終わったら答案用紙を提出
- 解答例は今日の授業資料に掲載

小テストの解答例 問題1

■ 下の図のグラフについて答えよ.

- a. AとFの距離を求めよ.
- 2
- b. グラフの直径(の長さ)を求めよ.
- 4
- c. ブリッジはどれか, すべて答えよ.
- D-A, A-B, F-G



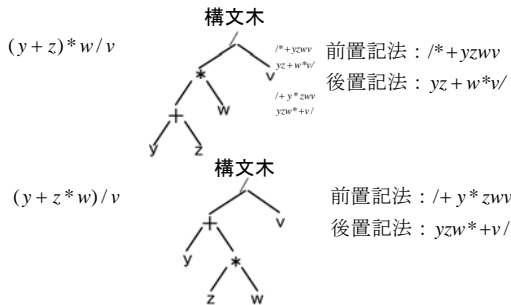
小テスト解答例 問題2

■ 下の用語を説明せよ.

- a: 順路
 - 節点が重複しない径路(径路:ある二つの節点を結ぶ節点と辺の列)
- b: 切断点(カットポイント)
 - その節点と節点につながっている辺を取り除くと連結グラフでなくなる節点
- c: 完全グラフ
 - 任意の節点がグラフ内の他の全ての節点と辺で直接つながっているグラフ
- d: 補グラフ
 - 完全グラフからある部分グラフの辺を取り除いたグラフ

小テスト 解答例 問題3

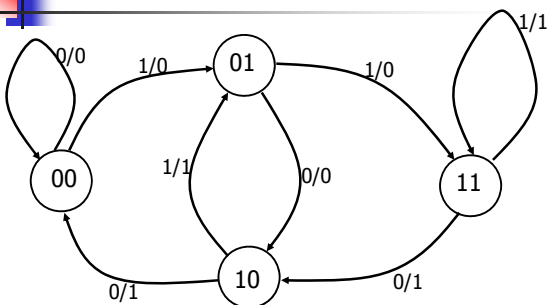
■ 次の数式(中置記法)の構文木を描け. また前置記法と後置記法に変換せよ. 但し「*」は乗算演算子とする.



例題4.4 a (p.87)

- 入力系列を2時点だけ遅らせて出力する
- 例) 011010...が入力されると,
.. 011010...を出力する
- ヒント: 直前の2入力を状態のラベルとすると考えやすい
- 00, 01, 10, 11

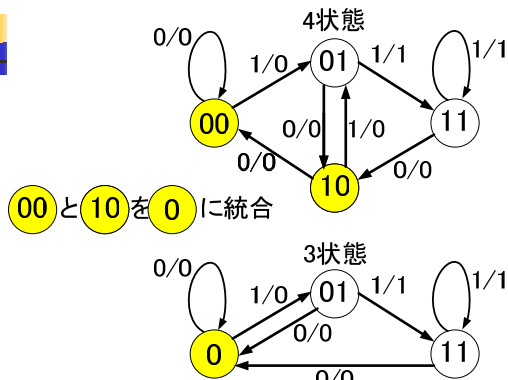
例題4.4 a 答え



演習問題2 例題4.4 b

- 入力系列に1が2個続いたときに1を出力し, それ以外のとき0を出力する状態遷移図
- 例) 入力系列が0011111011100...に対して,
.. 001111001100...を出力する
- ヒント: 直前の2つの入力を状態のラベルとすると考えやすい

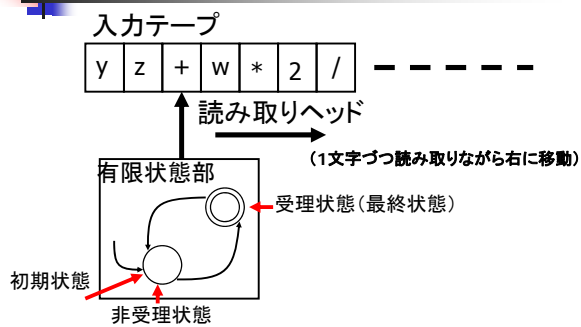
演習問題 例題4.4 b 答え



4.2 有限オートマトン

- 有限オートマトン
- 有限オートマトンと言語の認識

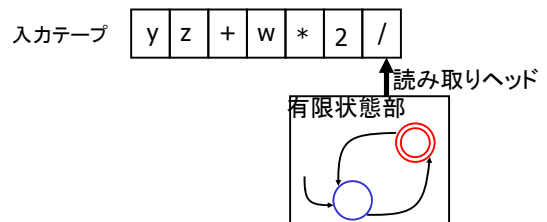
有限オートマトン (FA) のモデル



入力語の受理, 不受理

(コンパイラの字句解析を想像してください)

- 入力文字列が終わったとき
 - 状態が**受理状態**なら 入力語を**受理する**
 - 状態が**非受理状態**なら 入力語を受理しない



有限オートマトン (FA) の定義

- 内部状態の有限集合 Q
- 入力アルファベット Σ
- 状態遷移関数 $\delta: Q \times \Sigma \rightarrow Q$
- 初期状態 $S \in Q$
- 最終状態 $F \subset Q$

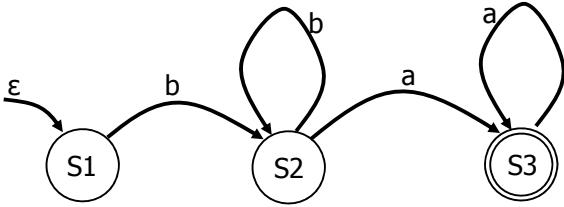
上の5項が決まると有限オートマトン (FA) は一意に決まる。

4.2.1 有限オートマトン

- 状態遷移図
 - 状態を節点に対応させ、状態遷移を有向辺に対応させたグラフ
- 状態遷移関数
 - $\delta(q, x)$ で遷移を記述
- FAの停止と受理およびエラー状態
 - 最後の文字を読み込んで停止した時の状態が最終状態ならその語は受理される
 - 入力記号以外の文字を読み込んだり状態遷移関数で未定義の組み合わせの場合: エラー状態になる
- 受理言語
 - $L(M)$ (M の受理する語): 有限オートマトン M の受理するすべての語の集合

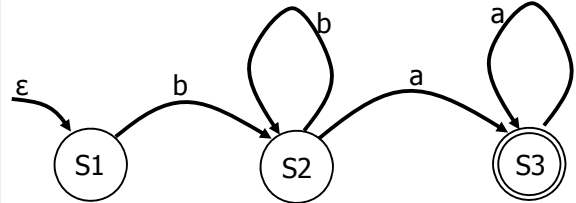
例題4.7 a (p.90)

- 受理する語の例
- 受理する語の説明



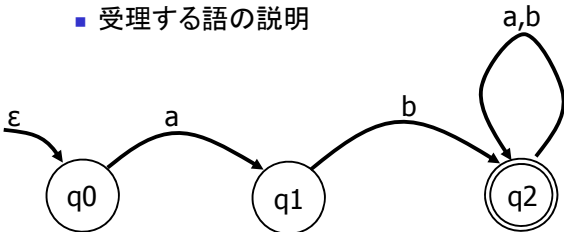
例題4.7 a 答え

- 例1 ϵba
- 例2 $\epsilon bbaaa$
- $b\{b\}a\{a\}$: 拡張BNF
- bb^*aa^* : 正規表現



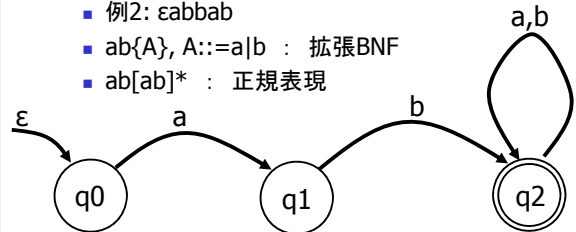
例題4.7 b

- 受理する語の例
- 受理する語の説明



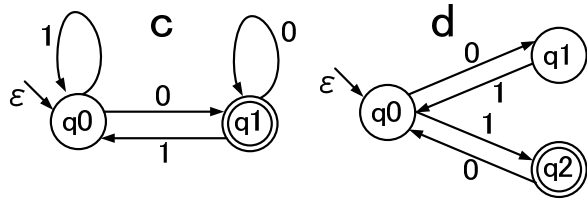
例題4.7 b 答え

- 例1: ϵab
- 例2: $\epsilon abbab$
- $ab\{A\}$, $A ::= a|b$: 拡張BNF
- $ab[ab]^*$: 正規表現



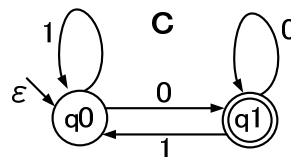
演習問題1 例題4.7 c,d

- 受理する語の例
- 受理する語の説明



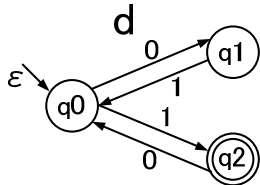
演習問題1 例題4.7 c 答え

- 例1: $\epsilon 100 \rightarrow 100$
- 例2: $\epsilon 11010 \rightarrow 11010$
- 説明: 0,1からなる文字列で最後が0



演習問題1 例題4.7 d 答え

- 例1: $\epsilon 011 \rightarrow 011$
- 例2: $\epsilon 1 \rightarrow 1$
- 説明: 01または10を0回以上繰り返す, 最後の文字が1

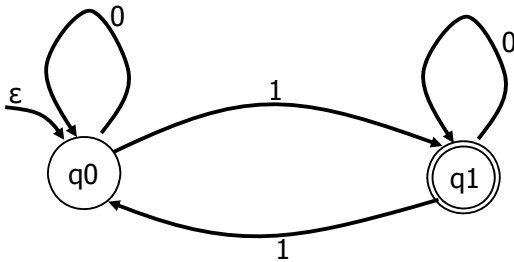


4.2.2 有限オートマトンと言語の認識 言語の受理

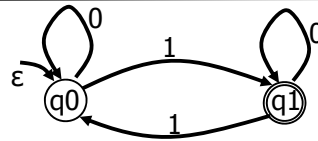
- ある言語 L に対して,
 - あるオートマトンが L に属するすべての語を受理し,
 - かつ, L に属さないすべての語を受理しないとき,
- **そのオートマトンは言語 L を受理する**
- (理想的な)C++のコンパイラは言語C++を受理する

例題4.10 a (p.93)

- (1) $M(Q, \Sigma, \delta, S, F)$ を明示
- (2) $w_1=001011, w_2=101101110$ の時の状態の遷移状況
- (3) w_1, w_2 は受理されるか
- (4) M が受理する言語



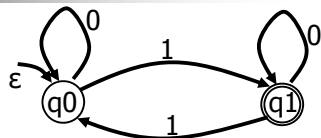
例題4.10 a-1 答え



- $Q=\{q_0, q_1\}, \Sigma=\{0,1\}, S=q_0, F=\{q_1\}$

δ	0	1
q0	q0	q1
q1	q1	q0

例題4.10 a-2,3,4 答え



w_1 のとき, 0 0 1 0 1 1
 $q_0 \rightarrow q_0 \rightarrow q_0 \rightarrow q_1 \rightarrow q_1 \rightarrow q_0 \rightarrow q_1$

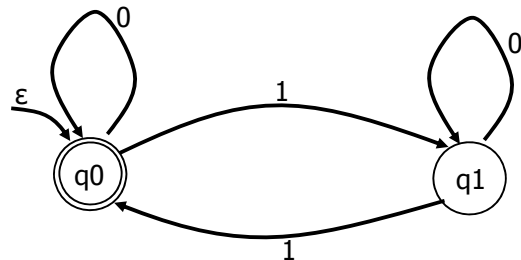
w_2 のとき, 1 0 1 1 0 1 1 1 0
 $q_0 \rightarrow q_1 \rightarrow q_1 \rightarrow q_0 \rightarrow q_1 \rightarrow q_1 \rightarrow q_0 \rightarrow q_1 \rightarrow q_0 \rightarrow q_0$

w_1 は受理される w_2 は受理されない

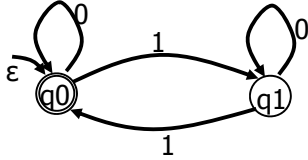
1を奇数個含む語からなる言語

演習問題2 例題4.10 b

- (1) $M(Q, \Sigma, \delta, S, F)$ を明示
- (2) $w_1=001011, w_2=101101110$ の時の状態の遷移状況
- (3) w_1, w_2 は受理されるか
- (4) M が受理する言語



演習問題2 例題4.10 b-1 答え

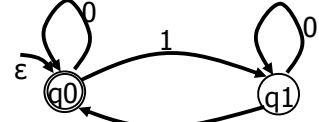


- $Q = \{q_0, q_1\}, \Sigma = \{0, 1\}, S = q_0, F = \{q_0\}$

δ	0	1
q0	q0	q1
q1	q1	q0

演習問題2

例題4.10 b-2,3,4 答え



w1のとき, 0 0 1 0 1 1 1
 $q_0 \rightarrow q_0 \rightarrow q_0 \rightarrow q_1 \rightarrow q_1 \rightarrow q_0 \rightarrow q_1$

w2のとき, 1 0 1 1 0 1 1 1 0
 $q_0 \rightarrow q_1 \rightarrow q_1 \rightarrow q_0 \rightarrow q_1 \rightarrow q_1 \rightarrow q_0 \rightarrow q_1 \rightarrow q_0 \rightarrow q_0$

w1は受理されない w2は受理される
 1を偶数個含む語からなる言語

これからの授業

- 正規表現
 - ↓
 - ε動作を含む非決定性有限オートマトン
 - ↓
 - ε動作を含まない非決定性有限オートマトン
 - ↓
 - 決定性有限オートマトン
 - ↓
 - 決定性有限オートマトンの最小化

4.3 正規表現

- 正規表現
 - 言語にどのような語が含まれているかを比較的わかりやすく表現する表現形式
- 正規表現の定義と性質
 - 正規表現の定義
 - 正規表現の性質
- 有限オートマトンの受理する言語の正規表現

正規表現の定義

- ① ϕ は正規表現である $L(\phi) = \text{空言語}$
- ϵ は正規表現である $L(\epsilon) = \{\epsilon\}$
- $a \in \Sigma$ ならば, a は正規表現である. $L(a) = \{a\}$
- ② r, s が正規表現ならば $L(r) = R, L(s) = S$ とすると
 - 和 $(r+s)$ は正規表現である. $L((r+s)) = R \cup S$: 言語の和
 - 積 (rs) は正規表現である. $L((rs)) = RS$: 言語の積
 - 閉包 (r^*) は正規表現である. $L((r^*)) = R^*$: 言語の閉包
- ③ 以上の手続きによって得られるものだけが正規表現である
 - 正規表現は, 文字記号と和, 積, 閉包の演算からなる数式表現
 - 演算子の優先順位: 閉包 \rightarrow 積 \rightarrow 和

正規表現の例

- $\{aab\} \rightarrow aab$
- $\{aab, aa\} \rightarrow aab + aa$
- $\{\epsilon, a, aa, aaa, \dots\} \rightarrow a^*$
- $\{aab\} \cup \{b, ba, baa, baaa, \dots\} \rightarrow aab + ba^*$

演習問題1 例題4.15 d e f k 言葉で説明

- d: $(0+1)^*$
- e: 0^*1^*
- f: $(0^*1^*)^*$
- k: $0(0+1+2)^*2$

演習問題1の解答例 例題4.15 d e f k

ここまで

- d: $(0+1)^*$
 - 0 または1を0回以上繰り返す
- e: 0^*1^*
 - 最初に0を0回以上繰り返す, 次に1を0回以上繰り返す
- f: $(0^*1^*)^*$
 - 最初に0を0回以上繰り返す, 次に1を0回以上繰り返す文字列を0回以上繰り返す
- k: $0(0+1+2)^*2$
 - 最初に0 次に0か1か2を0回以上繰り返す最後に2

正規言語の性質 1

- r, s, t を正規表現として
 - 交換律 $r + s = s + r$
 - 結合律 $(r + s) + t = r + (s + t)$
 - $(rs)t = r(st)$
 - べき等律 $r + r = r$
 - 分配律
 - 左分配律 $r(s + t) = rs + rt$
 - 右分配律 $(r + s)t = rt + st$

正規言語の性質 2

単位元

$$r + \phi = r \quad r\epsilon = \epsilon r = r \quad r\phi = \phi r = \phi$$

スター演算に関する性質

$$\phi^* = \epsilon$$

$$r^* = (r^*)^* = (\epsilon + r)^* = \epsilon + rr^* = \epsilon + r^*r = r^*r^*$$

$$(r+s)^* = (r^*s^*)^* = (r^*s^*)r^* = r^*(s^*)^*$$

$$= r^* + r^*s^*(r+s)^* = (r^*s^*)^* + (s^*r^*)^*$$

$$(rs)^*r = r(sr)^*$$

$$(rs+r)^*r = r(sr+r)^*$$

$$(r^*s^*)^* = \epsilon + (r+s)^*s$$

$$(rs^*t)^* = \epsilon + r(s+tr)^*t$$

正規言語の性質 3 (p98 (4.7),(4.8))

重要

■ x, s, t が正規表現のとき,

$\epsilon \notin L(t)$ とすると

つぎの性質が成立する.

$x = s + xt$ ならば $x = st^*$

$x = s + tx$ ならば $x = t^*s$

ちょっと脱線

■ $x = s + tx$ ならば $x = t^*s$ とはどのような意味か

- x: 額面500円の図書券(現在は発行終了)
- s: 500円の本
- t: 15円の本(無いと思いますが)
- $x = s + tx$: 500円の図書券を持っていれば, 500円の本か「15円の本+500円の図書券」を手に入れられる
 - 500円の図書券で15円の本を買くと485円おつりをもらえる.
 - 500円の図書券は金券ショップで485円で購入可能
- $x = t^*s$: 500円の図書券を持っていると0冊以上の15円本と1冊の500円の本を買うことができる

演習問題2 例題4.20 a, b, c

- a: $x=0+x0$
- b: $x=00+11+x0+x1$
- c: $x=01+x0^*1$

演習問題2の解答例

例題4.20 a, b, c

- a :
 - $x=0+x0$
 - $\rightarrow x=00^*$
- b :
 - $x=00+11+x0+x1$
 - $\rightarrow x=(00+11)+x(0+1)$
 - $\rightarrow x=(00+11)(0+1)^*$
- c :
 - $x=01+x0^*1$
 - $\rightarrow x=01(0^*1)^*$

$x=s+xt$ ならば $x=st^*$
 $x=s+tx$ ならば $x=t^*s$

今日のまとめ

- 小テストの解説
- 正規表現
 - 正規表現の性質
 - 状態 q_i で受理する言語
 - 有限オートマトン(状態遷移図) \rightarrow 正規表現
 - 漸化式 \rightarrow 連立方程式 \rightarrow 正規表現
- 非決定性有限オートマトン