

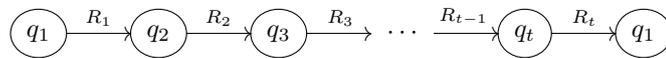
10 形式言語論 (4) – 有限オートマトンから正規表現へ –

前回, 任意の正規表現に対して, それは表す言語と同じものを受理する有限オートマトンを構成した. 今回は逆に, 任意に与えられた有限オートマトンが認識する言語と同じものを表す正規表現を構成する手法を述べる. 大雑把に, 「正規表現」から「有限オートマトン」を構成するときには, 「正規表現をラベルにもつ矢印」を細かいピースに分解し, その度に「新たな状態」を加えて有限オートマトン構成したが, これの逆を行う. つまり, 「有限オートマトン」の状態を減らしながら, その度に「ラベルに正規表現」をのせていく.

● 10-1 : 基本的なアイデア

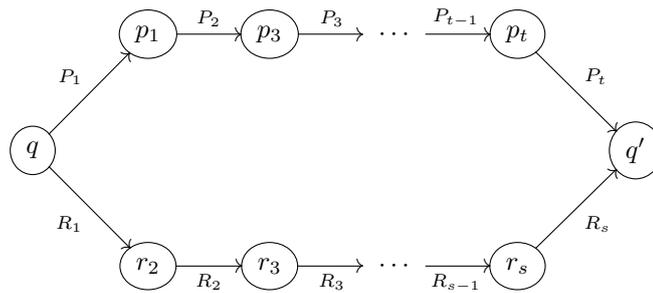
与えられた ε -規則を許した非決定性有限オートマトンを正規表現への変換に用いる基本的なアイデアは次のとおりである.

- (i) 状態 q_1 から, q_1 自身へと戻って来る矢印の列



があったとき, $(R_1 R_2 \cdots R_t)^*$ をとる.

- (ii) 状態 q から状態 q' へ至る 2 つの矢印の列



があったとき, $(P_1 P_2 \cdots P_t + R_1 R_2 \cdots R_s)$ をとる.

● 10-2 : 一般状態の消去と矢印ラベルの書き換え

$M = (K, \Sigma, \delta, q_0, F)$ を ε -規則を許す非決定性有限オートマトンであるとする. $q \in K$ を M のひとつの状態とする. M の状態遷移図において, q の周辺に存在する状態と, その間にある矢印の状況を書き出すことにしよう. 状態 q の周りに存在する矢印を列挙する.

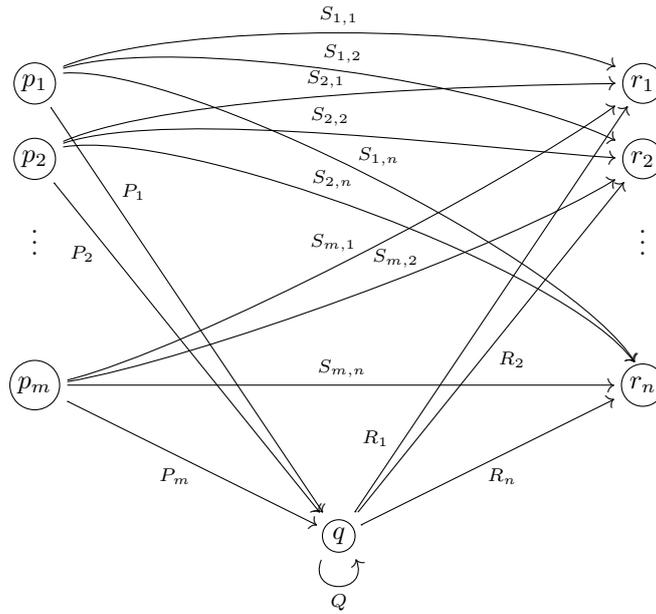
- (a) q から q 自身への, ラベルが Q の矢印 $q \xrightarrow{Q} q$.
- (b) q とは異なる状態 p_1, \dots, p_m であって, p_i から q への, ラベルが P_i の矢印 $p_i \xrightarrow{P_i} q$.
- (c) q とは異なる状態 r_1, \dots, r_n であって, q から r_i への, ラベルが R_i の矢印 $q \xrightarrow{R_i} r_i$.
- (d) 各 $i = 1, 2, \dots, m$ と $j = 1, \dots, n$ に対して, p_i から r_j への, ラベルが $S_{i,j}$ の矢印 $p_i \xrightarrow{S_{i,j}} r_j$.

実際には, q から q への矢印や, p_i から r_j への矢印は存在するとは限らない. そこで, 次の操作を行う.

(一般状態化) 状態 q をひとつ固定する. これに対して,

- (0-1) q から q 自身への矢印がなければ, ラベルが \emptyset である矢印 $q \xrightarrow{\emptyset} q$ を追加する.
- (0-2) p_i から r_j への矢印がなければ, ラベルが \emptyset である矢印 $p_i \xrightarrow{\emptyset} r_j$ を追加する.

すると, 状態 q の周辺では, 以下のような図が描かれていることになる.



(状態消去) (一般状態化) で着目した状態 q を消去し, q へ向かう矢印と q から出ていく矢印を全て消去する. その後, p_i から r_j へのラベルを次のように書き換える.

$$p_i \xrightarrow{S_{i,j} + P_i(Q)^* R_j} r_j \quad (i = 1, 2, \dots, m, j = 1, 2, \dots, n)$$

注意 10.1. 形式的には上記のようにラベルを付け替えればよいが, 正規表現が簡単になる場合は簡単にする.

- (a) q から q 自身への矢印がなければ, $\emptyset^* = \varepsilon$ であるから, 全ての i, j に対して $p_i \xrightarrow{S_{i,j} + P_i R_j} r_j$ と書き換えれば良い.
- (b) p_i から r_j 自身への矢印がなければ, \emptyset との和演算は意味がないので, この i, j については $p_i \xrightarrow{P_i(Q)^* R_j} r_j$ と書き換えれば良い.

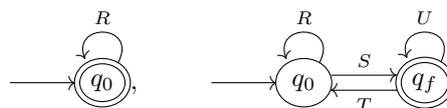
● 10-2 : 正規表現への変換

では, 有限オートマトンを正規表現へ変換する手順を示そう.

(手順 1) 各受理状態 $q, q' \in F$ ($q = q'$ でもよい) に対して, q から q' への矢印 $q \xrightarrow{a_1 a_2 \dots a_s} q'$ が存在している場合, この矢印を次のように書き換える.

$$q \xrightarrow{a_1 + a_2 + \dots + a_s} q'$$

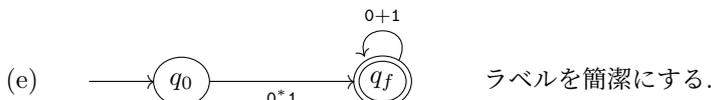
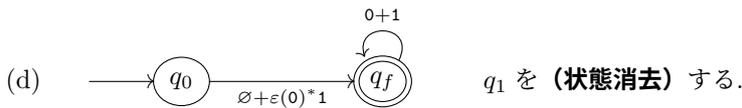
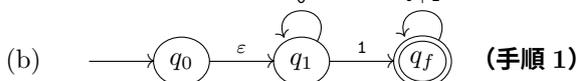
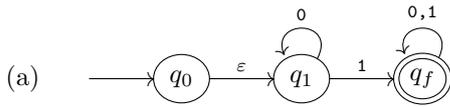
(手順 2) 各受理状態 q_f に対して, 開始状態 q_0 以外の全ての状態 q の (一般状態化) と (状態消去) を行う. その結果, 以下で示す有限オートマトンを得る. ただし, 矢印がない場合はラベル \emptyset の矢印があるとみなす.



(手順 3) (手順 2) で得た全ての有限オートマトンに対して, 1 状態であれば R^* , 2 状態であれば $(R + S(U)^* T)^* S(U)^*$ を構成する.

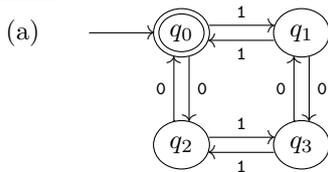
(手順 4) (手順 3) で得た全ての正規表現の和演算をとる.

例 10-1 次の (a) で与えられた有限オートマトンを正規表現へ変換する.

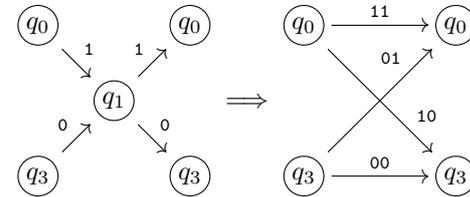
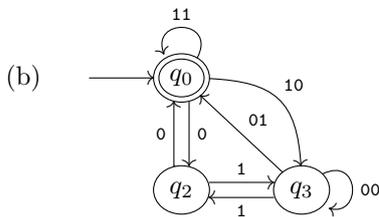


(f) $0^*1(0+1)^*$ (手順 3)

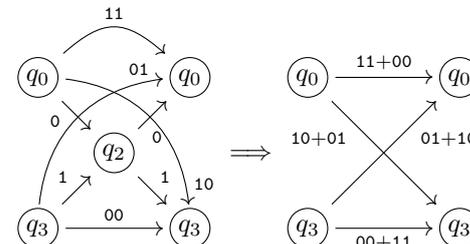
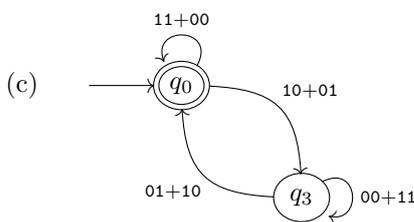
例 10-2 次の (a) で与えられた有限オートマトンを正規表現へ変換する.



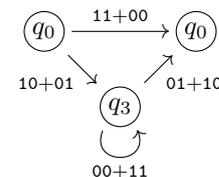
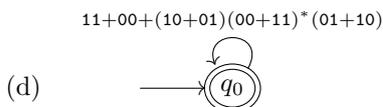
(手順 2) により, q1 を (一般状態化) (状態消去) する.



(手順 2) により, q2 を (一般状態化) (状態消去) する.

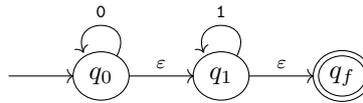


(手順 2) により, q3 を (一般状態化) (状態消去) する.

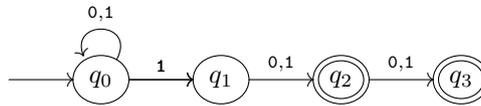


(e) $(11+00+(10+01)(00+11)^*(01+10))^*$ (手順 3)

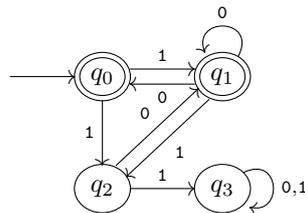
小テスト 10-1 次の非決定性有限オートマトンが認識する言語を受理する言語を表す正規表現を求めよ.



小テスト 10-2 次の非決定性有限オートマトンが認識する言語を受理する言語を表す正規表現を求めよ.



演習問題 10-1 下図で表される有限オートマトンが認識する言語を表現する正規表現を構成せよ.



演習問題 10-2 下図で表される有限オートマトンが認識する言語を表現する正規表現を構成せよ.

