DFAの最小化

今日の授業内容

- DFAの状態数を最小化する
- 最小化したDFAの状態数が最小であることの証明
- 状態数最小のDFAはどれも同じであることの証明

DFA の遷移関数の全関数化

ある関数がどの引数に対しても値が定義されているとき全関数と呼び、そうでないとき部分関数と呼ぶ。

DFA の遷移関数には、一般に値が未定義の部分がある。これはある状態である記号を読んだときに次の状態が決まっていないことに対応する。

ダミー状態を付け加えて、遷移関数の未定義の値をダミー状態にすることを全関数化と呼ぶ。この操作は前回の授業で補集合を受理するオートマトンの構成のときに使った

DFA の最小化 — 区別できる状態

$M = (K, T, t, q_0, F)$ を最小化する。遷移関数は全関数とする。

状態の組 q_1, q_2 について、片方の状態を最終状態に遷移させもう片方の状態を非最終状態に遷移させる記号列があるときに、q_1 と q_2 を区別できると言う。

区別できない状態をまとめても、オートマトンが受理する言語は変わらないから、区別できない状態をひとまとめにすることにより、状態を減らす

区別できる状態の組の集合を以下のように求める
$D_0 = \{(q, q') \mid q \in F, q \in K \setminus F\} = F \times (K \setminus F)$
$D_{i+1} = D_i \cup \{(q, q') \in K \times K \mid (t(q, a), t(q', a)) \in D_i$ となる $a \in T$ が存在する $\}$
D_i は長さ i の記号列で区別できる状態の組の集合
そのうち $D_{i+1} = D_i$ になるので、それを D とする

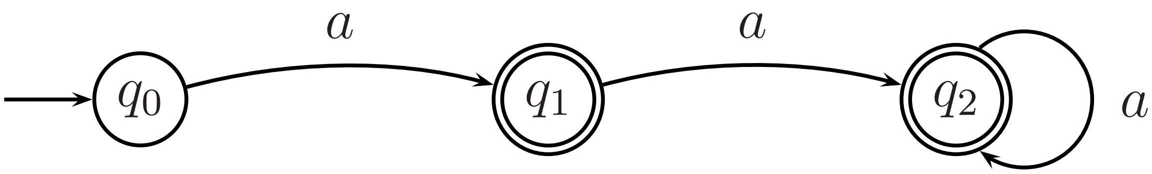
DFA の最小化 2 — 区別できない状態をまとめる

状態 q と q' は $(q, q') \notin D$ のとき区別できない
$[q]$ を状態 q と区別できない状態からなる**集合**とする

新しいオートマトン $M' = (K', T', t', q'_0, F')$ を以下のように作る
$K' = \{[q] \mid q \in K\}$
$T' = T$
$t([q], a) = [t(q, a)]$
$q'_0 = [q_0]$
$F' = \{[q] \mid q \in F\}$

さらに出発状態から到達できない状態、および最終状態に到達できない状態を除く。遷移関数を全関数にするためにダミー状態を付け加えた場合、ダミー状態からは最終状態に到達できないため、ダミー状態は削除される。

最小化の例



$D_0 = F \times (K \backslash F) = \{q_1, q_2\} \times \{q_0\} = \{(q_1, q_0), (q_2, q_0)\}$

次に D_1 を求めたい

$\{(q, q') \mid (t(q, a), t(q', a)) \in D_0$となる $a \in K$ が存在する $\} = \emptyset$ である。

従って $D_1 = D_0 = D$.

表を用いた D の求め方

q_0	■	■	■
q_1	$\times$	■	■
q_2	$\times$		■
	q_0	q_1	q_2

1. 上のような表を書き、D_0 に対応する部分に $\times$ を書く

2. $\times$が書いていない部分に対応する状態の組を適当な記号で遷移させて、$\times$がついている状態の組に移せるかかどうか調べ、移せる場合は$\times$を記入する。

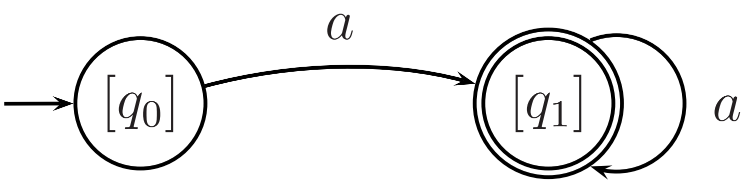
3. $\times$を記入できるところが無くなるまで 2 を繰り返す

4. $\times$が記入されているところが D

最小化の例 つづき

$[q_0] = \{q_0\}$
$[q_1] = \{q_1, q_2\}$
$[q_2] = \{q_1, q_2\}$

$t([q_0], a) = [q_1]$
$t([q_1], a) = [q_2] = [q_1]$



なぜ最小なのか

あるオートマトンについて、状態 q にあるとき記号列 w を入力すると状態 q' に移るとき

$$q \xrightarrow{w} q'$$

と書く

$M = (K, T, t, q_0, F)$ 最小化したオートマトン
$M' = (K', T, t', q_0', F')$: $T(M) = T(M')$ となるオートマトン
$K = \{q_0, q_1, \ldots, q_n\}$ とする
$q_0 \xrightarrow{w_i} q_i$ となる記号列 w_i を選んでおく

$q_0' \xrightarrow{w_i} q_i'$ とすると
$q_0', \ldots, q_n'$ は相異なるので、K' は K より多くの状態を持つ

$q_i' \neq q_j'$ の証明：記号列 w を

$$\begin{array}{lcl} q_i & \xrightarrow{w} & \text{最終状態} \\ q_j & \xrightarrow{w} & \text{未定義または非最終状態} \end{array}$$

とする。$w_i w \in T(M)$ かつ $w_j w \notin T(M)$ である
もし $q_i' = q_j'$ となる $i \neq j$ があったとすると、$w_i w$ と $w_j w$ の両方を同時に受理または拒絶するので、$T(M) \neq T(M')$ となり矛盾が起きる

最小オートマトンは本質的に一つ

$T(M) = T(M')$, M, M' は状態数最小、
$M = (K, T, t, q_0, F)$, $M' = (K', T, t', q'_0, F')$

$$
\begin{aligned}
L_i &= \{w \mid q_0 \xrightarrow{w} q_i\}, \\
L'_i &= \{w \mid q'_0 \xrightarrow{w} q'_i\}
\end{aligned}
$$

すべての L'_k について ある i が存在して $L_i \supseteq L'_k$。そのような i が無かったとして矛盾を導くことができる (付録参照)。同様にすべての L_k について ある i が存在して $L'_i \supseteq L_k$。したがってすべての i について $L_i = L'_k$ となる k が存在する。添字を付け直して $L_i = L'_i$ とする。

$t(q_i, a) = q_j$ の場合、$w \in L_i$ として $wa \in L_j = L'_j$ だから、$t'(q'_i, a) = q'_j$ である。

付録

「すべての L'_k について ある i が存在して $L_i \supseteq L'_k$」の証明

$i \neq j$ ならば、

$$\begin{array}{lcl} q_i & \xrightarrow{w} & \text{最終状態} \\ q_j & \xrightarrow{w} & \text{未定義または非最終状態} \end{array}$$

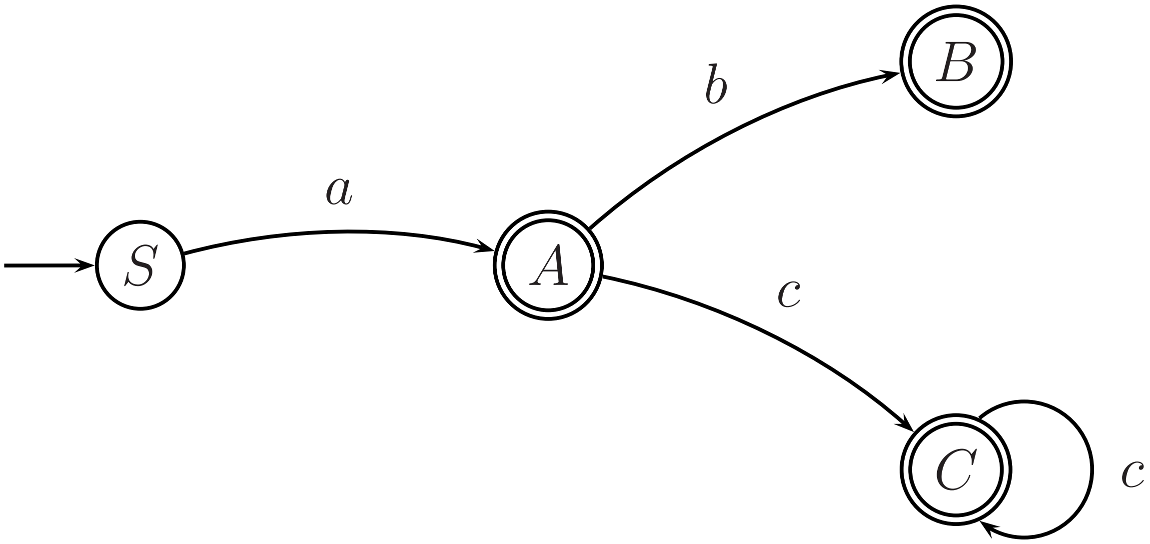
となる記号列 w が存在する (もし無ければ q_i と q_j をまとめられる)

$L_i \cap L'_k \neq \emptyset$ とする。もし L'_k がどの L_i とも交わりが無ければ、q'_k からは最終状態に到達できないことになり、q'_k を削除できるので矛盾。

$w_k \in L'_k$ かつ $w_k \notin L_i$ とする。もし $w_k \in L_j$ となる j が存在した場合、上記の w と $w_i \in L'_k \cap L_i$ を用いて $w_i w \in T(M)$ かつ $w_k w \notin T(M)$ となるが、オートマトン M' は $w_i w$ と $w_k w$ を区別できないので矛盾。もし $w_k \in L_j$ となる j が存在しない場合、オートマトン M は w_k を入力したときの状態が未定義なので、いかなる記号列 w についても $w_k w$ を受理しない。上記の w と $w_i \in L'_k \cap L_i$ を用いて $w_i w \in T(M)$ かつ $w_k w \notin T(M)$ となるが、オートマトン M' は $w_i w$ と $w_k w$ を区別できないので矛盾。

演習問題

演習問題 34 以下のオートマトンを最小化し、グラフと (K, T, t, q_0, F) の両方の形式で書け。いきなり答えだけ書かないで解答過程を書くこと。



演習問題 35 今日の授業でわかりにくい所や要望を書いて下さい

前回の演習の採点基準：
グラフまたは記号によるオートマトンの記述が正しければ 0.5 点。小数点切り上げ。満点は 8 点