

# 「オートマトンと形式言語」補足資料(5)

## 1 DFAから正規表現への変換に関する補足(教科書p.80-88)

DFAから正規表現へ変換する手続きについて、整理しておく。

Step 1. DFAをGNFAへ変換する。このとき、DFAの状態数を  $k$  とすると、GNFAの状態数は  $k + 2$  個になる。

Step 2. GNFAの状態数が2個になるまで、以下を繰り返す：GNFAの等価変換により、GNFAの状態数を1つ減らす。

Step 3. 状態数が2個のGNFAから、正規表現を取り出す。

Step 3は簡単。状態数が2個のGNFAには、辺が1つしかないので、その辺のラベルである正規表現を持ってくればよい。

### DFAからGNFAへの変換

まず、変換の手続きは以下の通り。

Step 1. DFAに、新しい状態  $q_{start}$  と  $q_{accept}$  を追加する。

Step 2. 開始状態を  $q_{start}$  に変更し、 $q_{start}$  からの辺を以下のように与える：

- $q_{start}$  から元の開始状態へは、 $\varepsilon$  の辺を与える。
- それ以外の状態へ、 $\emptyset$  の辺を与える。

Step 3. 終了状態を  $q_{accept}$  に変更し、 $q_{accept}$  への辺を以下のように与える：

- 元々の終了状態からは、 $\varepsilon$  の辺を与える。
- それ以外の状態からは、 $\emptyset$  の辺を与える。

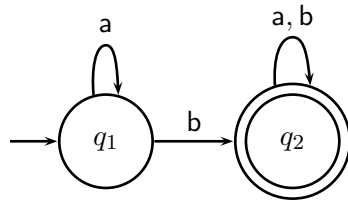
なお、 $q_{start}$  から  $q_{accept}$  への辺は、Step 2で与えてもStep 3で与えてもよいが、 $\emptyset$  の辺を与えることになる。

Step 4. 元の状態同士の辺(自分自身同士も含む)を、以下のように場合によっては変更する。

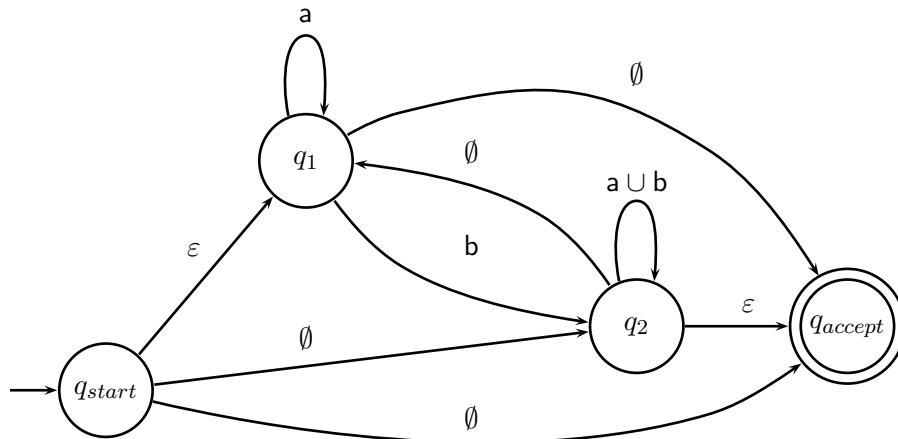
- 元の辺がないときは、 $\emptyset$  の辺を与える。
- 元の辺のラベルが複数あるとき、それを  $\cup$  で繋いだ正規表現に変更する。例えば、ラベルが  $a_1, \dots, a_k$  であれば、 $a_1 \cup \dots \cup a_k$  に直す。
- それ以外の場合はそのまま。

例 1. (教科書 図 1.67)

変換前の DFA  $M_1$ :



変換後の GNFA  $N_1$ :



なお、教科書では、ラベルが  $\emptyset$  の辺は省略しているので注意.

### GNFA の等価変換

GNFA の状態を 1 つ減らす等価変換手続きは以下の通り.

Step 1.  $q_{start}$ ,  $q_{accept}$  以外の状態から, 削除する状態を 1 つ選ぶ. ここでは, 選んだ状態を  $q_{rip}$  とおく.

Step 2. 状態  $q_{rip}$  と  $q_{rip}$  につながっている辺をすべて取り除く.

Step 3. 残った辺のそれぞれについて, ラベルを以下のように更新する:

(a) 考える辺を  $q_i$  から  $q_j$  への辺とする.

(b) 正規表現  $R_1, \dots, R_4$  を以下のようにおく:

$R_1$ : 元の GNFA の,  $q_i$  から  $q_{rip}$  への辺のラベル

$R_2$ : 元の GNFA の,  $q_{rip}$  から  $q_{rip}$  への辺のラベル

$R_3$ : 元の GNFA の,  $q_{rip}$  から  $q_j$  への辺のラベル

$R_4$ : 元の GNFA の,  $q_i$  から  $q_j$  への辺のラベル

(c)  $q_i$  から  $q_j$  への辺のラベルを,  $R_1 R_2^* R_3 \cup R_4$  に更新する.

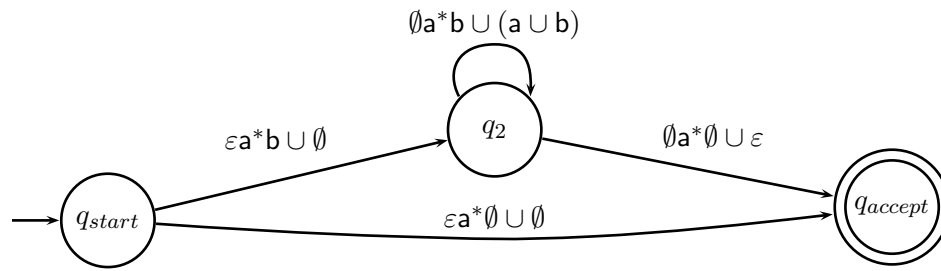
(もちろん,  $R_1, \dots, R_4$  は, 考える辺によってそれぞれ異なるので注意.) なお, 正規表現は, 簡略化できるなら, した方がわかりやすい.

例 2.

さきほどの GNFA  $N_1$  を変換して, GNFA  $N_2$  を構成する.

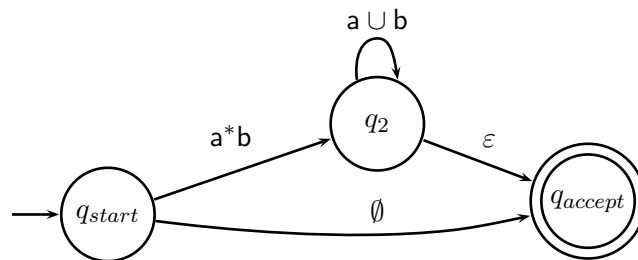
削除する状態に  $q_1$  を選んだとしよう.

素直に, それぞれの辺を,  $R_1R_2^*R_3 \cup R_4$  の形で更新すると, 以下ようになる.



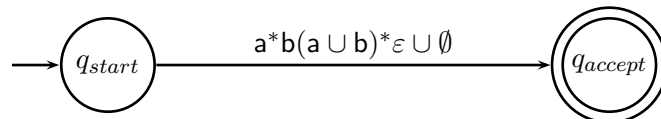
それぞれの正規表現を簡略化すると以下ようになる.

変換後の GNFA  $N_2$ :



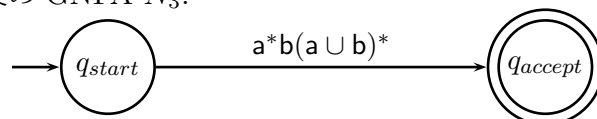
次に, GNFA  $N_2$  を変換して, GNFA  $N_3$  を構成する. 次に削除する状態としては,  $q_2$  しか選択肢はない.

素直に, 辺を更新すると, 以下ようになる.



正規表現を簡略化すると以下ようになる.

変換後の GNFA  $N_3$ :

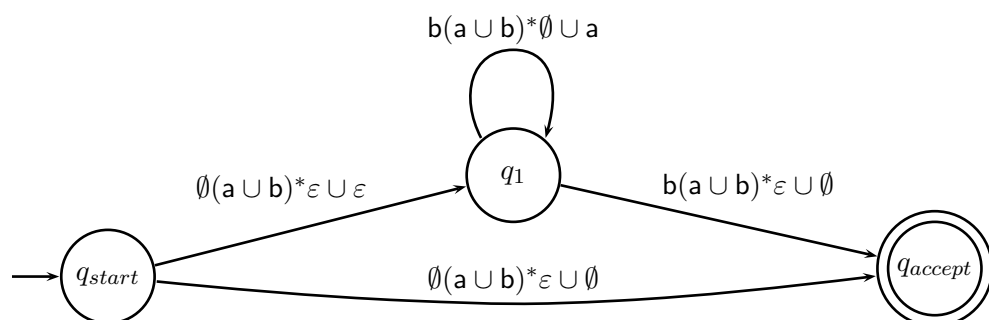


従って, DFA  $M_1$  と等価な正規表現として,  $a^* b (a \cup b)^*$  が得られた.

### 例 3.

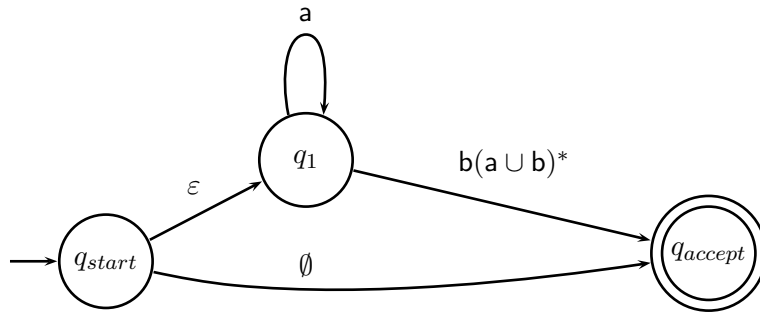
GNFA  $N_1$  を変換する際に, 削除する状態に  $q_2$  を選ぶことも出来る.

素直に, それぞれの辺を,  $R_1R_2^*R_3 \cup R_4$  の形で更新すると, 以下ようになる.



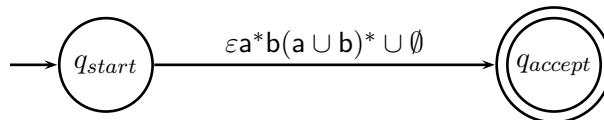
それぞれの正規表現を簡略化すると以下ようになる.

変換後の GNFA  $N_4$ :



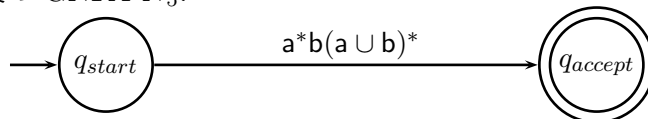
次に、GNFA  $N_4$  を変換して、GNFA  $N_5$  を構成する。次に削除する状態としては、 $q_1$  しか選択肢がない。

素直に、辺を更新すると、以下のようになる。



正規表現を簡略化すると以下のようになる。

変換後の GNFA  $N_5$ :



従って、DFA  $M_1$  と等価な正規表現として、 $a^* b(a \cup b)^*$  が得られた。

この例の場合には、結果として、どちらの順序で状態を削除しても、同じ正規表現が得られたが、場合によっては、同じ正規表現が得られるとは限らない。しかし、複数の正規表現が得られたとしても、それらは、実は等価となっている。したがって、どのような順序で、状態を削除しても構わない。