第 10 講 DFA と NFA

林　恒俊

DFA と NFA が定義する言語

- DFA と NFA の定義を単純に比較すると相違点は状態遷移表だけでありその他の要素は全く同一である。NFA の状態遷移表は DFA のそれを制限を少なくして拡張したものになっている。以上の点から直感的には NFA が高度の能力を備えるように見える。すなわち受理できる言語がより広範囲なように見える。
- 事実はそうではない。DFA が定義する言語のクラスと NFA が定義する言語のクラスは同一である。すなわち

$$\mathcal{L}_{\mathbf{DFA}} = \mathcal{L}_{\mathbf{NFA}}$$

  である。
- これを証明するためには次の 2 点をいえばよい。
  - $\mathcal{L}_{\mathbf{DFA}} \subseteq \mathcal{L}_{\mathbf{NFA}}$ すなわちある DFA について同じ言語を受理する NFA が存在する
  - $\mathcal{L}_{\mathbf{DFA}} \supseteq \mathcal{L}_{\mathbf{NFA}}$ すなわちある NFA について同じ言語を受理する DFA が存在する

DFA から NFA

- この証明は自明である。NFA の定義は DFA の定義を包含しているためすべての DFA は同時に NFA でもある。したがって

$$\mathcal{L}_{\mathbf{DFA}} \subseteq \mathcal{L}_{\mathbf{NFA}}$$

  が成立する。

NFA から DFA

- 与えられた NFA に対して同一言語を受理する DFA を次の順で構成することにより $\mathcal{L}_{\mathbf{NFA}} \subseteq \mathcal{L}_{\mathbf{DFA}}$ を証明する。
  1. 記号列遷移の 1 記号遷移化
  2. 空遷移の除去
  3. 部分集合構成法により DFA の構成

  与えられた NFA にこれらの各段階にしたがって変更を加えて新しい NFA を構成する。同時に変更を行っても受理する言語が同一であることを検証する。最後の段階で最終的な DFA を構成する。
- 当初の NFA を $M = (Q, \Sigma, \Delta, q_0, F)$ とし各段階で得られる NFA をそれぞれ M_1, M_2, M_3 とする。

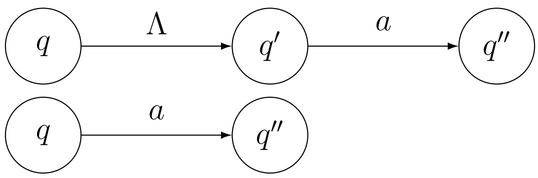
記号列遷移の 1 記号遷移化

- これが最も容易な変形である。$\bigcirc\xrightarrow{abc}\bigcirc$ という遷移があれば、状態を追加して $\bigcirc\xrightarrow{a}\bigcirc\xrightarrow{b}\bigcirc\xrightarrow{c}\bigcirc$ という遷移で置換えることにより 1 個の記号で遷移するようにできる。
- 形式的には次のような方法で構成する。
  1. $M_1 = (Q_1, \Sigma, \Delta_1, q_0, F)$ として定義し Q_1, Δ_1 をそれぞれ Q, Δ と等しくする
  2. すべての $(q, \sigma, q') \in \Delta$ ただし $|\sigma| > 1$ について次の操作を行う
     - Δ_1 から (q, σ, q') を取除く
     - $|\sigma| = k$ なら新規状態 $q_1, q_2, \ldots, q_{k-1}$ を Q_1 に追加する
     - Δ_1 に遷移 $(q, \sigma(1), q_1)$, $(q_1, \sigma(2), q_2), \ldots, (q_{k-1}, \sigma(k), q')$ を追加する
- M_1 のアルファベット、初期状態、終了状態集合は M と同一であり M の持っている長さが 1 以下の記号列に対する遷移はすべて M_1 にも含まれる。長さが 2 以上の記号列の M の遷移について M_1 は実質

的に同じ遷移を行う (遷移先の最後の状態が同一)。これらの点から M_1 の受理する記号列は M が受理する記号列と同じである。すなわち M_1 と M は同一言語を定義する。

空遷移の除去

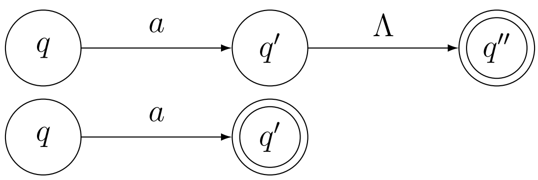
- 次のように状態 q から空遷移で q' に遷移しさらに a で q'' に遷移する場合、空遷移は入力記号を消費しないため状態 q から記号 a で q'' に遷移するのと等価である。



  すなわち空遷移と 1 記号による遷移を含む複数段階の遷移を 1 記号だけの遷移で置換えると空遷移を除去できる。これにはある状態から

  - 空遷移で遷移可能なすべての状態をしらべて
  - そこから 1 記号による遷移を求め
  - これらを元状態からの 1 記号による遷移とすればよい。

- 空遷移の遷移先が終了状態の場合記号を消費せずに終了状態をとる。したがって空遷移を除去した後その遷移元を終了状態に加えておかなければならない。



- 次のように M_2 を構成しそれに次のアルゴリズムにしたがって状態遷移を追加すればよい。

  1. $M_2 = (Q_2, \Sigma, \Delta_2, q_0, F_2)$ として定義し $\Delta_2 \leftarrow \emptyset$ とする。また Q_2, F_2 は後で計算する

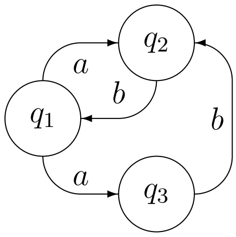
2. Q_1 の各状態 $q \in Q_1$ について
   - Σ の各記号 $a \in \Sigma$ について
   - 状態 q から記号 a による状態 q' への遷移が存在する場合すなわち $(q, [1, a]) \vdash^*_{M_1} (q', [2, a])$ なら
   - $\Delta_2 \leftarrow \Delta_2 \cup \{(q, a, q')\}$
3. Q_2 を Δ_2 の下で q_0 から到達可能な状態の集合とする
4. Q_2 中から次の条件を満たす状態 q をすべて選びその集合を F_2 とする
   - q から入力記号を消費せずに F_1 のいずれかの要素 f に到達できるもの (この中には F_1 の要素も含まれる) すなわち

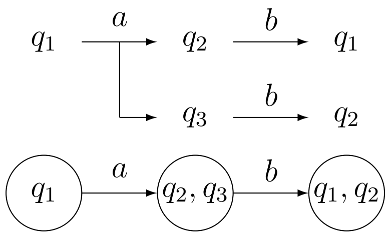
$$F_2 = \{q \mid (q, [1, \Lambda]) \vdash^*_{M_1} (f, [1, \Lambda])\} \ (f \in F_1)$$

- M_1 と M_2 が同一の言語を受理することを証明するためには M_1 が受理する記号列 x は M_2 で受理されること及び M_2 が受理する記号列 x は M_1 でも受理されることをいえばよい。前半については x 中のある記号 a の M_1 での (空遷移を含む) 状態遷移が M_2 の状態遷移と最終的に同じになることから説明できる。後半についても同様に説明できる。詳細は省略する。

部分集合構成

- M_2 での非決定性は入力記号に対してある状態から複数の状態に遷移することが原因で生じる。M_3 で複数状態への遷移を 1 状態への遷移でエミュレートすることにより決定的動作を行わせることができる。すなわち複数個の遷移先状態を統合して 1 状態と見なすことにより非決定性を解消する。





上図の状態 q_1 で a が入力なら非決定的に q_2 または q_3 へ遷移する。これを状態 q_1 から a で状態 q_2, q_3 へ遷移するものと考える訳である。さらに状態 q_2, q_3 で入力が b なら q_2 は q_1 へ、q_3 は q_2 へ遷移するので、状態 q_2, q_3 から b で状態 q_1, q_2 に遷移すると考える。

- この考察から M_3 を構成するためには M_2 の状態の部分集合を M_3 の状態として採用しその上に M_2 のすべての状態遷移を含むようにすればよい。これを**部分集合構成法** (**subset construction**) と呼ぶ。
- $M_3 = (Q_3, \Sigma, \delta_3, \{q_0\}, F_3)$ として定義する。
  - Q_3 は集合を要素とする集合で Q_2 の冪集合の部分集合である
$$Q_3 \subseteq 2^{Q_2}$$
  - 初期状態は M_2 の初期状態 q_0 のみを要素とする集合 $\{q_0\}$ である
  - F_3 は F_2 の要素を含む Q_3 の要素から構成される
$$F_3 = \{Q' \mid Q' \in Q_3, Q' \cap F_2 \neq \emptyset\}$$

  なお Q_3 は集合を要素とする集合であるが要素の名前を適切につけ直すことにより普通の状態集合とすることができる。
- δ_3 は組織的に構成する方法と順次に構成する方法が考えられる。

組織的な DFA の構成

- $Q_2 = \{q_0, q_1, \ldots, q_k\}$, $\Sigma = \{a_1, a_2, \ldots, a_n\}$ とする。δ_3 を次のように状態として Q_2 のすべての部分集合を含む遷移表として作成する。

	a_1	a_2	$\cdots$	a_{n-1}	a_n
$\emptyset$	$\emptyset$	$\emptyset$	$\cdots$	$\emptyset$	$\emptyset$
$\{q_0\}$	$\emptyset$	$\emptyset$	$\cdots$	$\emptyset$	$\emptyset$
$\{q_1\}$	$\emptyset$	$\emptyset$	$\cdots$	$\emptyset$	$\emptyset$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\{q_0, q_1, \ldots, q_k\}$	$\emptyset$	$\emptyset$	$\cdots$	$\emptyset$	$\emptyset$

- すべての $Q' \in Q_3$ について次の計算を行う。

	a_1	$\cdots$	a_i	$\cdots$	a_n
$\emptyset$	$\emptyset$	$\cdots$	$\emptyset$	$\cdots$	$\emptyset$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
Q'	$\emptyset$	$\cdots$	Q''	$\cdots$	$\emptyset$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
$\{q_0, q_1, \ldots, q_k\}$	$\emptyset$	$\cdots$	$\emptyset$	$\cdots$	$\emptyset$

  状態 $q' \in Q'$ と記号 $a_i \in \Sigma$ のすべての組合せ (q', a_i) について遷移 (q', a_i, q'') が Δ_2 に含まれている場合 δ_3 の遷移 (Q', a_i, Q'') の Q'' に q'' を追加する。

- なお Δ_2 で遷移先が未定義の場合 δ_3 では空集合が遷移先になっている。

- 完成した遷移表にどのような入力記号列に対しても $\{q_0\}$ から到達できない状態を含む場合がある。最後に到達できない状態を削除する。

順次的な DFA の構成

- $Q_2 = \{q_0,\, q_1, \ldots, q_k\}$, $\Sigma = \{a_1,\, a_2, \ldots, a_n\}$ とする。次のような Q_3 の初期状態だけからなる遷移表を初期値として δ_3 を構成する。次状態が求まればそれを追加することにより遷移表を拡張する。

	a_1	a_2	$\cdots$	a_{n-1}	a_n
$\{q_0\}$	$\emptyset$	$\emptyset$	$\cdots$	$\emptyset$	$\emptyset$

- 遷移 (q_0, a_i, q') が Δ_2 に含まれていれば δ_3 の遷移 $(\{q_0\}, a_i, Q')$ の遷移先 Q' に q' を追加する。

	a_1	$\cdots$	a_i	$\cdots$	a_n
$\{q_0\}$	$\emptyset$	$\cdots$	Q'	$\cdots$	$\emptyset$

- 状態 Q' が δ_3 に含まれていなければ表に状態 Q' の行を追加してで拡

張する。

	a_1	$\cdots$	a_i	$\cdots$	a_n
$\{q_0\}$	$\emptyset$	$\cdots$	Q'	$\ldots$	$\emptyset$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
Q'	$\emptyset$	$\cdots$	$\emptyset$	$\cdots$	$\emptyset$

- 追加された状態 Q' のすべての要素とアルファベットのすべての組合せについて状態遷移を調べ表に追加する。表に新規の状態が追加されなくなれば遷移表 δ_3 は完成する。
- M_2 と M_3 が受理する言語の同一性について証明することはそれほど難しくはない。各自試みるとよい。詳細は省略する。

FSA の能力は決定的・非決定的に依存しない

- 以上の説明から与えられた NFA に対して同一言語を受理する DFA を構成できることが示された。したがって

$$\mathcal{L}_{\mathbf{NFA}} \subseteq \mathcal{L}_{\mathbf{DFA}}$$

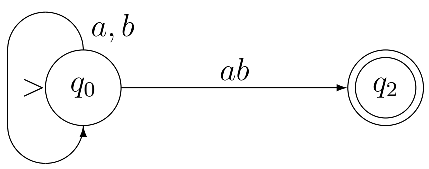
  が成立する。そして最終的に

$$\mathcal{L}_{\mathbf{NFA}} = \mathcal{L}_{\mathbf{DFA}}$$

  が証明された。
- FSA では決定的あるいは非決定的のいずれであっても定義する言語クラスは同じである。
- 以下で NFA から DFA を構成する例題をいくつか考える。

単純な NFA

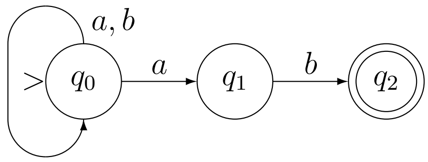
- 次の NFA M_{ab} を考察する。



$$
\begin{aligned}
M_{ab} &= (\{q_0, q_2\}, \{a, b\}, \Delta_{ab}, q_0, \{q_2\}) \\
\Delta_{ab} &= \{(q_0, a, q_0), (q_0, b, q_0), (q_0, ab, q_2)\}
\end{aligned}
$$

M_{ab} は言語 $\{a, b\}^* ab$ を受理する。

- M_{ab} を 1 記号のみで遷移するようにした NFA M'_{ab} は次の図に示される。



$$
\begin{aligned}
M'_{ab} &= (\{q_0, q_1, q_2\}, \{a, b\}, \Delta'_{ab}, q_0, \{q_2\}) \\
\Delta'_{ab} &= \{(q_0, a, q_0), (q_0, b, q_0), (q_0, a, q_1), (q_1, b, q_2)\}
\end{aligned}
$$

- M'_{ab} から部分集合構成法で DFA M''_{ab} を構成する。
  - まず M''_{ab} の初期状態は $\{q_0\}$ である。
  - M'_{ab} で q_0 から a で遷移する先の状態は q_0 と q_1 であり b で遷移する先の状態は q_0 である。
  - したがって M''_{ab} の遷移関数には

    $$(\{q_0\}, a, \{q_0, q_1\}), (\{q_0\}, b, \{q_0\})$$

    という遷移が含まれる
- 状態 $\{q_0, q_1\}$ からからの遷移は
  - a による遷移は q_0 が寄与するもののみで $(\{q_0, q_1\},\ a,\ \{q_0, q_1\})$
  - b による遷移には q_0 が寄与する q_0 と q_1 が寄与する q_2 が含まれるため $(\{q_0, q_1\},\ b,\ \{q_0, q_2\})$ が M''_{ab} の遷移関数に追加される
- 状態 $\{q_0, q_2\}$ からの遷移は q_0 が寄与する分のみなので

  $$(\{q_0, q_2\}, a, \{q_0, q_1\}), (\{q_0, q_2\}, b, \{q_0\})$$

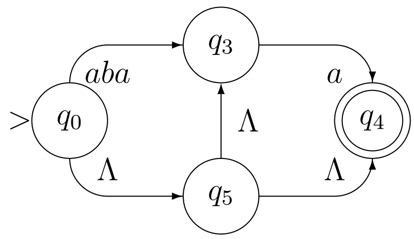
  が追加される。
- 終了状態は q_2 を含む M''_{ab} の状態 $\{q_0, q_2\}$ のみである。

- M''_{ab} の状態遷移表は次のようになる。

	a	b
$\{q_0\}$	$\{q_0, q_1\}$	$\{q_0\}$
$\{q_0, q_1\}$	$\{q_0, q_1\}$	$\{q_0, q_2\}$
$\{q_0, q_2\}$	$\{q_0, q_1\}$	$\{q_0\}$

少し複雑な NFA

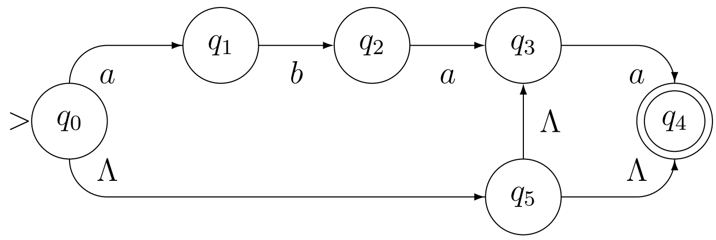
- ここでは次の NFA M_0 から等価な (同一言語を受理する) DFA を構成する。



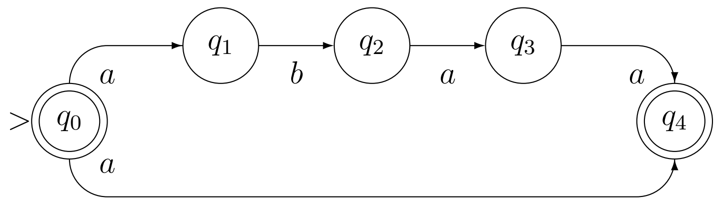
$$\begin{aligned} M_0 &= (\{q_0, q_3, q_4, q_5\}, \{a, b\}, \Delta_0, q_0, \{q_4\}) \\ \Delta_0 &= \{(q_0, aba, q_3), (q_0, \Lambda, q_5), (q_5, \Lambda, q_3), (q_3, a, q_4), (q_5, \Lambda, q_4)\} \end{aligned}$$

M_0 は $\{\Lambda, a, abaa\}$ の有限な言語を受理する。

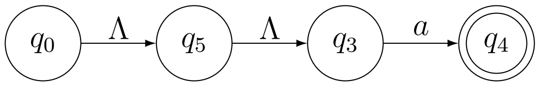
- M_0 の記号列遷移を 1 記号遷移に変換すれば次の NFA M_1 が得られる。



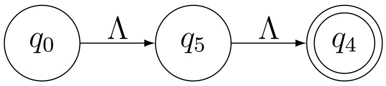
- M_1 の空遷移を除去すると次の NFA M_2 が得られる。



M_1 の

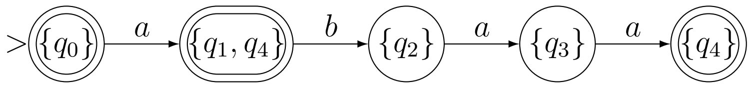


という状態遷移が M_2 の状態遷移 (q_0, a, q_4) に寄与している。また次のように



q_0 から q_4 に空遷移で到達できるため q_0 が終了状態に追加される。

- 部分集合構成により DFA M_3 は次のように構成される。



- この DFA の状態遷移表は次のようになる。

	a	b
$\{q_0\}$	$\{q_1, q_4\}$	
$\{q_1, q_4\}$		$\{q_2\}$
$\{q_2\}$	$\{q_3\}$	
$\{q_3\}$	$\{q_4\}$	
$\{q_4\}$		

考察

DFA の状態遷移表の空欄は空集合で埋められていると考えられる。さらに空集合も DFA の状態の一つと考えると DFA は未定義遷移のない完全 FSA になる。ただし空集合に遷移するような記号列は決して受理されることがない。空集合には元の終了状態を含まないからである。