帰納推論による出力付き有限オートマトンの生成

大泉しげ美　　裴 東善　　榊原康文　　横森 貴
（富士通株式会社）

1. はじめに

字句解析や順序回路の記述など、有限オートマトンは様々な分野で利用されている。有限オートマトンには、ある入力ストリングが受理されるか否かのみを問題にするものの他に、その入力に対する出力を持つ出力付きの有限オートマトンなどがあり、それらは用途によって使い分けられる。

この有限オートマトンを生成する方法として、例からの帰納的推論が考えられる。中でも Angluinによるアルゴリズム〔1〕は、有限個の具体例から状態数最少の決定性有限オートマトンが多項式時間内で得られる点で優れている。しかし出力付き有限オートマトンを得たい場合にはそのアルゴリズムはそのままでは使えない。そこでアルゴリズムを拡張して、入力例とそれに対応する出力例を与えることで、出力付きの有限オートマトンを帰納的に推論する方法について検討する。

2. 出力付き有限オートマトン

出力付き有限オートマトンには Moore型と Mealy型がある。 Moore型は各状態が出力を持つもので、その状態に到達したときに出力が出される。一方 Mealy型の場合は、出力は状態遷移に付随して出される。

Moore型と Mealy型は互いに変換可能である。また Mealy型は Moore型に比べて少ない状態数で記述できることがある。

Moore型に関しては、通常の有限オートマトンで、受理状態を出力１、非受理状態を出力０と考えると、Angluinのアルゴリズムを容易に拡張できる。しかし Mealy型の方が状態数が少なくて済むというメリットがあるため、本稿では Mealy型への拡張について述べる。

Mealy型有限オートマトンは（Q，Σ，Δ，δ，λ，q_0）で定義される。ここでQは状態の有限集合、q_0はQの要素で初期状態、Σは入力記号、Δは出力記号、δはQ×ΣからQへの写像で遷移関数、またλはQ×ΣからΔへの写像で各状態での出力を定める出力関数である。

3. Angluinによる推論方法

Angluinのアルゴリズムでは、ユーザとの対話を繰り返すことにより学習を進めていく。対話の結果、ある学習状況に到達した時点で、システムは推測として有限オートマトンを出力する。推測が目的のものであれば、ユーザはYES を返答しアルゴリズムを終了する。もし間違っていれば反例〔１〕を与えて、さらに推論を進めていく。

この方法ではユーザには以下の責任が課される。

①ある入力記号列が受理されるか否かの判断。

②推測が目的のものか否かの判断。否のときはその反例。

4. Mealy型出力付きオートマトンの推論アルゴリズム

Angluinの方法では、内部に観測表を持っており、その観測表を組み立てることで推論が行われる。推測は観測表が閉じて、無矛盾になったときに出される。

この観測表を拡張することで、 Mealy型を推論することができる。以下に定義とアルゴリズムを示す。

観測表

観測表は$(S,E\cdot A,T)$ で構成される。 Aは入力記号である。 Sは接頭語をとる演算に関して閉じている入力記号列の空でない有限集合、 Eは接尾語をとる演算に関して閉じている入力記号列の空でない有限集合であり、$E\cdot A$ は Eと Aとの連接である。また Tは $((S\cup S\cdot A)\cdot(E\cdot A))$ からΔへの有限関数である。この観測表は図1.のような形で表される。

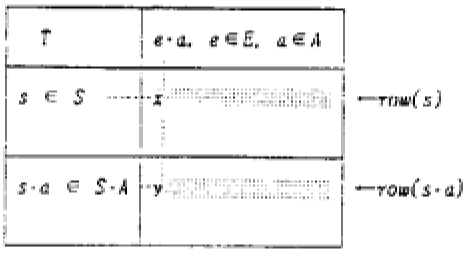

図1.　観測表

図2.　入出力関係

この表の解釈は、記号列$s\cdot e\cdot a$ が入力されたとき出力される記号列のうち、最後に出力される記号がx である場合、$T(s\cdot e\cdot a) = x$である。（図2.）

また、$s \in S\cup S\cdot A$ のとき，$row(s)$は$f(e\cdot a) = T(s\cdot e\cdot a)$ で定義される$E\cdot A$ からΔへの有限関数f を表す。

閉じた、無矛盾な観測表

$S\cdot A$ の任意の要素t に対し、$row(t) = row(s)$ なるSの要素 sが存在するとき、その観測表は閉じている

Learning Finite Automata with Outputs from Examples
Shigemi OOIZUMI, Touzen HAI, Yasubumi SAKAKIBARA, Takashi YOKOMORI
FUJITSU LIMITED

という。また、S の任意の要素$s1, s2$ が$row(s1) = row(s2)$ を満たし、A の任意の要素a に対し$row(s1 \cdot a) = row(s2 \cdot a)$ であるとき、その観測表は無矛盾であるという。

$(S, E \cdot A, T)$ が閉じた、無矛盾な観測表であるとき、それに対応するMealy 型出力付き有限オートマトン $Me(S, E \cdot A, T)$ が以下のように定義できる。

状態集合　$Q = \{ row(s) : s \in S \}$
初期状態　$q0 = \{ row(nl) \}$ （nlは空列）
出力関数　$\lambda(row(s), a) = T(s \cdot nl \cdot a) = T(s \cdot a)$
　　　　　$(s \in S, \ nl \cdot a \in E \cdot A)$
遷移関数　$\delta(row(s), a) = row(s \cdot a)$

アルゴリズム

Sと Eを $\{nl\}$ に初期化する。各 $a \in A$ について、$S \cdot (E \cdot A)$ と $S \cdot A \cdot (E \cdot A)$に対する出力記号を質問して最初の観測表を組み立てる。

```
REPEAT:
  WHILE (S,E·A,T)が閉じていない、または無矛盾でない DO
    IF 無矛盾でない THEN
      row(s1) = row(s2) かつ
      T(s1·a1·e·a2) ≠ T(s2·a1·e·a2)であるような
       S中のs1とs2, a1,a2∈A, e∈E を見つける。
      すべての a∈A について a1·e·a をE·A に加える。
      各出力記号を質問して、Tを ((S∪S·A)·(E·A))
      に拡張する。
    IF 閉じていない THEN
      row(s1·a) が全ての s∈S に対するrow(s)と異
      なるようなs1∈S と a∈A を見つける。
      s1·aをS に加える。
      各出力記号を質問して、T を ((S∪S·A)·(E·A))
      に拡張する。
  観測表が閉じて無矛盾になったら、Me = (S,E·A,T)
  とし推測Meを作る。
  IF ユーザが反例t (= s·a , s∈S, a∈A ) を返答
  する THEN
     sとそのすべての接頭語を Sに加える。
     そして Tを ((S∪S·A)·(E·A)) に拡張する。
UNTIL ユーザが推測Meに対して YESと返答する。
停止そしてMeを出力。
```

ここで反例とは、ある入力記号列 tに対してユーザが考えている出力記号列と、推測された有限オートマトンが出力する出力記号列とが異なっているような t をいう。

ここではユーザには以下の責任が課される。
①ある入力記号列に対する最後の出力記号を答える。
②推測が目的のものか否かの判断。否のときはその反例。

5. 実行例

０と１の列で最後に同じ記号が２つ続くものの全体を考える。この言語は状態数が５未満の有限オートマトンでは表現できない。しかし Mealy型オートマトンなら状態数３で構成できる。〔２〕

入力記号を 0と 1、出力記号を Yと Nとする。まず最初の観測表を組み立てる。（図３.）

しかしこれは閉じていないので記号列 0と 1を Sに移動して、観測表を拡張する。（図４.）この観測表は閉じて無矛盾であるので、システムは推測Meを作り出力する。（図５.）

T	E·A	nl·0	nl·1
S	nl	N	N
S·A	0	Y	N
	1	N	Y

図３.
初期の観測表

T	E·A	nl·0	nl·1	
S	nl	N	N	(q0)
	0	Y	N	(q1)
	1	N	Y	(q2)
S·A	00	Y	N	
	01	N	Y	
	10	Y	N	
	11	N	Y	

図４.
閉じて無矛盾な観測表

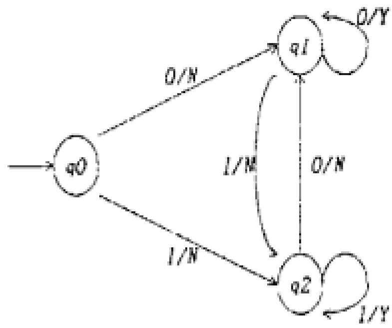

図５.　推測されたMealy 型オートマトン

6. まとめ

Mealy 型の出力付き有限オートマトンを具体例から帰納的に推論する方法について示した。状態遷移を記述するような場合に入力列に対する出力が分かっていれば、その出力付きオートマトンを生成できる。

なお、本研究は新世代コンピュータの開発の一環としてICOTの委託によって行ったものである。

参考文献

〔１〕Angluin,D. "Learning regular sets from queries and counter-examples", Inform. and Computation 75, 87-106, 1987
〔２〕J.Hopcroft, J.Ullman "オートマトン　言語理論　計算論Ｉ", サイエンス社, 1984