KAIST 전산학과 최광무

(개요) FA와 받아들이는(accept) 언어를 나타내는(denote) 정규식을 구하는 연립방정식을 이용한 방법

DFA가 받아들이는 언어와 같은 regular expression을 구하는 이론적이고 어려운(따라서 좀 느린) 방법을 앞에서 recursion을 이용하여 구해보았다. 좀 더 분석적이고 간명한 방법을 생각하여 보자.

확장된 FA[1]를 생각하자, 상태 q에서 문자열 $x_1, x_2, \cdots, x_n (1 \le \forall i \le n{:}\ x_i \in \Sigma^*)$을 보고 상태 $p_1, p_2, \cdots, p_n$으로 상태변환을 한다면 이것은

$$q \ = \ x_1 p_1 + x_2 p_2 + \ \cdots \ + x_n p_n$$

으로 표시할 수 있다. 상태 q가 끝나는 상태 F에 속한다면 $q \ = \ \epsilon$이 추가(+)된다.

$$q \ = \ x_1 p_1 + x_2 p_2 + \ \cdots \ + x_n p_n + \epsilon \qquad \text{if } q \ \in \ F,$$
$$q \ = \ x_1 p_1 + x_2 p_2 + \ \cdots \ + x_n p_n \qquad \text{otherwise}(q \ \not\in \ F).$$

만일 상태가 n개 있다면, 상태에 관한 등식 n개를 가질 수 있다. 상태를 미지수로 보고 상태의 1차 상수인 x_i와 y_i을 기지수로 본다면 이는 미지수 n개에 대한 1차 n원 연립방정식으로 볼 수 있다.

$$1 \le \forall i \le n{:}\ q_i \ = \ x_1 q_1 + x_2 q_2 + \ \cdots \ + x_n q_n + y_i$$
$$\textbf{if } \ q_i \in F \to \ y_i = \epsilon \ \mid \ q_i \not\in F \to \ y_i = \varnothing \ \ \textbf{fi}$$

이 n차 연립방정식을 초기상태에 관하여 풀면 우리는 오토마타가 받아들이는 언어를 표현하는 regular expression을 구할 수 있다[2].

연립방정식을 푸는 방법을 생각해 보자. 연립방정식을 푸는 기본적인 해법은 미지수(변수; 문자)를 차례로 제거하는 방법[3]이다. 그런데 등식의 좌변에 있는 변수가 우변에도 나타났을 때[4] 우변에 있는 변수를 좌변으로 이항한 후 나누어서 해결한다.

$$q_i \ = \ x_i q_i + x_{i+1} q_{i+1} + \ \cdots \ + x_n q_n + y_i{}^{5)} \qquad (1)$$

$$q_i \ = \ \frac{1}{1 - x_i} (x_{i+1} q_{i+1} + \ \cdots \ + x_n q_n + y_i) \qquad (2)$$

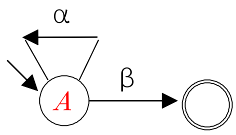
그러나 regular repression은 **빼기**(이항)와 **나누기** 연산은 정의되어 있지 않으므로 (2)의 해법은 답이 아니다. 식 (1)을 아래 (3)과 같이 간단히 써보자.

$$A \ = \ \alpha A \ + \ \beta^{6)} \qquad (3)$$

---

1) $\delta : Q \times \Sigma^* \to 2^Q$
2) 최종상태 F는 ϵ으로 바뀌었음에 주의하라.
3) 대수학에서 Gauss-Jordan elimination을 배웠을 것이다.
4) Gauss-Jordan에서 non-zero diagonal element.
5) q_i보다 작은 변수들($q_1, \cdots, q_{i-1}$)은 Gauss-Jordan에 의하여 이미 제거되었다.
6) 식 (1)에서 q_i는 A로, x_i는 α로, $x_{i+1} q_{i+1} + \ \cdots \ + x_n q_n + y_i$는 β로 각각 바꾸었다.

(3)식을 오토마타로 표현하여 보면,



이때 regular expression (3)의 해는

$$A = \alpha^* \beta \qquad (4)$$

이다. 아래에 두 번 틀려서 맞은(?), 재미있는(?) 증명 방법을 생각해보자.

(이상한 증명) $A = \alpha A + \beta$

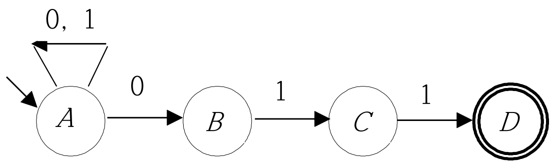
$\neq$ $A = \dfrac{1}{1-\alpha}\beta$ 사칙연산의 경우

$=$ $A = (1 + \alpha + \alpha^2 + \cdots)\beta$ Taylor expansion

$\neq$ $A = \alpha^* \beta$ *의 정의

위의 (**이상한 증명**)은 수학의 유연성을 보여주는 재미있는 예[7]다.

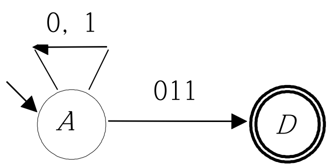
(예 1) $(0+1)^*011$를 받아들이는 NFA ⇒ RE



$A = (0+1)A + 0B$ ①
$B = 1C$ ②
$C = 1D$ ③
$D = \epsilon$ ④

식 ④에 D를 식 ③에 대입 $C = 1$ ⑤
식 ⑤에 C를 식 ②에 대입 $B = 11$ ⑥
식 ⑥에 B를 식 ①에 대입 $A = (0+1)A + 011$ ⑦
식 ⑦에 공식 (3), (4) 적용하여 A를 소거 $A = (0+1)^*011$ ⑧
(예 1)은 우리가 구성한 NFA가 맞는다는 증명이 되기도 한다.

(예 2) $(0+1)^*011$를 받아들이는 XFA ⇒ RE



7) 두 번 틀리면($\neq$) 맞는가?

$A \;=\; (0+1)A + 011D$ ①

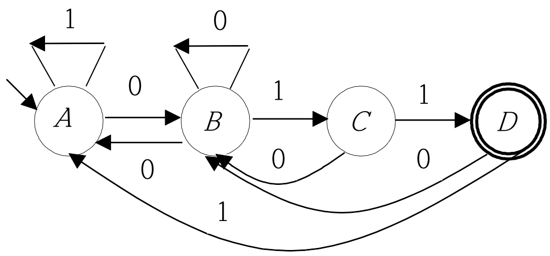
$D \;=\; \epsilon$ ②

식 ②에 D를 식 ①에 대입 $A \;=\; (0+1)A + 011$ ③

식 ③에 공식 (4) 적용하여 A를 소거 $A \;=\; (0+1)^{*}011$ ④

(예 2)는 확장한 XFA가 (예 1)의 NFA와 같다는 증명이 되기도 한다.

(예 3-1) $(0+1)^{*}011$를 받아들이는 DFA ⇒ RE_1



$A \;=\; 0B + 1A$ ①

$B \;=\; 0B + 1C$ ②

$C \;=\; 0B + 1D$ ③

$D \;=\; 0B + 1A + \epsilon$ ④

식 ④에 D를 식 ③에 대입 $C \;=\; 0B + 1(0B + 1A + \epsilon) \;=\; 0B + 10B + 11A + 1$

$= (0+10)B + (11A + 1)$ ⑤

식 ⑤에 C를 식 ②에 대입 $B \;=\; 0B \;+\; 1[(0+10)B + (11A+1)]$

$= (0+10+110)B \;+\; (111A + 11)$ ⑥

식 ⑥에 공식 (3), (4) 적용 $B \;=\; (0+10+110)^{*}(111A + 11)$ ⑦

식 ⑦에 B를 식 ①에 대입 $A \;=\; 0(0+10+110)^{*}(111A + 11) \;+\; 1A$

$= 0(0+10+110)^{*}111A \;+\; 1A \;+\; 0(0+10+110)^{*}11$

$= [0(0+10+110)^{*}111+1]A \;+\; 0(0+10+110)^{*}11$

$= [0(0+10+110)^{*}111+1]^{*} \cdot 0(0+10+110)^{*}11$

$= [0(0+10+110)^{*}111+1]^{*}0(0+10+110)^{*}11$

$= \mathrm{RE}_1$ ⑧

(예 3-2) $(0+1)^{*}011$를 받아들이는 DFA ⇒ RE_2

$D \;=\; \underline{0B+1A} \;+\; \epsilon$ ④

식 ①에 $\underline{A} \;=\; \underline{0B+1A}$를 식 ④에 대입 $D \;=\; \underline{A} \;+\; \epsilon$ ⑨

식 ⑨에 D를 식 ③에 대입 $C \;=\; \underline{0B+1A} \;+\; 1$ ⑩

식 ①에 $\underline{A} \;=\; \underline{0B+1A}$를 식 ⑩에 대입 $= \underline{A} \;+\; 1$ ⑪

식 ⑪에 C를 식을 ②에 대입 $B \;=\; \underline{0B+1A} \;+\; 11$ ⑫

식 ①에 $\underline{A} \;=\; \underline{0B+1A}$를 식 ⑫에 대입 $= \underline{A} \;+\; 11$ ⑬

식 ⑬에 B를 식 ①에 대입 $A \;=\; 0A + 011 + 1A \;=\; (0+1)A + 011$

식 ⑩에 공식 (3), (4) 적용 $= (0+1)^{*}011 \;=\; \mathrm{RE}_2$ ⑭

(예 3)과 (예 4)는 식 ⑧에 $\mathrm{RE}_1 \;=\; [0(0+10+110)^{*}111+1]^{*}0(0+10+110)^{*}11$과
식 ⑭에 $\mathrm{RE}_2 \;=\; (0+1)^{*}011$이 서로 같다는 증명이기도 하다.