KAIST 전산학과 최광무

수학기초

(정의) **Equivalent** relation $R \subseteq A \times A$이 정의한 equivalent class[1]의 집합이 **Partition**,

$$Par_R(A) = \{[a]_R \mid a \in A\}$$이고,

이 Partition의 크기[2] $|Par_R(A)|$를 equivalent relation R의 **index**라 부른다.

(정리) 두 개의 equivalent relation $R_1, R_2 \subseteq A \times A$이 $R_1 \subseteq R_2$이면 equivalent class

$$[a]_{R_1} \subseteq [a]_{R_2}$$이고,

equivalent relation R의 index가 $|Par_{R_1}(A)| \geq$[3] $|Par_{R_2}(A)|$이므로,

equivalent relation R_1이 equivalent relation R_2보다 **더 정교한**[4] partition을 가지고 equivalent relation R_2가 equivalent relation R_1보다 **더 성근** partition을 가진다고 한다.

(예 1) 집합 $A = \{a, b, c, d\}$에 관한 두 개의 equivalent relation

$R_1 / id_A = \{(b,c), (c,b)\}$

$R_2 / id_A = \{(a,b), (b,a), (a,c), (c,a), (b,c), (c,b)\}$를 생각하자.

$R_1 \subset R_2$이고

$Par_{R_1}(A) = \{\{a\}, \{b,c\}, \{d\}\}$, $Par_{R_2}(A) = \{\{a,b,c\}, \{d\}\}$이며

$\{a\}, \{b,c\} \subset \{a,b,c\}$, $\{d\} = \{d\}$이다.

$|Par_{R_1}(A)| = 3 > |Par_{R_2}(A)| = 2$이다.

(예 2) 두 개의 **극단적인** equivalent relation $id_A, A \times A \subseteq A \times A$을 생각하자. 각각의 equivalent class는

$$[a]_{id_A} = \{a\} \subseteq [a]_{A \times A} = A$$이므로,

equivalent relation R의 index가

$$|Par_{id_A}(A)| = |A| \geq |Par_{A \times A}(A)| = 1$$이다.

가장 작은[5] equivalent relation id_A가 **가장 정교한** partition을, 가장 큰 equivalent relation $A \times A$가 **가장 성근** partition을 가진다.

(정의) Equivalent relation $R \subseteq A \times A$가

$$\forall a, b, c \in A:\ a\,R\,b \Rightarrow ac\,R\,bc$$이면

Equivalent relation R은 **right invariant**하다고 한다.

(정의) 임의의 오토마타 M에 관련된 relation $R_M \subseteq \Sigma^* \times \Sigma^*$을 아래와 같이 정의한다.

$$x, y \in \Sigma^*,\ x\,R_M\,y,\ \text{iff}\ \delta(q_o, x) = \delta(q_0, y).$$

(사실) R_M은 **right invariant** equivalent relation이다.

---

1) equivalent class는 $[a]_R = \{b \in A \mid a\,R\,b\}$로 정의하였다.

2) 또는 equivalent class의 수

3) 부분집합 $\subseteq$의 방향과 부등호 $\geq$의 방향이 반대임에 주의하자.

4) R_1 is a **refinement** of R_2

5) $id_A \subseteq A \times A$의 관점에서(with respect to $\subseteq$) 작다는 뜻이다.

오토마타의 **상태**는 스트링의 **equivalent class**이다. 오토마타의 상태는 이 상태로 오기에 필요한 스트링들의 집합과 동동하다.

(정의) 임의의 언어 L에 관련된 relation $R_L \subseteq \Sigma^* \times \Sigma^*$을 아래와 같이 정의한다.

$$x, y \in \Sigma^*,\ xR_L y,\ \text{iff}\ \ \forall z \in \Sigma^*:\ xz \in L \Leftrightarrow yz \in L$$

(사실) R_L도 **right invariant** equivalent relation이다.

(정리) (Myhill-Nerode Theorem) 다음 세 문장은 같다.

(1) 언어 $L \subseteq \Sigma^*$이 regular하다.

(2) 언어 L이 오토마타 M에 관련된 right invariant한 equivalent relation R_M의 **유한한** equivalent class들의 합이다.

(3) 언어 L에 관련된 R_L의 right invariant한 equivalent relation의 index가 **유한**하다.

(증명)

(1) → (2) 언어 L이 regular하므로 언어 L을 받아들이는 오토마타 M이 존재한다.
오토마타 M과 관련된 relation $R_M \subseteq \Sigma^* \times \Sigma^*$을 보자.
R_M은 =와 관련된 relation이므로 reflexive하고 symmetric하며 transitive하다.
즉 **equivalent**하다. 또한 $x\, R_M\, y$, 즉 $\delta(q_0, x) = \delta(q_0, y)$이면

$\delta(q_0, xz) = \delta(\delta(q_0, x), z) = \delta(\delta(q_0, y), z) = \delta(q_0, yz)$이므로 $xz\, R_M\, yz$이다. 즉 **right invariant**하다. 따라서 언어 L은 right invariant equivalent relation R_M의 **유한한** equivalent class들의 합이다[6].

(2) → (3) 언어 L이 어떤 오토마타 M과 관련된 right invariant equivalent relation R_M의 **유한한** equivalent class들의 합이라고 하자.
R_M이 right invariant하므로 $\forall z \in \Sigma^*:\ x\, R_M\, y \Rightarrow xz\, R_M\, yz$이고,

$xz\, R_M\, yz$는 $\delta(q_0, xz) = \delta(q_0, yz)$이므로 $xz \in L \Leftrightarrow yz \in L$이다. 즉 $xR_L\, y$이다.

정리하면 $x\, R_M\, y \Rightarrow xR_L\, y$이다. 즉 R_L의 equivalent index가 유한하다.
$R_M \subseteq R_L$이므로 오토마타 M과 관련된 equivalent relation R_M이 **더 정교한** partition을 언어 L과 관련된 R_L이 **더 성근**[7] partition을 가진다.

(3) → (1) 언어 L에 관하여 정의된 equivalent relation R_L이 유한한 index를 가진다 하자.
오토마타 M_L를

$$M_L \ = \ (\{[x]_L | x \in \Sigma^*\}, \Sigma, \{\delta([x]_L, a) = [xa]_L | x \in \Sigma^*, a \in \Sigma\}, [\epsilon]_L, \{[x]_L | x \in L\})$$

이라 하자.

---

6) 상태가 유한하므로 최종상태도 유한하다.

7) Regular 언어 L을 받아들이는 DFA의 상태 수의 **최소값**이 $|Par_{R_L}(\Sigma^*)|$이다.

R_L이 **right invariant**하므로 $\forall y \in [x]_L$, $\forall a \in \Sigma$: $ya \in [xa]_L$이고, $\delta([x]_L, a) = [xa]_L$이 적절하다[8].
또한 $\forall x \in \Sigma^*$: $\delta([\epsilon]_L, x) = [x]_L$이고 $[x]_L \in F$, iff $x \in L$이므로 $L = L(M_L)$이다.
즉 언어 L을 받아들이는 임의의 DFA M에 대하여 $R_M \subseteq R_L$이므로, 오토마타 M_L이 $L = L(M)$인 DFA중 **가장 성근** index[9]를 가진다,
즉 state의 수가 **가장 작다(Minimal state DFA)**.

8) Deterministic FA이다.
9) $Par_{R_L}(\Sigma^*) = Par_{R_{M_L}}(\Sigma^*)$