امید یعقوبی
جلسه‌ی 5
گروه دوم حل تمرین

رشته‌های تمایزپذیر (Distinguishable Strings)

می‌گوییم دو رشته‌ی $x,y \in \Sigma^*$ در زبان L تمایزپذیرند، اگر وجود داشته باشد، رشته‌ی $z \in \Sigma^*$ به شکلی که $xz \in L$ و $yz \notin L$ یا $xz \notin L$ و $yz \in L$

مثال: $L=\{0^n1^n : n \geq 0\}$

$x=0$ $\quad$ $y=00$

$z=1 \rightsquigarrow xz=01 \in L$ و $yz=001 \notin L$

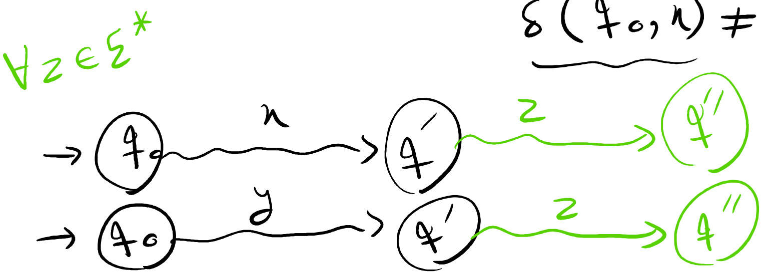
تمایزناپذیر ← $\forall z \in \Sigma^*.\ xz \in L \iff yz \in L$

در زبان بالا $x=10$ و $y=0110$

چون برای هر $z \in \Sigma^*$ ، $yz \notin L$ و $xz \notin L$ پس x و y تمایزناپذیرند.

لم: اگر D یک DFA باشد. و x و y در زبان $L(D)$ تمایزپذیر باشند آنگاه: $\hat{\delta}(q_0,x) \neq \hat{\delta}(q_0,y)$

اثبات.
فرض کنیم این طور نباشد.



اثبات. با برهان خلف اگر نباشد آنگاه $q' = \hat{\delta}(q_0, x) = \hat{\delta}(q_0, y)$

و براساس تعریف تمایزپذیری $\exists z \in \Sigma^* \big((xz \in L \wedge yz \notin L) \vee (xz \notin L \wedge yz \in L) \big)$

① بدون از دست دادن عمومیت (wlog) فرض کنیم

$$\hat{\delta}(q_0, x) = q' \Rightarrow$$

بگذاریم

قضیه‌های جلسات قبل

$$\hat{\delta}(q_0, xz) = \hat{\delta}(\hat{\delta}(q_0, x), z) = \hat{\delta}(q', z) = \hat{\delta}(\hat{\delta}(q_0, y), z)$$

$$\hat{\delta}(\hat{\delta}(q_0, x), z) \in F \Leftrightarrow \hat{\delta}(\hat{\delta}(q_0, y), z) \in F$$

$$\hat{\delta}(q', z) \in F \Leftrightarrow$$

که تناقض است. X

Distinguishing Set

مجموعه‌ی رشته‌های دو به دو قابل تمایز: برای زبان L ، مجموعه‌ی S به شکل $S=\{x_1, x_2, \dots, x_k\}$ یک مجموعه‌ی دو به دو قابل تمایز است اگر برای هر x_i, x_j که $i \neq j$ ، x_i و x_j تمایزپذیر در L باشند.

مثال در زبان $L=\{0^n1^n : n \geq 0\}$

$S_1=\{0, 00, 000\}$

$0011 \in L$

$00011 \notin L$

لم: برای زبان منظم L ، برای هر DFA با نام $D=(Q, \Sigma, \delta, q_0, F)$ که $L(D)=L$ و هر مجموعه‌ی رشته‌های دو به دو قابل تمایز با نام S داریم:

$$|S| \leq |Q|$$

اثبات. فرض کنیم $S=\{x_1, \dots x_k\}$ و D یک DFA باشد. برای هر $x_i \in S$ و $j \neq i$ ، $1 \leq j \leq k$ بر اساس لم قبل $\hat{\delta}(q_0, x_i) \neq \hat{\delta}(q_0, x_j)$ پس $|S| \leq |Q|$. ∎

نتیجه : $L=\{0^n1^n : n\geq 0\}$ منظم نیست.

اثبات. فرض کنیم منظم باشد. پس برای آن DFA با نام D وجود دارد که $L(D)=D$. بگذاریم تعداد حالت های D برابر K باشد. مجموعه ی $S_K=\{0,00,\dots,0^{K+1}\}$ را در نظر بگیرید. برای هر $x_i=0^i$ که $1\leq i\leq K+1$ بگذاریم $z=1^i$، پس $x_iz\in L$ $(0^i1^i\in L)$ ولی برای هر $1\leq j\leq K+1$ که $j\neq i$ $x_jz\notin L$ $(0^i1^j, i\neq j)$. پس S_K یک مجموعه‌ی دو به دو قابل تمایز است. طبق لم قبل هر DFA با نام D' که $L(D')=L$ باید دست کم $K+1$ حالت داشته باشد که تناقض است. ✗

Notation : اگر x و y در زبان L تمایز ناپذیر باشند می‌نویسیم $x\approx_L y$

$$x\approx_L y \quad \text{iff} \quad \forall z\in\Sigma^*\left(xz\in L \iff yz\in L\right)$$

Fact : $\approx_L$ هم‌ارزی است.

Reflexive ⟵ از تعریف.

Symmetric ⟵ از تعریف.

هدف

Transitive: $x \approx_L y \wedge y \approx_L z \Rightarrow x \approx_L z$

① $\forall w \quad xw \in L \iff yw \in L$

② $\forall w \quad yw \in L \iff zw \in L$

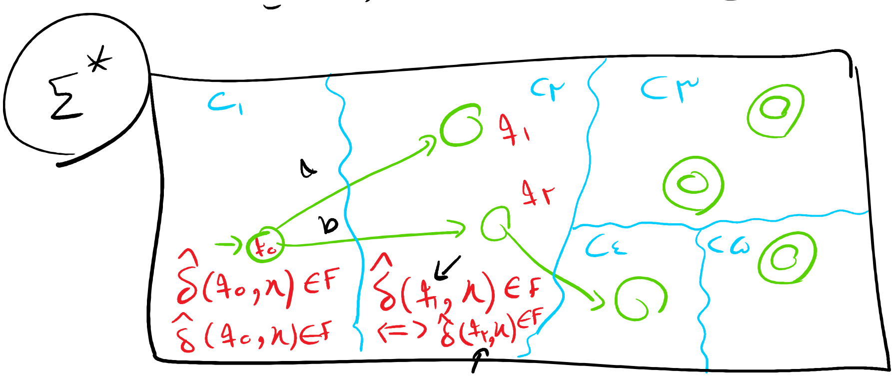
③ برای w دلخواه فرض کنیم $xw \in L$ باشد. آنگاه طبق ① $yw \in L$ و طبق ② $zw \in L$. چون w دلخواه بود پس:

$$\forall w \quad (xw \in L \iff zw \in L)$$

④ $x \approx_L z$ تعریف

$[x] \rightsquigarrow$ کلاس هم‌ارزی x روی $\approx_L$

* برای هر w که $w \in [x]$، $x \approx_L w$. پس اعضای کلاس هم‌ارزی مشخص C_1 تمایزناپذیر در L هستند. و برای هر $w' \notin [x]$ داریم $x \not\approx_L w'$. یعنی تمام اعضای $[x]$ نسبت به رشته‌هایی که در آن نیستند تمایزپذیرند.



Myhill-Nerode Thm : برای زبان L، اگر تعداد کلاس‌های هم‌ارزی Σ^* نسبت به $\sim_L$ بی‌نهایت باشد، آن‌گاه L منظم نیست.

① در غیر این صورت L دارای یک DFA است که تعداد حالت‌های آن دقیقاً به تعداد کلاس‌های هم‌ارزی Σ^* نسبت به $\sim_L$ است. ②

① ← نتیجه‌ی قضیه‌ها و لم‌های قبل.

② ← فرض کنیم تعداد کلاس‌های هم‌ارزی $\sim_L$ متناهی باشد. DFA با نام M را طوری می‌سازیم که برای هر کلاس دقیقاً یک حالت دارد.

* ← حالت آغازین $q_{[\varepsilon]}$
* ← برای هر $x \in L$ ، $q_{[x]} \in F$
* ←

$$q_{[x]} \xrightarrow{a} q_{[xa]}$$

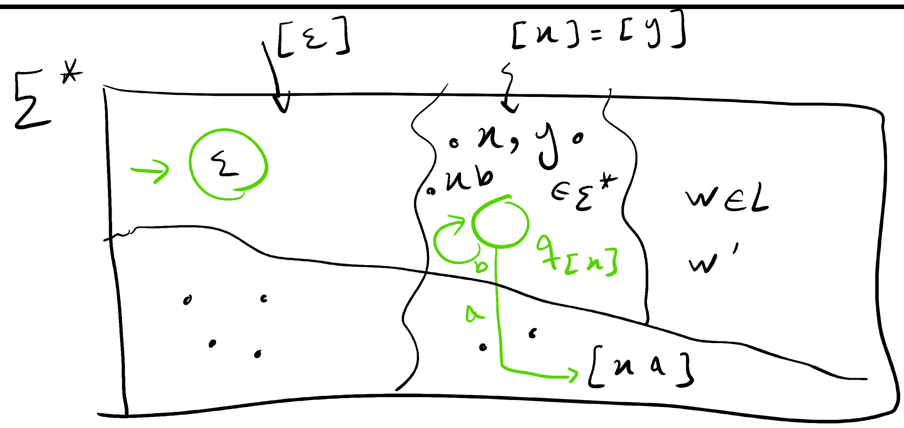
$$\hookrightarrow \delta(q_{[x]}, a) = q_{[xa]}$$

$$[x] = [y] \Longrightarrow q_{[x]} = q_{[y]}$$

اگر $w \in L$ ، $w = w_1 w_2 \cdots w_n$ به شکلی که $w_i \in \Sigma$

$$q_{[\varepsilon]} \xrightarrow{w_1} q_{[w_1]} \xrightarrow{w_2} q_{[w_1 w_2]} \cdots \to q_{[w_1 w_2 \cdots w_n]} \in F$$

* ممکن است $q_{[w'']} = q_{[w']}$ (در بالا)



$q_{[x]} \stackrel{?}{\equiv} q_{[y]}$

$x \approx_L y \Rightarrow q_{[x]} = q_{[y]}$

حالت های هم‌ارز: می‌گوییم دو حالت p و q هم ارزند $p \equiv_M q$

اگر $\forall x \in \Sigma^* \left(\hat{\delta}(q,x) \in F \iff \hat{\delta}(p,x) \in F \right)$

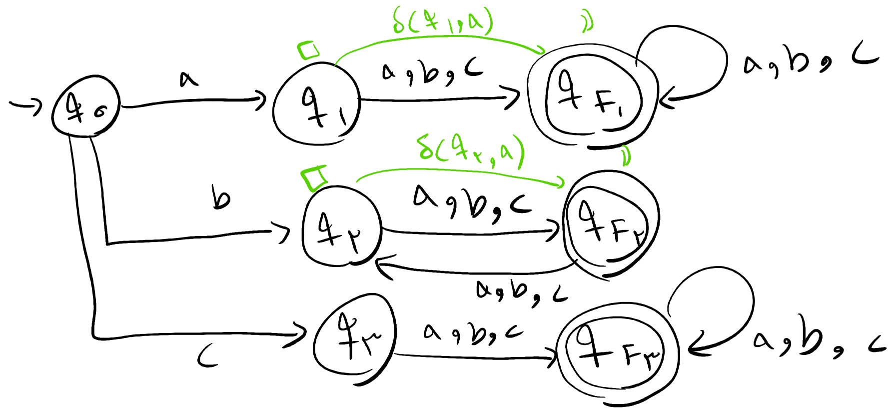
تمرین ساده: نشان دهید $\equiv_M$ هم‌ارزی است.

حالت های هم‌ارز از طول n: می‌گوییم دو حالت p و q هم‌ارز از طول n اند و

می‌نویسیم: $p \equiv_M^n q$ اگر

$$\forall i \in \{0,1,\dots,n\} \ \forall x \in \Sigma^i \left(\hat{\delta}(q,x) \in F \iff \hat{\delta}(p,x) \in F \right)$$

مثال :



$q_1 \equiv_M^1 q_2$ *

$q_1 \not\equiv_M^2 q_2$ *

$q_1 \equiv_M q_3$

$q_{F_1} \equiv_M q_{F_2}$

$\hat{\delta}(q_1, ab) = q_{F_1} \in F$

$\hat{\delta}(q_2, ab) = q_2 \notin F$

* $q_1 \not\equiv_M^2 q_2$ چون $q_1 \equiv_M^1 q_2$ ولی
$\delta(q_1, a) \not\equiv_M^1 \delta(q_2, a)$
$q_{F_1} \not\equiv_M^1 q_{F_2}$ $(z = a)$

* می توان $\equiv_M^n$ را به شکل بازگشتی محاسبه کرد :

(1) Base $(n=0)$: $p \not\equiv_M^0 q$ اگر و تنها اگر $p \in F$ و $q \notin F$
یا $p \notin F$ و $q \in F$

کلاس های هم ارزی $\equiv_M^0$ برابر $C_1 = F$, $C_2 = Q \setminus F$ (set Difference)

(2) محاسبه ی $\equiv_M^n$ از روی $\equiv_M^{n-1}$: [دو شرط زیر اساسی است]

(A) $p \equiv_M^{n-1} q$ و

(B) $\forall a \in \Sigma : \delta(p, a) \equiv_M^{n-1} \delta(q, a)$?

آنگاه : $p \equiv_M^n q$

B :

لم : $p \not\equiv_M^n q$ اگر و تنها اگر $p \not\equiv_M^{n-1} q$ یا وجود داشته باشد $a \in \Sigma$ به شکلی که $\delta(q,a) \not\equiv_M^{n-1} \delta(p,a)$

$\Leftarrow$: اگر $p \not\equiv_M^n q$ پس وجود دارد رشته‌ی x که $|x| \leq n$ و $\hat{\delta}(q,x) \in F$, $\hat{\delta}(p,x) \notin F$ (1) یا $\hat{\delta}(q,x) \notin F$, $\hat{\delta}(p,x) \in F$ ،

wlog ، فرض کنیم (1) ۲ حالت داریم $\underbrace{|x| = n}_{\alpha}$ یا $\underbrace{|x| < n}_{\beta}$

در حالت β که $|x| < n$ به روشنی : $p \not\equiv_M^{n-1} q$

در حالت α که $|x| = n$ بگذاریم $x = a x'$ که a اولین نماد x باشد.

$$\begin{cases} \hat{\delta}(p,x) = \hat{\delta}(\delta(p,a), x') \notin F \\ \hat{\delta}(q,x) = \hat{\delta}(\delta(q,a), x') \in F \end{cases}$$

اگر دقت کنید $|x'| = n-1$. پس اگر $\delta(p,a) = p'$, $\delta(q,a) = q'$ $p' \not\equiv_M^{n-1} q'$.

$\boxtimes$

$\Longrightarrow$

برای طرف دیگر: اگر $p \not\equiv_M^{n-1} q$ به روشنی از تعریف داریم $p \not\equiv_M^n q$

در غیر این صورت وجود دارد $a \in \Sigma$ که

$$\delta(q,a) \not\equiv_M^{n-1} \delta(p,a)$$

پس وجود دارد x' که $|x'| \leqslant n-1$ و $\hat{\delta}(\delta(q,a),x') \in F$

$\hat{\delta}(\delta(p,a),x') \notin F$

یا برعکس

حال بگذاریم $x = ax'$، داریم: $\hat{\delta}(q,ax') \in F$

و $\hat{\delta}(p,ax') \notin F$

یا برعکس

که چون $|x| = n$ پس طبق تعریف $p \not\equiv_M^n q$

[x شاهدی برای $p \not\equiv_M^n q$ است]

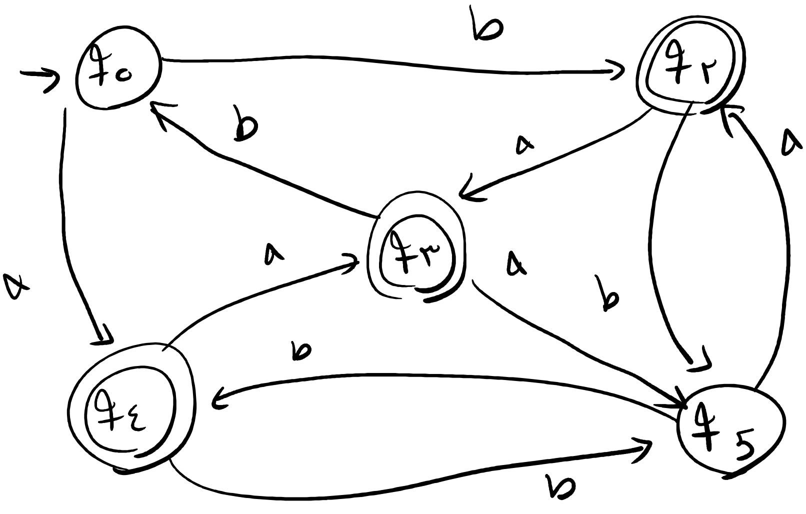
مشاهده: اگر وجود داشته باشد k به شکلی که $\equiv_M^k = \equiv_M^{k+1}$ آن‌گاه برای هر $n \geqslant k$ داریم $\equiv_M^k = \equiv_M^n$

محاسبه‌ی $\equiv_M^{k+1}$ از $\equiv_M^k$ مستقل از مقدار k است.

نتیجه: از آن جا که تعداد کلاس های هم ارزی برای DFA با نام D
دست بالا $|Q|$ است، تعداد مراحل الگوریتم از $|Q|$
کمتر مساوی خواهد بود.

note: با ۲ کلاس هم ارزی شروع می کنیم و در هر مرحله تعداد کلاس ها
بیشتر می شوند یا تغییر نمی کنند و نهایتاً به $|Q|$ کلاس هم ارزی
می رسیم. [اگر تعداد کلاس ها در یک مرحله نسبت به قبل تغییر
الگوریتم می تواند پایان یابد.]

مثال



$\equiv^0_M$:

$F = \{q_4, q_3, q_2\}$ ①

$Q \setminus F = \{q_0, q_5\}$ ②

①

$\delta(q_0, a) \in F$, $\delta(q_5, a) \in F$

$\delta(q_0, b) \in F$, $\delta(q_5, b) \in F$

$\equiv^1_M$	a	b
q_0	$q_4 \in F$	$q_2 \in F$
q_5	$q_2 \in F$	$q_4 \in F$

$\equiv^1_M$	a	b
q_2	$q_3 \in F$	$q_5 \notin F$
q_3	$q_5 \notin F$	$q_0 \notin F$
q_4	$q_3 \in F$	$q_5 \notin F$

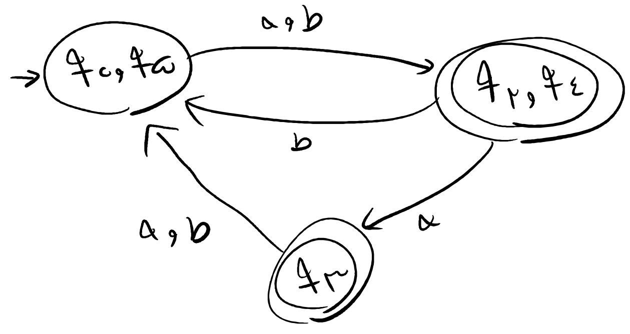
$\not\equiv^1_M$ $\not\equiv$ $\not\equiv^1_M$

$\equiv^1_M$ ✓

① $\{q_0, q_5\}$

② $\{q_3\}$ ⟷ شامل یک عضو است $\equiv$

③ $\{q_2, q_4\}$



Minimal DFA

$\equiv^2_M$	a	b
q_2		
q_4		

$$q_2 \not\equiv^2_M q_4 \begin{cases} q_2 \not\equiv^1_M q_4 \longrightarrow \times \\ \exists a \quad \underbrace{\delta(q_2, a) \not\equiv^1_M \delta(q_4, a)} \end{cases}$$

تناقض است چون برای $a \in \Sigma$ داریم

$$\delta(q_2, a) = \delta(q_4, a)$$

Note: مساوی بودن از $\equiv^n_M$ بسیار قوی‌تر است. پس

$$q_2 \equiv^2_M q_4 \xRightarrow{\text{از قبل}} q_2 \equiv_M q_4$$

گرامر: (V, Σ, R, S)

V ← مجموعه‌ی متغیرها

Σ ← الفبا (به اعضای آن اصطلاحاً ترمینال می‌گوییم)

$R \subseteq V \times (V \cup \Sigma)^*$ ← قوانین تولید

مثال $\underbrace{\overset{\in V}{A} \longrightarrow \overset{\in (V \cup \Sigma)^*}{AaaA}}$

$S \in V$

مثال:

G_1:

$$\left.\begin{array}{l} S \longrightarrow aSb \\ S \longrightarrow \varepsilon \end{array}\right\} \quad S \longrightarrow aSb \mid \varepsilon$$

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow \overset{\in \Sigma^*}{aabb} \in L$$

اشتقاق
derivation

$$L(G_1) = \{a^n b^n : n \geq 0\}$$

$$L(G) = \{w \in \Sigma^* \mid S \overset{*}{\Longrightarrow} w\}$$

* قدرت توصیفی گرامرها از DFA بیشتر است.

* می‌توانیم با گذاشتن محدودیت روی قوانین تولید قدرت توصیفی آن را با DFAها برابر کنیم.

گرامر اکیدا (Strictly) خطی راست گرامری است که قوانین تولید در آن به شکل زیر هستند:

$$A \to xB \qquad A, B \in V,$$
$$A \to \varepsilon \qquad x \in \Sigma \cup \{\varepsilon\}$$

مشاهده: اشتقاق (derivation) در این نوع گرامر به شکل زیر است:

$$S \Rightarrow x_1 A_1 \Rightarrow x_1 x_2 A_2 \Rightarrow \cdots \Rightarrow x_1 \cdots x_n A_n \Rightarrow x_1 \cdots x_n$$

مثال: $L_2 = \{ a^{2K} \mid K \geq 0 \}$

$$G_2 = S \to aA$$
$$A \to aS$$
$$S \to \varepsilon$$

تمرین: claim $L(G_2) = L_2$,

لم: اگر G گرامر اکیدا خطی راست باشد آنگاه $L(G)$ منظم است.

هیئت: از روی G می‌توان NFA با نام N ساخت که

$$L(N) = L(G)$$

* اگر $(A, xB) \in R$ که $A, B \in V$ و $x \in \Sigma \cup \{\varepsilon\}$ آنگاه

سوال: آیا: $B \overset{?}{\in} \delta(A, x)$

هینت: فکر روی این مساله ممکن است شما را به جواب نزدیک ترکند.

* سوال: حالت های پایانی کدامند؟ آیا به قواعد به شکل (C, ε) مرتبطند؟

Note: بین تعریف گرامر خطی راست و گرامر الگو خطی راست ممکن است تفاوت هایی وجود داشته باشد.

مرحله ی آخر:

اثبات: $L(G) \subseteq L(N)$ و $L(N) \subseteq L(G)$ ،

هینت اثبات: با استقرا.