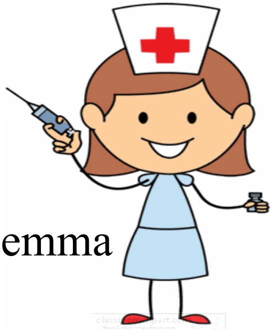
درس مبانی نظریه محاسبه

جلسه هفتم

لم تزریق

The pumping lemma

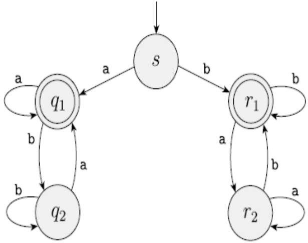


زبانهای منظم، عبارات منظم، ماشینهای متناهی

قضیه: زبان A منظم است اگر و فقط اگر ماشین متناهی M وجود داشته باشد بطوریکه $L(M) = A$

قضیه: زبان A منظم است اگر و فقط اگر عبارت منظم R وجود داشته باشد بطوریکه $L(R) = A$

ماشینهای متناهی $\Leftrightarrow$ زبانهای منظم $\Leftrightarrow$ عبارات منظم

ماشین	زبان	توصیف
	L	$a\Sigma^*a + b\Sigma^*b + a + b$

زبانهای غیر منظم

 علارغم اینکه دانش خوبی نسبت به ساختار زبانهای منظم داریم، متاسفانه این زبانها دایره بسیار محدودی از مجموعه‌ها را شامل می‌شوند. مجموعه‌هایی با الگوهایی بسیار ساده و متداول جزو زبانهای منظم قرار نمی‌گیرند. برای مثال قبلا در این کلاس توضیح داده شد که زبان زیر نمی‌تواند منظم باشد

$$A = \{a^n b^n \mid n \geq 0\}$$

به دلایلی می‌توان ثابت کرد که زبان زیر نیز نمی‌تواند منظم باشد.

$$B = \{w \in \{0,1\}^* \mid n_w(1) = n_w(0)\}$$

در توصیف بالا $n_w(a)$ به معنی تعداد رخداد a در رشته w می‌باشد.

اما چگونه می‌توان ثابت کرد یک زبانی منظم نیست؟

لم تزریق : اگر L یک زبان منظم باشد آنگاه عدد صحیح p وجود دارد بطوریکه هر رشته $w \in L$ که طول حداقل p داشته باشد را می‌توان بصورت $w = xyz$ نوشت بطوریکه رشته‌های x و y و z خواص زیر را دارند:

- برای هر $i \geq 0$ رشته xy^iz هم جزو زبان L است.
- $|y| > 0$. یعنی y رشته تهی نیست.
- $|xy| \leq p$.

توضیح مهم: لم تزریق برای هر زبان منظمی صادق است. اگر شرایط لم تزریق در مورد زبانی صدق نکند می‌توان نتیجه گرفت که زبان مورد نظر منظم نیست. اما عکس لم برقرار نیست. یعنی اگر شرایط لم تزریق در مورد زبان L صدق کند، به این معنا نیست که زبان L حتما منظم است.

بیایید درستی لم تزریق را برای چند زبان منظم چک کنیم.

- $L = \Sigma^* a \Sigma^*$

می‌توانیم p را برابر با 2 انتخاب کنیم.

رشته w در زبان L را در نظر بگیرید که طول حداقل دو دارد. می‌دانیم رشته w حتما شامل کاراکتر a است. رشته w می‌تواند بصورت یکی دو حالت زیر باشد:

(۱) حرف اول رشته a است. در این حالت x را حرف اول رشته، y را حرف دوم رشته و z را بقیه رشته می‌گیریم. می‌بینید که با تکرار y (یا حذف آن) رشته حاصل باز هم در زبان L است.

(۲) حرف اول رشته b است. در این حالت قرار می‌دهیم $x = \epsilon$. رشته y را حرف اول می‌گیریم و z بقیه رشته خواهد بود. می‌بینید که با تکرار y (یا حذف آن) رشته حاصل باز هم در زبان L است.

- $L = a\Sigma^*a + b\Sigma^*b + a + b$

می‌توانیم p را برابر با 3 انتخاب کنیم. رشته w در زبان L را در نظر بگیرید که طول حداقل 3 دارد. در این حالت بصورت زیر عمل می‌کنیم:

x را حرف اول رشته، y را حرف دوم رشته و z را بقیه رشته می‌گیریم. می‌بینید که با تکرار y (یا حذف آن) رشته حاصل باز هم در زبان L است.

- $L = ab + ba + bba + abba$

کافی است p را برابر با 5 انتخاب کنیم. هیچ رشته‌ای در زبان L با طول بیشتر یا مساوی 5 وجود ندارد. پس لم تزریق در مورد این زبان هم برقرار است.

اثبات لم تزریق

اگر L یک زبان متناهی باشد (به تعداد متناهی عضو داشته باشد) آنگاه شرایط لم تزریق بصورت بدیهی در مورد L برقرار است. کافی است عدد p را یکی بیشتر از طول بزرگترین رشته در L قرار دهیم.

$$p = \max\{|w|\} + 1 \text{ where } w \in L$$

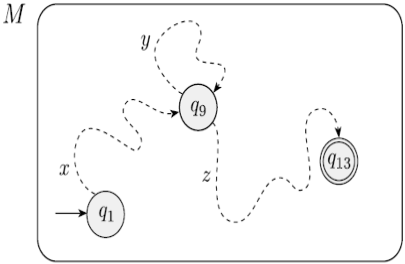
پس می‌توانیم فرض کنیم که L نامتناهی است. اگر زبان L منظم باشد پس ماشین متناهی قطعی $M = (Q, \Sigma, \delta, q_0, F)$ وجود دارد که معادل زبان L است. عدد p را برابر با تعداد وضعیتهای ماشین M قرار می‌دهیم.

$$p = |Q|$$

از آنجا که L به تعداد نامتناهی رشته دارد، پس حتما رشته‌ای به طول حداقل p دارد. فرض کنید w چنین رشته‌ای باشد. فرض کنید لیست زیر دنباله وضعیتهایی باشد که رشته w موقع پردازش در ماشین M آنها را ملاقات می‌کند.

$$S = q_0, \ldots, q_{13}$$

دقت کنید تعداد وضعیتها در دنباله S حداقل $p + 1$ است. (چون طول رشته بیشتر یا مساوی p است.) از آنجا که $|Q| = p$، حتما یکی از وضعیتها در دنباله S تکرار می‌شود (اصل لانه کبوتر). فرض کنید q_9 وضعیتی باشد که تکرار می‌شود. به شکل زیر دقت کنید.



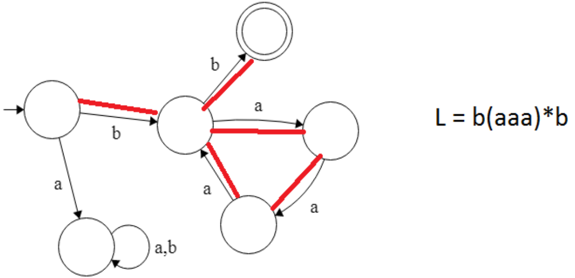
$$S = q_0, \ldots, q_9, \ldots, q_9, \ldots, q_{13}$$

آن قسمت از رشته w که قبل از ورود به q_9 خوانده می‌شود را x می‌نامیم. قسمت داخل دور ایجاد شده را y می‌نامیم و بقیه رشته را z می‌نامیم.

دقت کنید، بنا به تعریف بالا کافی است داشته باشیم $|xy| \leq p$.

علاوه بر این، همه رشته‌هایی مثل xy^iz که از تکرار (تزریق) یا حذف قسمت y بدست می‌آیند، توسط ماشین M پذیرفته می‌شوند. پس شرایط لم تزریق در مورد زبان L صادق است. $\square$

یک مثال



$L = \{a^n b^n \mid n \geq 0\}$ منظم نیست

اثبات با برهان خلف: فرض کنید L منظم باشد. پس شرایط لم تزریق در مورد L باید صادق باشد. حال فرض کنید عدد p در صورت لم را انتخاب کرده‌ایم. از رشته زیر برای ایجاد تناقض استفاده می‌کنیم.

$$w = a^p b^p$$

روشن است که w عضوی از L است. طول w هم حداقل p است.

نشان می‌دهیم x و y و z هر جور انتخاب شوند، تناقض پیش می‌آید. چون $|xy| \leq p$ پس x و y به ناچار باید از قسمت a^p انتخاب شوند. اما تکرار y باعث می‌شود که رشته حاصل از L بیرون بیافتد (به تعداد مساوی a و b نداشته باشد.) تناقض!

$$\overbrace{\underbrace{aaa\ldots aa}_{x}\,\underbrace{aa\ldots a}_{y}}^{p}\,\underbrace{\overbrace{bbb\ldots bbbbbb}^{p}}_{z}$$

منظم نیست $L = \{w \in \{a,b\}^* \mid n_a(w) = n_b(w)\}$

اثبات با برهان خلف: فرض کنید L منظم باشد. پس شرایط لم تزریق در مورد L باید صادق باشد. حال فرض کنید عدد p در صورت لم را انتخاب کرده‌ایم.
مانند اثبات قبل می‌توانیم از رشته زیر برای ایجاد تناقض استفاده می‌کنیم.

$$w = a^p b^p$$

x و y و z هر جور انتخاب شوند، تکرار y باعث می‌شود که رشته حاصل از L بیرون بیافتد (به تعداد مساوی a و b نداشته باشد.) تناقض!

چند نکته در مورد اثبات قبل

- دقت کنید در اثبات قبل، شرط سوم یعنی شرط $|xy| \leq p$ خیلی به ما کمک کرد. اگر این شرط وجود نداشت و می‌توانستیم قسمت y را از هرجای رشته انتخاب کنیم، تناقض مورد نظر ایجاد نمی‌شد. مانند حالت زیر

$$\underbrace{aaa\ldots aa}_{x}\underbrace{\color{red}{ab}}_{y}\underbrace{bbb\ldots bbb}_{z}$$

- انتخاب رشته خوب برای ایجاد تناقض مهم است. برای مثال اگر برای ایجاد تناقض از رشته $w = (ab)^p$ استفاده می‌کردیم با مشکل مواجه می‌شدیم. مانند موقعیت زیر

$$\underbrace{ababab\ldots ab}_{x}\underbrace{ab}_{y}\underbrace{ababab\ldots ab}_{z}$$

منظم نیست $L = \{ww \mid w \in \{a,b\}^*\}$

اثبات با برهان خلف: زبان L حاوی همه رشته‌هایی است که از تکرار یک رشته از الفبای $\{a,b\}$ بدست می‌آید. رشته‌های زیر در زبان L قرار دارند.

$$aaaa, abab, aaabaaab, bbbabbbbab$$

فرض کنید L منظم باشد. پس شرایط لم تزریق در مورد L باید صادق باشد. حال فرض کنید عدد p در صورت لم را انتخاب کرده‌ایم.

چه رشته‌ای را ایجاد تناقض انتخاب کنیم؟

یک انتخاب بد: $w = a^{2p}$

چرا؟

$$\underbrace{aa\ldots a}_{x}\underbrace{aa}_{y}\underbrace{aa\ldots a}_{z}$$

تزریق y تناقضی ایجاد نمی‌کند.

یک انتخاب خوب:

$$w = a^p b^p a^p b^p$$

اگر از شرط سوم $|xy| \leq p$ استفاده کنیم براحتی قابل مشاهده است که تزریق y باعث بهم ریختن نظم رشته می‌شود و آن را از دایره زبان خارج می‌کند.