بسم الله الرحمن الرحیم

نظریه زبان‌ها و ماشین‌ها

جلسه ۱۲

مجتبی خلیلی
دانشکده برق و کامپیوتر
دانشگاه صنعتی اصفهان

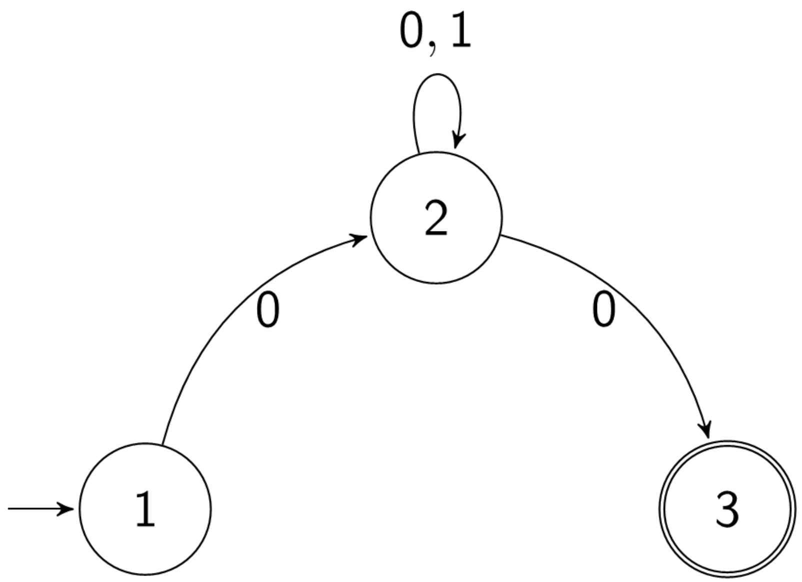
لم تزریق برای زبان‌های منظم

- روشی برای تشخیص زبانهایی که منظم نیستند.

لم تزریق برای زبان‌های منظم

- مثال:

$0(0 \cup 1)^*0$:



لم تزریق برای زبان‌های منظم

THEOREM 1.70

Pumping lemma If A is a regular language, then there is a number p (the pumping length) where if s is any string in A of length at least p, then s may be divided into three pieces, $s = xyz$, satisfying the following conditions:

1. for each $i \geq 0$, $xy^iz \in A$,
2. $|y| > 0$, and
3. $|xy| \leq p$.

لم تزریق برای زبان‌های منظم

- این لم بیان می‌کند همه زبان‌های منظم یک ویژگی خاص (تزریق) دارند.
  - عکس آن لزوما درست نیست؛ داشتن این ویژگی لزوما به معنای منظم بودن زبان نیست.
  - اگر زبانی این ویژگی را نداشت، نامنظم است (اثبات از طریق تناقض).
  - از این لم برای اثبات نامنظم بودن یک زبان استفاده می‌شود.

لم تزریق برای زبان‌های منظم

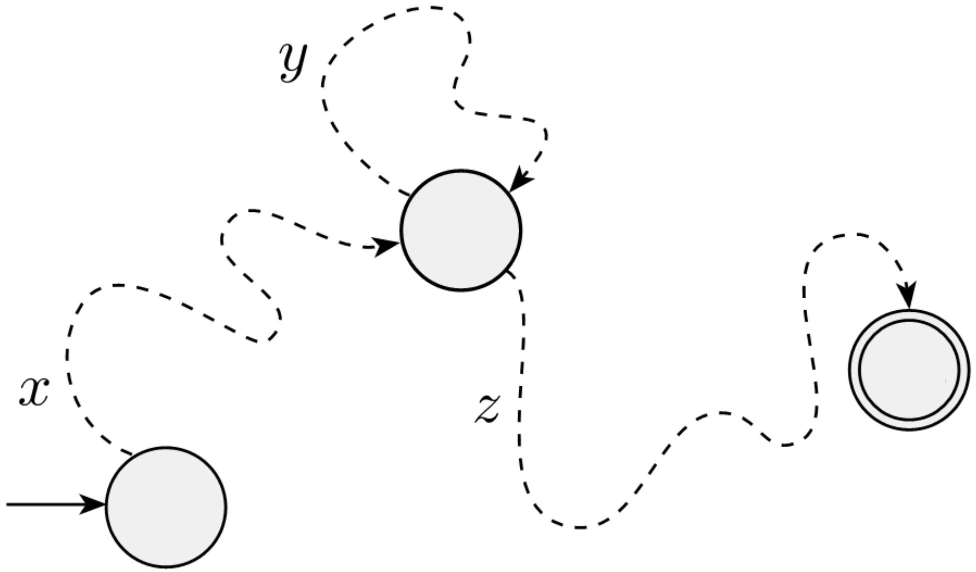
IUT-ECE

- این لم بیان می‌کند دست کم یک طول تزریق وجود دارد که به ازای آن لم برقرار است.
  - چنانچه برای یک مقدار مشخص طول تزریق، لم برقرار نبود نمی‌توان نتیجه گرفت زبان منظم نیست.
- لم باید برای همه رشته‌های با طول دست کم p برقرار باشد.
  - از اینکه برای برخی رشته‌ها برقرار است نمیتوان نتیجه گرفت زبان می‌تواند منظم باشد.
- برای یک رشته معلوم، باید دست کم یک روش ممکن برای جداسازی در xyz باشد چنانچه xy^iz در زبان باشد.
  - از اینکه نتوان با یک روش به چنین جداسازی رسید نمیتوان نتیجه گرفت که زبان نامنظم است.

لم تزریق برای زبان‌های منظم

- هر رشته در یک زبان منظم L، با طول بزرگتر از طول تزریق p، میتواند تزریق شود.

- این بدین معنی است که هر $s \in L$ یک بخش (y) را شامل می‌شود که می‌تواند بارها (به کمک حلقه) تکرار شود (تزریق) و رشته‌های جدید نیز باید در زبان L باشند.



لم تزریق برای زبان‌های منظم

THEOREM 1.70

Pumping lemma If A is a regular language, then there is a number p (the pumping length) where if s is any string in A of length at least p, then s may be divided into three pieces, $s = xyz$, satisfying the following conditions:

1. for each $i \geq 0$, $xy^iz \in A$,
2. $|y| > 0$, and
3. $|xy| \leq p$.

لم تزریق برای زبان‌های منظم

IUT-ECE

When s is divided into xyz, either x or z may be ε, but condition 2 says that $y \neq \varepsilon$. Observe that without condition 2 the theorem would be trivially true. Condition 3 states that the pieces x and y together have length at most p. It is an extra technical condition that we occasionally find useful when proving certain languages to be nonregular. See Example 1.74 for an application of condition 3.

ایده اثبات لم تزریق

PROOF IDEA Let $M = (Q, \Sigma, \delta, q_1, F)$ be a DFA that recognizes A. We assign the pumping length p to be the number of states of M. We show that any string s in A of length at least p may be broken into the three pieces xyz, satisfying our three conditions. What if no strings in A are of length at least p? Then our task is even easier because the theorem becomes *vacuously* true: Obviously the three conditions hold for all strings of length at least p if there aren't any such strings.

ایده اثبات لم تزریق

IUT-ECE

If s in A has length at least p, consider the sequence of states that M goes through when computing with input s.

$$\begin{array}{cccccccccccccccccc}
s = & & s_1 & & s_2 & & s_3 & & s_4 & & s_5 & & s_6 & & \cdots & & s_n & \\
& \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
& q_1 & & q_3 & & q_{20} & & q_9 & & q_{17} & & q_9 & & q_6 & & q_{35} & & q_{13}
\end{array}$$

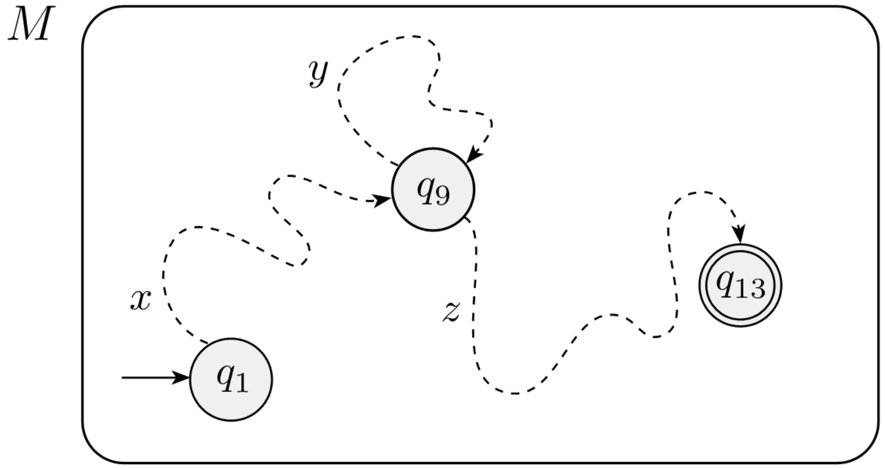
If we let n be the length of s, the sequence of states $q_1, q_3, q_{20}, q_9, \ldots, q_{13}$ has length $n + 1$. Because n is at least p, we know that $n + 1$ is greater than p, the number of states of M. Therefore, the sequence must contain a repeated state.

pigeonhole principle

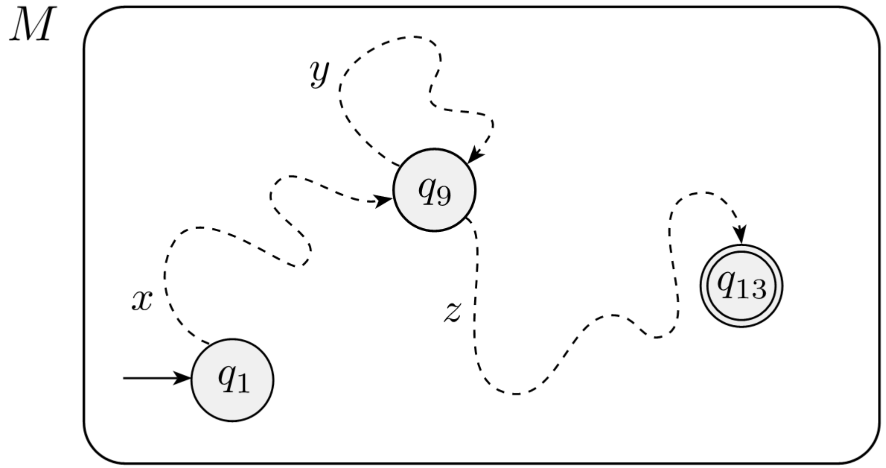
ایده اثبات لم تزریق

We now divide s into the three pieces x, y, and z.

$$
\begin{array}{ccccccccccccccccccccc}
s = & & s_1 & & s_2 & & s_3 & & s_4 & & s_5 & & s_6 & & \cdots & & s_n & \\
& \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
& q_1 & & q_3 & & q_{20} & & q_9 & & q_{17} & & q_9 & & q_6 & & q_{35} & & q_{13}
\end{array}
$$

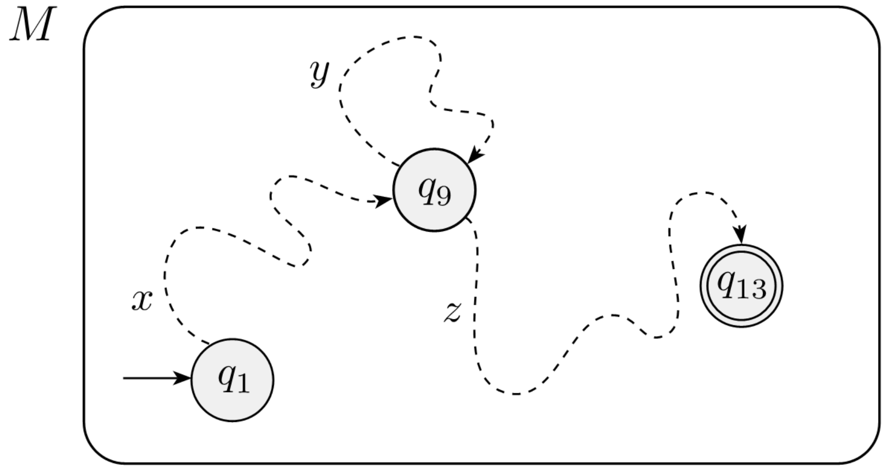


ایده اثبات لم تزریق



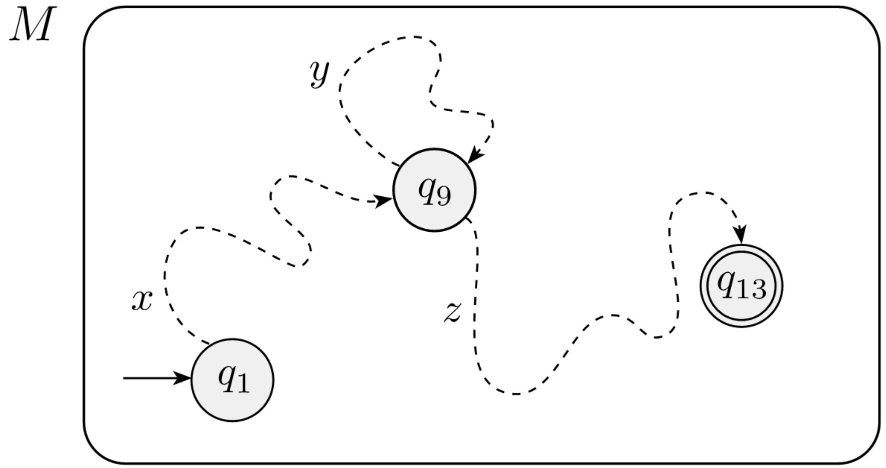
Let's see why this division of s satisfies the three conditions. Suppose that we run M on input $xyyz$. We know that x takes M from q_1 to q_9, and then the first y takes it from q_9 back to q_9, as does the second y, and then z takes it to q_{13}. With q_{13} being an accept state, M accepts input $xyyz$. Similarly, it will accept xy^iz for any $i > 0$. For the case $i = 0$, $xy^iz = xz$, which is accepted for similar reasons. That establishes condition 1.

ایده اثبات لم تزریق



Checking condition 2, we see that $|y| > 0$, as it was the part of s that occurred between two different occurrences of state q_9.

ایده اثبات لم تزریق



In order to get condition 3, we make sure that q_9 is the first repetition in the sequence. By the pigeonhole principle, the first $p+1$ states in the sequence must contain a repetition. Therefore, $|xy| \leq p$.

اثبات لم تزریق

IUT-ECE

PROOF Let $M = (Q, \Sigma, \delta, q_1, F)$ be a DFA recognizing A and p be the number of states of M.

Let $s = s_1 s_2 \cdots s_n$ be a string in A of length n, where $n \geq p$. Let $r_1, \ldots, r_{n+1}$ be the sequence of states that M enters while processing s, so $r_{i+1} = \delta(r_i, s_i)$ for $1 \leq i \leq n$. This sequence has length $n+1$, which is at least $p+1$. Among the first $p+1$ elements in the sequence, two must be the same state, by the pigeonhole principle. We call the first of these r_j and the second r_l. Because r_l occurs among the first $p+1$ places in a sequence starting at r_1, we have $l \leq p+1$. Now let $x = s_1 \cdots s_{j-1}$, $y = s_j \cdots s_{l-1}$, and $z = s_l \cdots s_n$.

As x takes M from r_1 to r_j, y takes M from r_j to r_j, and z takes M from r_j to r_{n+1}, which is an accept state, M must accept $xy^i z$ for $i \geq 0$. We know that $j \neq l$, so $|y| > 0$; and $l \leq p+1$, so $|xy| \leq p$. Thus we have satisfied all conditions of the pumping lemma.

مثال

EXAMPLE 1.73

Let B be the language $\{0^n1^n | n \geq 0\}$. We use the pumping lemma to prove that B is not regular. The proof is by contradiction.

مثال

IUT-ECE

$$B = \{0^n1^n : n \geq 0\}$$

- فرض کنیم منظم است.
- فرض کنید p طول پمپ برای زبان B
- انتخاب $w = 0^p1^p$

$$w = \underbrace{000000000\ldots0}_{p}\underbrace{111111111\ldots1}_{p}$$

مثال

w = xyz, with |y| > 0 and |xy| ≤ p

$$w = \underbrace{000000000\ldots0}_{p}\underbrace{111111111\ldots1}_{p}$$

- داریم

$$w = \underbrace{000000}_{x}\underbrace{000}_{y}\underbrace{\ldots0111111111\ldots1}_{z}$$

- بنابراین به تناقض میرسیم و زبان نامنظم است.

مثال

EXAMPLE 1.74

Let $C = \{w|\ w$ has an equal number of 0s and 1s$\}$. We use the pumping lemma to prove that C is not regular. The proof is by contradiction.

Here condition 3 in the pumping lemma is useful. It stipulates that when pumping s, it must be divided so that $|xy| \leq p$. That restriction on the way that s may be divided makes it easier to show that the string $s = 0^p 1^p$ we selected cannot be pumped. If $|xy| \leq p$, then y must consist only of 0s, so $xyyz \notin C$. Therefore, s cannot be pumped. That gives us the desired contradiction.

EXAMPLE 4.9

Let $\Sigma = \{a, b\}$. The language

$$L = \{w \in \Sigma^* : n_a(w) < n_b(w)\}$$

is not regular.

Suppose we are given m. Since we have complete freedom in choosing w, we pick $w = a^m b^{m+1}$. Now, because $|xy|$ cannot be greater than m, the opponent cannot do anything but pick a y with all a's, that is

$$y = a^k, \qquad 1 \leq k \leq m.$$

We now pump up, using $i = 2$. The resulting string

$$w_2 = a^{m+k} b^{m+1}$$

is not in L. Therefore, the pumping lemma is violated, and L is not regular.

EXAMPLE 1.75

Let $F = \{ww \mid w \in \{0,1\}^*\}$. We show that F is nonregular, using the pumping lemma.

Assume to the contrary that F is regular. Let p be the pumping length given by the pumping lemma. Let s be the string 0^p10^p1. Because s is a member of F and s has length more than p, the pumping lemma guarantees that s can be split into three pieces, $s = xyz$, satisfying the three conditions of the lemma. We show that this outcome is impossible.

Condition 3 is once again crucial because without it we could pump s if we let x and z be the empty string. With condition 3 the proof follows because y must consist only of 0s, so $xyyz \notin F$.

Observe that we chose $s = 0^p10^p1$ to be a string that exhibits the "essence" of the nonregularity of F, as opposed to, say, the string 0^p0^p. Even though 0^p0^p is a member of F, it fails to demonstrate a contradiction because it can be pumped.

EXAMPLE 1.77

Sometimes "pumping down" is useful when we apply the pumping lemma. We use the pumping lemma to show that $E = \{0^i 1^j | \ i > j\}$ is not regular. The proof is by contradiction.

Assume that E is regular. Let p be the pumping length for E given by the pumping lemma. Let $s = 0^{p+1}1^p$. Then s can be split into xyz, satisfying the conditions of the pumping lemma. By condition 3, y consists only of 0s. Let's examine the string $xyyz$ to see whether it can be in E. Adding an extra copy of y increases the number of 0s. But, E contains all strings in 0^*1^* that have more 0s than 1s, so increasing the number of 0s will still give a string in E. No contradiction occurs. We need to try something else.

The pumping lemma states that $xy^iz \in E$ even when $i = 0$, so let's consider the string $xy^0z = xz$. Removing string y decreases the number of 0s in s. Recall that s has just one more 0 than 1. Therefore, xz cannot have more 0s than 1s, so it cannot be a member of E. Thus we obtain a contradiction. ■