زبان های نامنظم

- برای درک قدرت ماشین ها باید محدودیت های آن ها را بشناسیم.
- برخی از زبان ها را نمی توان با ماشین حالت متناهی تشخیص داد.
- مثلاً برای تشخیص زبان زیر باید در حالت ها تعداد 0 ها را حفظ کنیم.

$$B = \{0^n 1^n \mid n \geq 0\}$$

- اما حالت های متناهی یک DFA پاسخگوی تعداد نامتناهی حالات ممکن نیست.
- نیاز به یک مبنای نظری برای اثبات منظم نبودن داریم: *لم پامپینگ*

لم پامپینگ

THEOREM

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:

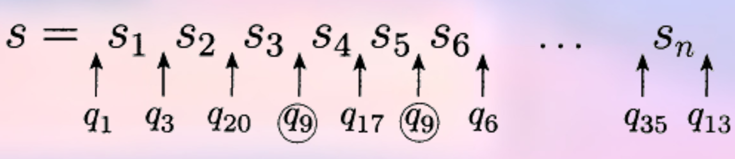
1. for each $i \geq 0$, $xy^iz \in A$,
2. $|y| > 0$, and
3. $|xy| \leq p$.

ایده اثبات

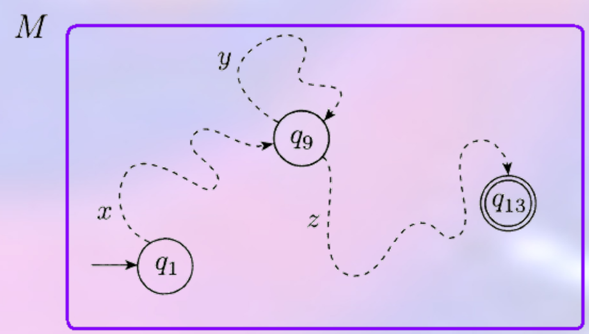
PROOF IDEA

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.

- بنا به اصل لانه کبوتری، برای رشته هایی با بیش از این طول، حداقل یک حالت در دنباله حالات تکرار می شود.



state q_9 repeating when M reads s



اثبات

PROOF

Let $M = (Q, \Sigma, \delta, q_1, F)$ be a DFA recognizing A.

p: the number of states of M.

$s = s_1 s_2 \cdots s_n$ be a string in A of length $n \geq p$.

$r_1, \ldots, r_{n+1}$ the sequence of states that M enters while processing s.

$r_{i+1} = \delta(r_i, s_i) \quad$ for $1 \leq i \leq n$.

pigeonhole principle $\Rightarrow \exists j, l \leq n+1 : r_j = r_l$

Let $x = s_1 \cdots s_{j-1}$, $y = s_j \cdots s_{l-1}$, and $z = s_l \cdots s_n$.

x takes M from r_1 to r_j,
y takes M from r_j to r_j, $\quad \Rightarrow$ $\boxed{M \text{ must accept } xy^i z \text{ for } i \geq 0.}$
z takes M from r_j to r_{n+1},

$j \neq l \Rightarrow \boxed{|y| > 0;}$ and $l \leq p+1$, so $\boxed{|xy| \leq p.}$

مثال

- با استفاده از لم پامپینگ و برهان خلف ثابت می کنیم که زبان زیر منظم نیست:

$B = \{0^n 1^n | n \geq 0\}$.
Let p be the pumping length given by the pumping lemma.
Choose $s = 0^p 1^p$.
pumping lemma => $s = xyz$, where for any $i \geq 0$ the string $xy^i z$ is in B.

1. The string y consists only of 0s.
2. The string y consists only of 1s.
3. The string y consists of both 0s and 1s.

- *چرا هیچ یک از سه حالت فوق نمی تواند درست باشد؟*

مثال

- با استفاده از لم پامپینگ و برهان خلف ثابت می کنیم که زبان زیر منظم نیست:

$C = \{w|\ w$ has an equal number of 0s and 1s$\}$.
Let p be the pumping length given by the pumping lemma.
Choose $s = 0^p 1^p$.
pumping lemma => $s = xyz$, where for any $i \geq 0$ the string xy^iz is in B.
If $|xy| \leq p$, then y must consist only of 0s, so $xyyz \notin C$.

- روش دیگر:

If C were regular, $C \cap 0^*1^*$ also would be regular.
But $C \cap 0^*1^*$ equals B, and we know that B is nonregular.

مثال

زبان $D=\{1^{n^2} \mid n \geq 0\}$. را در نظر بگیرید. می توان نشان داد که این زبان نامنظم است.

اگر p طول پمپ کردن باشد رشته مورد نظر می تواند به صورت $s=xyz=1^{p^2}$ باشد. آنگاه $|xyz| \leq p^2$ و $|y| \leq p$.

آنگاه $|xy^2z| \leq p+p^2$ که به وضوح مطلقا از $(1+p)^2$ کوچکتر است و نمی تواند با آن برابر باشد.

مثال

- زبان کپی $F=\{ww:\ w\in\{0,1\}^*\}$ را در نظر بگیرید. با استفاده از لم پمپینگ می توان نشان داد که این زبان نامنظم است. فرض کنید p طول پمپ کردن باشد.
- اگر رشته مورد نظر را $s=0^p10^p1$ بگیریم جواب خواهد داد.

مثال

- فرض کنید $E = \{0^i1^j: i>j\}$
- با استفاده از لم پمپینگ می توان ثابت کرد که زبان فوق نیز نامنظم است. با در نظر گرفتن طول پمپ برابر p، رشته مورد استفاده می تواند $s=0^{p+1}1^p$ باشد.