L5. 2022-09-16

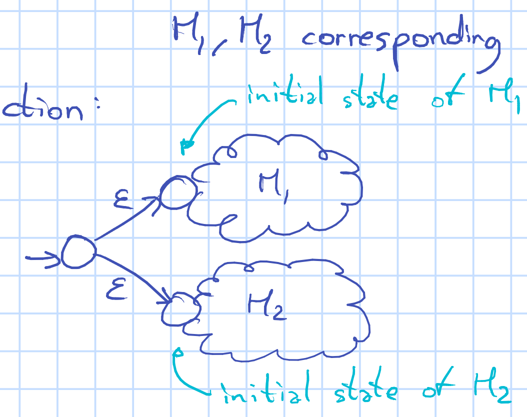
Regular Operations and Languages.

We saw how to build automaton for the complement of a language. Which other operations can we do?

Union: $L_1 \cup L_2 = L_1 | L_2 = \{ s : s \in L_1 \text{ or } s \in L_2 \}$

M_1, M_2 corresponding automata.

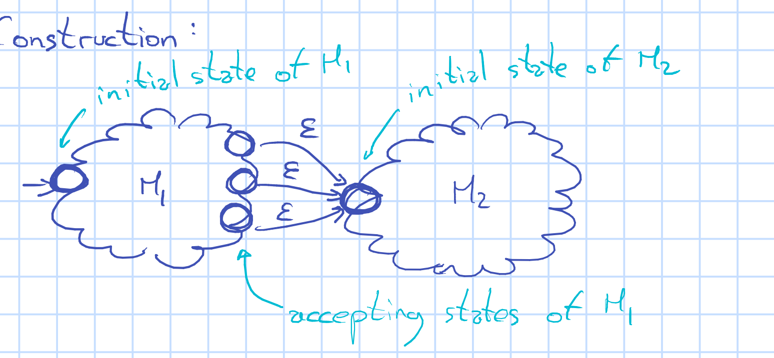
Construction:



* Note: $L_1 \cap L_2 = \overline{(\overline{L_1} \cup \overline{L_2})}$ (de Morgan).

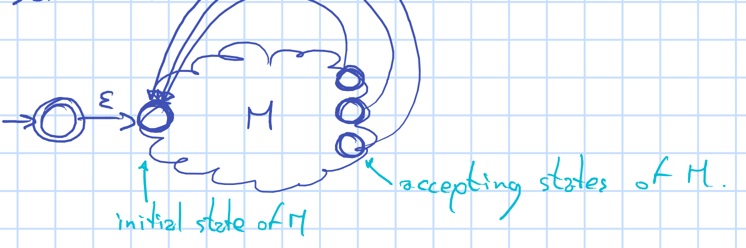
Concatenation: $L_1 L_2 = \{ s_1 s_2 : s_1 \in L_1, s_2 \in L_2 \}$

Construction:

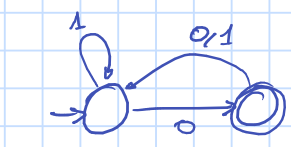


Kleene Star: $L^* = \{ s^k : s \in L, k \in \mathbb{Z}, k \geq 0 \}$
$= \{$any number of repetitions of a string in the language$\}$.

Construction:



We can use these operations to describe languages more compactly.

e.g. Language of 

is

$$((0(0|1))|1)^* 0 \qquad \text{equiv.} \quad (0.|1)^* 0$$

In programming we use "." as a shorthand for $(\cup \Sigma)$ (e.g. $(0|1)$).

These expressions are called regular expressions, and the languages they describe are regular languages.

We have seen every regex can be recognized by an NFA. The converse is also true.

thm: L can be described by a regex iff L can be recognized by an NFA (DFA).

proof: We've seen $\Rightarrow$]. See Sipser 1.3 for $\Leftarrow$].

Examples of Regexes.

$a^* b^*$ contains ε, a, aaa, bb, aab, ...
does not contain ba, ...

$a^* | b^*$ contains ε, a, aaa, bb, ...
does not contain ab, ba, ...

$aba | bab$ contains exactly aba, bab.

$(aaa)^*$ contains ε, aaa, aaaaaa, ...
does not contain a, aa, aaaa, ...

Are all languages regular? No! e.g. palindromes.

To prove that we'll need auxilliary fact:

lem (Pumping) If L regular then $\exists p \in \mathbb{N}$ st if $s \in L$ and $|s| \geq p$, then $s = xyz$ st:

- → $xy^i z \in L \quad \forall i \geq 0$
- → $y \neq \varepsilon$
- → $|xy| \leq p$

proof idea: an automaton is finite, so for large enough strings it will have to repeat states. Let y be the fragment of s we see during such a loop.

Using pumping lemma to show a language is not regular:

$L_1 = \{\text{palindromes}\} = \{ s : s = s^R \}$

Proof: Assume L_1 regular. Then pumping lemma applies.

Let p given by lemma (we don't know its value).

Let $s = 0^p 1 0^p \in L$ (we choose this)

Then $y = 0^j$ (because $|xy| \leq p$, $xy \leq 0^p$)

Then $xy^2 z = 0^{p+i} 1 0^p \notin L$.

Contradiction!! Hence L_1 not regular.

$L_2 = \{ 0^n 1^n \}$. Choose $s = 0^p 1^p$. Then $y = 0^j$.

Then $xy^2 z = 0^{p+i} 1 0^p \notin L$!!

$L_3 = \{ |s_i = 0| = |s_i = 1| \}$. Choose $s = 0^p 1^p$. Then $y = 0^j$.

Then $xy^2 z = 0^{p+i} 1 0^p \notin L$!!

(alt, $L_2 = L_3 \cap (0^* 1^*)$. If L_3 were regular, so would be L_2)

$L_4 = \{ ss \}$. $s = 0^p 1 0^p 1$.

$L_5 = \{ 1^{(n^2)} \}$. $s = 1^{(p^2)}$.

Note $p^2 + i$ not a square for $0 \leq i \leq p$.

$L_6 = \{ 0^i 1^j : i \geq j \}$. $s = 0^p 1^{p-1}$.