From lecture 2:

Theorem

Suppose L is a language over the alphabet Σ. If L is accepted by a finite automaton M, and if n is the number of states of M, then

$\forall$ *for every* $x \in L$
satisfying $|x| \geqslant n$

$\exists$ *there are three strings* u, v, *and* w,
such that $x = uvw$ *and the following three conditions are true:*
(1) $|uv| \leqslant n$,
(2) $|v| \geqslant 1$

$\forall$ *and (3) for all* $m \geqslant 0$, $uv^m w$ *belongs to* L

[M] Thm. 2.29

From lecture 2:

Theorem

If L *is a regular language, then*

$\exists$ *there exists a constant* $n \geqslant 1$

such that

$\forall$ *for every* $x \in L$

with $|x| \geqslant n$

$\exists$ *there exists a decomposition* $x = uvw$

with (1) $|uv| \leqslant n$,

and (2) $|v| \geqslant 1$

such that

$\forall$ *(3) for all* $m \geqslant 0$, $uv^m w \in L$

From lecture 2: To contradict the pumping lemma, we prove the negation:

Theorem

If

$\forall$ *for every* $n \geqslant 1$

$\exists$ *there exists* $x \in L$

with $|x| \geqslant n$

such that

$\forall$ *for every decomposition* $x = uvw$

with (1) $|uv| \leqslant n$,

and (2) $|v| \geqslant 1$

$\exists$ *(3) there exists* $m \geqslant 0$,

such that

$uv^m w \notin L$

then L *is not a regular language.*

Given a language L, to prove L is not a regular language:

1. Opponent picks n.
2. We choose a string $x \in L$ with $|x| \geqslant n$.
3. Opponent picks u, v, w with $x = uvw$, $|uv| \leqslant n$, $|v| \geqslant 1$.
4. If we can find $m \geqslant 0$ such that $uv^m w \notin L$, then we win.

If we can always win, then L does not fulfil the pumping lemma.

≈[VU Automata & Complexity] L3

Example

$L = \{ x \in \{a, b\}^* \mid n_a(x) > n_b(x) \}$ is not accepted by FA

[M] E 2.31

$L \subseteq \{a\}^*$

Example

$L = \{\ a^{i^2} \mid i \geqslant 0\ \}$ is not accepted by FA

$L = \{\Lambda, a, aaaa, aaaaaaaaa, \ldots\}$

[M] E 2.32

Fun fact

$L^4 = \{a\}^*$

Lagrange's four-square theorem

The length of uv^2w cannot be a square: we will show it is strictly in between two consecutive squares.
$|uv^2w| = |z| + |v| > |z| = n^2$.
$|uv^2w| = |z| + |v| \leqslant n^2 + n < (n+1)^2$.

Let L be the set of legal C programs.
$x = \texttt{main}()\{\{\{\ldots\}\}\}$

[M] E 2.33

Prove the following generalization of the pumping lemma, which can sometimes make it unnecessary to break the proof into cases.

If L can be accepted by an FA,
then there is an integer n
such that for any $x \in L$ with $|x| \geqslant n$
and for any way of writing x as $x_1 x_2 x_3$ with $|x_2| = n$,
there are strings u, v and w such that

a. $x_2 = uvw$
b. $|v| \geqslant 1$
c. For every $m \geqslant 0$, $x_1 uv^m wx_3 \in L$

$L = \{ a^i b^j c^j \mid i \geqslant 1 \text{ and } j \geqslant 0 \} \cup \{ b^j c^k \mid j, k \geqslant 0 \}$

– can be pumped, as in the pumping lemma

– is not accepted by FA

[M] E 2.39

Let $n = 1$ be the number of states of M.
For every $x \in L$, $|x| \geqslant 1$, it is either the case that (a) $x = a^i b^j c^j$ where $i \geqslant 1$ and $j \geqslant 0$ or (b) $x = b^j c^k$ where $j \geqslant 1, k \geqslant 0$ or $j \geqslant 0, k \geqslant 1$.
For both cases $\exists u, v, w, x = uvw, |uv| \leqslant 1, |v| \geqslant 1$. This implies $|u| = 0$ and $u = \Lambda$, while either (a) $v = a$ or (b) $v = b$ or $v = c$.
$\forall m \geqslant 0$, $uv^m w \in L$:
(a) $uv^m w = \Lambda a^m a^{i-1} b^j c^j$
$m = 0$, $i = 1$: $b^j c^j \in L$ (rhs), $i > 1$ $a^{i-1} b^j c^j \in L$ (lhs)
$m \geqslant 1$, $i \geqslant 1$: $a^m a^{i-1} b^j c^j \in L$ (lhs)
(b) $uv^m w$
$m \geqslant 0$, $j \geqslant 1$: $b^m b^{j-1} c^k \in L$ (rhs)
$m \geqslant 0$, $j = 0, k \geqslant 1$: $c^m c^{k-1} \in L$ (rhs)

Analogous for other n.

Remark: L does not fulfil the generalized pumping lemma, e.g., take $x = ab^n c^n$ and $x_2 = b^n$.

Decision problem: problem for which the answer is 'yes' or 'no':

Given . . . , is it true that . . . ?

Given an undirected graph $G = (V, E)$,
does G contain a Hamiltonian path?

Given a list of integers $x_1, x_2, \ldots, x_n$,
is the list sorted?

decidable $\iff$ $\exists$ algorithm that decides

$M = (Q, \Sigma, \delta, q_0, A)$
membership problem $\quad x \in L(M)$?

Specific to M: Given $x \in \Sigma^*$, is $x \in L(M)$?

Arbitrary M: Given FA M with alphabet Σ, and $x \in \Sigma^*$, is $x \in L(M)$?

Decidable, easy

[M] E 2.34

Theorem

The following two problems are decidable

1. *Given an* FA M, *is* $L(M)$ *nonempty?*
2. *Given an* FA M, *is* $L(M)$ *infinite?*

[M] E 2.34

Lemma

Let M be an FA with n states and let $L = L(M)$.

L is nonempty,
if and only if L contains an element x with $|x| < n$
(at least one such element).

Theorem

The following two problems are decidable

1. *Given an* FA M, *is* $L(M)$ *nonempty?*
2. *Given an* FA M, *is* $L(M)$ *infinite?*

[M] E 2.34

Lemma

Let M be an FA with n states and let $L = L(M)$.

L is infinite,
if and only if L contains an element x with $|x| \geqslant n$
(at least one such element).

cf. [M] Exercise 2.26

Lemma

Let M be an FA with n states and let $L = L(M)$.

L is infinite,
if and only if L contains an element x with $|x| \geqslant n$
(at least one such element).

Lemma

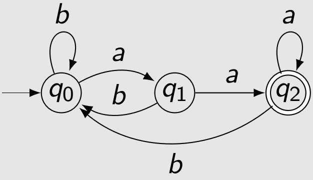
Let M be an FA with n states and let $L = L(M)$.

L contains an element x with $|x| \geqslant n$ (at least one such element)
if and only if L contains an element x with $n \leqslant |x| < 2n$
(at least one such element).

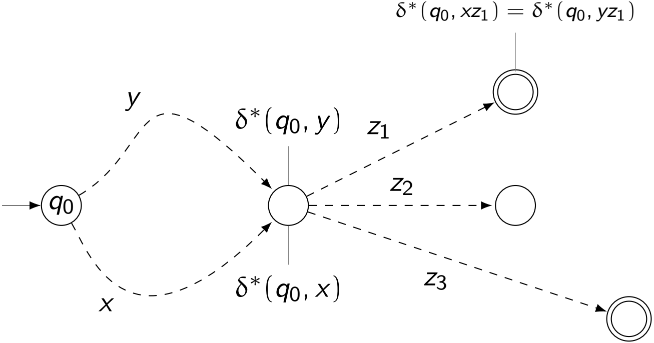
From lecture 1:

Example

$L_1 = \{\ x \in \{a, b\}^* \mid x \text{ ends with } aa\ \}$



[M] E. 2.1


$\delta^*(q_0, xz_1) = \delta^*(q_0, yz_1)$
y
$\delta^*(q_0, y)$
z_1
q_0
z_2
z_3
x
$\delta^*(q_0, x)$

Definition

Let L be language over Σ, and let $x, y \in \Sigma^*$.
Then x, y are *distinguishable* wrt L (*L-distinguishable*),
if there exists $z \in \Sigma^*$ with

$$xz \in L \text{ and } yz \notin L \quad \text{or} \quad xz \notin L \text{ and } yz \in L$$

Such z *distinguishes* x and y wrt L.

Equivalent definition:
let $L/x = \{ z \in \Sigma^* \mid xz \in L \}$

x and y are L-*distinguishable* if $L/x \neq L/y$.
Otherwise, they are L-*indistinguishable*.

The strings in a set $S \subseteq \Sigma^*$ are *pairwise L-distinguishable*, if for every pair x, y of distinct strings in S, x and y are L-distinguishable.

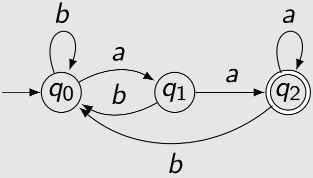
Definition independent of FAs

[M] D 2.20

From lecture 1:

Example

$L_1 = \{ x \in \{a, b\}^* \mid x \text{ ends with } aa \}$



$S = \{\Lambda, a, aa\}$

Example

$L_1 = \{ x \in \{a, b\}^* \mid x \text{ ends with } aa \}$

L_1/x for $x = \Lambda, a, b, aa \ldots$

Theorem

Suppose $M = (Q, \Sigma, q_0, A, \delta)$ is an FA accepting $L \subseteq \Sigma^*$.

If $x, y \in \Sigma^*$ are L-distinguishable, then $\delta^*(q_0, x) \neq \delta^*(q_0, y)$.

For every $n \geqslant 2$, if there is a set of n pairwise L-distinguishable strings in Σ^*, then Q must contain at least n states.

Hence, indeed: if $\delta^*(q_0, x) = \delta^*(q_0, y)$, then x and y are not L-distinguishable.

Proof. . .

[M] Thm 2.21

Exercise 2.5.

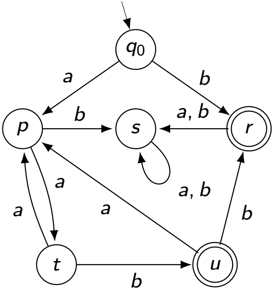
Suppose $M = (Q, \Sigma, q_0, A, \delta)$ is an FA, q is an element of Q, and x and y are strings in Σ^*. Using stuctural induction on y, prove the formula

$$\delta^*(q, xy) = \delta^*(\delta^*(q, x), y)$$

$L = \{aa, aab\}^*\{b\}$

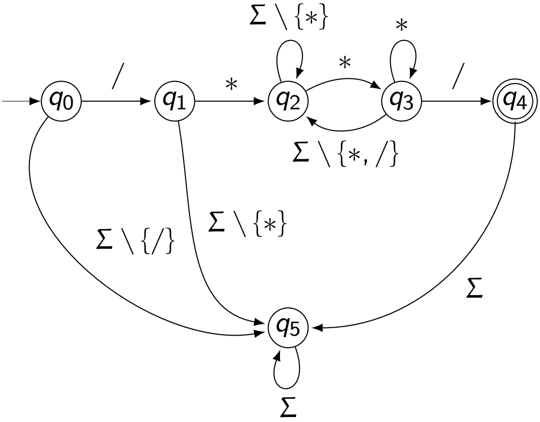
[M] E 2.22

$L = \{aa, aab\}^*\{b\}$



[M] E 2.22

$\Sigma = \{ \, l, d, \_, *, / \, \}$
$L_2 = \{ \, w \mid w \text{ is a (multi-line) C-style comment} \, \}$



$\approx$ [Stanford CS103] L13

Next challenge?

Can you find a language that satisfies the generalized version of the pumping lemma but is not accepted by a finite automaton?