Automata Theory

Mark van den Bergh / Rudy van Vliet

Bachelor Informatica
Data Science and Artificial Intelligence
Universiteit Leiden

Fall 2024

Universiteit Leiden
Leiden Institute of Advanced Computer Science

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 structural induction on y, prove the formula

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

Proof. Induction on $|y|$. If $|y| = 0$, i.e., $y = \Lambda$, then

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

is (almost trivially) true. Next, suppose the statement is true for $|y| < n$ and consider y such that $y = z\sigma$, $|y| = n$. Then

$$\begin{aligned} \delta^*(q, xy) &= \delta(\delta^*(q, xz), \sigma) && \text{defn. } \delta^*, y = z\sigma \\ &= \delta(\delta^*(\delta^*(q, x), z), \sigma) && \text{ind. hypothesis} \\ &= \delta^*(\delta^*(q, x), y) && \text{defn. } \delta^*, y = z\sigma \end{aligned}$$

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

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

Proof. Suppose L is accepted by an FA of n states. Consider $x = a^{n^2}$. Let u, v, w be such that $x = uvw$, $|uv| \leqslant n$ and $|v| \geqslant 1$. Hence, $u = a^i$, $v = a^j$ and $w = a^k$ such that $i + j + k = n^2$, $i + j \leqslant n$ and $j \geqslant 1$.
Consider uv^2w. Note that

$$|uv^2w| = i + 2j + k = n^2 + j > n^2.$$

Moreover,

$$n^2 + j \leqslant n^2 + n < n^2 + 2n + 1 = (n+1)^2.$$

Hence, the length of uv^2w lies in between two consecutive squares, so it cannot be a square.

[M] E 2.32

$L = \{\ a^{i^2} \mid i \geqslant 0\ \}$

Fun fact

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

Lagrange's four-square theorem

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

[M] E 2.33

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

Suppose L is accepted by an FA of n states. Let $x \in L$ be such that $|x| \geqslant n$. If $x = a^i b^j c^j$ with $i \geqslant 1$ and $j \geqslant 0$, take $u = \Lambda$, $v = a$ and w the rest of the string. Then $|uv| = 1 \leqslant n$ and $|v| = 1$. Moreover, $uv^m w = a^{m+i-1} b^j c^j \in L$ for all $m \geqslant 0$.
If $x = b^j c^k$ with $j, k \geqslant 0$, take $u = \Lambda$, v the first symbol of x, and w the rest. Then $|uv| = 1 \leqslant n$, $|v| = 1$ and, for all $m \geqslant 0$, we see that $uv^m w$ is either of the form $b^\ell c^k$ or c^ℓ with some $\ell \geqslant 0$, so in any case $uv^m w \in L$.
So L satisfies the pumping lemma. However, L is not regular. We will prove this later.
Note: L does not satisfy the generalized pumping lemma as discussed in Exercise 2.24.

[M] E 2.39

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

Given alphabet Σ with $|\Sigma| > 1$, and two lists of words $x_1, \ldots, x_n \in \Sigma^*$ and $y_1, \ldots, y_n \in \Sigma^*$.

Does there exist a sequence of indices $i_1, \ldots, i_m \in \{1, \ldots, n\}$ such that $x_{i_1} x_{i_2} \cdots x_{i_m} = y_{i_1} y_{i_2} \ldots y_{i_m}$?

Undecidable (see Computability).

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).

Proof. If L contains an element x with $|x| < n$, it is clear that L is nonempty.
Conversely, suppose L is nonempty, and suppose for the sake of contradiction that $|x| \geqslant n$ for all $x \in L$. Consider arbitrary such x. By the pumping lemma, there exist u, v, w such that $x = uvw$, $|v| \geqslant 1$ and $uw \in L$. Note that $|uw| < |x|$. If $|uw| < n$, we are done; otherwise, we can repeat the argument, applying the pumping lemma to uw, until we find a string in L of length smaller than n.

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

Proof. Suppose L is infinite. Note that $\{x \in L \mid |x| < n$ is finite, consisting of at most $\sum_{k=0}^{n} |\Sigma|^k$ words. Hence, L must indeed contain at least one x with $|x| \geqslant n$.
Conversely, suppose that L contains an element x with $|x| \geqslant n$. By the pumping lemma, there exist u, v, w such that $x = uvw$, $|v| \geqslant 1$, and $uv^m w \in L$ for all $m \geqslant 0$. Hence, this gives an infinite sequence of words that are elements of L.

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).

Proof. The reverse implication is trivial.
Suppose that L contains an element x with $|x| \geqslant n$, and suppose that $|x| \geqslant 2n$ By the pumping lemma, there exist u, v, w such that $x = uvw$, $1 \leqslant |v| < n$ and $uw \in L$. Hence, $n = 2n - n \leqslant |x| - n < |uw| < |x|$. If $|uw| < 2n$, we are done. Otherwise, we repeat the argument.

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?*

Proof. To check whether $L(M)$ is nonempty, we need to check whether $\{x \in L \mid |x| < n\}$ is nonempty, which is a finite amount of words to be checked.

To check whether $L(M)$ is infinite, we need to check whether $\{x \in L \mid n \leqslant |x| < 2n\}$ is nonempty, which is again a finite amount of words to be checked.

[M] E 2.34

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*.

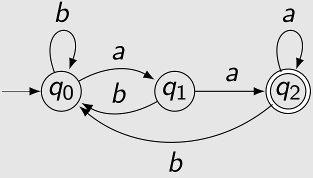
The strings in a set $S \subseteq \Sigma^*$ are *pairwise L-distinguishable*, if for every pair x, y of distinct strings in S, x and y are L-distinguishable.

Definition independent of FAs

[M] D 2.20

Example

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



$S = \{\Lambda, a, aa\}$ are pairwise L_1-distinguishable:
$\Lambda \cdot a = a \notin L_1$, but $a \cdot a \in L_1$;
$\Lambda \cdot \Lambda = \Lambda \notin L_1$, but $aa \cdot \Lambda = aa \in L_1$;
$a \cdot \Lambda = a \notin L_1$, but $aa \cdot \Lambda = aa \in L_1$.

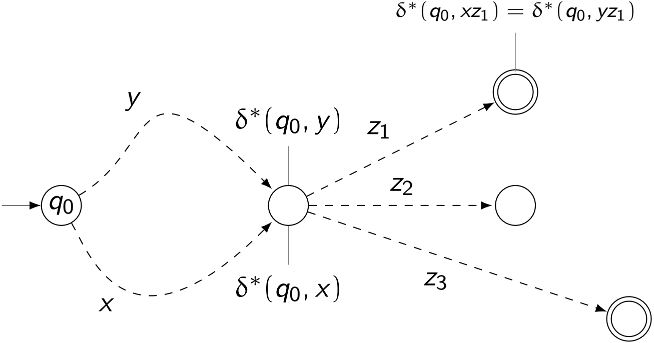
Example

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

Using alternative definition:
$L_1/\Lambda = L_1$;
$L_1/a = L_1 \cup \{a\}$;
$L_1/aa = L_1 \cup \{\Lambda, a\}$.

All unequal, so $\{\Lambda, a, aa\}$ pairwise L_1-distinguishable.

Same state, same future



Theorem

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

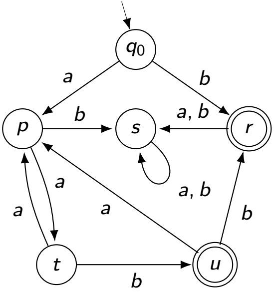
If $x, y \in \Sigma^*$ *are L-distinguishable, then* $\delta^*(q_0, x) \neq \delta^*(q_0, y)$.

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.

[M] Thm 2.21

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



[M] E 2.22