IX

AUTOMATA & FORMAL LANGUAGES

NINTH LECTURE : 25 October 2022

RECAP

Decision Problems

WORD PROBLEM: solved for noncontracting grammars

EMPTINESS PROBLEM: solved for regular grammars using the PUMPING LEMMA

EQUIVALENCE PROBLEM ?

PUMPING LEMMA

If G is regular, there is some n such that every $w \in \mathcal{L}(G)$ with $|w| \geq n$ can be pumped up or down.

§ 2.7 Minimisation of deterministic automata

GOAL FOR TODAY: Prove that each regular language has a unique minimal automaton which can be algorithmically constructed.

HOMOMORPHISMS

Lectures V & VI

If $D = (\Sigma, Q, \delta, q_0, F)$ and $D' = (\Sigma, Q', \delta', q_0', F')$ are deterministic automata over the same alphabet Σ, we say that a map $f : Q \to Q'$ is a *homomorphism from D to D'* if

(i) for all $q \in Q$ and $a \in \Sigma$, we have that $\delta'(f(q), a) = f(\delta(q, a))$,

(ii) we have $f(q_0) = q_0'$, and

(iii) for all $q \in Q$, we have that $q \in F$ if and only if $f(q) \in F'$.

-?

PROPOSITION 2.5
If there is a homomorphism from D to D', then $\mathcal{L}(D) = \mathcal{L}(D')$.

REFLECTION ON FAILURE TO BE BIJECTIVE

Non-surjective: By (*), if f is not surjective, then no state that can be reached from q_0' can fail to be in the range of f. All **ACCESSIBLE** states are in the range of f.

Non-injective: If $f(p) = f(q)$ then by (*) they are either both accept or both reject states. Also $f(\hat{\delta}(q,w)) = \hat{\delta}(f(q),w) = \hat{\delta}(f(p),w) = f(\hat{\delta}(p,w))$

This property will be called INDISTINGUISHABLE [For every w, $\hat{\delta}(q,w) \in F \Leftrightarrow \hat{\delta}(p,w) \in F$]

So, we observe that failure to be surjective or injective only affect INACCESSIBLE states or pairs of INDISTINGUISHABLE state. These do not change the language accepted by D.

[More detail in §2.7.]

Lecture VI

Def. A state q is called ACCESSIBLE if there is a word w s.t.
$$q = \hat{\delta}(q_0, w).$$
If not, it's called INACCESSIBLE.

Two states q, q' are DISTINGUISHED by a word w if
$$\hat{\delta}(q,w) \in F \ \& \ \hat{\delta}(q',w) \notin F$$
or vice versa

We say q, q' DISTINGUISHABLE if there is a word w that distinguishes them.

If not: INDISTINGUISHABLE. $\sim$.

Reminder (Lecture VI):

If $f: Q \longrightarrow Q'$ is a homomorphism, then

(1) p, q distinguishable $\Longrightarrow f(p) \neq f(q)$

(2) $q' \in Q'$ accessible $\Longrightarrow q' \in \mathrm{ran}(f)$

So: ✪

$$\boxed{\begin{array}{l}\text{If } f \text{ is a homom. from } D \text{ to } D' \text{ and} \\ \bullet \text{ all distinct states in } D \text{ are distinguishable, then } f \text{ is injective} \\ \bullet \text{ all states in } D' \text{ are accessible, then } f \text{ is surjective.}\end{array}}$$

Def. An automaton D is called irreducible if for $q \neq q'$ q and q' are distinguishable and all state are accessible.

✪ means: any homom. between irreducible automata is an iso.

Note: $q \sim q' \;:\Longleftrightarrow$ q and q' are indistinguishable

$\sim$ is an equivalence relation

Write $[q] := \{q' \in Q \,;\, q \sim q'\}$

EQUIVALENCE CLASS

Define the <u>QUOTIENT</u> <u>AUTOMATON</u>

$$D/_{\sim} := (\Sigma, Q/{\sim}, [\delta], [q_0], [F])$$

with $[\delta]([q], a) := [\delta(q,a)]$

$[F] := \{[q] \,;\, q \in F\}$

<u>Properties</u>

① Clearly, if $[q] \cap F \neq \emptyset$, then $[q] \subseteq F$.

② Welldefinitness: Suppose $q \sim q'$

$$\begin{aligned} [\delta]([q],a) &= [\delta(q,a)] \\ &= [\delta(q',a)] \\ &= [\delta]([q'],a). \end{aligned}$$

[If w dist. $\delta(q,a)$ & $\delta(q',a)$, then aw dist. q and q'.]

③ If $q \not\sim q'$, i.e., $[q] \neq [q']$,
then they are distinguished in $D/\sim$:

By induction: $\hat{[\delta]}([q], w) \overset{(*)}{=} [\hat{\delta}(q, w)]$

Suppose w.l.o.g. $\hat{\delta}(q, w) \in F$ & $\hat{\delta}(q', w) \notin F$

$$\hat{[\delta]}([q], w) \overset{(*)}{=} [\hat{\delta}(q, w)] \in [F].$$

$$\hat{[\delta]}([q'], w) \overset{(*)}{=} [\hat{\delta}(q', w)] \notin [F].$$

④ $\mathcal{L}(D) = \mathcal{L}(D/\sim)$

[Just because $q \longmapsto [q]$ is a homomorphism.]

⑤ If D had no inaccessible states, then $D/\sim$ has no inaccessible states.

[The quotient map is a surjection.]

Theorem 2.22

For every $\mathcal{D}$ there is $\mathcal{I}$ irreducible s.t. $\mathcal{L}(\mathcal{D}) = \mathcal{L}(\mathcal{I})$.

Proof. Let $A \subseteq Q$ be the accessible states in $\mathcal{D}$:

$$\mathcal{D}^* := (\Sigma, A, \delta \restriction A \times \Sigma, q_0, F \cap A)$$

Clearly, $id : A \longrightarrow Q$ is a homomorphism between $\mathcal{D}^*$ and $\mathcal{D}$.

$$\Longrightarrow \mathcal{L}(\mathcal{D}^*) = \mathcal{L}(\mathcal{D}).$$

Let $\mathcal{I} := \mathcal{D}^*/\sim$

By (3), (4), and (5),

$\mathcal{I}$ is irreducible and

$$\mathcal{L}(\mathcal{D}^*/\sim) = \mathcal{L}(\mathcal{D}^*) = \mathcal{L}(\mathcal{D}).$$

q.e.d.

Remark The # of states in $\mathcal{I}$ is $\leq$ to the # of states in $\mathcal{D}$.

<u>Theorem 2.23</u> If I, I' are irreducible and $\mathcal{L}(I) = \mathcal{L}(I')$, then I and I' are isomorphic.

<u>Proof.</u> By the earlier remark, we only need to show that there is a homom. from I to I'.

$I = (\Sigma, Q, \delta, q_0, F)$ $I' = (\Sigma, Q', \delta', q_0', F')$

W.l.o.g. $Q \cap Q' = \emptyset$.

Extend $\sim$ to $Q \cup Q'$.

$q \in Q, q' \in Q'$ Say $q \sim q'$ if for all w, $\hat{\delta}(q, w) \in F \Longleftrightarrow \hat{\delta}'(q', w) \in F.$

By assumption we have $q_0 \sim q_0'$.

<u>Claim</u> For all $q \in Q$, there is $q' \in Q'$ s.t. $q \sim q'$.

Define $sp(q) :=$ length of shortest path from q_0 to q

Since I is irreducible, $sp(q)$ is defined for all $q \in Q$.

Prove claim by induction on $sp(q)$:

$sp(q) = 0 \implies q = q_0 \sim q_0'$. Done!

$sp(q) = k+1$ Find $p \in Q$, $a \in \Sigma$ s.t.

$\delta(p, a) = q$ Then $sp(p) = k$.

$\xRightarrow{IH}$ find $p' \in Q'$ s.t. $p' \sim p$.

Get $q' := \delta'(p', a) \sim \delta(p, a) = q$.

q.e.d. (Claim).

<u>Claim</u> $q' \sim q \sim p'$, then $q' = p'$.

[By transitivity, $q' \sim p'$. Irreducibility of I' gives the claim.]

So, using claims, define

$f(q) :=$ the unique $q' \in Q'$ s.t. $q \sim q'$.

<u>Claim</u> f is homom.

(ii) $q_0 \sim q_0'$, so $f(q_0) = q_0'$.

(iii) $q \in F \iff f(q) \in F'$

[Just from definition of $\sim$.]

(i) Fix q and $q' = f(q)$, i.e., $q \sim q'$.

Clearly, $\delta(q,a) \sim \delta'(q',a)$.

Therefore:

$$f(\delta(q,a)) = \delta'(q',a)$$
$$= \delta'(f(q),a)$$

q.e.d.

<u>Remark.</u> This means that there is an (up to iso) unique irreducible automaton for any given regular language.

Furthermore, the size of it is $\leq$ the size of any automaton accepting the same language.

THE MINIMAL AUTOMATON

§2.8 Decision Problems for regular grammars

Decision Problems

WORD PROBLEM: solved for noncontracting grammars

EMPTINESS PROBLEM: solved for regular grammars using the PUMPING LEMMA

EQUIVALENCE PROBLEM: ?

FROM OUR RECAP ON PAGE 1.

Idea Given G & G', produce minimal automata $\mathfrak{I}, \mathfrak{I}'$ for $\mathcal{L}(G)$ and $\mathcal{L}(G')$ and check whether

$$\mathfrak{I} \cong \mathfrak{I}'.$$

If so, $\mathcal{L}(G) = \mathcal{L}(G')$;
if not, $\mathcal{L}(G) \neq \mathcal{L}(G')$.

Proposition 2.27 Given det. automaton D and a state q, there is an algorithm that determines whether q is accessible.

Proof. If there is w s.t. $\hat{\delta}(q_0, w) = q$, then the shortest such word must have length $\leq |Q|$.

Why? By the proof of the PL, any path of length $> |Q|$ will repeat states (PIGEONHOLE) and thus can be shortened.

Thus it's enough to check all possible paths of length $\leq |Q|$.

q.e.d.