X

TENTH LECTURE
(start of week 4)

27 October 2023

AUTOMATA & FORMAL LANGUAGES

Lecture $\overline{IX}$ page 7

ACCESSIBLE : $\hat{\delta}(q_0, w) = q$ some w

INDISTINGUISHABLE : $A_q = A_{q'}$

$A_q := \{w;\ \hat{\delta}(q,w) \in F\}$

<u>Def.</u> A det. automaton D is called **IRREDUCIBLE** if it has no inacc. or indistinguishable states.

<u>Corollary</u> (to our observations)
if I, I' are irreducible and f is a homom. between I & I', then f is an isomorphism.

Lecture $\overline{IX}$ page 9

<u>CONSEQUENCE</u>
For any reducible automaton D, there is an irreducible automaton with strictly fewer states.

therefore if n is ~~the~~ minimal # of states, every automaton of size n is irreducible.

Lecture IX
page 10

THEOREM Any two irreducible automata accepting the same language are isomorphic.

This implies that all minimal automata accepting L are isomorphic and of minimal size.

Proof. $\mathcal{I} = (\Sigma, Q, \delta, q_0, F)$ irreducible
$\mathcal{I}' = (\Sigma, Q', \delta', q_0', F')$

If $q \in Q$, $q' \in Q'$, I write

$$q \sim q' \quad :\text{iff} \quad \underbrace{A_q}_{= \{w;\, \hat{\delta}(q,w) \in F\}} = \underbrace{A'_{q'}}_{= \{w;\, \hat{\delta}'(q',w) \in F'\}}$$

Claim 1 Each $q \in Q$ has $q' \in Q'$ s.t. $q \sim q'$.

PROVED IN LECTURE IX.

Lecture IX
page 11

Claim 2 If $q \sim q'$ and $q \sim p'$, then $q' = p'$.

Still to be proved

Proof of Claim 2

$$\begin{array}{ccc} q \sim q' & & q \sim p' \\ \Updownarrow & & \Updownarrow \\ A_q = A'_{q'} & & A_q = A'_{p'} \\ & \searrow\!\!\!\Downarrow \quad \swarrow & \\ & A'_{q'} = A'_{p'} & \end{array}$$

$A'_{q'} = A'_{p'} \iff q'$ and p' are indistinguishable

$\Longrightarrow$ $q' = p'$.
[$\mathcal{I}'$ irreducible]

q.e.d. (Claim 2)

Define

$$f(q) := \text{the unique } q' \text{ s.t. } q \sim q'$$
[Exists by Claims 1 & 2].

We'll show:

f is homomorphism
[by the Corr. from $\overline{IX}$, p. 7, this is enough to show it's iso].

① We already proved in Claim 1 that $q_0 \sim q_0'$, so $f(q_0) = q_0'$.

② $f(q) = q'$ [i.e., $q \sim q'$]

Then
$$\begin{aligned} q \in F &\iff \delta(q, \varepsilon) \in F \\ &\iff \delta'(q', \varepsilon) \in F' \quad [\text{by } q \sim q'] \\ &\iff q' \in F'. \end{aligned}$$

③ Suppose have
$$\overline{q} := \delta(q, a)$$
$$\overline{q'} := \delta'(q', a)$$

If $\overline{q}, \overline{q'}$ are distinguishable by w, so w.l.o.g.
$$\hat{\delta}(\overline{q}, w) \in F \ \& \ \hat{\delta}'(\overline{q'}, w) \notin F',$$

then $\hat{\delta}(q, aw) \in F \ \& \ \hat{\delta}'(q', aw) \notin F'$,

contradicting $q \sim q'$.

q.e.d.

SUMMARY Let L be a regular language with det. automaton D s.t. $L = \mathcal{L}(D)$.

Let

$$n := \min \{ |Q'| ; \text{ there is } D' \text{ with } \mathcal{L}(D') = L \}$$

MINIMISATION THEOREM

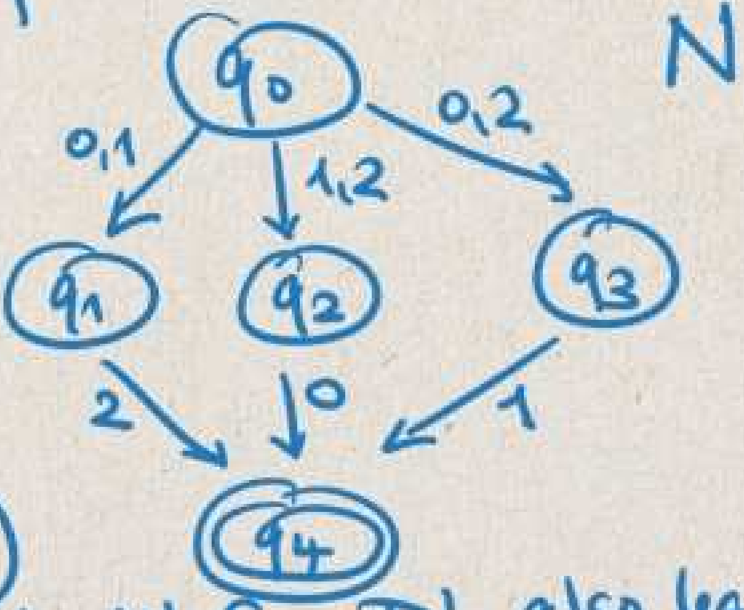
There is an irreducible automaton I such that $\mathcal{L}(I) = L$ and TFAE for any D':

(i) D' is irreducible

(ii) $|Q'| = n$.

Follows directly from what we proved.

Remark This is not true for non-deterministic automata.



Food for thought:

$R = (0+1)2 + (1+2)0 + (0+2)1$

$\mathcal{L}(N) = \mathcal{L}(R)$

$R' = 0(1+2) + 1(0+2) + 2(0+1)$

The natural nondet. automaton N' for R' also has five states, but is not isomorphic to N.

DECISION PROBLEMS

	Type 0	Type 1	Type 2	Type 3
WORD P.	?	✓	✓	✓
EMPTINESS P.	?	?	?	✓ → ?
EQUIVALENCE P.	?	?	?	?

NOW

Emptiness Problem: $\{G;\ \mathcal{L}(G) \neq \emptyset\}$

Equivalence Problem: $\{(G, G');\ \mathcal{L}(G) = \mathcal{L}(G')\}$

Theorem The emptiness problem for regular grammar is solvable.

Lemma If L is a language satisfying the (R)PL with pumping # n, then if $L \neq \emptyset$, it contains a word of length $< n$.

Proof. Every word of length $\geq n$ can be pumped down and therefore cannot a word of minimal length in L. Thus if $L \neq \emptyset$, any word of minimal length must be shorter than n.

q.e.d.

Proof of Theorem

Take G, transform it to det. automaton.
Count its states: n.
Then look at all words $w \in \mathbb{W}$ with $|w| < n$ and check [SOLVABILITY OF THE WORD PROBLEM] whether $w \in \mathcal{L}(G)$.

uses

If any of them is: NO!
If all of them aren't: YES!
[by lemma]

q.e.d.

Remarks

1. Still haven't defined "algorithm".
2. This algorithm is terribly inefficient!

Equivalence problem.

Prop. 1 The isomorphism problem for det. aut. is solvable:

$$\{(D, D') ; D \cong D'\}$$

Proof. If $|Q| \neq |Q'|$, answer NO! If $|Q| = |Q'| = n$, there are n^n functions. Check for each of them whether they are isos. q.e.d.

Prop. 2 The accessibility problem for det. automata is solvable:

$$\{(D,q)\ ;\ \exists w\ \hat{\delta}(q_0,w)=q\}$$

$\Longleftrightarrow$ w is accessible

Proof. By (the proof of) the (R)PL, we know that if q is accessible, there is a word w of $|w|<|Q|$ such that

$$\hat{\delta}(q_0,w)=q.$$

Check all w with $|w|<|Q|$.
If none of them produce q, NO!
Otherwise, YES!

q.e.d.

Prop. 3 The indistinguishability problem for det. automata is solvable.

$$\{(D,q,q')\ ;\ q\sim q'\}$$

Proof. We'll give an algorithm that produces a complete answer to the entire "distinguishability situation" in D:

(q_1, q_{n-1}) marked by ε
(q_2, q_n) marked by ε
$\delta(q_1, a) = q_2 \quad \delta(q_1, a) = q_n$

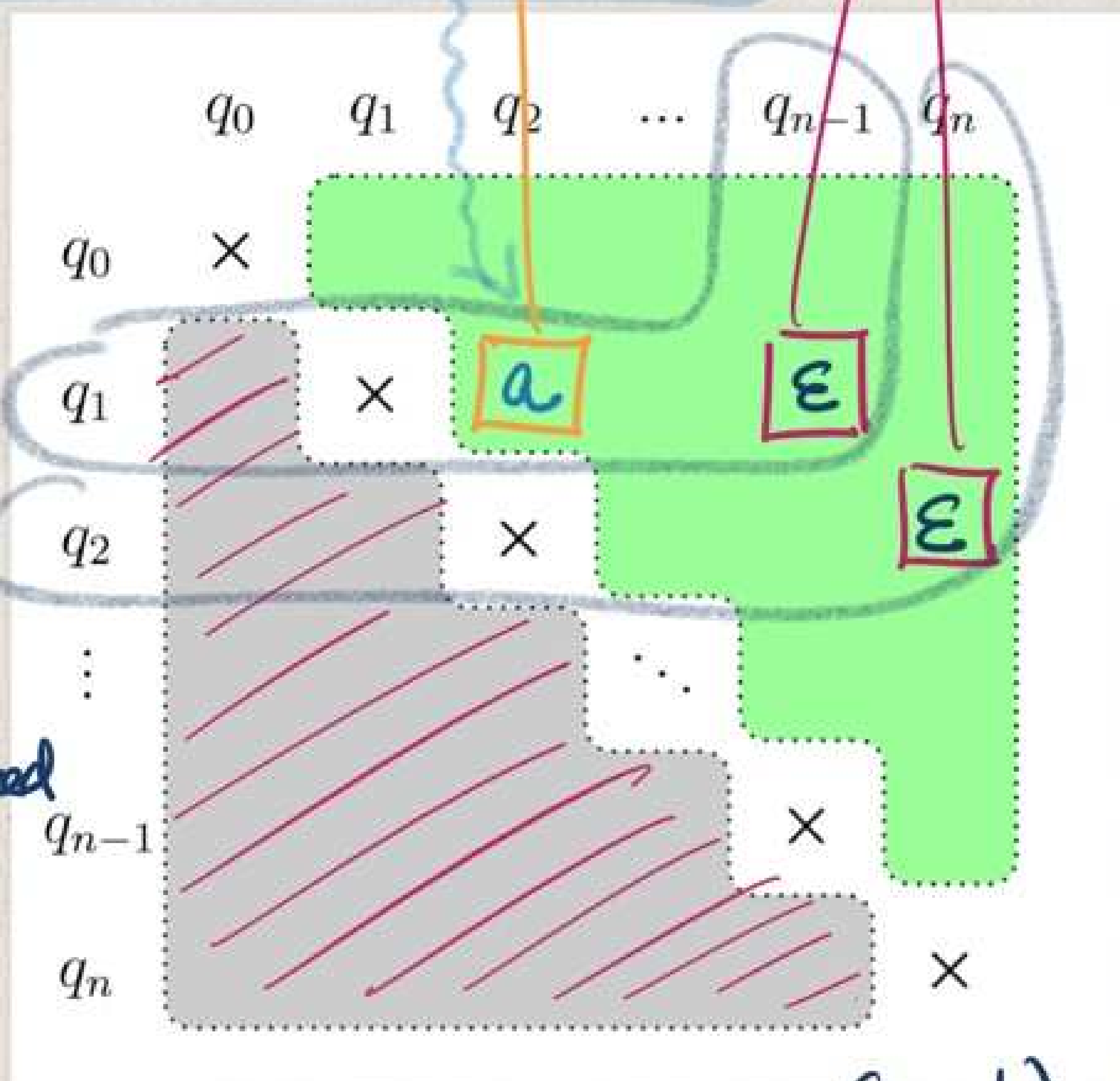
EXAMPLE

STEP 2

STEP 1.

TABLE FILLING ALGORTHM

Only need to fill half of the table.



<u>STEP 1</u>

If $q \in F$ & $q' \notin F$ or vice versa, they are distinguished by ε, so we write ε in the cell (q, q')

<u>STEP k+1</u> Consider all unmarked pairs (q, q') and for each $a \in \Sigma$, define

$$q_* := \delta(q, a)$$
$$q'_* := \delta(q', a)$$

If (q_*, q'_*) is marked by v in the table, then mark (q, q') by av. Otherwise, do nothing.

The algorithm terminates if in STEP k, no new is marked.

<u>Claim</u>: $q \sim q' \iff (q, q')$ is unmarked by the algorithm

$\Longrightarrow$ obvious by construction.

$\Longleftarrow$ We say (q, q') is a bad pair with witness length n if (q, q') is unmarked but there is a word w of length n distinguishing q & q'.

We'll show that there are no bad pairs.
Suppose towards a contradiction that (q, q') is a bad pair with minimal witness length n.

Note that $n > 0$, since all pairs distinguished by ε were labelled in STEP 1.
So some word aw distinguishes q, q'.

$$q_* := \delta(q, a)$$
$$q_*' := \delta(q', a).$$

By construction, q_*, q_*' are distinguished by w.
Since $|w| < n$, (q_*, q_*') is marked in some step of the algorithm.

So, by construction, in the next step, (q, q') was marked. Contradiction!

q.e.d.

PROPOSITION 1 The isomorphism problem for deterministic automata is solvable.

PROPOSITION 2 The accessibility problem for deterministic automata is solvable.

PROPOSITION 3 The indistinguishability problem for deterministic automata is solvable.

Theorem The equivalence problem for regular grammars is solvable.

Proof. Start with G, G'.
Transform them into det. aut. D, D'.
Use Props 2 & 3 to produce irreducible automata I, I' s.t.

$$\mathcal{L}(D) = \mathcal{L}(I)$$
$$\mathcal{L}(D') = \mathcal{L}(I')$$

Use Prop 1 to check whether I, I' are isomorphic.

If so: YES.
If not: NO.

q.e.d.

DECISION PROBLEMS

	Type 0	Type 1	Type 2	Type 3
WORD P.	?	✓	✓	✓
EMPTINESS P.	?	?	?	~~?~~ ✓
EQUIVALENCE P.	?	?	?	~~?~~ ✓

CLOSURE PROPERTIES

Proposition on page 7

	Type 0	Type 1	Type 2	Type 3
Concatenation	✓	✓	✓	~~?~~ ✓
Union	✓	✓	✓	~~?~~ ✓
Intersection	?	?	?	~~?~~ ✓
Complement	?	?	?	~~?~~ ✓
Kleene plus	?	?	?	✓

Added after the proof in the section on Regular Expressions.