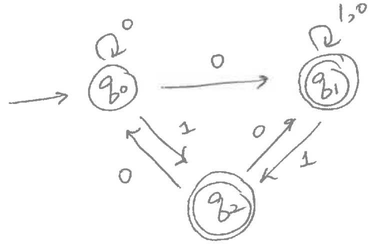
# MATH 2301

09/10/2023

* Recap : Given an NFA M, can we construct a regex r such that

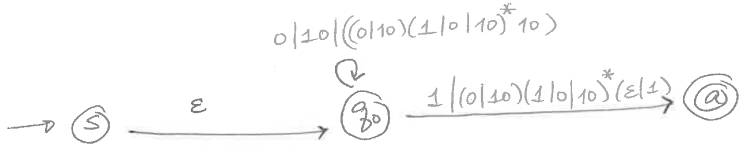
$L(M) = L(r)$?

* Last time : Described a procedure for this.

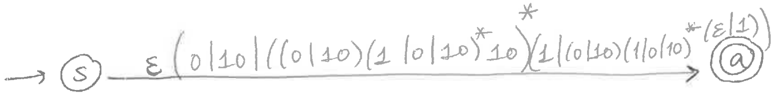


(original machine M)

We converted it to a generalised NFA (GNFA):



Last step : Remove $q_0$



Call the final label r

By construction, $L(M) = L(r)$ !

* Summary

Step 0 : Sanitise M by pulling out a separate start state Ⓢ and a separate (single) accept state @, redirecting all previous accepting states to @ via ε-arrows, and Ⓢ to the previous start state via an ε-arrow.

Step 1 : Choose some state ⓠ (different from Ⓢ or @) to delete, and delete it. For each pair x, y such that { x can be equal to y, but x ≠ q and y ≠ q

$$ⓧ \xrightarrow{l_1} \overset{l_2}{ⓠ} \xrightarrow{l_3} ⓨ$$

were previous labels, add an arrow ⓧ ⟶ ⓨ, labelled by

$$l_1\,(l_2^*)\,l_3.$$

[If there is no loop $l_2$, just drop it to make $l_1 l_3$.]

Step 2 : Repeat step 1 until there is no other state left, other than Ⓢ and @. Let r = label on the arrow Ⓢ ⟶ @. That is your answer!

When adding $l_1(l_2^*)l_3$ to ⓧ ⟶ ⓨ, if there is already some label $l$ on ⓧ ⟶ ⓨ, the new label becomes

$$l \mid l_1(l_2^*)\,l_3$$

(Two options on one label get joined by | )

* Def: We say that a language L is regular if: one of the following (equivalent) conditions holds:

(1) There is a DFA M such that $L(M) = L$,

(2) There is an NFA M such that $L(M) = L$,

(3) There is a regex r such that $L(r) = L$.

(Our discussion so far has shown that any one of these implies the other two.)

Q: Does this mean that some languages are not regular??

E.g.: $L = \{0^n 1^n \mid n \in \mathbb{N}\}$

$L = \{\varepsilon, 01, 0011, 000111, \ldots\}$

I claim that L is not regular.

An answer: countability. [ASIDE]

Beyond the scope of this class

Each regex is finite in size. (finitely many symbols in r)
One could enumerate all possible regexes in dictionary order: $r_1, r_2, r_3, r_4, \ldots$

Consider $L(r_1), L(r_2), L(r_3), \ldots$ ← List of all possible regular languages (with repeats)

Now on the other hand, the set of all possible languages is uncountable. [NOT LISTABLE]

⟹ There are non-regular languages

Back to

$L = \{0^n 1^n \mid n \in \mathbb{N}\}$ ← This description is not a regex.

Claim: There is no DFA M such that $L(M) = L$.

Pf. Suppose there existed some DFA M such that $L(M) = L$.

M has some number of states, say N states.

If w is long enough, i.e. $|w| > N$ then as w runs through M, it must repeat a state.

Path of $w = w_1 w_2 \cdots w_k$ ; say w accepted by M

This state repeated.

w can be split into three parts $xyz = w$

Note that $xyyz$, $xyyyz$, $xyyyyyyyz$ are also accepted by this DFA M.

M has N states,

$L = \{0^n 1^n \mid n \in \mathbb{N}\}$.

Q: What happens if you feed the string

$w = 0^N 1^N$ into M?

→ Can you produce a contradiction from this?