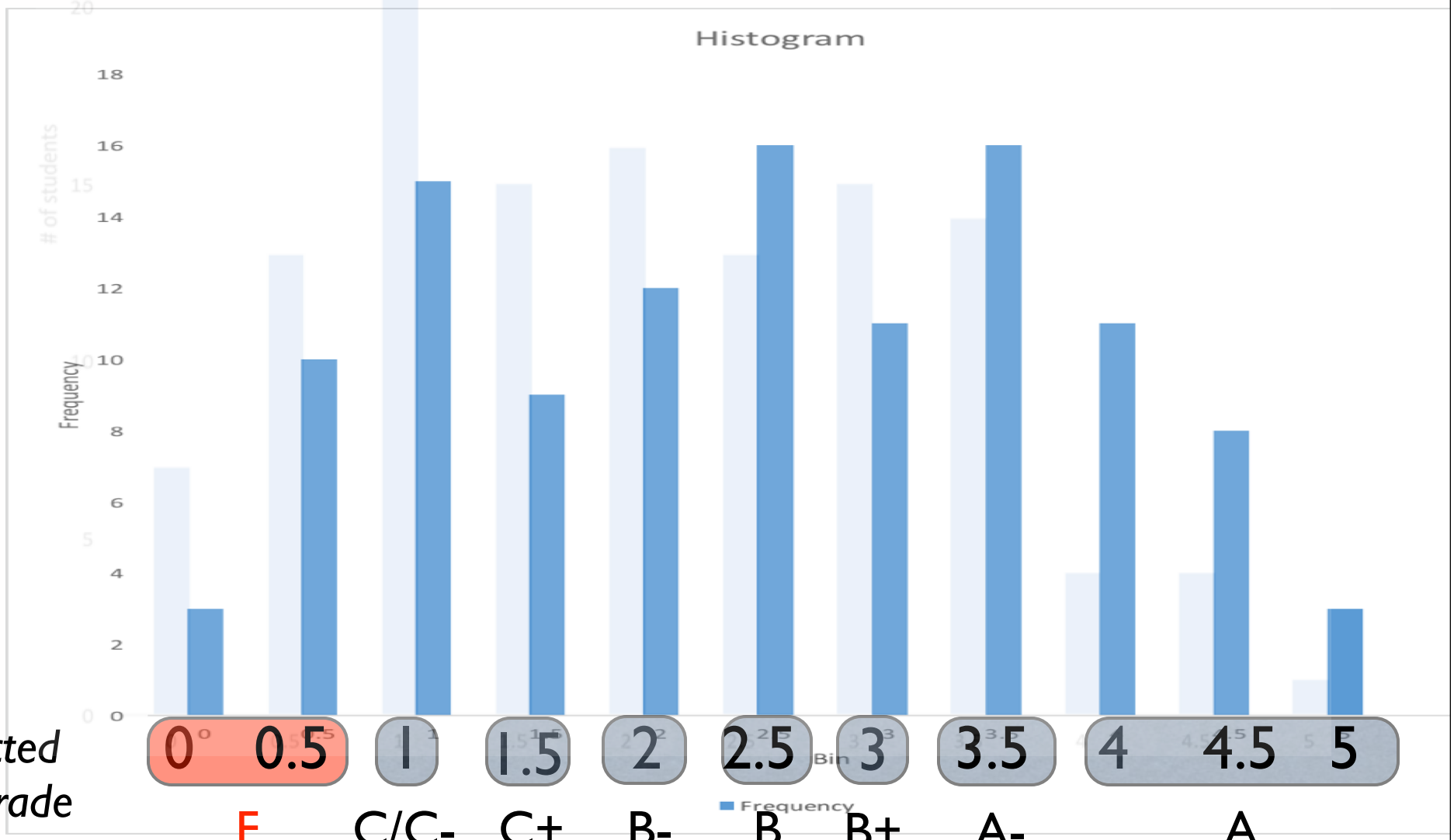
Quiz 2 much better than Quiz 1

- quiz 2 mean: 2.47; quiz 1 mean: 2.0



projected final grade

F C/C- C+ B- B B+ A- A

Concatenation - NFA

THEOREM

THE CLASS OF REGULAR LANGUAGES IS CLOSED UNDER CONCATENATION.

If L_1 and L_2 are regular then so is $L_1 \circ L_2$.

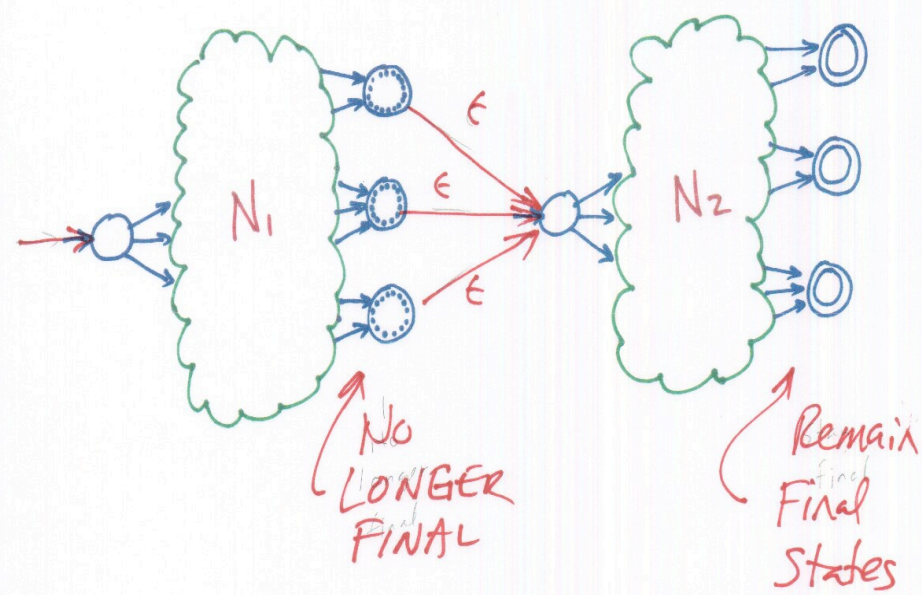
PROOF

CAN'T DO IT YET.

WE NEED...

NONDETERMINISM

- concatenation of regular languages is still regular



Non-Determinism

"GIVEN THE CURRENT STATE, THERE MAY BE MULTIPLE NEXT STATES."

- The next state is chosen at random.
- All next states are chosen in parallel and pursued simultaneously.

Same Thing

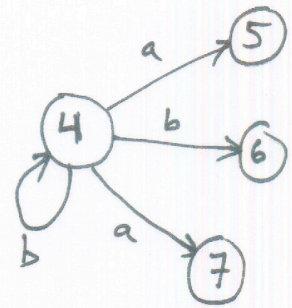
FSM means Deterministic Finite State Automaton/Machine

DFA means DETERMINISTIC FINITE STATE AUTOMATON

NFA means NONDETERMINISTIC FINITE STATE AUTOMATON

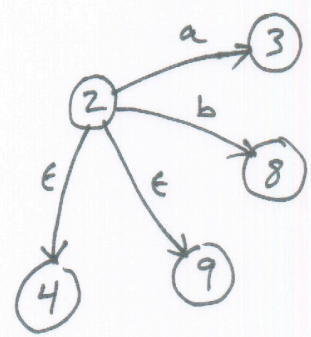
NOW WE WILL ALLOW

- MULTIPLE EDGES WITH THE SAME LABEL OUT OF A NODE.



Which edge should you take???

- EPSILON EDGES



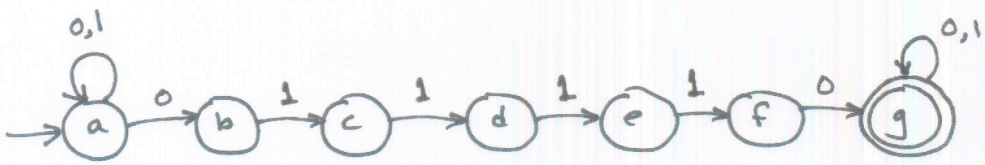
Can take an ε-edge without scanning a symbol.

It is "OPTIONAL"!

Any path to accept is accepted

EXAMPLE

All strings that contain 011110



EXAMPLE STRING: 010001111010101

Lots of bad choices
that dont works,
that dont reach an accept state

All we need is one way
to reach ACCEPT.

If there is any way to run
the machine that ends
with ACCEPT,
Then the NFA accepts.

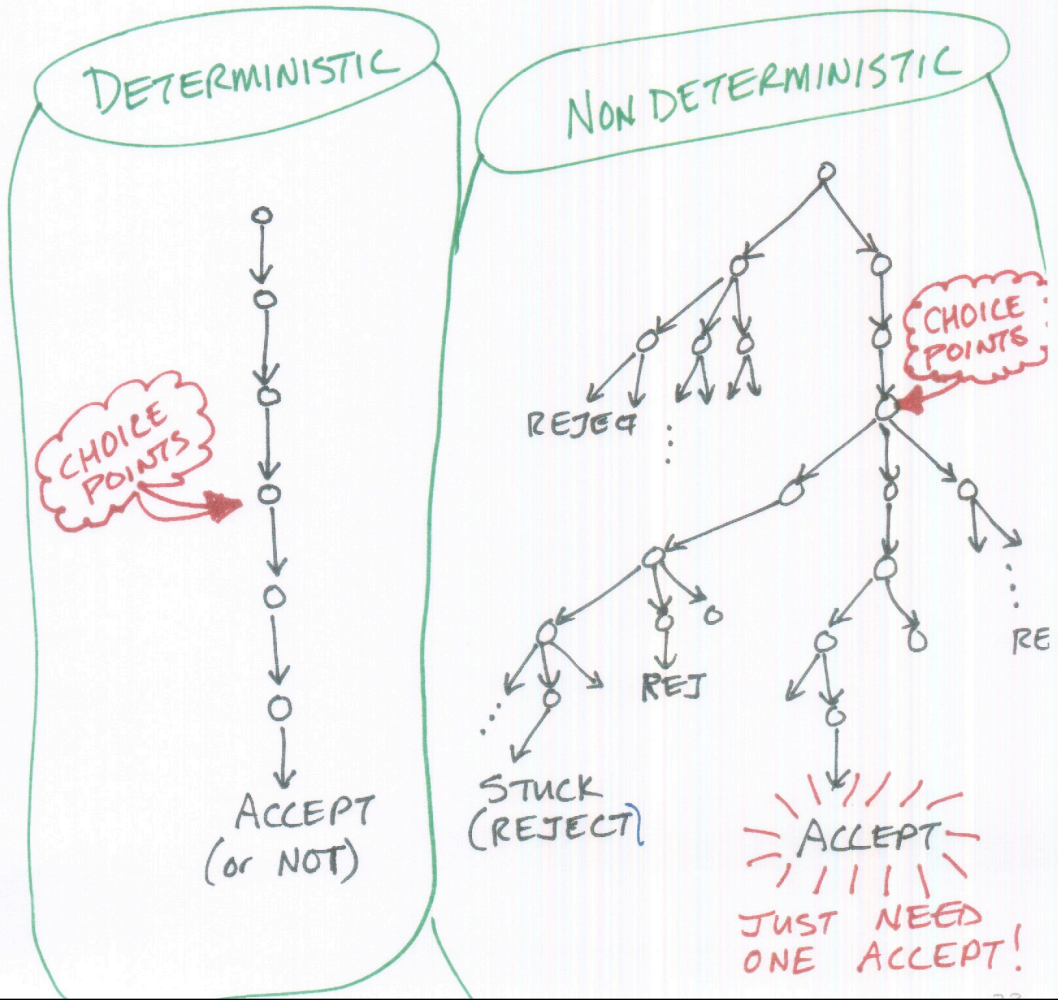
- DFA is much harder (see HW1)
- but does NFA help for the complement?

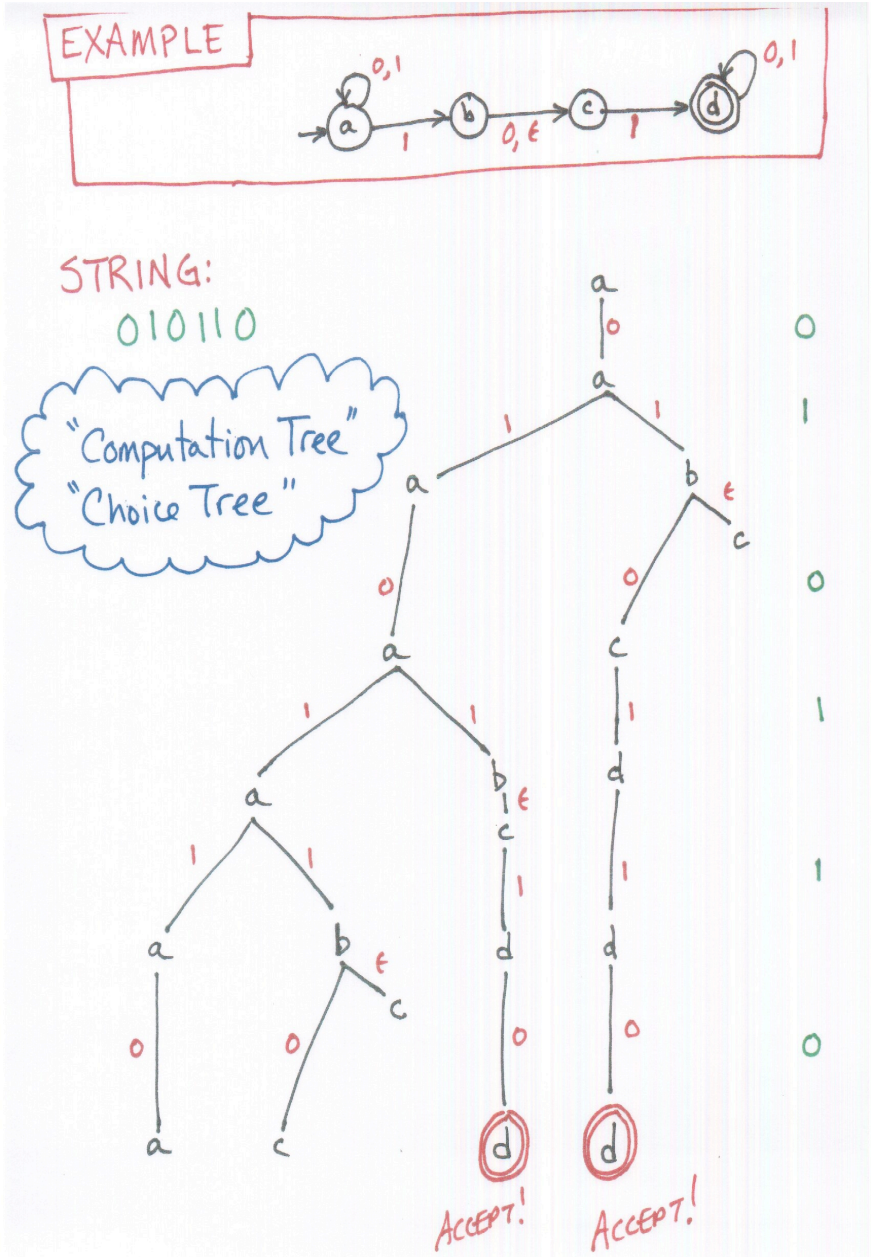
As long as there is (at least) one path...

LOTS OF CHOICES - WHICH ONE TO TRY?

* Try them all.
* Make the right choice at each point.

EQUIVALENT





NFA == DFA, but often a lot easier

THEOREM

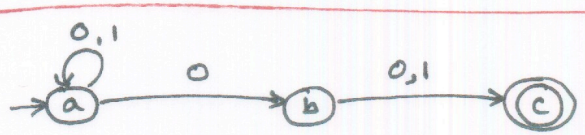
FOR EVERY
NONDETERMINISTIC F.S.M.
THERE IS AN EQUIVALENT
DETERMINISTIC F.S.M.

... But it may be large and hard to find!

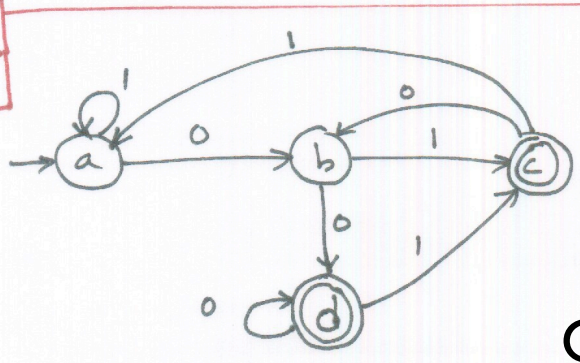
EXAMPLE

All strings over $\{0,1\}^*$ that have a "0" in the second to the last position.

NFA



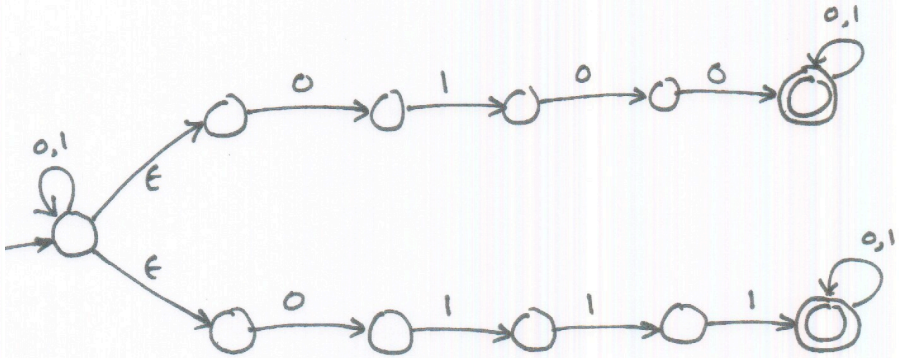
FSM
DFA



EXAMPLE

String contains either
... 0100...
or ... 0111...

When to start looking?
Which string to look for?
NONDETERMINISM!

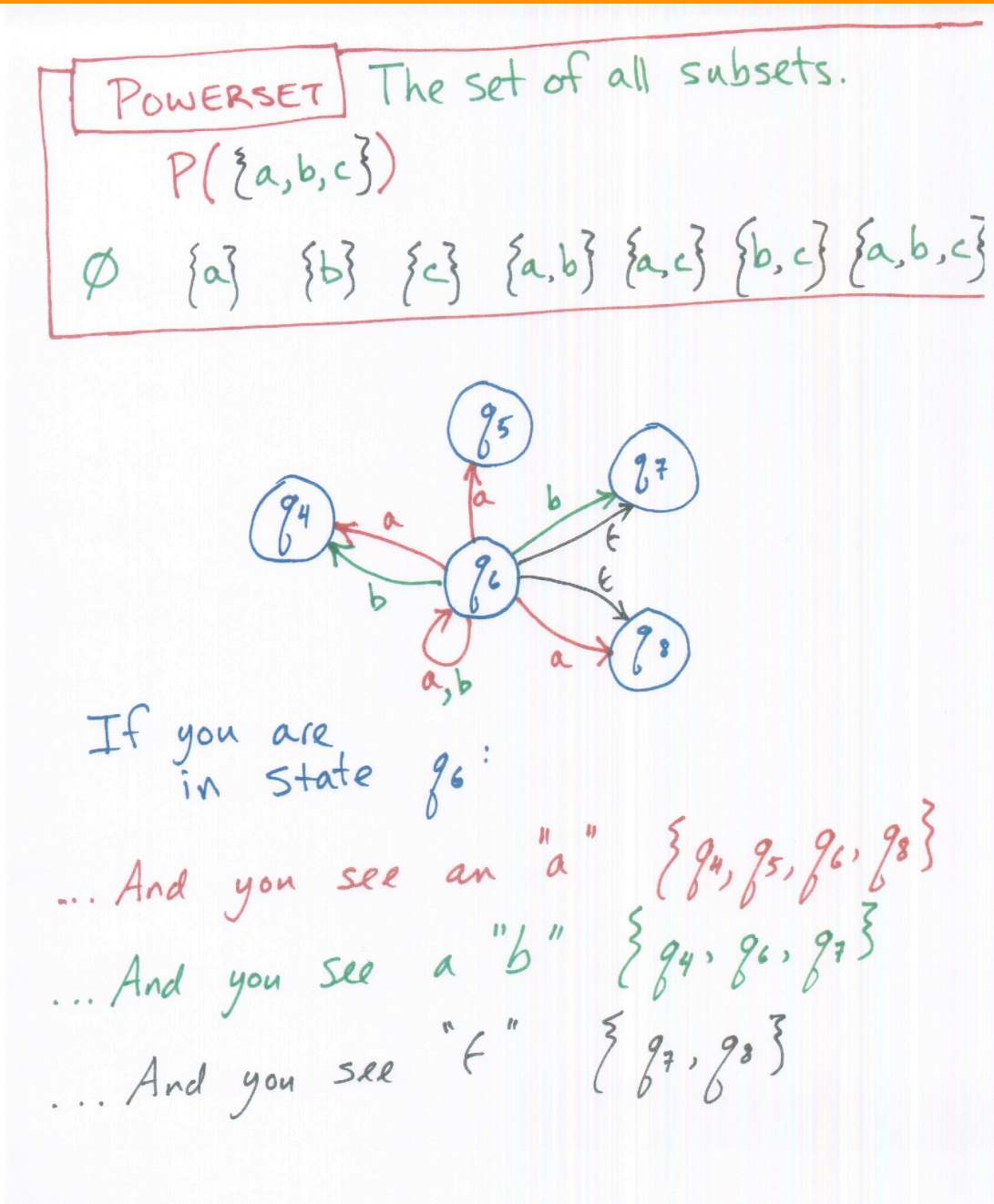


CHALLENGE:

Build/Design a DFA to recognize this language.

Q: how to use NFA for union? for intersection?

Power set (all subsets of a set)



- power set is one of the most important & beautiful mathematical concepts
- P(Q) is often written as 2^Q since $|2^Q| = 2^{|Q|}$

$$2^Q \triangleq \{A \mid A \subseteq Q\}$$

Powerset
NATURAL LANGUAGE SEARCH

Formal Definition of NFA

FORMAL DEFINITION OF NONDETERMINISTIC FINITE STATE MACHINE

Σ = Alphabet

$\Sigma_\epsilon = \Sigma \cup \{\epsilon\}$

$M = (Q, \Sigma, \delta, q_0, F)$

Q = States

Σ = Alphabet

q_0 = Start State, $q_0 \in Q$ (member of, not Epsilon.)

F = Accept States, $F \subseteq Q$

δ = Transition Function

$$\delta : Q \times \Sigma_\epsilon \rightarrow P(Q)$$

Q: *is every DFA also an NFA??*

A: technically, no!

$\delta_{NFA}(p, a) = \{\delta_{DFA}(p, a)\}$

- two differences on δ
  - $Q \times (\Sigma \cup \{\varepsilon\})$
  - $P(Q)$
- is δ total or partial function?
  - no difference, since $\varnothing \in P(Q)$
  - no need to define trap state

$$\delta : Q \times \Sigma_\epsilon \rightarrow P(Q)$$

BEFORE:

States	a	b	c
q_0	q_1	q_2	q_4
q_1	q_2	q_0	q_1
⋮	⋮	⋮	⋮

FOR NFA:

States	a	b	c	ϵ
q_0	$\{q_1, q_2\}$	$\varnothing$	$\{q_4, q_2\}$	$\{q_0\}$
q_1	$\{q_4, q_2, q_1\}$	$\{q_5, q_0\}$	$\varnothing$	$\{q_4, q_3\}$
⋮	⋮	⋮	⋮	⋮

Computation of NFA w/o epsilon

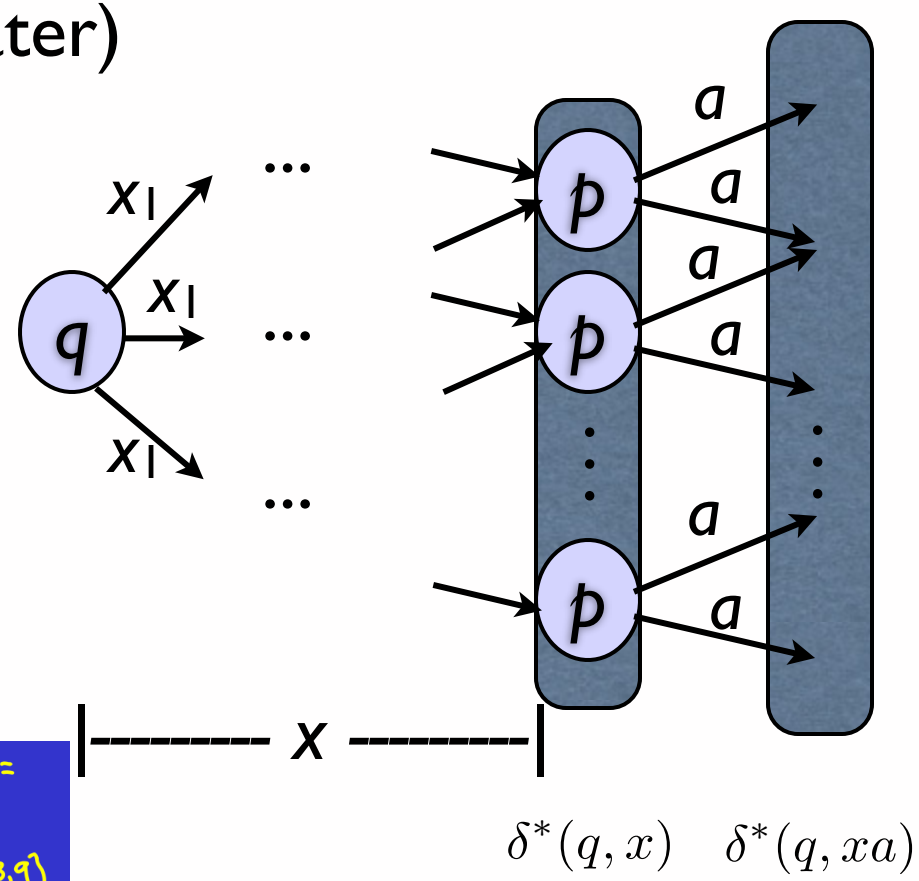
- string w is accepted iff. $\delta^*(q_0, w) \cap F \neq \varnothing$
- without epsilon transitions, δ^* is easy to define
- with epsilon transitions, you need to define epsilon-closure (very hard; will come back later)

$$\delta^*(q, \epsilon) = \{q\}$$

$$\delta^*(q, xa) = \bigcup_{p \in \delta^*(q,x)} \delta(p, a)$$

alternatively

$$\delta^*(q, ax) = \bigcup_{p \in \delta(q,a)} \delta^*(p, x)$$

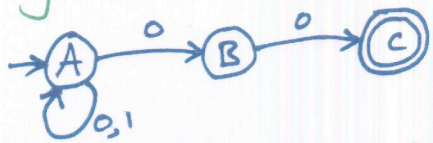


$\bigcup_{t \in \{1,2,3\}} \{t, t^2\} = \{1,1\} \cup \{2,4\} \cup \{3,9\}$

NFA => DFA w/o epsilon

EXAMPLE

Accept all strings over $\{0,1\}^*$ ending with "00."



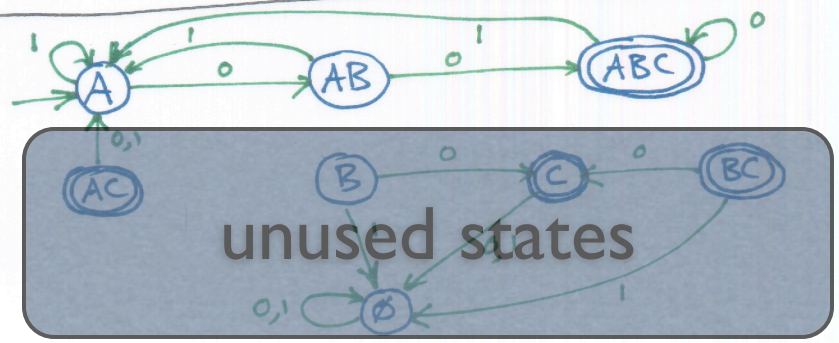
String to check: 00100

Simulate the execution. Put a finger on any state we could be in.

$\emptyset$ A B C AB BC AC ABC

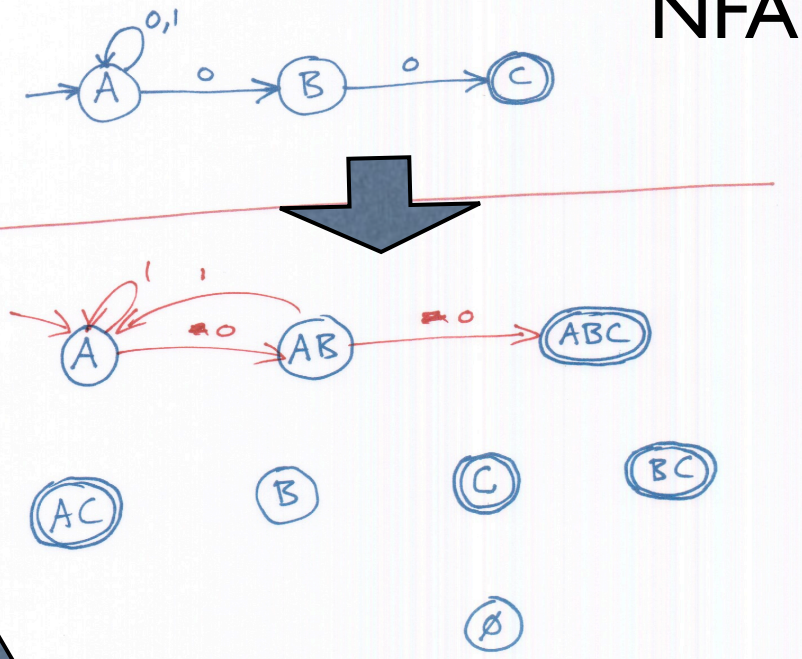
Let N = Number of states in NFA.
What is the (worst case) number of states in the equivalent DFA?

Use Next Slide



- how to convert NFA to DFA?

NFA



NFA => DFA w/o epsilon

THEOREM

EVERY NONDETERMINISTIC FSM HAS AN EQUIVALENT DETERMINISTIC FSM.

"EQUIVALENT" = Recognizes the same language.

PROOF BY CONSTRUCTION

Given a NFA, let's show how to build an equivalent DFA.

Let $M = (Q, \Sigma, \delta, q_0, F)$

This is the NFA we're given.

Construct

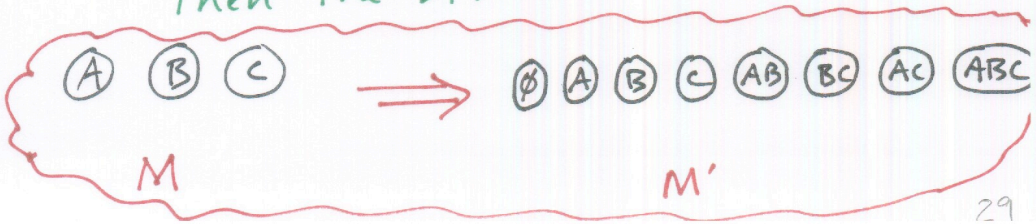
$M' = (Q', \Sigma, \delta', q_0', F')$

This is the DFA we're buiding.

Where...

$Q' = P(Q)$

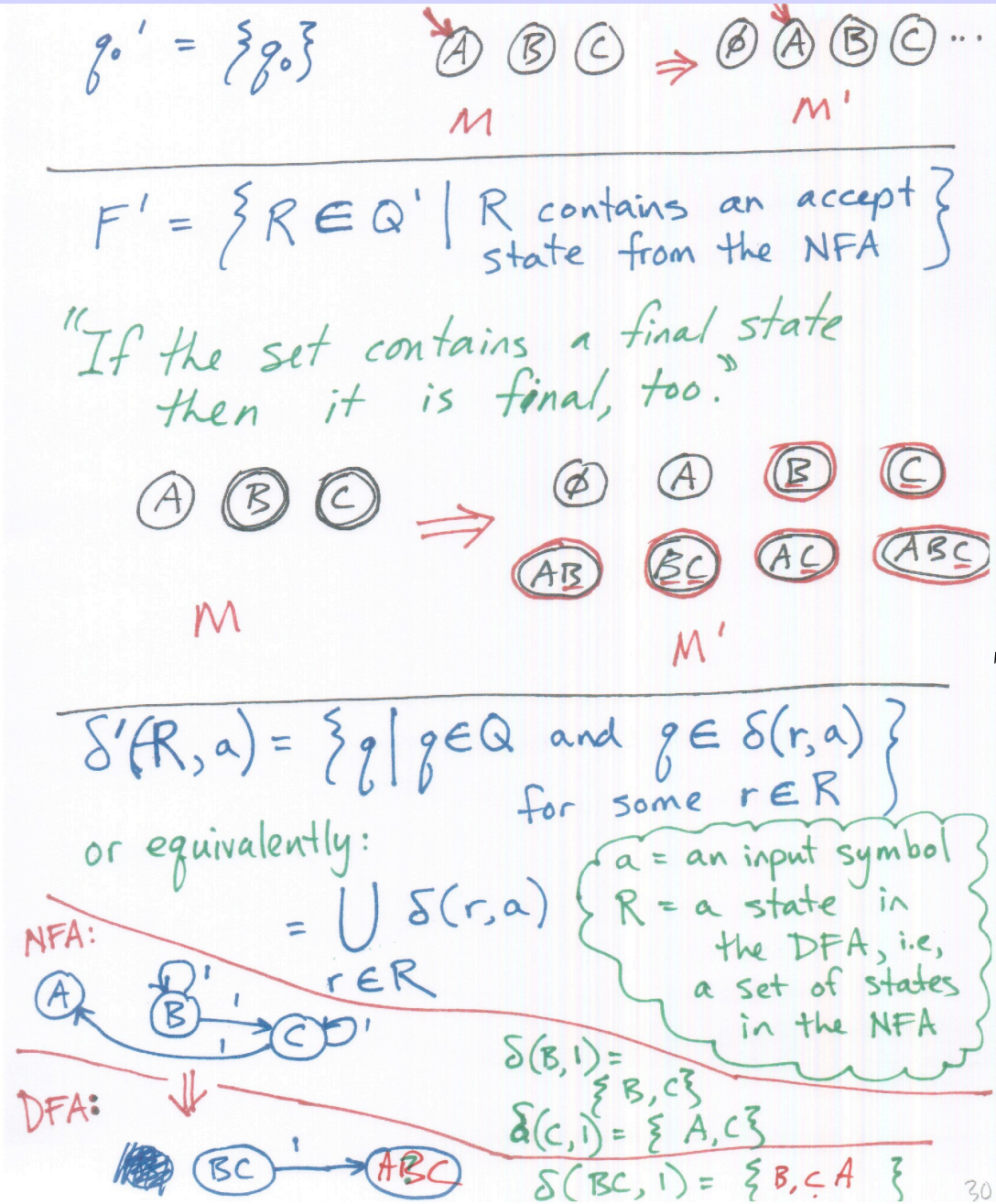
Assume NFA has K states.
Then the DFA will have 2^K states.



29

- try make a NFA for {ab, aba}* and then convert to DFA

Formally: subset construction



NFA $M = (Q, \Sigma, \delta, q_0, F)$

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

where $Q' = P(Q)$

$q_0' = \{q_0\}$

$F' = \{A \in P(Q) \mid A \cap F \neq \emptyset\}$

$\delta'(R, a) = \bigcup_{r \in R} \delta(r, a)$

Thm.: $L(M) = L(M')$

Idea: $\forall w \in L(M), \exists p \in F,$

s.t. $\delta^*(q_0, w) = p.$

want to show: $p \in \delta'^*(q_0, w)$

also need the other direction

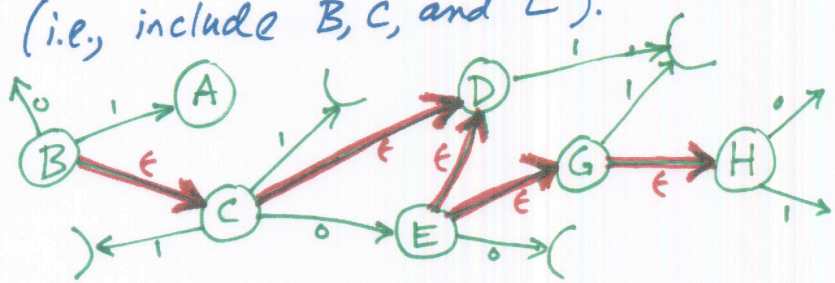
Epsilon-Closure

BUT WHAT ABOUT ϵ-EDGES?

Consider a state in the DFA that we're building.

BCE ← A state "R" in the DFA is a set of states from the NFA.

Look back at M, the NFA.
What states can we reach by going through ϵ-edges?
Also include the states we're in.
(i.e., include B, C, and E).



DEFINE "EPSILON-CLOSURE"

$$E(R) = \{ q \in Q \mid q \text{ can be reached from a state in } R \text{ by following zero or more } \epsilon\text{-edges.} \}$$

EXAMPLE

$$E(\text{BCE}) = \{B, C, D, E, G, H\} = \text{BCDEGH}$$

Modify the transition function:

$$\delta'(R, a) = \{ q \in Q \mid q \in E(\delta(r, a)) \text{ for some } r \in R \}$$

Also, modify the start state in the constructed DFA:

$$q_0' = E(\{q_0\})$$

END OF PROOF