CS103, Autumn 2026

Lecture 15:

Finite Automata

Part 2 of 3

Finite Automata

Part 2

1. **Recap from Last Time**
2. Tabular DFAs
3. Regular Languages
4. Complementing Regular Languages
5. Announcements
6. Beyond DFAs: Nondeterminism
7. ε-Transitions
8. Nondeterminism Intuitions
9. Designing NFAs
10. What’s Next?

Suppose you know the following:

$x \in \Sigma \qquad y \in \Sigma^*$

Which of the following options is correct?

(A) x is a character and y is a character.
(B) x is a character and y is a string.
(C) x is a string and y is a character.
(D) x is a string and y is a string.
(E) None of these

Answer at ***https://cs103.stanford.edu/pollev***

Formal Language Theory

- An ***alphabet*** is a set, usually denoted Σ, consisting of elements called ***characters***.
  - $a \in \Sigma$ means "a is a single character."
- A ***string over Σ*** is a finite sequence of zero or more characters taken from Σ.
- The ***empty string*** has no characters and is denoted ε.
- A ***language over Σ*** is a set of strings over Σ.
- The language **Σ*** is the set of all strings over Σ.
  - $w \in \Sigma^*$ means "w is a string of characters from Σ."

The Language of an Automaton

- If A is an automaton that processes strings over Σ, the ***language of A***, denoted $\mathcal{L}(A)$, is the set of all strings A accepts.
- Formally:

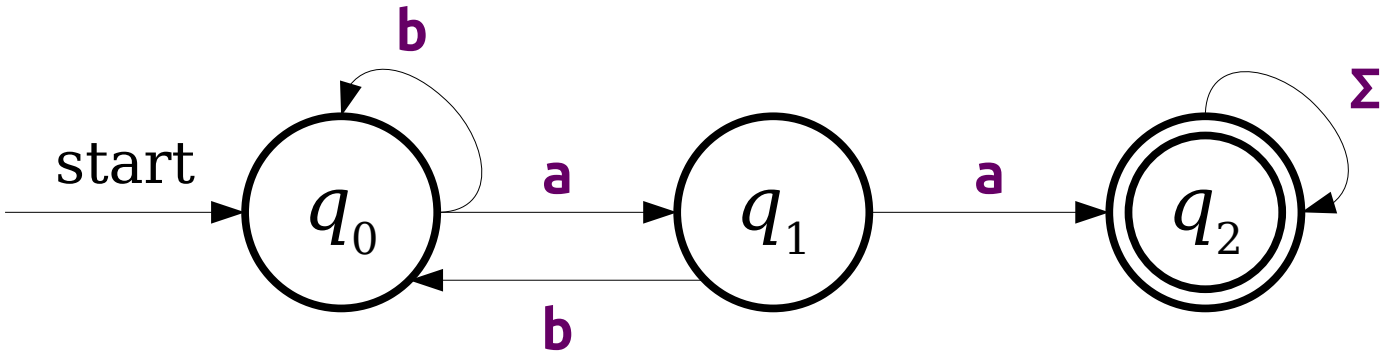
$$\mathcal{L}(A) = \{ w \in \Sigma^* \mid A \text{ accepts } w \}$$

DFAs

- A ***DFA*** is a
  - ***D***eterministic
  - ***F***inite
  - ***A***utomaton
- A DFA is defined relative to some alphabet Σ.
- For each state in the DFA, there must be ***exactly one*** transition defined for each symbol in Σ.
  - This is the “deterministic” part of DFA.
- There is a unique start state.
- There are zero or more accepting states.

Recognizing Languages with DFAs

$L = \{ w \in \{\texttt{a}, \texttt{b}\}^* \mid w$ contains $\texttt{aa}$ as a substring $\}$

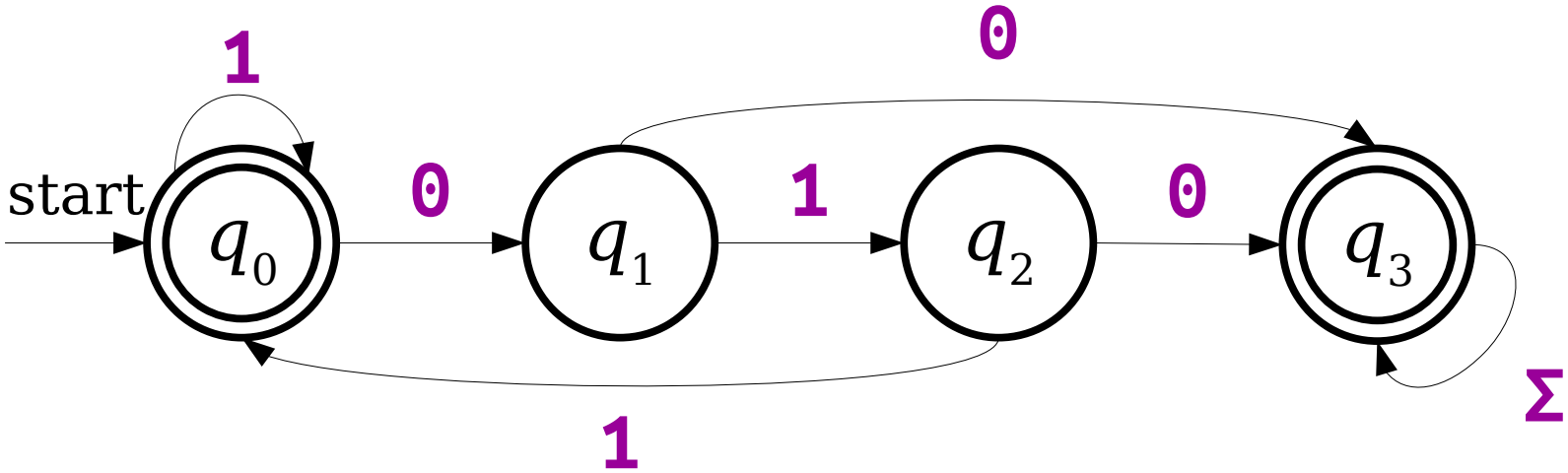


Finite Automata

Part 2

1. Recap from Last Time
2. **Tabular DFAs**
3. Regular Languages
4. Complementing Regular Languages
5. Announcements
6. Beyond DFAs: Nondeterminism
7. ε-Transitions
8. Nondeterminism Intuitions
9. Designing NFAs
10. What’s Next?

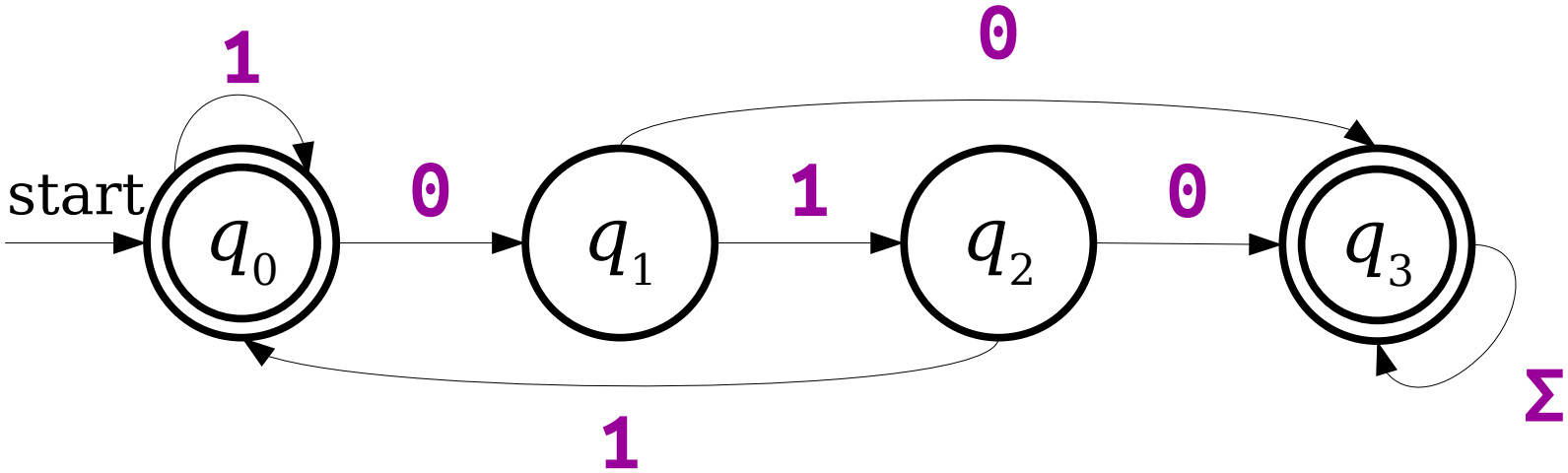
Tabular DFAs



	0	1
$*q_0$	q_1	q_0
q_1	q_3	q_2
q_2	q_3	q_0
$*q_3$	q_3	q_3

These stars indicate accepting states.

Tabular DFAs



	0	1
$*q_0$	q_1	q_0
q_1	q_3	q_2
q_2	q_3	q_0
$*q_3$	q_3	q_3

Since this is the first row, it's the start state.

Simulating a DFA

```
int kTransitionTable[kNumStates][kNumSymbols] = {
     {0, 0, 1, 3, 7, 1, …},
      …
};

bool kAcceptTable[kNumStates] = {
    false,
    true,
    true,
    …
};

bool accepts(string input) {
    int state = 0;
    for (char ch: input) {
        state = kTransitionTable[state][ch];
    }
    return kAcceptTable[state];
}
```

Finite Automata

Part 2

1. Recap from Last Time
2. Tabular DFAs
3. **Regular Languages**
4. Complementing Regular Languages
5. Announcements
6. Beyond DFAs: Nondeterminism
7. ε-Transitions
8. Nondeterminism Intuitions
9. Designing NFAs
10. What’s Next?

The Regular Languages

A language L is a ***regular language*** when there exists a DFA D such that $\mathcal{L}(D) = L$.

If L is a language and $\mathcal{L}(D) = L$, we say that D ***recognizes*** the language L.

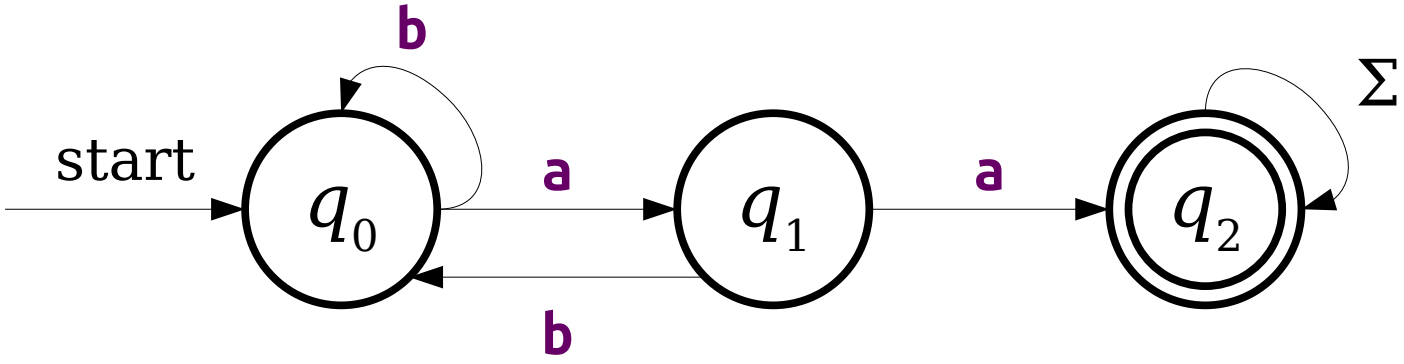
Finite Automata

Part 2

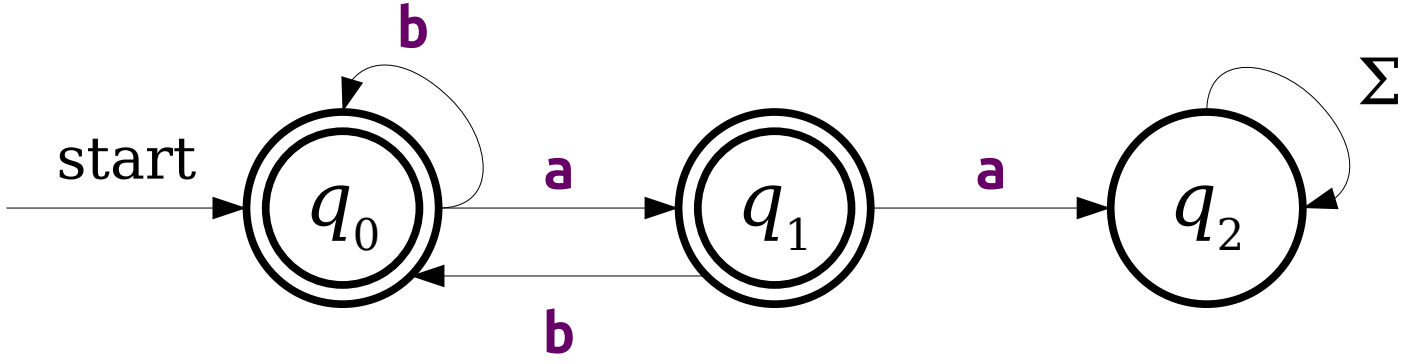
1. Recap from Last Time
2. Tabular DFAs
3. Regular Languages
4. **Complementing Regular Languages**
5. Announcements
6. Beyond DFAs: Nondeterminism
7. ε-Transitions
8. Nondeterminism Intuitions
9. Designing NFAs
10. What’s Next?

Complementing Regular Languages

$L = \{ w \in \{\texttt{a}, \texttt{b}\}^* \mid w$ contains **aa** as a substring $\}$

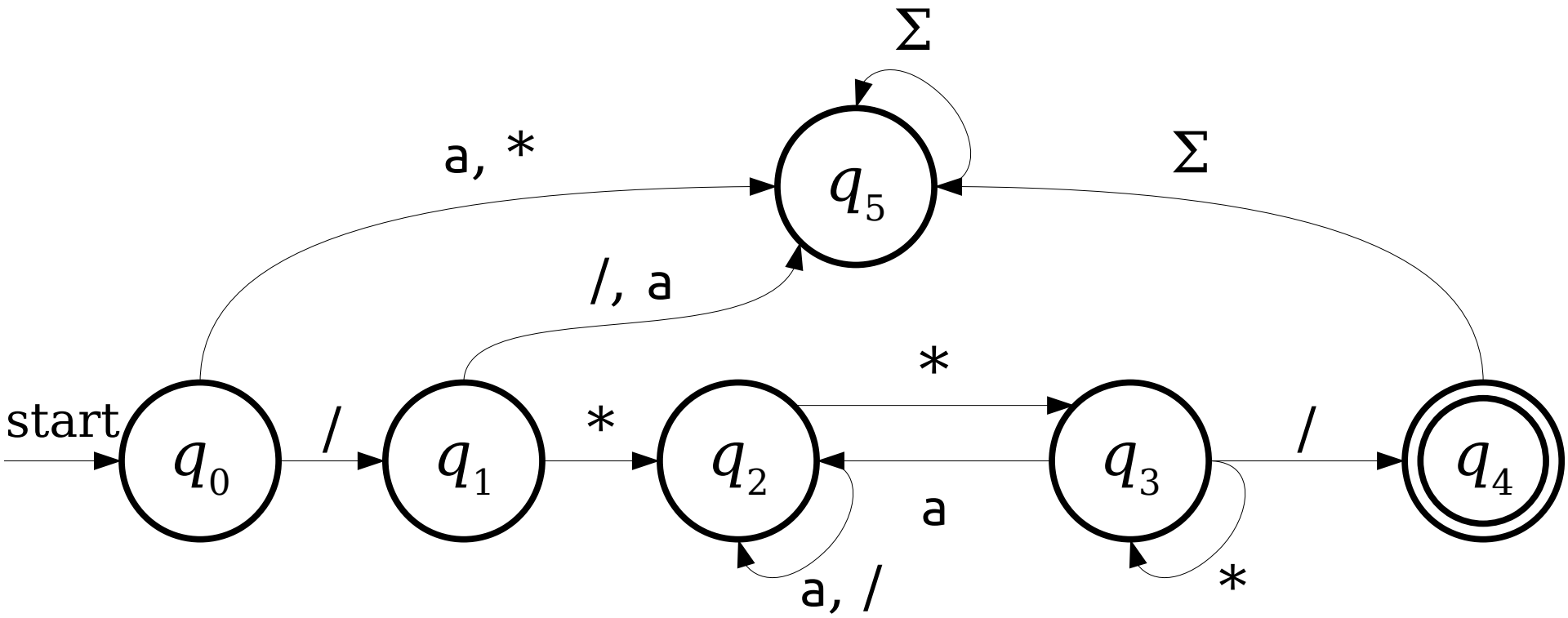


$\overline{L} = \{ w \in \{\texttt{a}, \texttt{b}\}^* \mid w$ ***does not*** contain **aa** as a substring $\}$



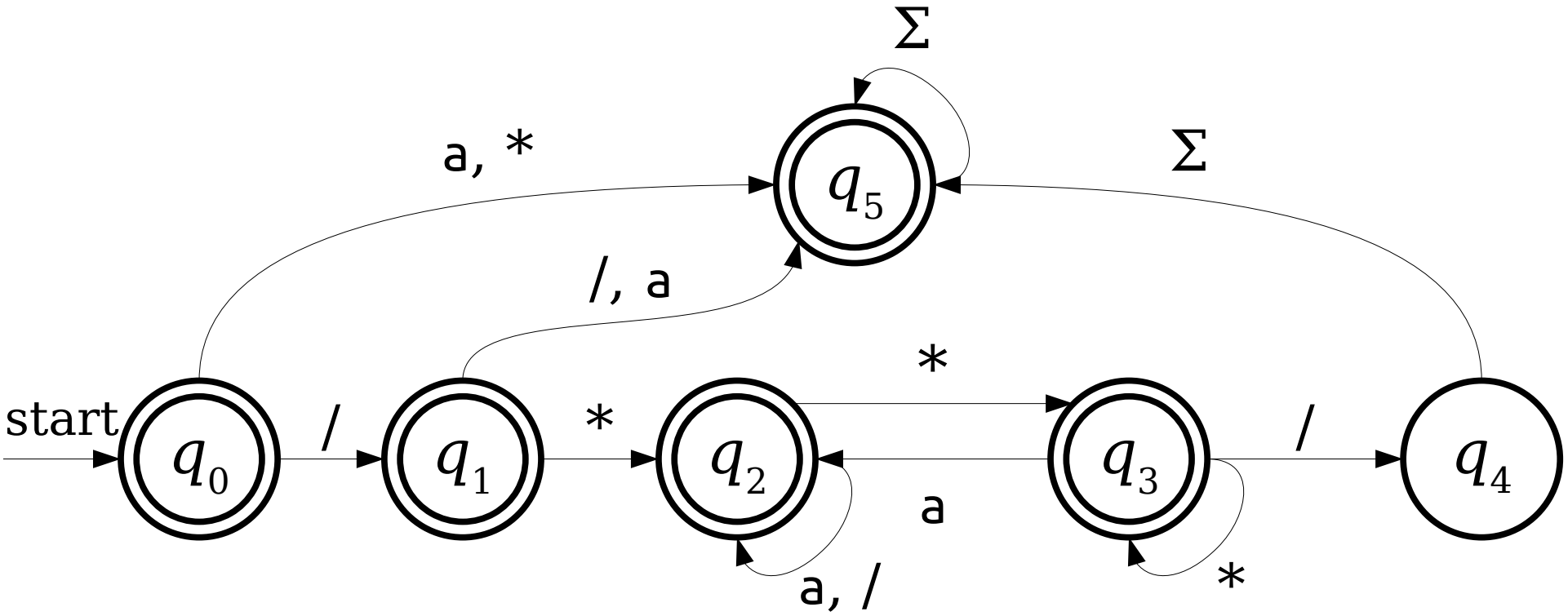
Complementing Regular Languages

$L = \{ w \in \{\texttt{a}, \texttt{*}, \texttt{/}\}^* \mid w$ represents a C-style comment $\}$



Complementing Regular Languages

$\overline{L}$ = { $w \in$ {a, *, /}* | w ***doesn't*** represent a C-style comment }



The Complement of a Language

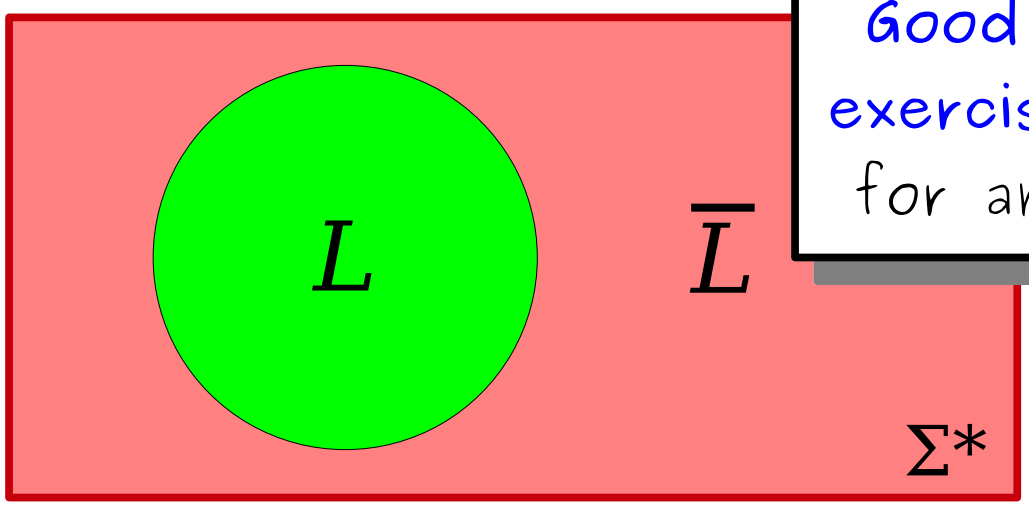
- Given a language $L \subseteq \Sigma^*$, the ***complement*** of that language (denoted $\overline{L}$) is the language of all strings in Σ^* that aren't in L.
- Formally:

$$\overline{L} = \Sigma^* - L$$

The Complement of a Language

- Given a language $L \subseteq \Sigma^*$, the ***complement*** of that language (denoted $\overline{L}$) is the language of all strings in Σ^* that aren't in L.
- Formally:

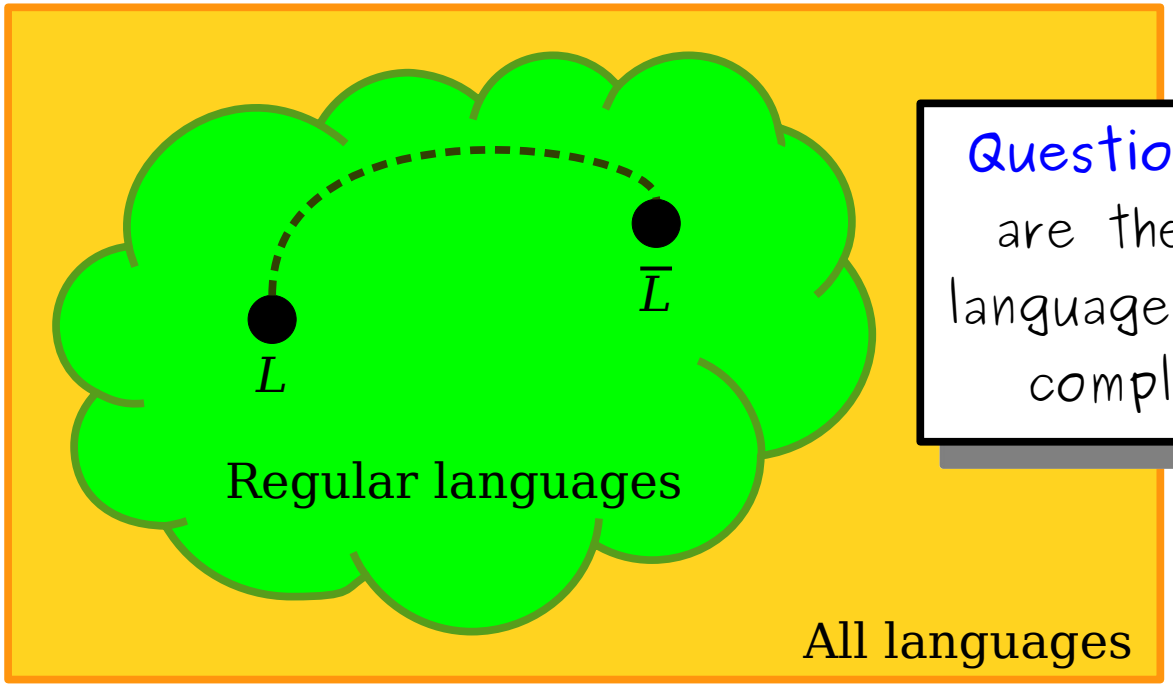
$$\overline{L} = \Sigma^* - L$$



Good proofwriting exercise: prove $\overline{\overline{L}} = L$ for any language L.

Closure Properties

- ***Theorem:*** If L is a regular language, then $\overline{L}$ is also a regular language.
  - ("The regular languages are ***closed under complementation***.")
- ***Proof idea:*** Show that swapping the accepting and rejecting states of a DFA for L gives a DFA for $\overline{L}$.



Question to ponder: are the *nonregular* languages closed under complementation?

Finite Automata

Part 2

1. Recap from Last Time
2. Tabular DFAs
3. Regular Languages
4. Complementing Regular Languages
5. **Announcements**
6. Beyond DFAs: Nondeterminism
7. ε-Transitions
8. Nondeterminism Intuitions
9. Designing NFAs
10. What’s Next?

Midterm Regrade Requests

- Regrade requests for the exam open at 8:00 PM today and close next Wednesday at 11:59 PM.

  ☞ ***Read the instructions before submitting regrade requests***. ☜

  - Regrade requests are intended as a way to correct grading errors. We're happy to correct mistakes we've made.
  - If you're looking to talk to someone about the exam, or discuss plans for the tail end of the quarter, or just get a pep talk, that's best handled outside the regrade request system.
  - To make sure regrade requests are used as intended, we're strict about their structure / formatting.

Other Midterm-Related Things

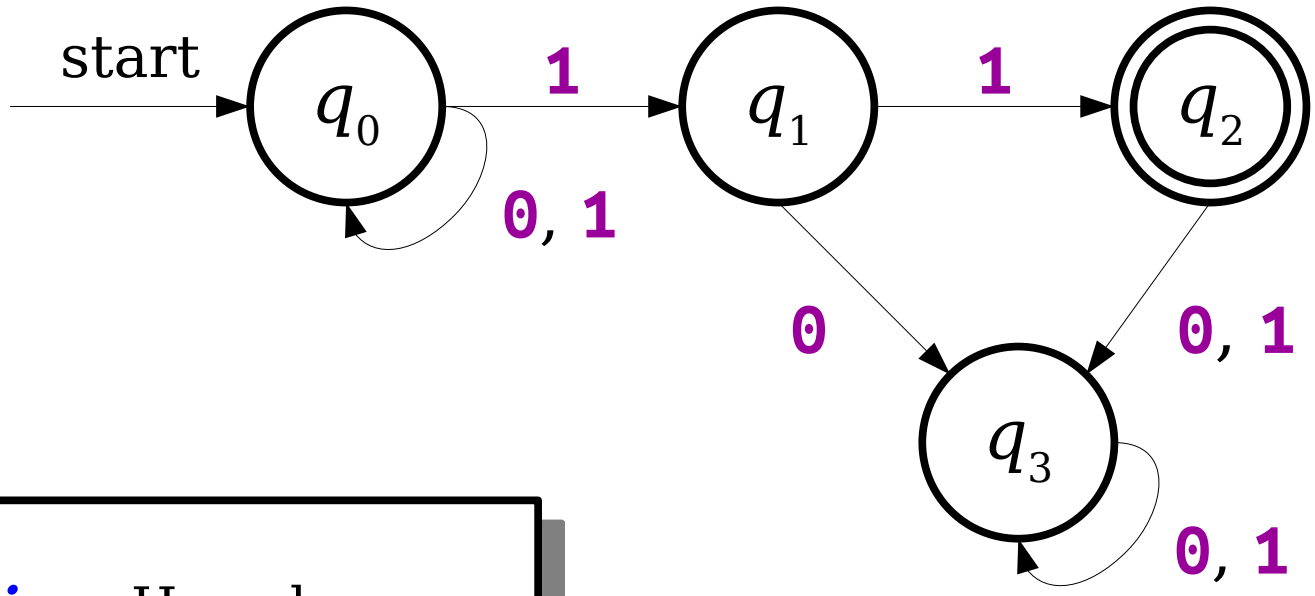
- Solutions are posted.
- See Joyce's announcement on Ed about setting up a 1-on-1 if you feel you need some extra support. Sign-up deadline: Friday, 1 PM.
- I've also posted a fancy Grade Cruncher.
  - See also: recitation dashboard.
- A note about the 50% weighted exam average policy.
- Problem Set 5 grades should be available Friday.
  - We extended our grading schedule to account for all the hours the TA team put into midterm proctoring, grading, and other activities last week.

Finite Automata

Part 2

1. Recap from Last Time
2. Tabular DFAs
3. Regular Languages
4. Complementing Regular Languages
5. Announcements
6. **Beyond DFAs: Nondeterminism**
7. ε-Transitions
8. Nondeterminism Intuitions
9. Designing NFAs
10. What’s Next?

The Motivation



Question: How do we interpret an automaton like this one?

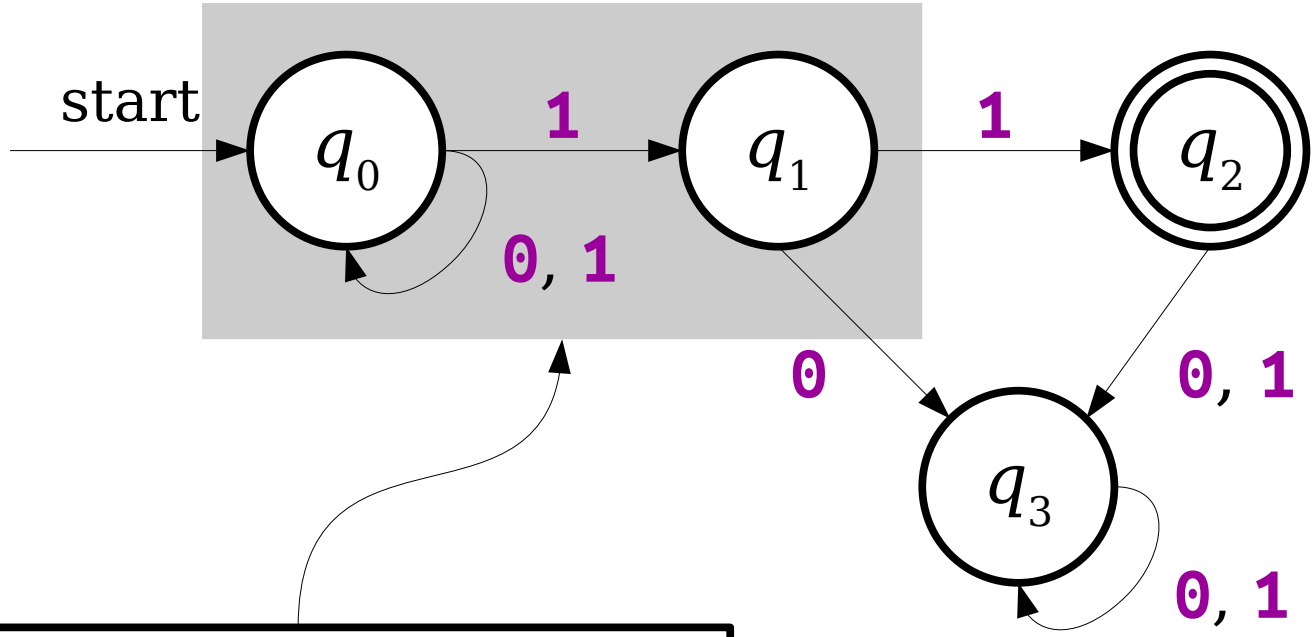
NFAs

- An ***NFA*** is a
  - ***N***ondeterministic
  - ***F***inite
  - ***A***utomaton
- NFAs are structurally similar to a DFA, but represents a fundamental shift in how we'll think about computation.

(Non)determinism

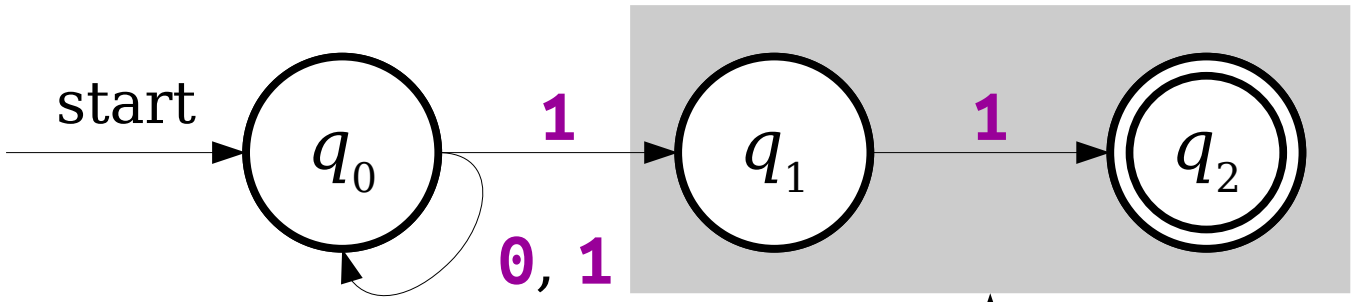
- A model of computation is ***deterministic*** if at every point in the computation, there is exactly one choice that can make.
  - The machine accepts if that series of choices leads to an accepting state.
- A model of computation is ***nondeterministic*** if the computing machine has a finite number of choices available to make at each point, possibly including zero.
- The machine accepts if ***any*** series of choices leads to an accepting state.
  - (This sort of nondeterminism is technically called ***existential nondeterminism***, the most philosophical-sounding term we'll introduce all quarter.)
- This idea was introduced by Michael Rabin and Data Scott as an internship project (!) at IBM in 1957. It won them the Turing Award (the "Nobel Prize of Computer Science") in 1976.

A Simple NFA



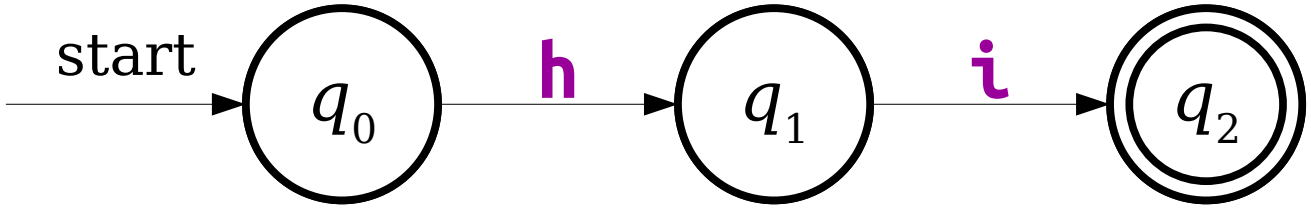
q_0 has two transitions defined on 1!

A More Complex NFA



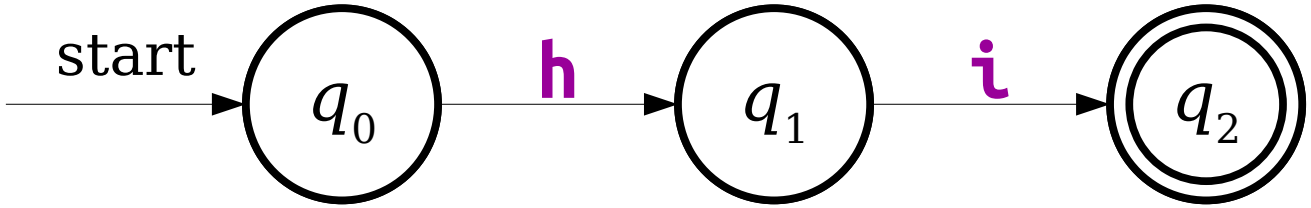
If a NFA needs to make a transition when no transition exists, the automaton dies and that particular path does not accept.

Hello, NFA!

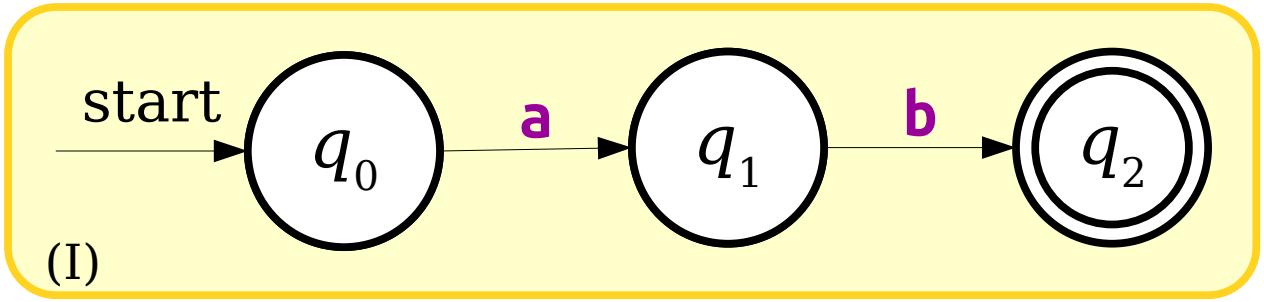


h	i

Tragedy in Paradise

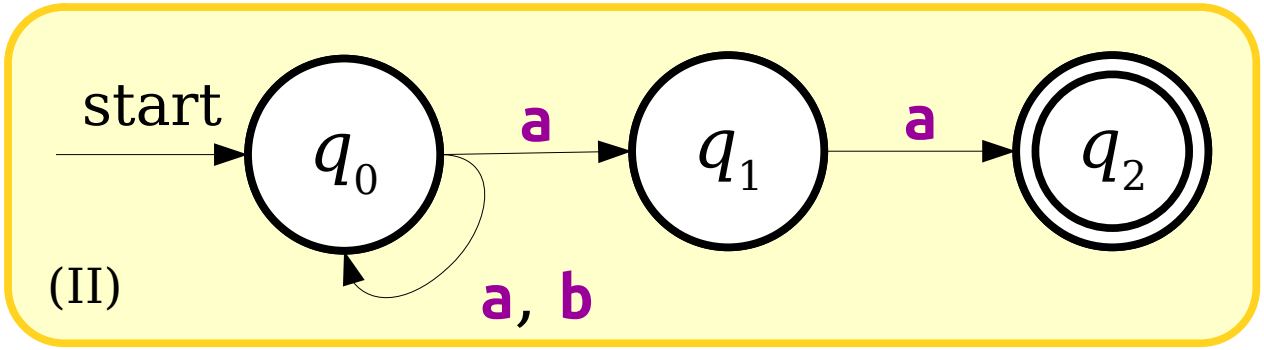


h	i	p

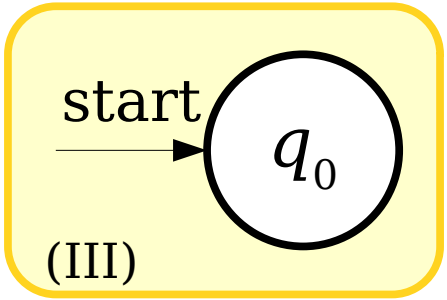


{ab}

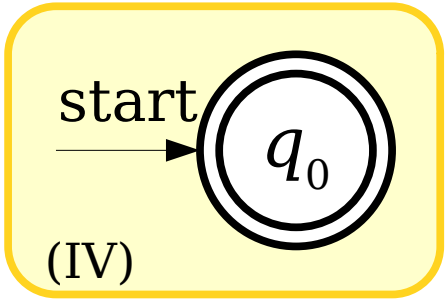
Question to ponder:
Why is the answer
{ $w \in \Sigma^*$ | w ends in aaa }
not correct?



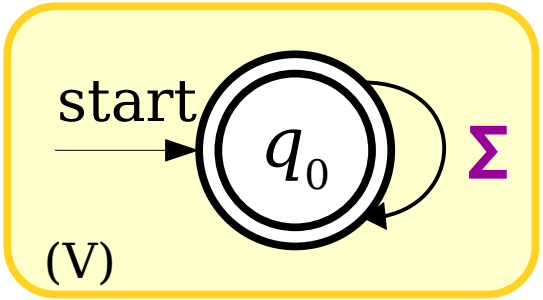
{ $w \in \Sigma^*$ | w ends in aa }



∅



{ε}



Σ*

The ***language of an NFA*** is

$$\mathcal{L}(N) = \{ w \in \Sigma^* \mid N \text{ accepts } w \}.$$

What is the language of each NFA? (Assume Σ = {a, b}.)
Answer at ***https://cs103.stanford.edu/pollev***

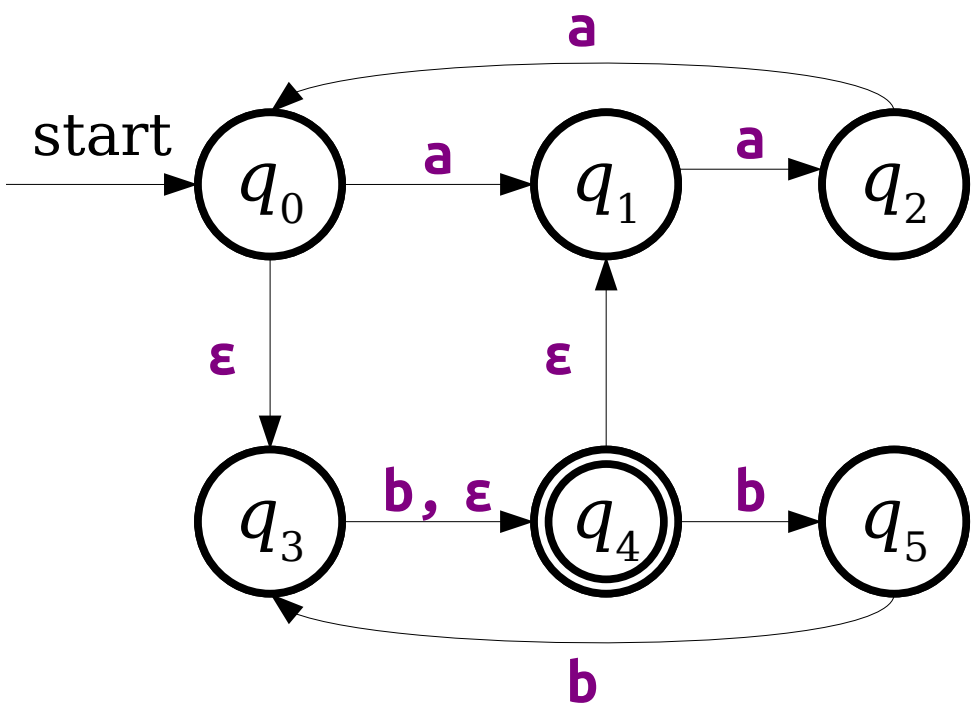
Finite Automata

Part 2

1. Recap from Last Time
2. Tabular DFAs
3. Regular Languages
4. Complementing Regular Languages
5. Announcements
6. Beyond DFAs: Nondeterminism
7. **ε-Transitions**
8. Nondeterminism Intuitions
9. Designing NFAs
10. What’s Next?

ε-Transitions

- NFAs have a special type of transition called the **ε-transition**.
- An NFA may follow any number of ε-transitions at any time without consuming any input.



Not at all fun or rewarding exercise: what is the language of this NFA?

(I actually don't know the answer. I made up this NFA just to show off ε-transitions.)

ε-Transitions

- NFAs have a special type of transition called the **ε-transition**.
- An NFA may follow any number of ε-transitions at any time without consuming any input.
- NFAs are not *required* to follow ε-transitions. It's simply another option at the machine's disposal.

NFAs

- An NFA is defined relative to some alphabet Σ.
- For each state in the NFA, there may be ***any number of*** transitions defined for each symbol in Σ, plus any number of
ε-transitions.
  - This is the “nondeterministic” part of NFA.
- There is a unique start state.
- There are zero or more accepting states.

DFAs

- A DFA is defined relative to some alphabet Σ.
- For each state in the DFA, there must be ***exactly one*** transition defined for each symbol in Σ. Additionally, ε-transitions are not allowed.
  - This is the "deterministic" part of DFA.
- There is a unique start state.
- There are zero or more accepting states.

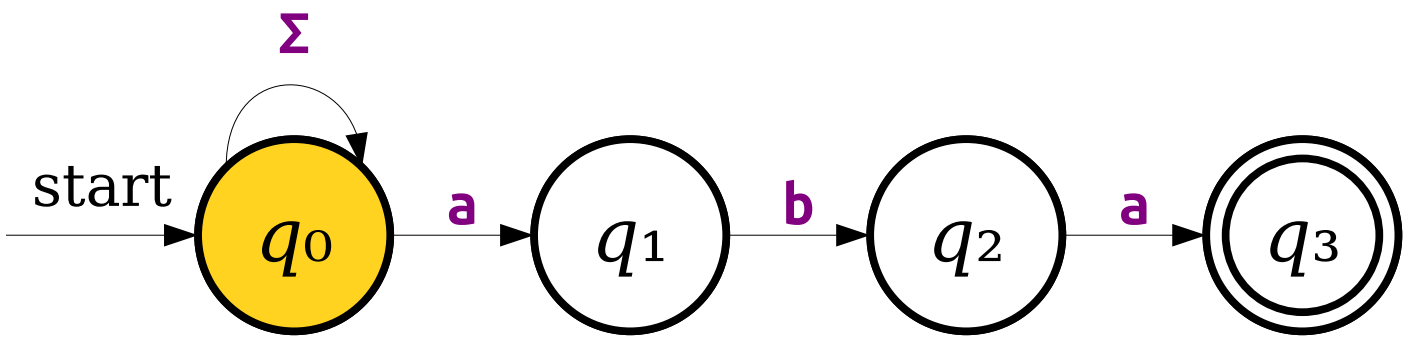
Finite Automata

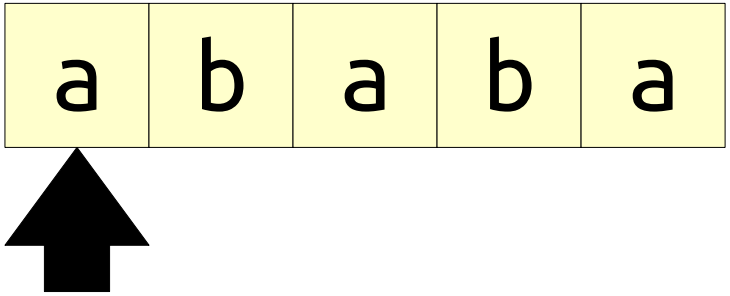
Part 2

1. Recap from Last Time
2. Tabular DFAs
3. Regular Languages
4. Complementing Regular Languages
5. Announcements
6. Beyond DFAs: Nondeterminism
7. ε-Transitions
8. **Nondeterminism Intuitions**
9. Designing NFAs
10. What’s Next?

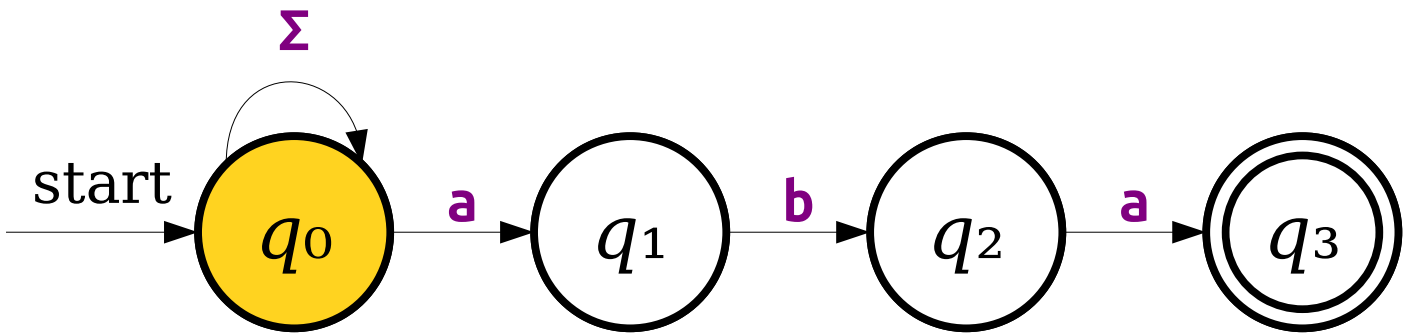
Intuition 1: Perfect Positive Guessing

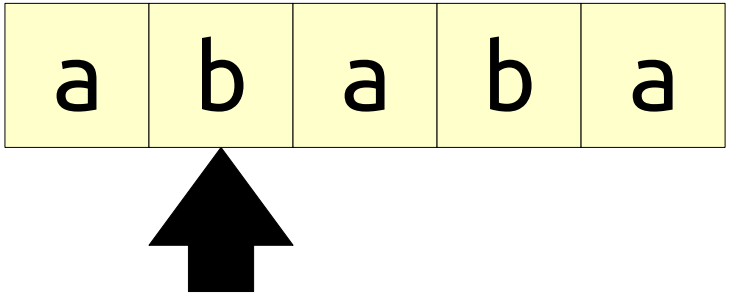
Perfect Positive Guessing



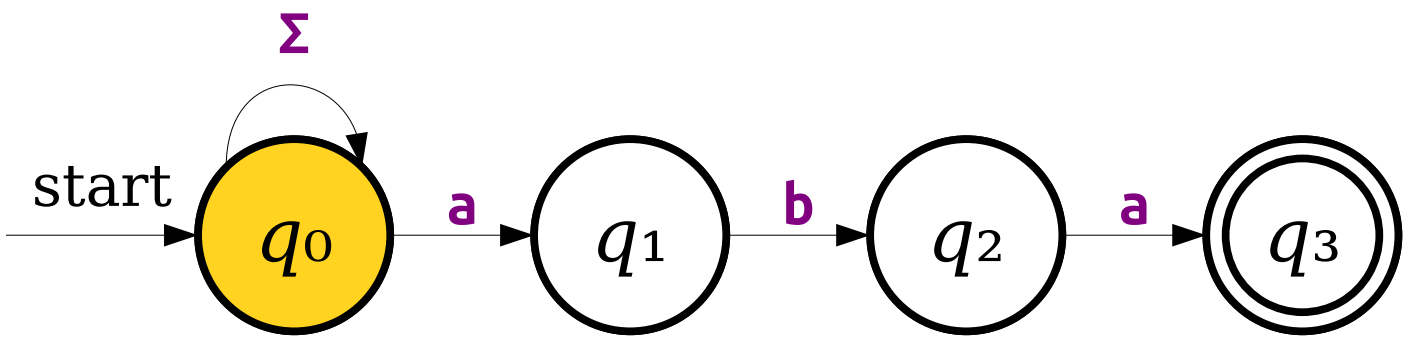


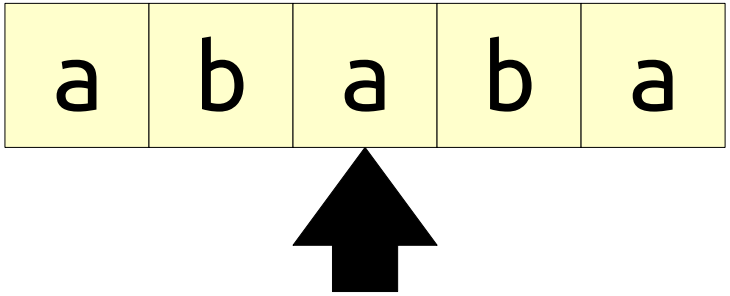
Perfect Positive Guessing



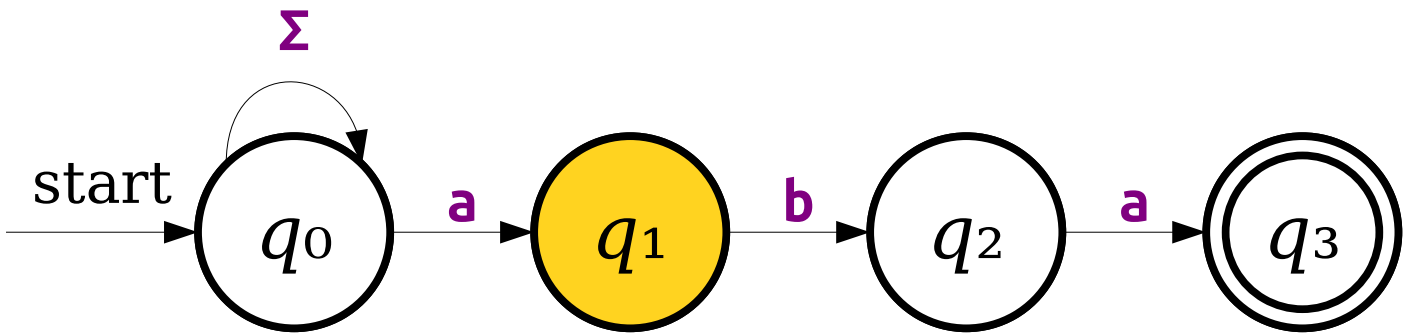


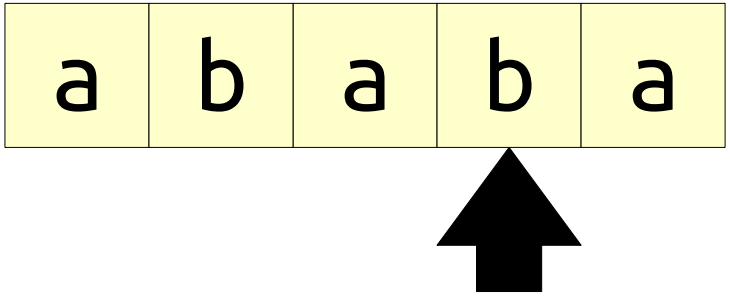
Perfect Positive Guessing



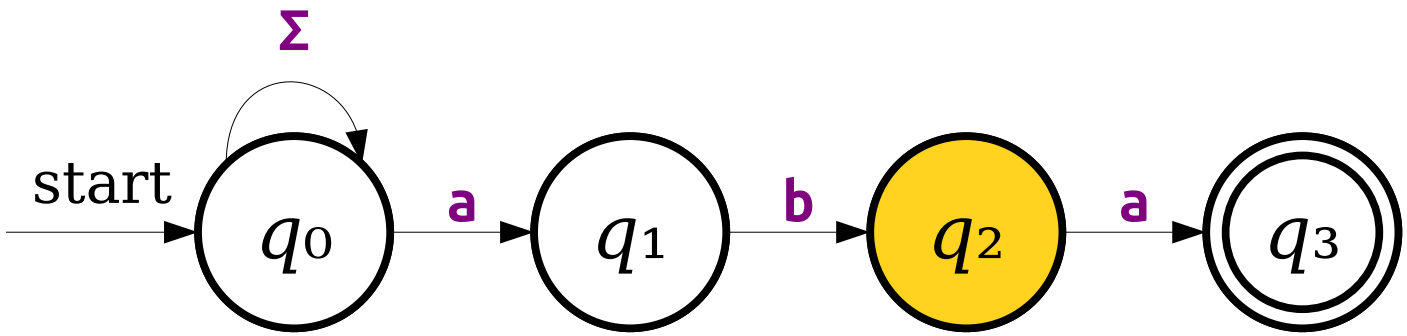


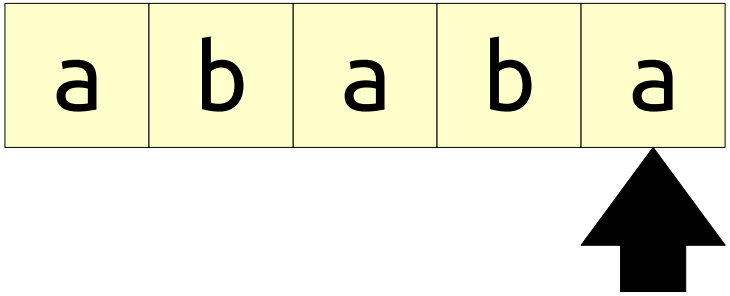
Perfect Positive Guessing



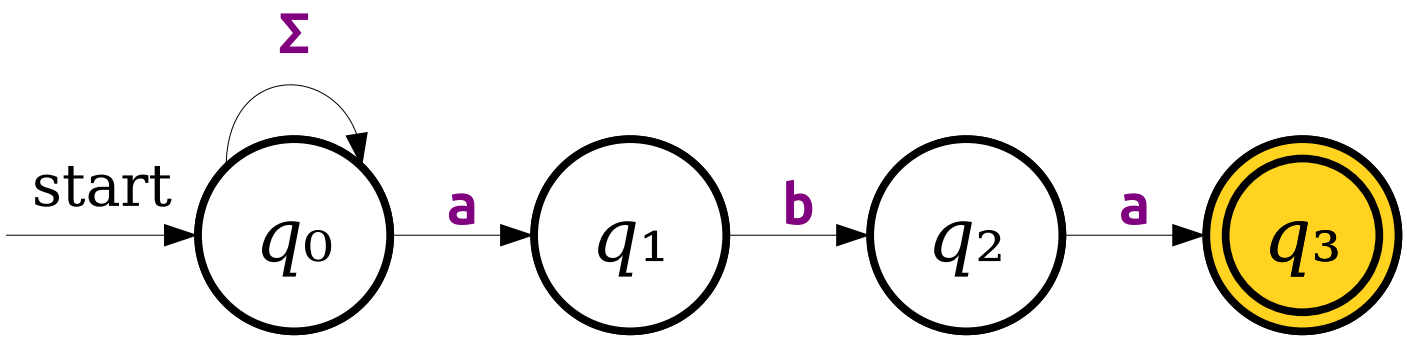


Perfect Positive Guessing





Perfect Positive Guessing



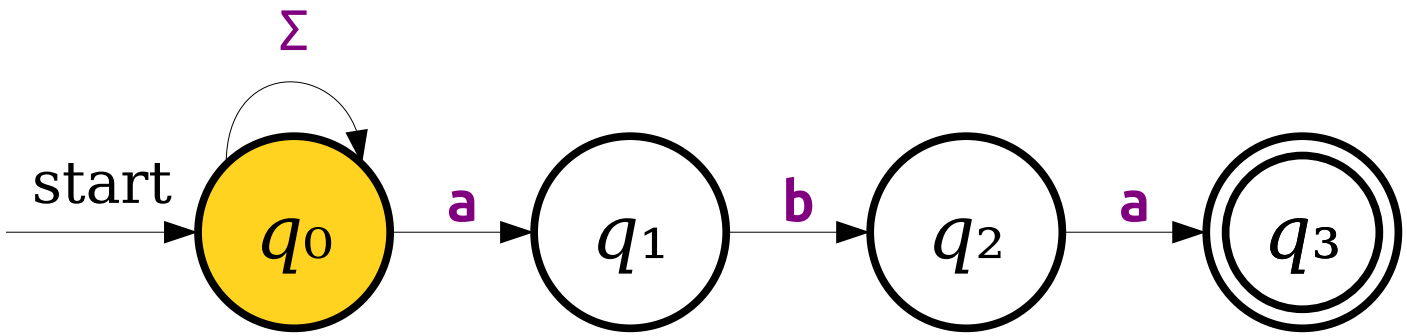
a	b	a	b	a

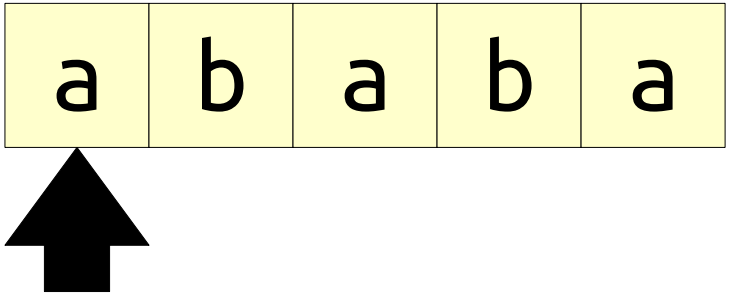
Perfect Positive Guessing

- We can view nondeterministic machines as having ***Magic Superpowers*** that enable them to guess choices that lead to an accepting state.
  - If there is at least one choice that leads to an accepting state, the machine will guess it.
  - If there are no choices, the machine guesses any one of the wrong guesses.
- There is no known way to physically model this intuition of nondeterminism – this is quite a departure from reality!
- (And no, this is not the same as a quantum computer. Come talk to me after class to learn why!)

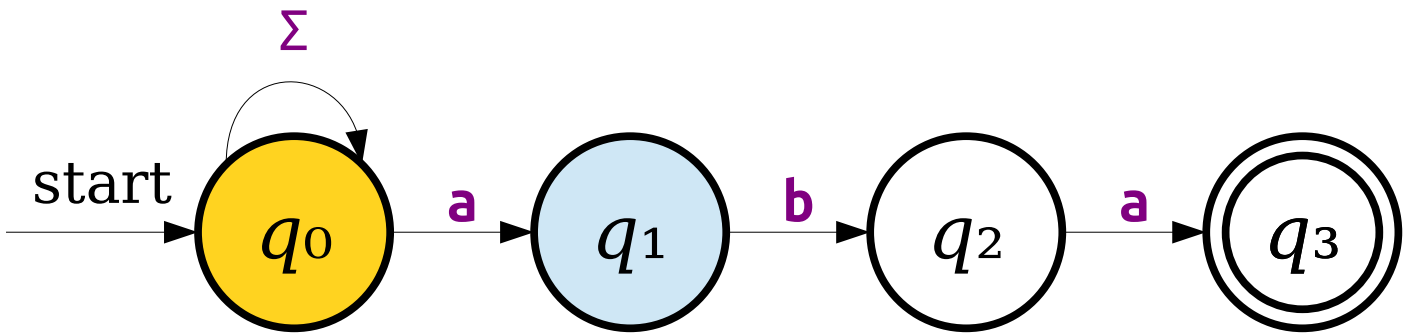
Intuition 2: Massive Parallelism

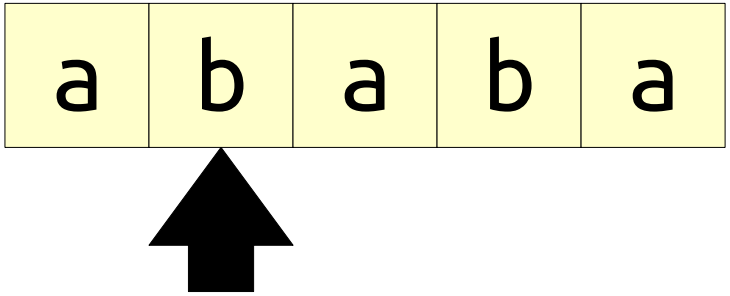
Massive Parallelism



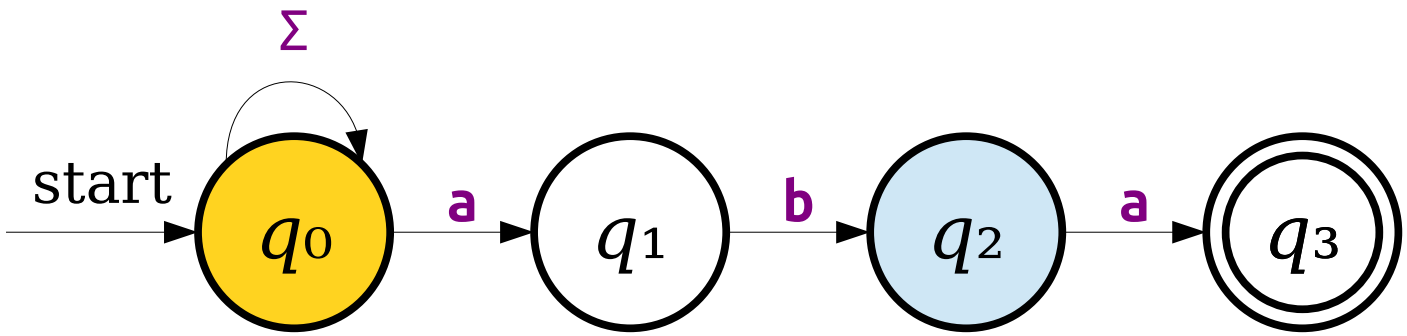


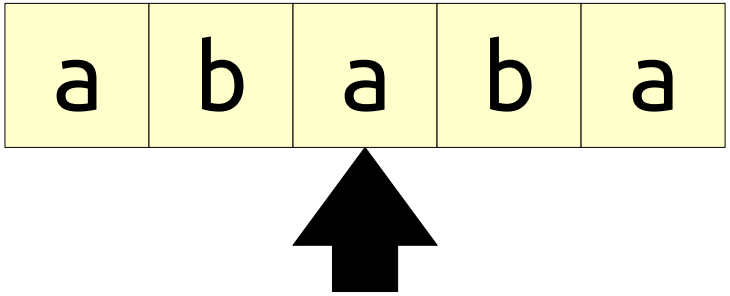
Massive Parallelism



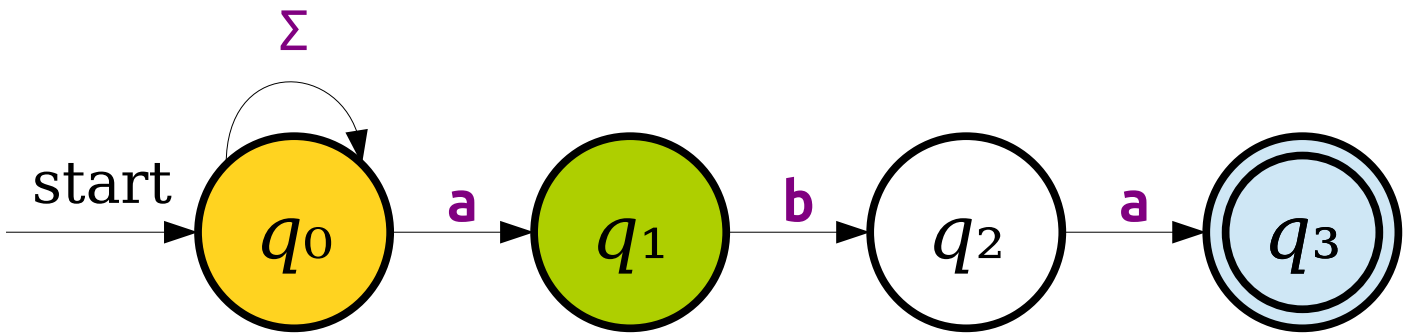


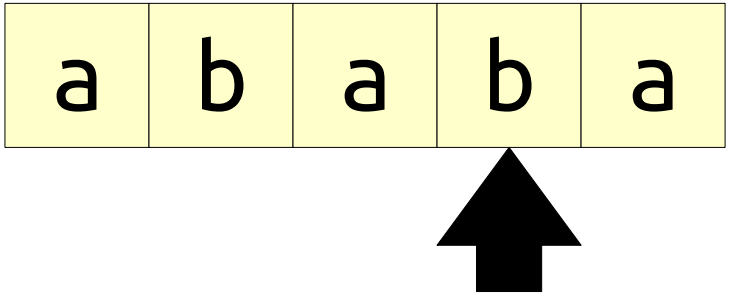
Massive Parallelism



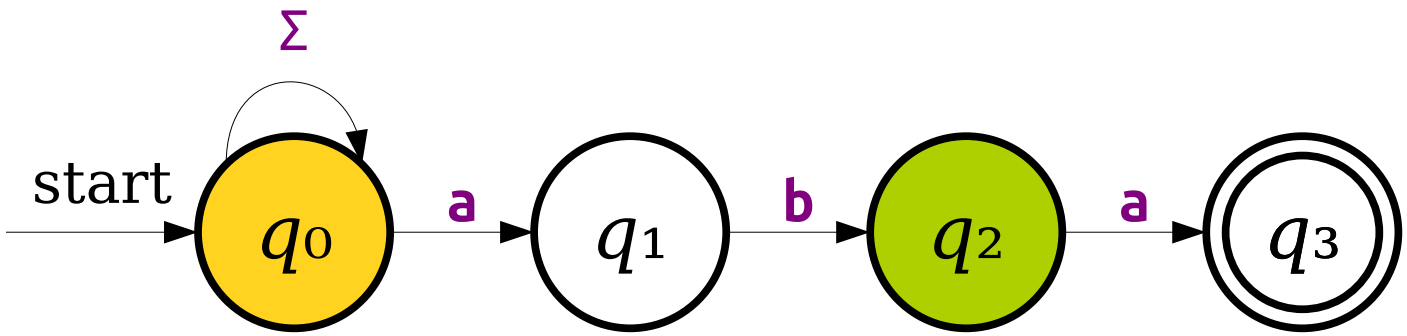


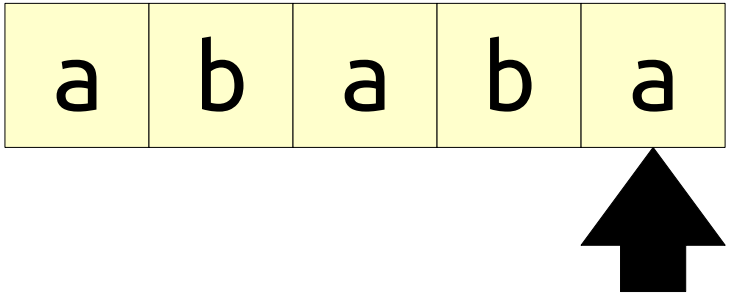
Massive Parallelism



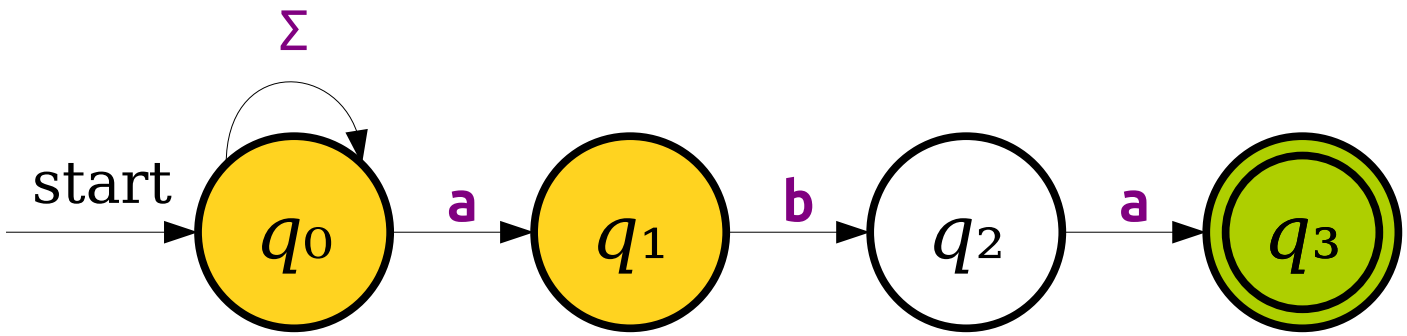


Massive Parallelism



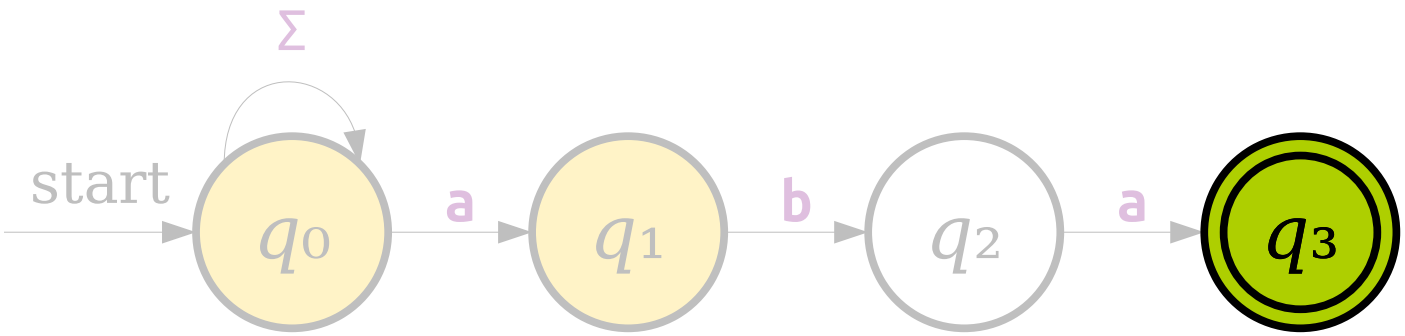


Massive Parallelism



a	b	a	b	a

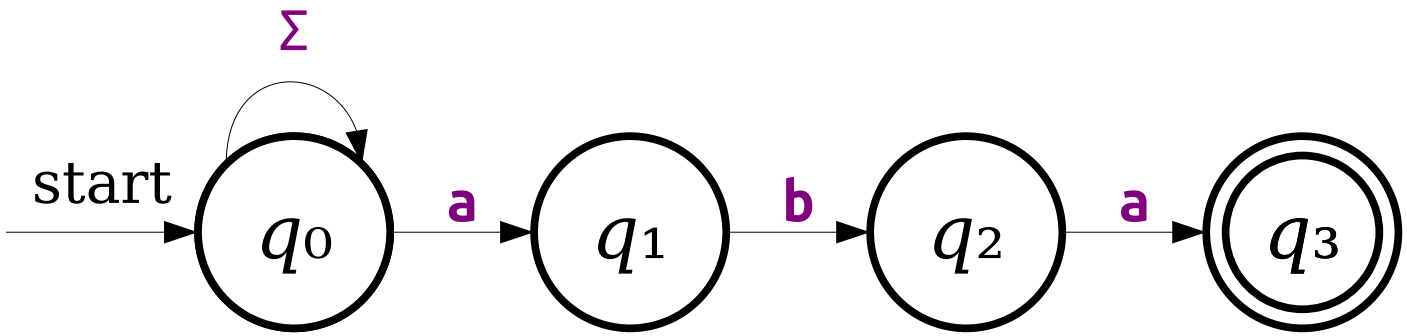
Massive Parallelism

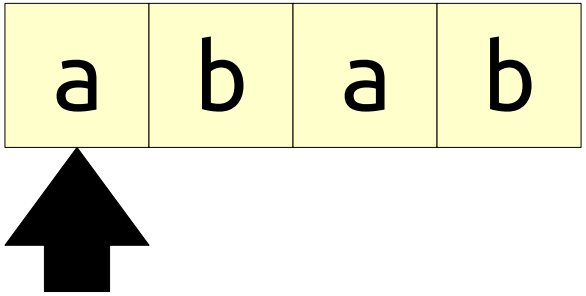


We're in at least one accepting state, so there's some path that gets us to an accepting state.

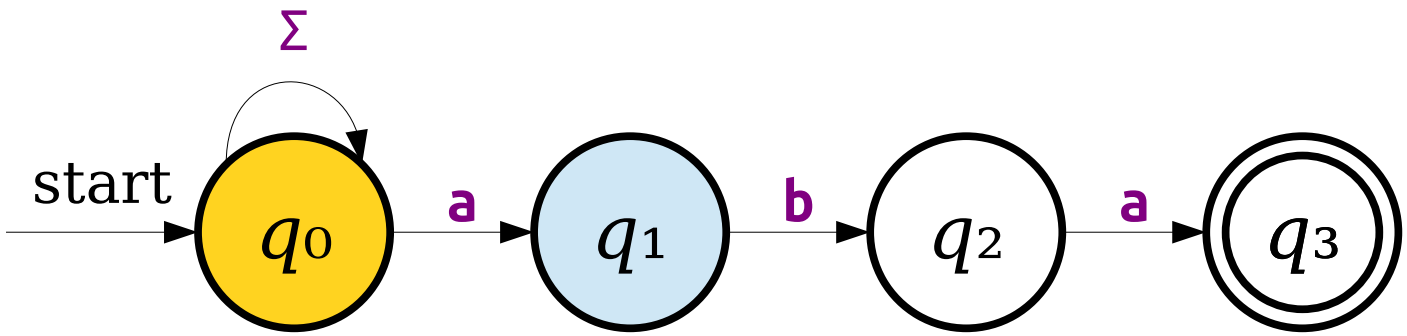
a	b	a	b	a

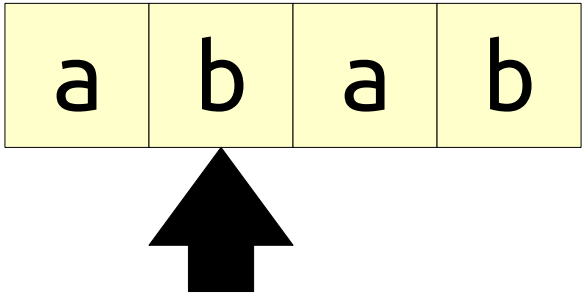
Massive Parallelism



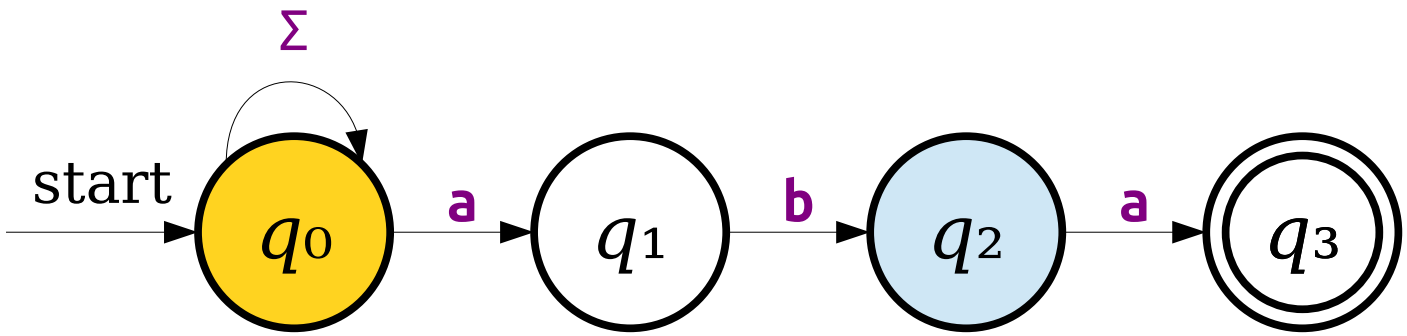


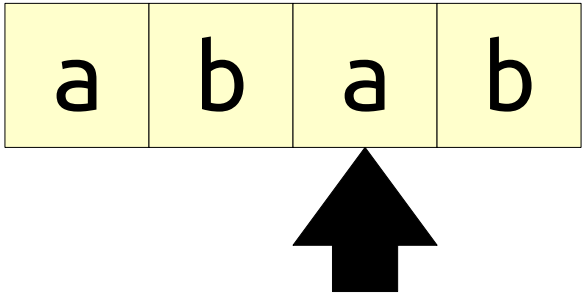
Massive Parallelism



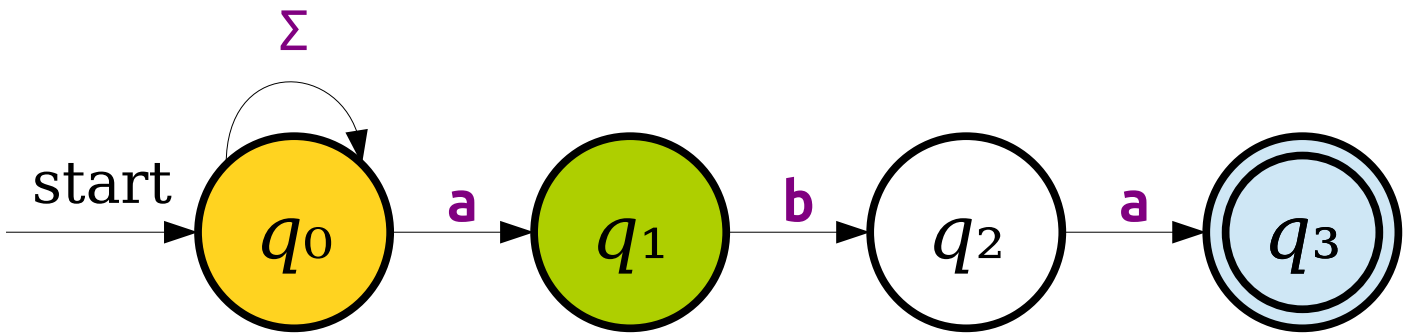


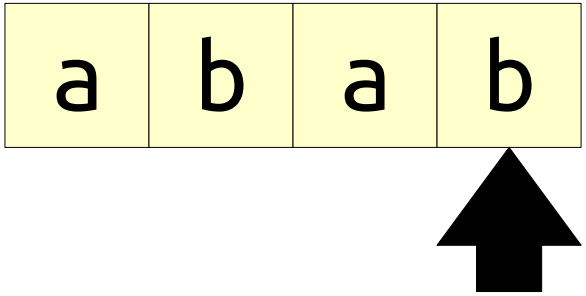
Massive Parallelism



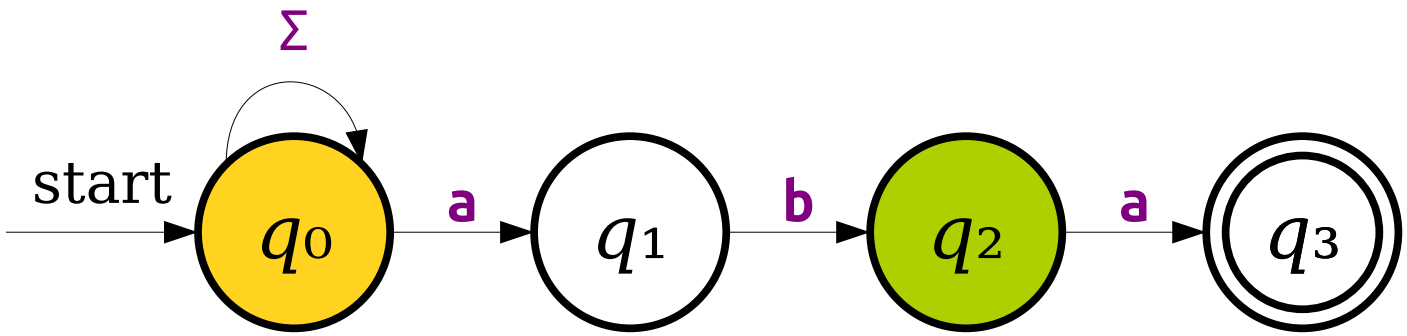


Massive Parallelism



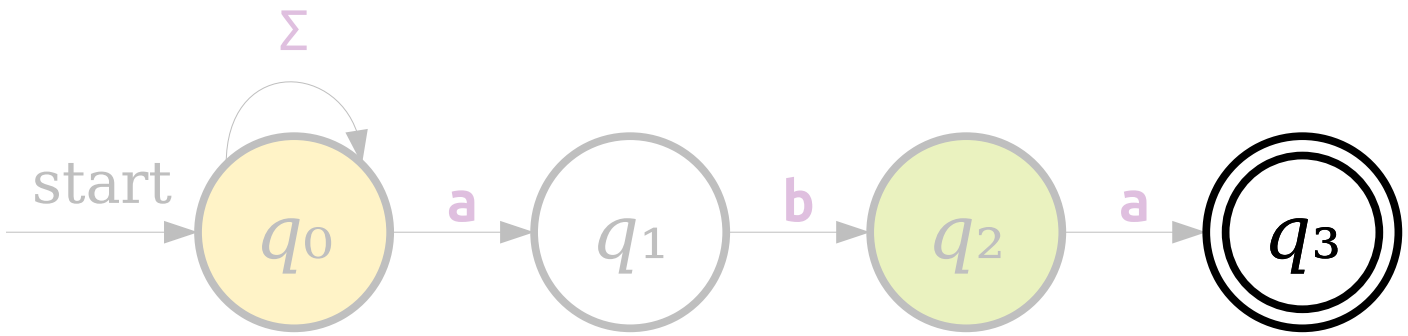


Massive Parallelism



a	b	a	b

Massive Parallelism



We're not in any accepting state, so no possible path accepts.

a	b	a	b

Massive Parallelism

- ***Key Idea:*** Imagine the NFA can be in many states at once. The NFA tries all possible transitions in parallel with one another.
- Here's a rigorous explanation about how this works; read this on your own time.
  - Start off with the start state active, plus all states that can be reached by zero or more ε-transitions.
  - When you read a symbol `a` in a set of states S:
    - Form the set S' of states that can be reached by following a single `a` transition from some state in S.
    - Your new set of states is the set of states in S', plus the states reachable from S' by following zero or more ε-transitions.

Finite Automata

Part 2

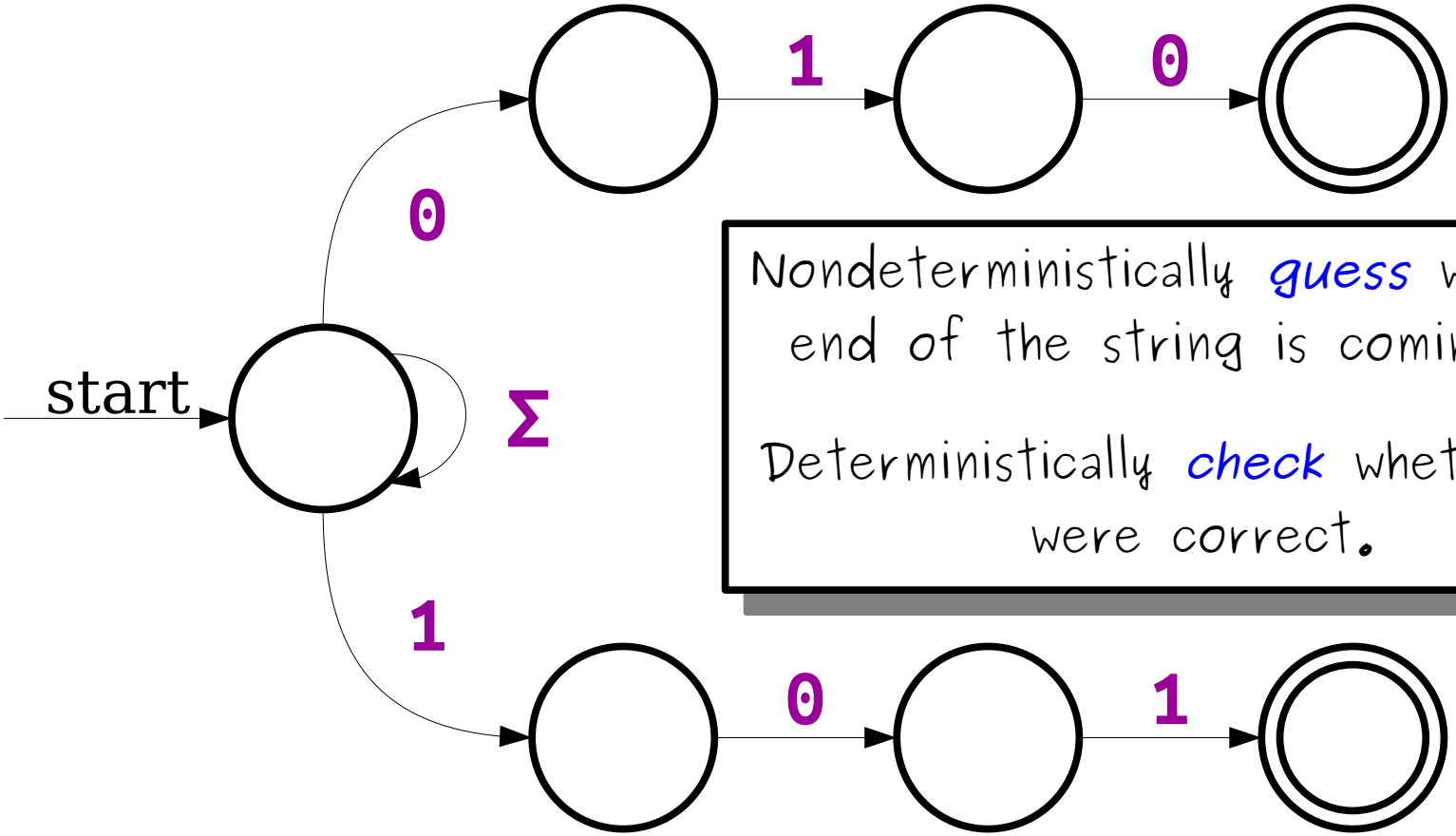
1. Recap from Last Time
2. Tabular DFAs
3. Regular Languages
4. Complementing Regular Languages
5. Announcements
6. Beyond DFAs: Nondeterminism
7. ε-Transitions
8. Nondeterminism Intuitions
9. **Designing NFAs**
10. What’s Next?

Designing NFAs

- ***Embrace the nondeterminism!***
- Good model: ***Guess-and-check***:
  - Is there some information that you'd really like to have? Have the machine *nondeterministically guess* that information.
  - Then, have the machine *deterministically check* that the choice was correct.
- The *guess* phase corresponds to trying lots of different options.
- The *check* phase corresponds to filtering out bad guesses or wrong options.

Guess-and-Check

$L = \{ w \in \{\mathtt{0}, \mathtt{1}\}^* \mid w \text{ ends in } \mathtt{010} \text{ or } \mathtt{101} \}$

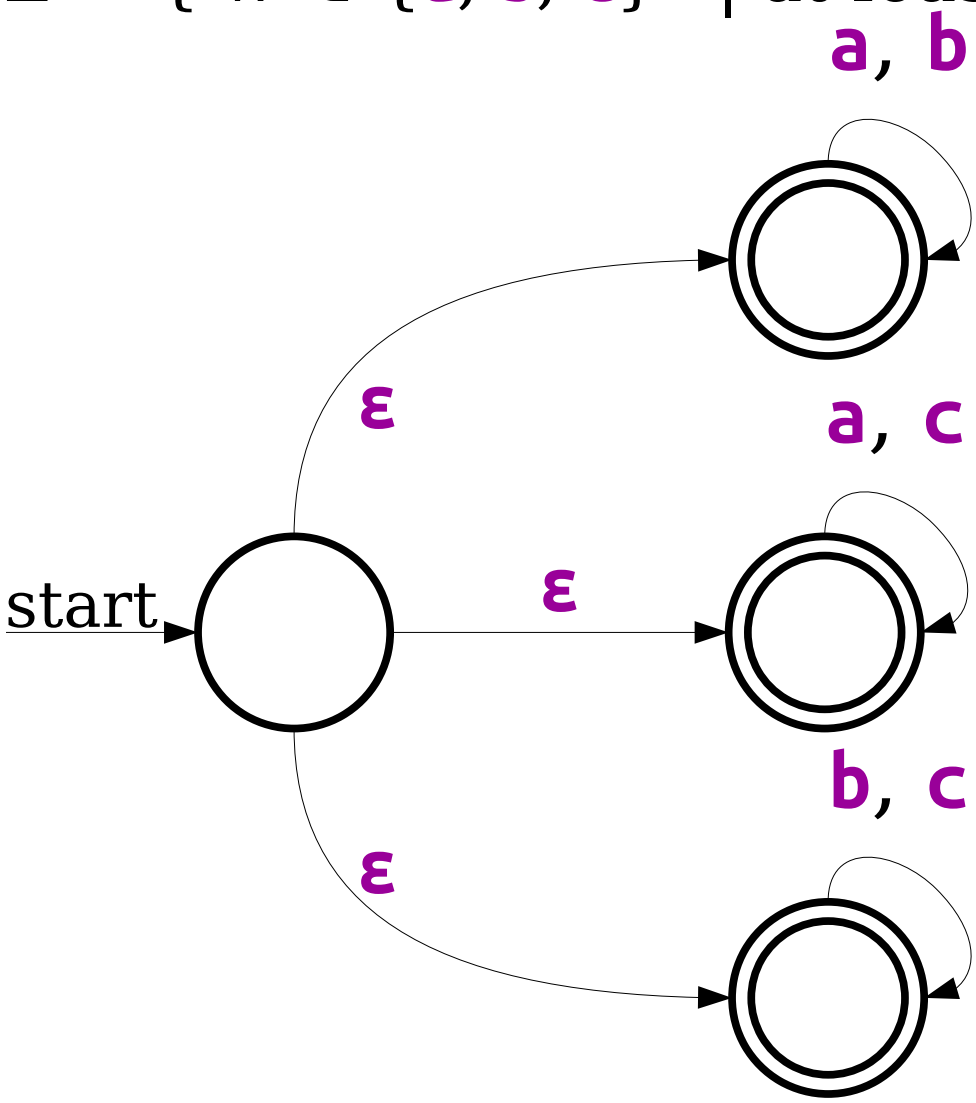


Nondeterministically *guess* when the end of the string is coming up.

Deterministically *check* whether you were correct.

Guess-and-Check

$L = \{ w \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}\}^* \mid$ at least one of **a**, **b**, or **c** is not in $w \}$



Nondeterministically *guess* which character is missing.

Deterministically *check* whether that character is indeed missing.

Just how powerful are NFAs?

Next Time

- ***The Subset Construction***
  - So beautiful. So elegant. So cool!
- ***More Closure Properties***
  - Transforming languages by transforming machines.
- ***The Kleene Closure***
  - What's the deal with the notation Σ^*?