

# Deterministic Finite Automaton

Week2

Textbook references: Within a chapter, each item is numbered consecutively. Figure 1.22 is the twenty-second numbered item in chapter one; it comes right after Example 1.21 and right before Definition 1.23.

In Computer Science, we operationalize “hardest” as “requires most resources”, where resources might be memory, time, parallelism, randomness, power, etc. To be able to compare “hardness” of problems, we use a consistent description of problems

**Input:** String

**Output:** Yes/ No, where Yes means that the input string matches the pattern or property described by the problem.

So far: we saw that regular expressions are convenient ways of describing patterns in strings. DFA give a model of computation for processing strings and and classifying them into Yes (accepted) or No (rejected). We will see that each set of strings is described by a regular expression if and only if there is a DFA that recognizes it. Another way of thinking about it: properties described by regular expressions require exactly the computational power of DFAs.

**Note:** Build a DFA that recognizes  $\{0, 10\}$  means building a DFA that accepts 0 and accepts 10 and rejects all other strings

Monday April 4

Build a regular expression that describes  $\{0, 10\}$  means only 0 and 10 match the pattern, i.e. the language of the regular expression has exactly these two elements.

**Review:** Formal definition of DFA:  $M = (Q, \Sigma, \delta, q_0, F)$

• Finite set of states  $Q$

**NOTE:** nonempty

• Start state  $q_0$

$q_0 \in Q$

• Alphabet  $\Sigma$

• Accept (final) states  $F$

$F \subseteq Q$

• Transition function  $\delta$

$\delta: Q \times \Sigma \rightarrow Q$

specify as part of machines

In the state diagram of  $M$ , how many outgoing arrows are there from each state?

$|\Sigma|$

$M = (\{q, r, s\}, \{a, b\}, \delta, q, \{s\})$  where  $\delta$  is (rows labelled by states and columns labelled by symbols):

$|\Sigma| = 2$

$\delta$	$\Sigma$	
	$a$	$b$
$q$	$r$	$q$
$r$	$r$	$s$
$s$	$s$	$s$

$\delta((q, a)) = r$

The state diagram for  $M$  is



so far only seen bs.

saw at least one a and it was last char we saw.

all extensions of the string that brought us here should be accepted.

Give two examples of strings that are accepted by  $M$  and two examples of strings that are rejected by  $M$ :

Accepts:  $q, q, r, s$   
 $bab$   $ab$

Rejects:  $\epsilon$   $a$   $b$

Add “labels” for states in the state diagram, e.g. “have not seen any of desired pattern yet” or “sink state”.

We can use the analysis of the roles of the states in the state diagram to describe the language recognized by the DFA.

$$L(M) = \{w \in \{a,b\}^* \mid w \text{ contains at least one } a \text{ followed by a } b\} = \{w \in \{a,b\}^* \mid \begin{matrix} ab \\ \text{is a} \\ \text{substring} \\ \text{of } w \end{matrix}\}$$

A regular expression describing  $L(M)$  is

Trace paths that lead from  $q_0$  to  $s$

$$b^* a a^* b (a \cup b)^*$$

$$= \{w_1 a b w_2 \mid \begin{matrix} w_1 \in \{a,b\}^* \\ \text{and} \\ w_2 \in \{a,b\}^* \end{matrix}\}$$

Using semantic structure

$$\Sigma^* a b \Sigma^*$$

remembering  
 $\Sigma$  is  
 shorthand  
 $(a \cup b)$

example  $aab$

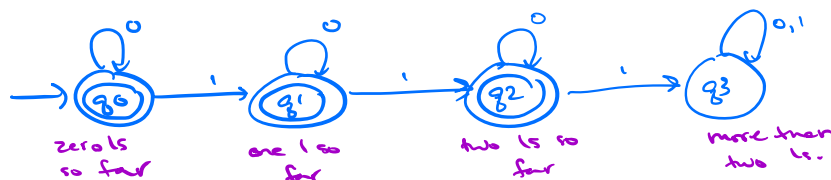
$$\underline{a}ab \in L(b^* \underline{a}a^* b (a \cup b)^*)$$

$$a \underline{a} b \in L(\Sigma^* a b \Sigma^*)$$

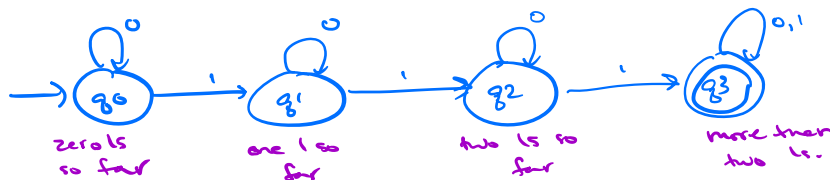
$$\begin{array}{ccccc} \varepsilon & a & a & b & \varepsilon \\ \hline & b^* & a & a^* & b (a \cup b)^* \\ \hline a & ab & & & \\ \hline \Sigma^* & ab & \Sigma^* & & \end{array}$$

Let the alphabet be  $\Sigma_1 = \{0, 1\}$ .

A state diagram for a DFA that recognizes  $\{w \mid w \text{ contains at most two 1's}\}$  is



A state diagram for a DFA that recognizes  $\{w \mid w \text{ contains more than two 1's}\}$  is



Notice similarities of these two state diagram

*Extra example:* A state diagram for DFA recognizing

$\{w \mid w \text{ is a string over } \{0, 1\} \text{ whose length is not a multiple of } 3\}$

Let  $n$  be an arbitrary positive integer. What is a formal definition for a DFA recognizing

$\{w \mid w \text{ is a string over } \{0, 1\} \text{ whose length is not a multiple of } n\}$ ?

## **Review: Week 2 Monday**

Please complete the review quiz questions on Gradescope about the languages recognized by DFAs.

Recall: Review quizzes based on class material are assigned each day. These quizzes will help you track and confirm your understanding of the concepts and examples we work in class. Quizzes can be submitted on Gradescope as many times (with no penalty) as you like until the quiz deadline: the three quizzes each week are all due on Friday (with no penalty late submission open until Sunday).

**Pre class reading for next time:** Pages 45-47.

## Wednesday April 6

Suppose  $A$  is a language over an alphabet  $\Sigma$ . By definition, this means  $A$  is a subset of  $\Sigma^*$ . **Claim:** if there is a DFA  $M$  such that  $L(M) = A$  then there is another DFA, let's call it  $M'$ , such that  $L(M') = \overline{A}$ , the complement of  $A$ , defined as  $\{w \in \Sigma^* \mid w \notin A\}$ .

**Proof idea:**

**Proof:**

A useful (optional) bit of terminology: the **iterated transition function** of a DFA  $M = (Q, \Sigma, \delta, q_0, F)$  is defined recursively by

$$\delta^*(q, w) = \begin{cases} q & \text{if } q \in Q, w = \varepsilon \\ \delta(q, a) & \text{if } q \in Q, w = a \in \Sigma \\ \delta(\delta^*(q, u), a) & \text{if } q \in Q, w = ua \text{ where } u \in \Sigma^* \text{ and } a \in \Sigma \end{cases}$$

Using this terminology,  $M$  accepts a string  $w$  over  $\Sigma$  if and only if  $\delta^*(q_0, w) \in F$ .

Fix  $\Sigma = \{a, b\}$ . A state diagram for a DFA that recognizes  $\{w \mid w \text{ has } ab \text{ as a substring and is of even length}\}$ :

Suppose  $A_1, A_2$  are languages over an alphabet  $\Sigma$ . **Claim:** if there is a DFA  $M_1$  such that  $L(M_1) = A_1$  and DFA  $M_2$  such that  $L(M_2) = A_2$ , then there is another DFA, let's call it  $M$ , such that  $L(M) = A_1 \cap A_2$ .

**Proof idea:**

**Formal construction:**

**Application:** When  $A_1 = \{w \mid w \text{ has } ab \text{ as a substring}\}$  and  $A_2 = \{w \mid w \text{ is of even length}\}$ .

Suppose  $A_1, A_2$  are languages over an alphabet  $\Sigma$ . **Claim:** if there is a DFA  $M_1$  such that  $L(M_1) = A_1$  and DFA  $M_2$  such that  $L(M_2) = A_2$ , then there is another DFA, let's call it  $M$ , such that  $L(M) = A_1 \cup A_2$ .  
*Sipser Theorem 1.25, page 45*

**Proof idea:**

**Formal construction:**

**Application:** A state diagram for a DFA that recognizes  $\{w \mid w \text{ has } ab \text{ as a substring or is of even length}\}$ :

## **Review: Week 2 Wednesday**

Please complete the review quiz questions on Gradescope about the languages recognized by DFAs.

Recall: Review quizzes based on class material are assigned each day. These quizzes will help you track and confirm your understanding of the concepts and examples we work in class. Quizzes can be submitted on Gradescope as many times (with no penalty) as you like until the quiz deadline: the three quizzes each week are all due on Friday (with no penalty late submission open until Sunday).

**Pre class reading for next time:** Introduction to Section 1.2



## Friday April 8

<b>Nondeterministic finite automaton</b> $M = (Q, \Sigma, \delta, q_0, F)$	
Finite set of states $Q$	Can be labelled by any collection of distinct names. Default: $q_0, q_1, \dots$
Alphabet $\Sigma$	Each input to the automaton is a string over $\Sigma$ .
Arrow labels $\Sigma_\epsilon$	$\Sigma_\epsilon = \Sigma \cup \{\epsilon\}$ .
Transition function $\delta$	Arrows in the state diagram are labelled either by symbols from $\Sigma$ or by $\epsilon$ . $\delta : Q \times \Sigma_\epsilon \rightarrow \mathcal{P}(Q)$ gives the <b>set of possible next states</b> for a transition from the current state upon reading a symbol or spontaneously moving.
Start state $q_0$	Element of $Q$ . Each computation of the machine starts at the start state.
Accept (final) states $F$	$F \subseteq Q$ .
$M$ accepts the input string	if and only if <b>there is</b> a computation of $M$ on the input string that processes the whole string and ends in an accept state.
Page 53	

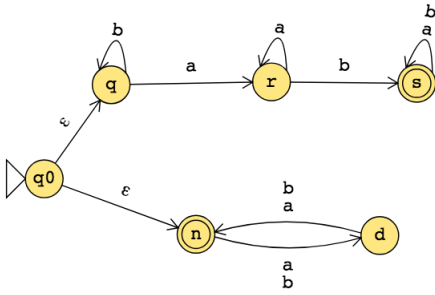
The formal definition of the NFA over  $\{0, 1\}$  given by this state diagram is:



The language over  $\{0, 1\}$  recognized by this NFA is:

Change the transition function to get a different NFA which accepts the empty string.

The state diagram of an NFA over  $\{a, b\}$  is below. The formal definition of this NFA is:



The language recognized by this NFA is:

## **Review: Week 2 Friday**

Please complete the review quiz questions on Gradescope about NFA.

**Pre class reading for next time:** Theorem 1.47, Theorem 1.49