Lecture 36

Introduction to Theory of Computation, Languages, DFA

Introduction

We will learn about various abstract computational machines: Deterministic Finite Automata, Context Free Grammars, and Turing Machines.

Why study these machines?

- DFAs are used in software for designing and checking the behaviour of digital circuits, In scanning large bodies of texts, verifying the correctness of programs.
- CFGs play a central role in Compilers.
- Turing machines are used to prove impossibility results.

Search vs Decision Problems

Search Problem: Problems where you have to find a solution or inform that no solution exists.

Example: $SEARCH\_PATH$: Given a graph G and vertices u and v, find a path from u to v or determine if no such paths exist.

Decision Problem: Problems that can be posed as a Yes or No question.

Example: $DECISION\_PATH$: Given a graph G and vertices u and v, find whether there is a path from u to v.

We will focus on decision problems as search problem can be posed as a collection of decision problem.

Example: $SEARCH\_PATH$ can be answered by answering the following decision problems:

Is P_1 a path from u to v? Is P_2 a path from u to v? Is P_3 a path from u to v?

Is P_4 a path from u to v? Is P_5 a path from u to v? $\ldots$

Formalising Problems

Definition: An **alphabet** is a finite, nonempty set of symbols, usually denoted by Σ.

$$\Sigma = \{0,1\},\ \Sigma = \{a, b, c, \ldots, z\}$$

Definition: A **string** is a finite sequence of symbols chosen from some alphabet.

0010101 is a string chosen from $\Sigma = \{0,1\}$

$abxdd$ is a string chosen from $\Sigma = \{a, b, c, \ldots, z\}$

ϵ denotes the empty string from any alphabet

Length of a string w, denoted by $|w|$, is the number of positions in w.

For instance, $|0010101| = 7,\ |abxdd| = 5,\ |\epsilon| = 0$

Σ^k denotes the set of all the strings of length k made from symbols of Σ.

$\Sigma^* = \Sigma^0 \cup \Sigma^1 \cup \Sigma^2 \cup \ldots$

Formalising Problems

Definition: A set of strings all of which are chosen from some Σ^*, where Σ is a particular alphabet, is called a **language**.

Examples: Some languages over $\Sigma = \{0,1\}$

$$L = \{0,1,11,101,1101\}$$

$$L = \{\epsilon,01,10,1001,1010,1100,0011,\ldots\}$$

$$L = \{\epsilon,10,11,101,111,1011,\ldots\}$$

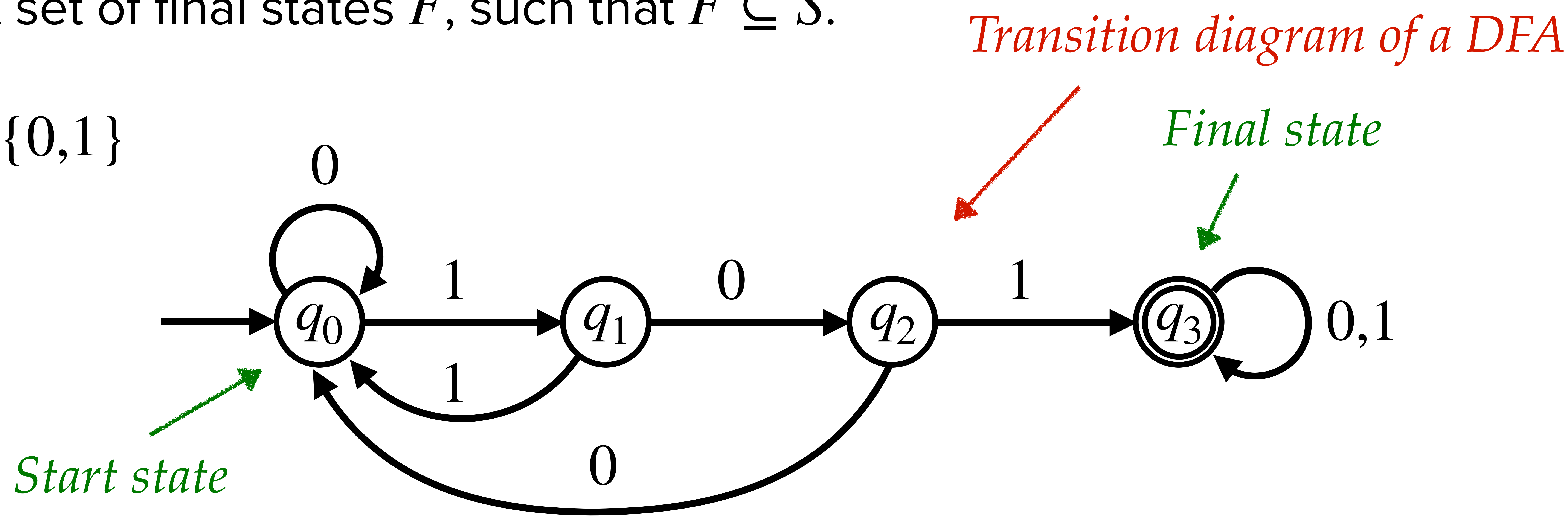
Definition: A **problem** is the question of deciding whether a given string is a member of some particular language.

Deterministic Finite Automaton

Definition: A **DFA** is a 5-tuple $\langle Q, \Sigma, \delta, q_0, F \rangle$

- A finite set of states, denoted by Q.
- A finite set of input symbols, denoted by Σ.
- A transition function, $\delta : Q \times \Sigma \rightarrow Q$
- A start state q_0, one of the states in Q.
- A set of final states F, such that $F \subseteq S$.

Example: $\Sigma = \{0,1\}$



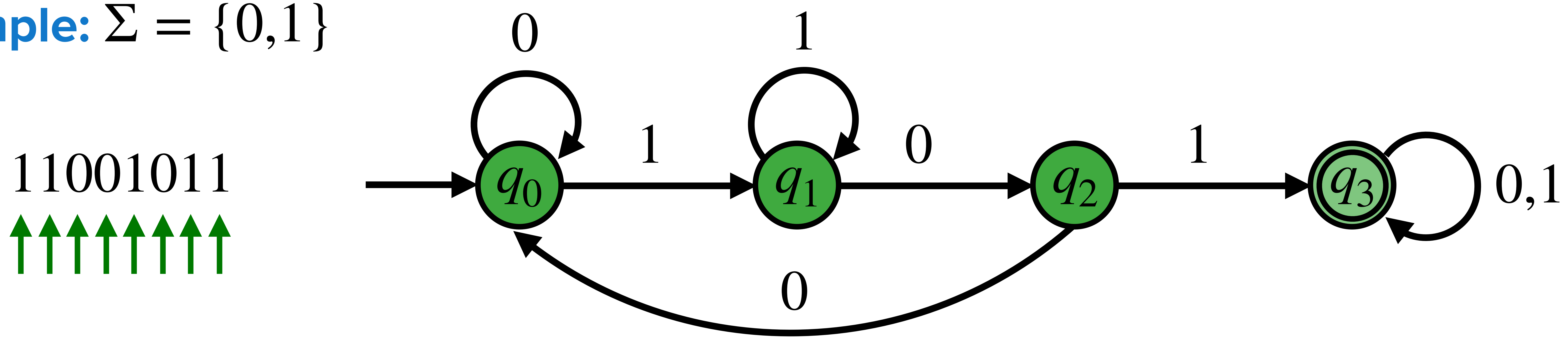
How DFA Processes Strings?

A DFA processes an input string in the following manner:

- It reads the input string one by one from left to right.
- It starts with the "start state" and moves from one state to another using δ.
- It **"accepts"** a string if after reading all the symbols it ends at a final state, else it **"rejects"**.

Definition: Language of a DFA M, denoted $L(M)$, is the set of strings that are accepted by it.

Example: $\Sigma = \{0,1\}$



Language of the above DFA is the set of binary strings that contain 101 as a substring.