

Prove the following languages over $\Sigma=\{0,1\}$ are regular by giving regular expressions for them:

1. $\{w \text{ contains two or more } 0\text{'s}\}$
2. $\{|w| = 3k \text{ for some integer } k\}$
3. $\{w \text{ has } 001 \text{ or } 1010 \text{ as a prefix}\}$
4. $\{w \text{ contains both } 00 \text{ and } 11 \text{ as substrings}\}$
5. $\{w \text{ contains } 01001 \text{ and } 0101 \text{ as substrings}\}$

Assignment #1 is due on Friday.

Any questions?

Recall you need a 50% average on the assignments. It is in your best interests to hand in as much as you can from the assignment when it is due.

Deterministic Finite Automata

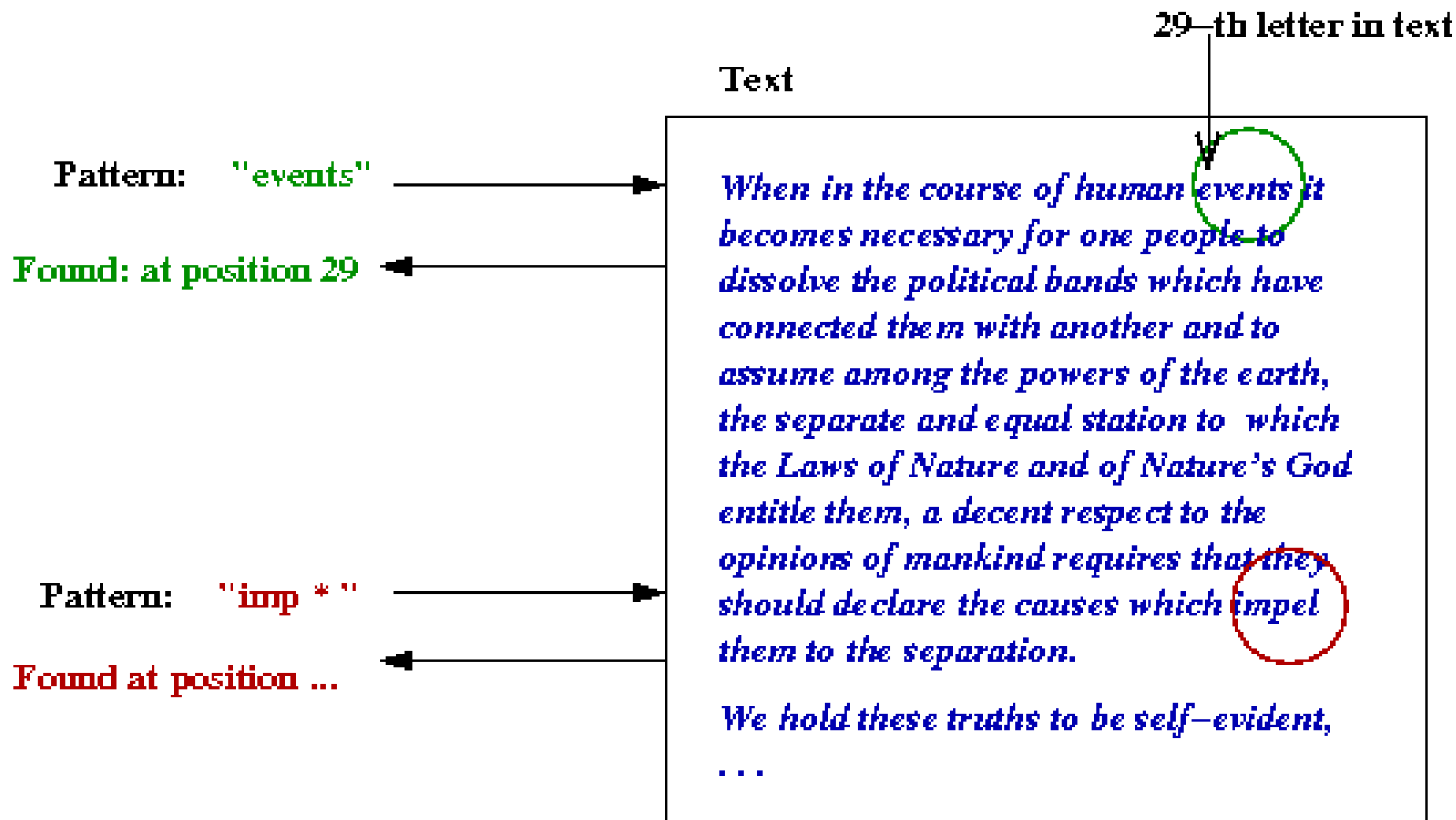
Given a language L , the **language recognition problem** is:

Given an input string w , is w in L ?

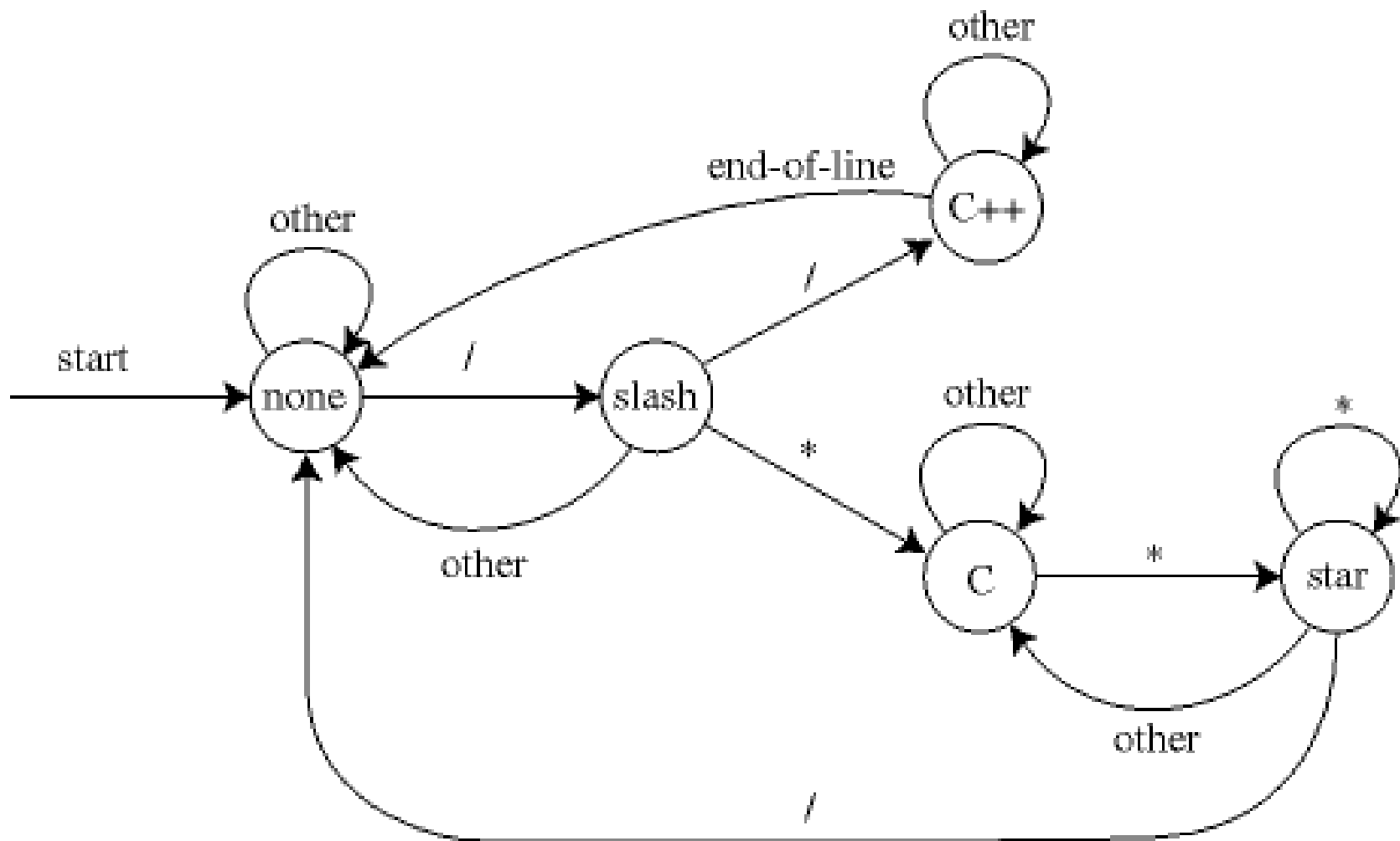
Answer: yes or no.

DFA's provide linear time language recognition algorithms for regular languages (languages which can be defined by a regular expression).

This class introduces DFA's, describes how they work, and defines the mathematical notation used to talk about them.



To strip comments from a program:



Applications of DFA's or DFA-like code:

[www.fbeedle.com/formallanguage/ch04.pdf]

Compilers and other language-processing systems- to divide an input program into tokens like identifiers, constants, and keywords and to remove comments and white space.

Typed commands- the command language can be quite complex, almost like a little programming language.

Speech-processing and other signal-processing systems- to transform an incoming signal.

More Applications

[www.fbeedle.com/formallanguage/ch04.pdf]

Text processors- to search a text file for strings that match a given pattern. This includes most versions of Unix tools like awk, egrep, and Procmail, along with a number of platform-independent systems such as MySQL.

Controllers- to track the current state of a wide variety of finite-state systems, from industrial processes to video games, implemented in hardware or in software.

$L = \{w \in \{a,b\}^* : \text{the number of } a\text{'s in } w \text{ is not divisible by } 3\}.$

1. Give a regular expression for L .
2. Design a DFA that accepts L .

A **Deterministic Finite Automaton** (DFA)

M is defined to be a quintuple

$(K, \Sigma, \delta, s, F)$ where

K is a finite set of **states**,

Σ is an alphabet,

δ is a **function** from $K \times \Sigma$ to K ,

$s \in K$ is the **start state**, and

$F \subseteq K$ is the set of **final states**.

Some examples:

1. $\{w: w \text{ has odd length}\}$
2. $\{w: w \text{ contains } 0011\}$
3. $\{w: w \text{ does not contain } 01\}$
4. $\{w: w \text{ starts and ends with the same symbol } (|w| \geq 1)\}$

$M = (K, \Sigma, \delta, s, F)$:

A **configuration** of a M is an element of $K \times \Sigma^*$.

For $\sigma \in \Sigma$, configuration $(q, \sigma w) \vdash (r, w)$

$[(q, \sigma w) \text{ yields in one step } (r, w)]$

if $\delta(q, \sigma) = r$.

The notation $\vdash^*$ means yields in zero or more steps.

$M = (K, \Sigma, \delta, s, F)$:

DFA M **accepts** input w if and only if
 $(s, w) \vdash^* (f, \varepsilon)$ for some $f \in F$.

$L(M)$, the **language accepted by M** is
 $\{ w \in \Sigma^* : (s, w) \vdash^* (f, \varepsilon) \text{ for some } f \in F \}$.

$L_1 = \{ w \in \{a, b\}^* : w \text{ begins with } aab \}$

$L_2 = \{ w \in \{0, 1\}^* : \text{has both } 00 \text{ and } 11 \text{ as substrings} \}$

$L_3 = \{ w \in \{a, b\}^* : w \text{ ends with } abab \}$

$L_4 = \{ w \in \{0, 1\}^* : \text{the number of 0's in } w \text{ is even and the number of 1's is congruent to 1 modulo 3} \}$

$L_5 = \{ w \in \{a, b\}^* : w \text{ contains } abaab \}$

What language does this DFA accept?

