

# Foundations of Computer Science

Notes from 02-26-2015

Kenny Roffo

March 26, 2015

## Contents

<b>1</b>	<b>Formal Languages</b>	<b>2</b>
1.1	Definitions . . . . .	2
1.2	Examples . . . . .	2
1.3	What is $\Sigma^*$ . . . . .	3
<b>2</b>	<b>Regular Expressions</b>	<b>3</b>
2.1	What is a Regular Expression? . . . . .	3
<b>3</b>	<b>Finite State Machines</b>	<b>4</b>
3.1	The Vending Machine . . . . .	4
3.2	Finite State machine . . . . .	5
3.3	The Turing Machine . . . . .	6
3.4	Non-deterministic Finite State Machines . . . . .	7

# 1 Formal Languages

## 1.1 Definitions

1. A symbol is the basic indivisible entity  
Natural Languages: words, not letters
2. An alphabet is a finite, nonempty set of symbols  
 $\Sigma$  is typically used as the name for the alphabet  
Natural languages:  $\Sigma = \{\text{all words in English}\}$  - Lexicon
3. A string (over  $\Sigma$ ) is a finite sequence of symbols (over  $\Sigma$ )  
Properties:
  - length
  - empty string denoted  $\lambda$
  - concatenation
  - $\lambda$  identity of concatenation
4.  $\Sigma^*$  is the set of all finite strings over  $\Sigma$   
e.g.  $\Sigma = \{0, 1\}$ ,  $\Sigma^* = \{\text{all strings of 0 and 1}\}$   
 $\Sigma$  is an alphabet  
 $\Sigma^*$  is defined recursively  
 $\lambda$  is an element of  $\Sigma^*$ , and if  $w \in \Sigma^*$ ,  $a \in \Sigma$ , then  $wa \in \Sigma^*$
5. A language  $L$  is any set of strings formed from a given alphabet  $\Sigma$   
 $\phi$  is a language over any alphabet  
 $\Sigma$  is a language over  $\Sigma$   
 $\Sigma^*$  is a language over  $\Sigma$

## 1.2 Examples

1.  $\Sigma = \{b\}$   
 $\Sigma^* = \{\lambda, b, bb, bbb, \dots\}$
2.  $\Sigma = \{a, b\}$   
 $\Sigma^* = \{\lambda, a, b, aa, ab, ba, bb, \dots\}$   
 $L$  is a language over  $\Sigma$  that is defined recursively: ex:  $L = \{w | w = a^n b^n, n \geq 1\}$   
( $a^n$  means  $aaa\dots a$   $n$  times)  
Thus  $L = \{ab, aabb, aaabbb, aaaabbbb, \dots\}$ , if  $w \in L$ , then  $awb \in L$

3. Let  $L_n = a^n b^n$ . Then  $L = L_0 \cup L_1 \cup L_2 \cup \dots$
4. Let  $\Sigma_i$  be the language of  $i$  length strings in the alphabet  $\Sigma = \{a, b\}$ .  
 Then  $\Sigma_0 = \{\lambda\}$ ,  $\Sigma_1 = \{a, b\}$ ,  $\Sigma_2 = \{aa, ab, ba, bb\}$ , ...  
 It follows that the language  $\Sigma^* = \Sigma_0 \cup \Sigma_1 \cup \Sigma_2 \cup \dots = \bigcup_{i=1}^{\infty} \Sigma_i$

### 1.3 What is $\Sigma^*$

Concatenation - Making new strings from existing strings. We can also concatenate strings with languages and languages with languages. If  $L_1$  and  $L_2$  are languages, then  $L_1 L_2 = \{w_1 w_2 | w_1 \in L_1, w_2 \in L_2\}$

ex: Let  $L_1 = \{\text{in, out}\}$  and  $L_2 = \{\text{law, door, ward}\}$ .  
 Then  $L_1 L_2 = \{\text{inlaw, outlaw, indoor, outdoor, inward, outward}\}$

$\Sigma^*$  is the set of all strings made from the alphabet  $\Sigma$ . But why  $\Sigma^*$ ?  
 $\Sigma^*$  is the result of concatenating  $\Sigma$  with itself zero or more times.  
 $\Sigma^+$  is the result of concatenating  $\Sigma$  with itself one or more times.  
 This is called the positive closure of  $\Sigma$ .

## 2 Regular Expressions

### 2.1 What is a Regular Expression?

A regular expression (regex) is a way to specify patterns for strings using union (or), concatenation, and  $*$

A regex over  $\Sigma$  is defined:

Basis: Every  $a \in \Sigma$  is a regex over  $\Sigma$

Recursive: If  $u$  and  $v$  are regex over  $\Sigma$  then  $u|v$ ,  $uv$ , and  $u^*$  are all regex over  $\Sigma$

Here,  $|$  means or and  $*$  means 0 or more. When in doubt, use parentheses.

grep - general regular expression parser - a Unix command which searches a file for a pattern defined by a regex.

Let  $X = \{a, ab, aba\}$  and  $Y = \{b, bb\}$ . Then

- $XY = \{ab, abb, abb, abbb, abab, ababb\}$  (Concatenation)
- $X|Y = \{a, ab, aba, b, bb\}$  (Like union)
- $X^* = \{a, aa, aaa, \dots, aab, ab, abab, ababab, \dots, aab, aaba, aaab, ababa, \dots\}$   
(All possible strings from 0 or more concatenations)
- $ababa \in X^*$
- $ababa \in XY^*$
- $ab(ab)^*a$  is a regex that matches  $ababa$

## 3 Finite State Machines

### 3.1 The Vending Machine

Consider a vending machine which contains Jelly beans and Gum. The Machine has inputs

- N - 5 cents
- D - 10 cents
- J - Jelly Bean (Costs 20 cents)
- G - Gum (Costs 15 cents)

These can be represented by  $\Sigma = \{N, D, J, G\}$ . The machine also has outputs

- b - beep when money is added
- j - jelly bean dispensed
- g - gum dispensed

Design a Finite State machine - a machine with a finite number of “things to remember”  
This vending machine has to “remember”:

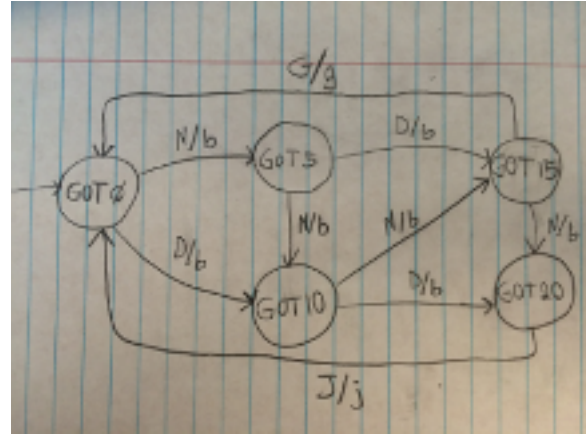
- total money deposited (but not the order in which coins are deposited)
- which product is selected

States are drawn with circles and named on the inside:

For our Vending Machine the states are described by how much money is in the machine, and the transitions represent money input, or a purchase or gum or a jelly bean. The states live in the set  $Q = \{GOT\emptyset, GOT5, GOT10, GOT15, GOT20\}$ . State transitions are defined by a function  $\delta : Q \times \Sigma \rightarrow Q$ . As an example,

$$(GOT5)(N) \rightarrow GOT10$$

Here is an example of a state transition diagram.



### 3.2 Finite State machine

A finite state machine has one or more states. For the vending machine in the previous section the initial and final states happen to be the same. Any path in the state transition diagram is called a computation. Any path in a finite state machine which starts in the initial state and ends in the final state is called an acceptable path.

A finite state machine must have:

- An input mechanism
- A computing mechanism
- An output mechanism

We use ordered pairs to reference what has happened in the fsm. Referring back to the vending machine, here is an example:

$$(GOT5, NNDG) \rightarrow (GOT10, NNDG) \rightarrow (GOT15, NDG) \rightarrow (GOT20, DG) \rightarrow (GOT20, G) \rightarrow (GOT\emptyset, \lambda)$$

The first part of the pair represents the current state while the second represents the input. Each pair is called an instantaneous configuration.

A finite state machine is defined by a 5-tuple:  $M = (Q, \Sigma, \delta, q_0, F)$ :

- $Q$ : finite, nonempty set of states
- $\Sigma$ : alphabet (finite, nonempty set of symbols)
- $\delta: Q \times \Sigma \rightarrow Q$  is the state transition function
- $q_0$ : initial state
- $F \subset Q$ : set of final states

### 3.3 The Turing Machine

Alan Turing was one of the most influential scientists of the 20<sup>th</sup> century. One of his contributions was the idea of a Turing Machine.

A Turing Machine is a machine that can solve any computable problem. In other words, if a problem is able to be computed, a Turing Machine can solve that problem.

The Wikipedia definition of a Turing machine:

A **Turing machine** is a hypothetical device that manipulates symbols on a strip of tape according to a table of rules. Despite its simplicity, a Turing machine can be adapted to simulate the logic of any computer algorithm, and is particularly useful in explaining the functions of a CPU inside a computer.

### The Man, the Wolf, the Goat and the Cabbage

Consider this classic problem:

A man once had to travel with a wolf, a goat and a cabbage. He had to take good care of them, since the wolf would like to taste a piece of goat if he would get the chance, while the goat appeared to long for a tasty cabbage. After some traveling, he suddenly stood before a river. This river could only be crossed using the small boat laying nearby at a shore. The boat was only good enough to take himself and one of his loads across the river. The other two subjects/objects he had to leave on their own. How must the man row across the river back and forth, to take himself as well as his luggage safe to the other side of the river, without having one eating another?

A table can be constructed similar to show the possible different states. States that are not allowed are crossed out. The route I chose to take is in bold, and the next possible states from there are listed on the next line. Note that there are infinitely many solutions. I chose my solution as it seemed the most logical to me on the fly.

	M	W	G	C
MWGC-	WGC-M	<del>GC-MW</del>	<b>WC-MG</b>	<del>WG-MC</del>
WC-MG	<b>MWC-G</b>	X	MWGC-	X
MWC-G	WC-MG	C-MWG	X	<b>W-MGC</b>
W-MGC	<del>MW-GC</del>	X	<b>MWG-C</b>	MWC-G
MWG-C	<del>WG-MC</del>	<b>G-MWC</b>	W-MGC	X
G-MWC	<b>MG-WC</b>	MWG-C	X	MGC-W
MG-WC	G-MWC	X	<b>-MWGC</b>	X

### 3.4 Non-deterministic Finite State Machines

Let  $\Sigma = \{a, b\}$ . Design a finite state machine that accepts all strings over  $\Sigma$  that contain the substring  $bab$ :

Regular Expression:  $(a|b)^*bab(a|b)^*$

$\delta$	$a$	$b$
$q_0$	$q_0$	$q_0, q_1$
$q_1$	$q_2$	$\dots$
$q_2$	$\dots$	$q_3$
$q_3$	$q_3$	$q_3$

As seen in the table, if the state is  $q_0$  and the input is  $a$ , then the machine has two options. We call this non-deterministic, and thus we call the machine a non-deterministic finite state machine. Let  $M$  be a non-deterministic finite state machine: Then

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

where

$$\delta : Q \times \Sigma \rightarrow \rho(Q)$$

Recall that  $\rho(S)$  is the power set of  $S$ . An alternate way of making the table above would be:

$\delta$	$a$	$b$
$q_0$	$\{q_0\}$	$\{q_0, q_1\}$
$q_1$	$\{q_2\}$	$\emptyset$
$q_2$	$\emptyset$	$\{q_3\}$
$q_3$	$\{q_3\}$	$\{q_3\}$

We define a new function,  $\delta' : \rho(Q) \times \Sigma \rightarrow \rho(Q)$

$\delta_1$	$a$	$b$
$\{q_0\}$	$\{q_0\}$	$\{q_0, q_1\}$
$\{q_0, q_1\}$	$\{q_0, q_2\}$	$\{q_0, q_1\}$
$\{q_0, q_2\}$	$\{q_0\}$	$\{q_0, q_1, q_3\}$
$\{q_0, q_1, q_3\}$	$\{q_0, q_2, q_3\}$	$\{q_0, q_1, q_3\}$
$\{q_0, q_2, q_3\}$	$\{q_0, q_3\}$	$\{q_0, q_1, q_3\}$
$\{q_0, q_3\}$	$\{q_0, q_3\}$	$\{q_0, q_1, q_3\}$

Now we let

$$\begin{aligned}
\{q_0\} &= 0 \\
\{q_0, q_1\} &= 01 \\
\{q_0, q_2\} &= 02 \\
\{q_0, q_1, q_3\} &= 013 \\
\{q_0, q_2, q_3\} &= 023 \\
\{q_0, q_3\} &= 03
\end{aligned}$$

Then our machine becomes  $M'(Q', \Sigma, \delta', 0, F)$  with

$$\begin{aligned}
Q' &= \{0, 01, 02, 013, 023, 03\} \\
\Sigma &= \{a, b\} \\
\delta' &: \rho(Q) \times \Sigma \rightarrow \rho(Q) \\
F &= \{013, 023, 03\}
\end{aligned}$$

$\delta'$	$a$	$b$
0	0	01
01	02	01
02	0	013
013	023	013
023	03	013
03	03	013

This may feel a little uneasy at first, but we now have a deterministic finite state machine (DFA).