Theory of Computation

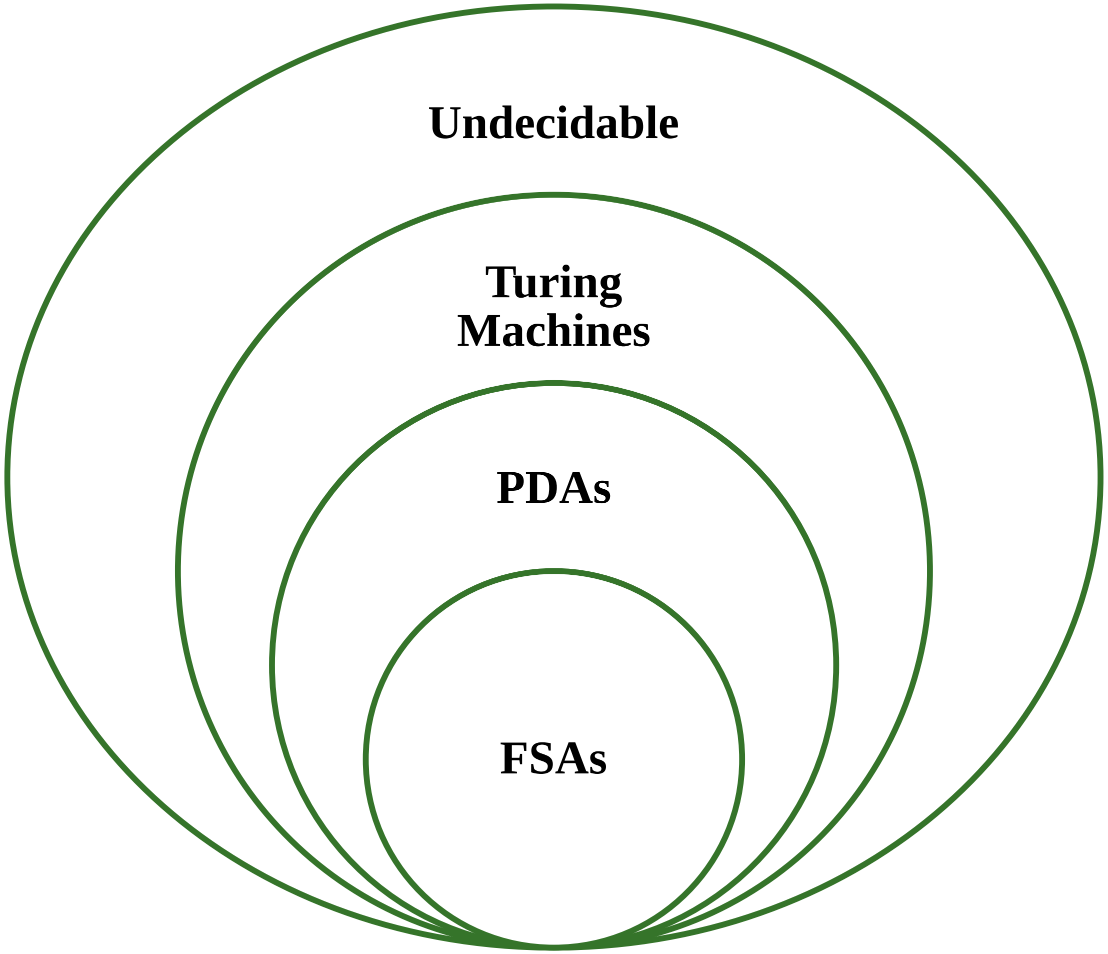
Lecture 2

More about Regular Expressions

Syllabus and Terminologies

- Regular Languages .. Regular Sets
  - REs (Regular Expressions)
  - FSMs (or FSA/FA) … Finite State Machines/Automata
    - DFA vs. NFA … Deterministic vs. Non-deterministic FSA
      - Comparison and conversion
    - Examples & Closure Operations
    - Pumping Lemma
- Context Free Languages
  - CFGs … Context Free Grammars
  - PDA … Push Down Automata
  - Parsing: CFG generating strings vs. PDA recognizing strings
- Turing Machine

Hierarchy of Machines



Example

- Regular expression: $(a+b)\cdot a^*$

$$
\begin{aligned}
L((a+b)\cdot a^*) &= L((a+b))\,L(a^*) \\
&= L(a+b)\,L(a^*) \\
&= (L(a)\cup L(b))\,(L(a))^* \\
&= (\{a\}\cup\{b\})\,(\{a\})^* \\
&= \{a,b\}\{\lambda,a,aa,aaa,\ldots\} \\
&= \{a,aa,aaa,\ldots,b,ba,baa,\ldots\}
\end{aligned}
$$

Example

- Regular expression $r = (aa)^*(bb)^*b$

- What is the regular languages that r describes?

$$L(r) = \{a^{2n}b^{2m}b: \quad n, m \geq 0\}$$

Equivalent Regular Expressions

- Definition:

Regular expressions r_1 and r_2

are **equivalent** if

$$L(r_1) = L(r_2)$$

Example

L = {all strings with no two consecutive 0's }

$$r_1 = (1 + 01)^*(0 + \lambda)$$

$$r_2 = (1^*011^*)^*(0 + \lambda) + 1^*(0 + \lambda)$$

$$L(r_1) = L(r_2) = L$$ $\Longrightarrow$ r_1 and r_2 are equivalent regular expr.

Example

- Regular expression $r = (0+1)^* 00 (0+1)^*$

$L(r)$ = { all strings with at least two consecutive 0's }

More RE Examples

RE-1

- Example 1

$\sum = \{a, b\}^*$

- Formally describe all words with a followed by any number of b's

$L = a\ b^* = ab^*$

```
cout<<'a';
i=0;
cin>>n;
while(i<n)
{
  cout<<'b';
  i++;
}
```

- Give examples for words in L

{a ab abb abbb …..}

RE-2

- Example

$\sum = \{a, b\}$

- Formally describe all words with a followed by one or more b's

$L = \cancel{a\,b}^{+} = abb^{*}$

- Give examples for words in L

{ab abb abbb …..}

RE-3

- Example

$\sum = \{a, b, c\}$

- Formally describe all words that start with an a followed by any number of b's and then end with c.

$L = a\, b^{*} c$

- Give examples for words in L

{ac abc abbc abbbc …..}

RE-4

$(ab)^* b^*$ abbbab

aa X

abbab ✓

- Example

$(a+b)^*$

$\sum = \{a, b\}$

- Formally describe the language that contains nothing and contains words where any <u>a</u> must be followed by one or more <u>b</u>'s

L = ~~b*~~(abb*)* OR (b+ab)*

① ab
② b
③ ab

abbab

- b abbbb Give examples for words in L

{Λ ab abb abababb ….. b bbb…}

RE-5

bbb

aaa

a b

- Example

aabbbb

abab X

$\Sigma = \{a, b\}$

- Formally describe all words where a's if any come before b's if any.

$L = a^* b^*$

- Give examples for words in L

$\{\Lambda$ a b aa ab bb aaa abb abbb bbb…..$\}$

NOTE: $a^* b^* \neq (ab)^*$

because first language does not contain abab but second language has.

Once single b is detected then no a's can be added

RE-6

- Example

$\sum = \{a\}$

- Formally describe all words where count of <u>a</u> is odd.

$L = a(aa)^* \text{ OR } (aa)^*a$

- Give examples for words in L

{a aaa aaaaa …..}

RE-7.1

- Example

Σ ={a, b, c}

- Formally describe all words where single a or c comes in the start then odd number of b's.

L = $(a+c)b(bb)^*$ $\quad a\,b(bb)^* + c\,b(bb)^*$

- Give examples for words in L

{ab cb abbb cbbb …..}

RE-7.2

- Example

$\sum$ ={a, b, c}

- Formally describe all words where single <u>a</u> or <u>c</u> comes in the start then odd number of <u>b</u>’s in case of <u>a</u> and zero or even number of <u>b</u>’s in case of <u>c</u>.

$L = ab(bb)^{*}+c(bb)^{*}$

- Give examples for words in L

{ab c abbb cbb abbbbb …..}

RE-8

- Example

$\sum = \{a, b, c\}$

- **Formally** describe all words where one or more a or one or more c comes in the start then one or more b's.

$aa^* bb^* + cc^* bb^*$

$L = \cancel{(a^+ + c^+)\, b^+} = (aa^*+cc^*)\ bb^*$

- Give examples for words in L

{ab cb aabb cbbb …..}

RE-9.1

- Example

$\Sigma = \{a, b\}$

- Formally describe all words with length three.

$L = \cancel{(a+b)^3} = (a+b)\,(a+b)\,(a+b)$

- List all words in L

  {aaa aab aba abb baa bab bba bbb}

- What is the count of words of length 4?

  $16 = 2^4$

- What is the count of words of length 44?

  2^{44}

RE-9.2

- Example

$\sum$ ={a, b, c, d}

- Formally describe all words with length three.

$L = \cancel{(a+b+c+d)^3} = (a+b+c+d)\ (a+b+c+d)\ (a+b+c+d)$

- First and last words in L:

  {aaa … ddd}

- What is the count of words?

  $4^3 = 64$

RE-10.1

- Example

Σ ={a, b}, What does L describe?

- $L = (a+b)^* a(a+b)^*$

Any string of a's and b's

Single a

Any string of a's and b's

- Give examples for words in L

{a ab aab bab abb …..}

RE-10.2

Ambiguity: ***abbaab can be parsed in 3 ways***

abbaab

Λ a bbaab

abbaab

abb a ab

abbaab

abba a b

- $L = (a+b)^{*}a(a+b)^{*}$

RE-11

- Example

∑ ={a, b}

- Formally describe all words with <u>*at least*</u> two <u>a</u>'s.

1) $L = b^*ab^*a(a + b)^*$

- Start with a jungle of <u>**b**</u>'s (or no <u>**b**</u>'s) until we find the first <u>**a**</u>, then more <u>**b**</u>'s (or no <u>**b**</u>'s), then the second <u>**a**</u>, then we finish up with anything.

- Give examples for words in L

{abbbabb aaaaa bbbabbbbabab…..}

RE-12

- Example

∑ ={a, b}

- Formally describe all words with <u>*exactly*</u> *two a's*.

  L = b*ab*ab*

  a a a a X

- Give examples for words in L

  {aa, aab, baba, and bbbabbbab …..}

- To make the word *aab, we let the first and second b* become Λ and the last becomes b*

RE-13.1

ab a bab b babba ✓

- Example

ba ✗

bbaa ✗

$\Sigma = \{a, b\}$

- Formally describe all words with at least one a and at least one b.

@ab

1) L = **(a + b)*a(a + b)* b(a + b)***

= (anything) a (anything) b (anything)

bab ✓

But **(a+b)*a(a+b)*b(a+b)*** expresses all words **except** words of the form some b's (at least one) followed by some a's (at least one). **bb*aa***

RE-13.2

2) L =**(a+b)*a(a+b)*b(a+b)* +   bb*aa***

Thus: **(a+b)*a(a+b)*b(a+b)* +   (a+b)*b(a+b)*a(a+b)***

= **(a+b)*a(a+b)*b(a+b)* +   bb*aa***

- Notice that it is necessary to write bb*aa* because b*a* will admit words we do not want, such as aaa.

Does this imply that

(a+b)*b(a+b)*a(a+b)*= bb*aa*??

NO!

Left side includes the word aba, which the expression on the right side does not.

RE-13.3

What about this RE? Does it Formally describe all words with at least one a and at least one b?

RE = **(a+b)*(ab+ba)(a+b)***