

Regular Expressions

Context Free Grammars  
(CFGs)

↑  
Formalism to describe patterns in strings

"Exactly This String"

↙ Or ↘  
"Exactly This" + "Exactly That"

"Exactly This" • "Then Exactly That"

"Exactly This" \*  
↙ Over and over again ↘  
0 or more repetitions of.

$[0-9] \cdot [0-9]^*$

$[0-9]^+$   
↙ 1 or more repetitions ↘  
of.

x.y

xy

"Hi"

"Hi"\* ← Kleene \*

( "Hi" + "Hello" ) \* ← Disjunction

( "Hi" + "Hello" ) \* • "Bye" \* ← Concatenation

Hi Hello Hi ||| Bye

[0-9][0-9][0-9][0-9][0-9]

[0-9]<sup>5</sup>

string • @ • string • " • "

char<sup>3</sup> [a-zA-Z]  
string [a-zA-Z0-9]

$[a-z]^* @ [a-z]^* \cdot (com + edu + org)$

somebody @ usc.edu

somebodyelse @ csci.usc.edu ✗

someone123 @ gmail.com ✗

someone @ github.io ✗

Hi + Hello

Hi ✓

Hello ✓

Bye ✗

$(\text{Hi + Hello})^+$   $(\text{Bye})^*$

$\epsilon$  ✗

Hi ✓

Bye ✗

Hello Bye ✗

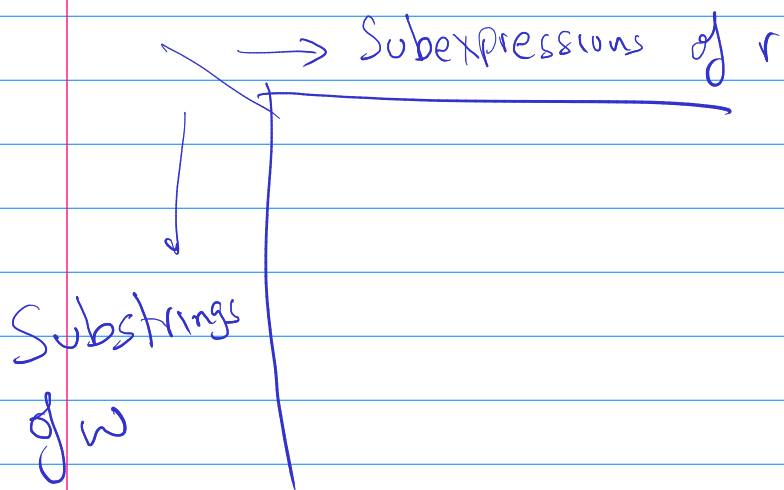
Hello ✓

Hello Hi Bye Bye ✓

$r ::= \varepsilon \mid w \mid r_1 + r_2 \mid r_1 \cdot r_2 \mid r_1^*$

Question: Given regex  $r$  string  $w$

determine whether  $w$  matches  $r$ .



$$\Theta(|r||w|^2)$$

$$O(|r||w|^3)$$

$r$   $w$   
 $H_i^*$   $Bye^*$   $H_i$   $Bye$

