

11 February 2025 11:34

→ Regular Expressions: Automata Theory

↳ Deterministic Finite Automata

→ a a a b b b b

⇒ DFA



⇒ N-DFA ~ Non-Deterministic Finite Automata

↓

Normal Form

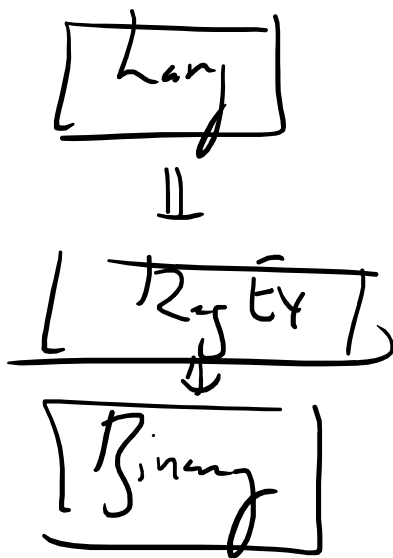
Chomsky  $NF$

Greibach  $NF$

Generalization

↓  
RegEx

⇒ Character ⇒ Words ⇒ Sentences ⇒ Corpus



- ⇒
- i) Compilation ~ Pattern
  - ii) Pattern ~ Matches

⇒ Regex Notations :

\d ⇒ Digits (0-9)

\d => Digits (0-9)

\D => Not a digit (0-9)

\w => Word characters (a-z, A-Z, 0-9)

\W => Not a word

\s => Whitespace (Tab, Space, Newline)

\S => Not whitespace.

=> Regex Boundaries:

\b => Word Boundary

\B => Not a word Boundary

^ => Start of string

\$ => End of string.