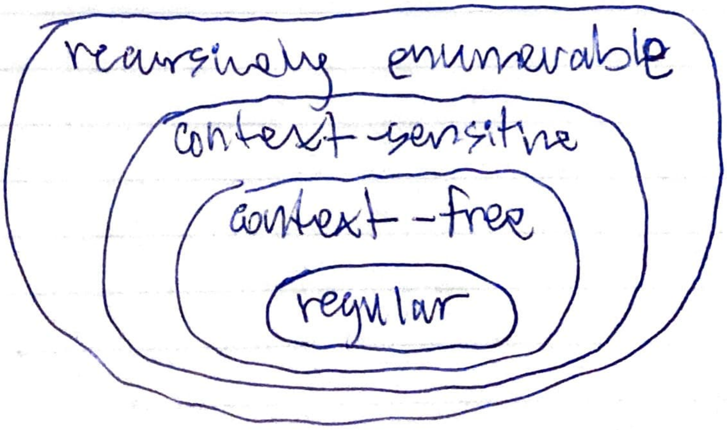
Regular Expression

- regex or regexp
- series of char, defines search pattern
- Syntax

-a	matches char itself	a	aab santa
-.	matches any char	.ed	jumped fixed
-+	matches 1+ occurences	o+d	food rode
-*	matches 0+ occurences	u*int	uint32_t int8_t
-?	1 or 0 occurences of 1-char expression	an?	an apple
-[]	any char enclosed	[aeiou]	computer
-[a-z]	any lowercase letter	[a-e]	computer
-()	group expressions	(pre)*y	Happy osprey
-{}	# of times expression is matched	[a-z]{3}	and Computer computer

- Automata Types



DFA:

Σ input alphabet

S finite nonempty set of states

S_0 start state

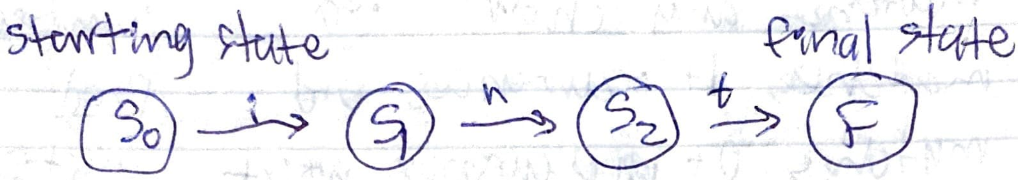
δ state trans func

F final states set

$$\left.\begin{array}{ll} a \in \Sigma & a = \text{letter} \\ x \in \Sigma^* & x = \text{str} \end{array}\right\} a \cdot x = \text{str}$$

$\lambda \in \Sigma^0$ λ = empty str

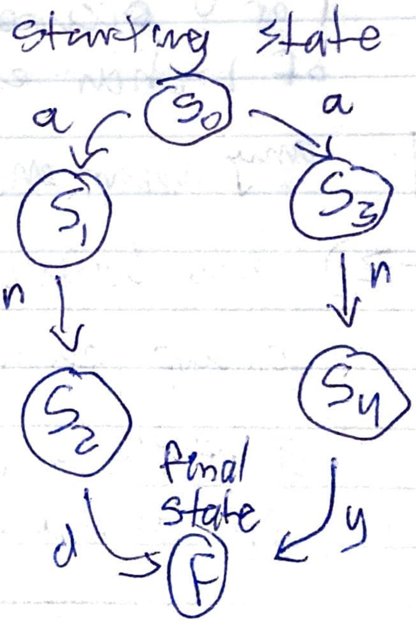
DFA int input



NFA regex "any | and"

input: and

must guess first


starting state
S0
a
a
S1
S3
n
n
S2
S4
final state
d
y
F