Notes of
Formal Laguage and Automata
CISC 3007

Box, ZHANG Huakang

November 23, 2021

1 Basic Definitions and Properties

Alphabets

- An alphabet is a finite set of symbols.
- Usually use Σ to represent an alphabet.

Strings

Definition

- A string is a finite sequence of symbols feom an alphabet.

String Operations

- Length: $|1100| = 4$
- Prefix
- Suffix
- Substring
- Concarenation: $\alpha = abd, \beta = ce, \alpha\beta = abdce$
- Exponentiation: $\alpha = abd, \alpha^3 = abdabdabd, \alpha^0 = \epsilon$
- Reversal: $\alpha = abd, \alpha^{Rev} = dba$
- Power of an alphabet: Σ^k is the set of all k-length strings formed by the alphabet in Σ. e.g., $\Sigma = \{a, b\}$, $\Sigma^2 = \{ab, aa, bb, ba\}$, $\Sigma^0 = \{\epsilon\}$
- Kleen Closure: $\Sigma^* = \Sigma^0 \cup \Sigma^1 ... = \cup_{k\geq 0}\Sigma^k$
- Kleen Plus: $\Sigma^+ = \Sigma^1 \cup \Sigma^2 ... = \cup_{k>0}\Sigma^k$

Languages

Definition A language is a set of strin gs over an alphabet.

2 Finite Automata

Deterministic Finite Automata

A DFA is a quintuple $(Q, \Sigma, \delta, q_0, F)$ where

- Q is a finite set of states
- Σ is a finite input alphabet

- δ is the transition function mapping $Q \times \Sigma$ to Q
- q_0 in Q is the initial state (only one)
- $F \subset Q$ is the set of final state(s) (zero or more)

Language of a DFA Giuven a DFA M, the language accepted (or recognized) by M is the set of all strings that start from the initial state, and reache one of the finnal states.

Non-deterministic Finite Automata

For each state, zero, one or more transitions are allowed on the same input symbol. A NFA is a quintuple $(Q, \Sigma, \delta, q_0, F)$ where

- Q is a finite set of states
- Σ is a finite input alphabet
- δ is the transition function mapping $Q \times \Sigma$ to a subset of Q
- q_0 in Q is the initial state (only one)
- $F \subset Q$ is the set of final state(s) (zero or more)

Noticed that the only difference between an NFA and a DFA is in the type of valu that δ returns: a set of states in the case of an NFA and a single state in the case of DFA.

Language of a NFA Given a NFA M, the language recognized by M is the set of all strings that start from the initial state, and has at least on path reaching a final state.

DFA and NFA NFA is equivalent to DFA

Constructing a DFA from a NFA

Given ang NFA $M = (Q, \Sigma, \delta, q_0, F)$ recognizing a language L over Σ, we can construct a DFA $N = (Q', \Sigma, \delta', q_0, F')$ which also recognizes L:

- Q' is the set of all subset of Q
- $q_0 = \{q_0\}$
- F' is the set of all states in Q' containing a finnal state of M
- $\delta'(\{q_1.q_2, ..\}, a) = \delta(q_1, a) \cup \delta(q_2, a)...$

NFA with ϵ-Transition

ϵ-Closures In an ϵ-NFA, the ϵCLOSE(q) of a state q is the set of states (including q) that can be reached from q by following a path whose edges are all labeled by ϵ

ϵ-NFA $\rightarrow$ NFA

Given any ϵ-NFA $M = (Q, \Sigma, \delta, q_0, F)$ recognizing a language L over Σ, we can construce its NFA $N = (Q, \Sigma, \delta', q_0, F')$ that also recognizes L:

- $\delta'(q_i, a) = q_j$ iff there is a path from q_i to q_j using exactly one arc labeled a and sero or more arcs labeled ϵ in M.
- $F' = F \cup \{q_0\}$ if a final state is reachable from q_0 using some ϵ-transitions in M. Otherwise, $F' = F$.