Regular Grammars

Chapter 7

Regular Grammars

A regular grammar G is a quadruple (V, Σ, R, S), where:

- V is the rule alphabet, which contains nonterminals and terminals,
- Σ (the set of terminals) is a subset of V,
- R is a finite set of rules of the form:

  $$X \rightarrow Y$$

- $S \in V - \Sigma$ -- the start symbol

Regular Grammars

In a regular grammar, all rules in R must:

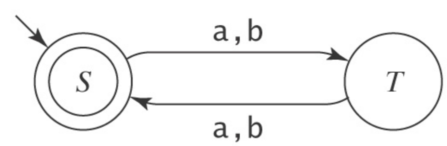
- have a left hand side that is a single nonterminal
- have a right hand side that is:
  - ε, or
  - a single terminal, or
  - a single terminal followed by a single nonterminal.

Legal: $S \rightarrow \mathtt{a}$, $S \rightarrow \varepsilon$, and $T \rightarrow \mathtt{a}S$
Not legal: $S \rightarrow \mathtt{a}S\mathtt{a}$ and $\mathtt{a}S\mathtt{a} \rightarrow T$

The language defined by a grammar: all terminal strings that can be obtained starting from S and applying the rules

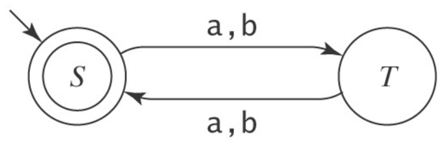
Regular Grammar Example

$L = \{w \in \{\mathtt{a}, \mathtt{b}\}^* : |w| \text{ is even}\}$ $\quad ((\mathtt{aa}) \cup (\mathtt{ab}) \cup (\mathtt{ba}) \cup (\mathtt{bb}))^*$



Regular Grammar Example

$L = \{w \in \{\mathtt{a}, \mathtt{b}\}^* : |w| \text{ is even}\}$ $((\mathtt{aa}) \cup (\mathtt{ab}) \cup (\mathtt{ba}) \cup (\mathtt{bb}))^*$



Grammar:

$S \to \varepsilon$
$S \to \mathtt{a}T$
$S \to \mathtt{b}T$
$T \to \mathtt{a}$
$T \to \mathtt{b}$
$T \to \mathtt{a}S$
$T \to \mathtt{b}S$

Language:

S → bT → bb
S → aT → abS → abbT
→ abbaS → abba
S → ε

Regular Languages and Regular Grammars

Theorem: The class of languages that can be defined with regular grammars is exactly the regular languages.

Proof: By two constructions.

Regular Languages and Regular Grammars

Regular grammar → FSM:

grammartofsm($G = (V, \Sigma, R, S)$) =

1. Create in M a separate state for each nonterminal in V.
2. Start state is the state corresponding to S .
3. If there are any rules in R of the form $X \to w$, for some $w \in \Sigma$, create a new state labeled #.
4. For each rule of the form $X \to w\ Y$, add a transition from X to Y labeled w.
5. For each rule of the form $X \to w$, add a transition from X to # labeled w.
6. For each rule of the form $X \to \varepsilon$, mark state X as accepting.
7. Mark state # as accepting.

FSM → Regular grammar: Similarly.

Strings that End with aaaa

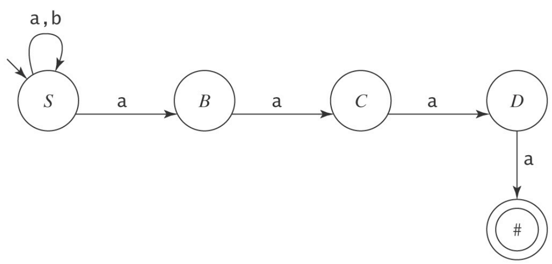
$L = \{w \in \{\mathtt{a}, \mathtt{b}\}^* : w \text{ ends with the pattern } \mathtt{aaaa}\}$.

$S \to \mathtt{a}S$
$S \to \mathtt{b}S$
$S \to \mathtt{a}B$
$B \to \mathtt{a}C$
$C \to \mathtt{a}D$
$D \to \mathtt{a}$

Strings that End with aaaa

$L = \{w \in \{\texttt{a}, \texttt{b}\}^* : w$ ends with the pattern $\texttt{aaaa}\}$.

$S \rightarrow \texttt{a}S$
$S \rightarrow \texttt{b}S$
$S \rightarrow \texttt{a}B$
$B \rightarrow \texttt{a}C$
$C \rightarrow \texttt{a}D$
$D \rightarrow \texttt{a}$

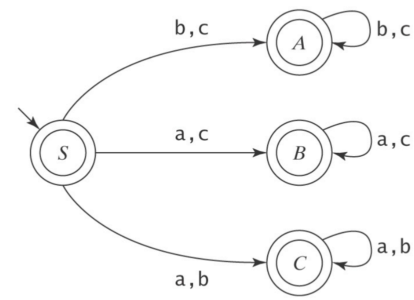


Example 2 – One Character Missing

$S \rightarrow \varepsilon$
$S \rightarrow \texttt{a}B$
$S \rightarrow \texttt{a}C$
$S \rightarrow \texttt{b}A$
$S \rightarrow \texttt{b}C$
$S \rightarrow \texttt{c}A$
$S \rightarrow \texttt{c}B$

$A \rightarrow \texttt{b}A$
$A \rightarrow \texttt{c}A$
$A \rightarrow \varepsilon$
$B \rightarrow \texttt{a}B$
$B \rightarrow \texttt{c}B$
$B \rightarrow \varepsilon$

$C \rightarrow \texttt{a}C$
$C \rightarrow \texttt{b}C$
$C \rightarrow \varepsilon$



Regular Languages