Regular Grammars

Chapter 7

Grammars

A 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:

$$X \rightarrow Y, \quad X, Y \in V^*$$

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

Regular Grammars

In a regular grammar, all rules in *R* must:

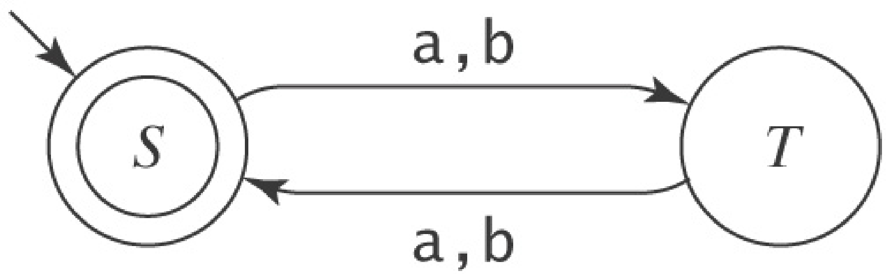
- have a left hand side that is a single nonterminal
- have a right hand side that is:
  - ε, or
  - a single terminal, or
  - a single terminal followed by a single nonterminal.

Legal: $S \rightarrow$ a, $S \rightarrow \varepsilon$, and $T \rightarrow$ aS
Not legal: $S \rightarrow$ aSa and aSa $\rightarrow T$

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

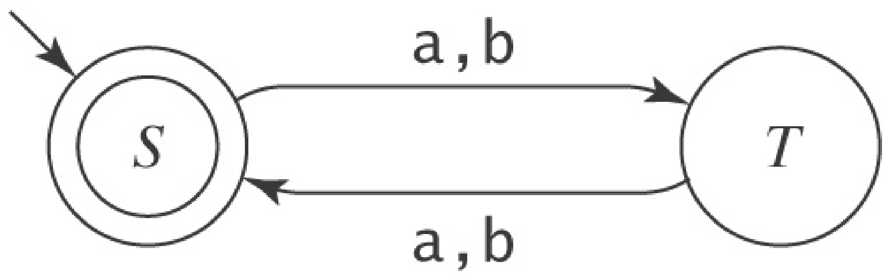
Regular Grammar Example

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



Regular Grammar Example

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



Grammar:

$S \rightarrow \varepsilon$
$S \rightarrow aT$
$S \rightarrow bT$
$T \rightarrow a$
$T \rightarrow b$
$T \rightarrow aS$
$T \rightarrow bS$

Language:

$S \Rightarrow bT \Rightarrow bb$
$S \Rightarrow aT \Rightarrow abS \Rightarrow abbT$
$\Rightarrow abbaS \Rightarrow abba$
$S \Rightarrow \varepsilon$

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* $\rightarrow$ *FSM:

grammartofsm(G = (V, Σ, 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 $X \rightarrow a$, for some $a \in \Sigma$, create a new state labeled #.
4. For each rule of the form $X \rightarrow a\ Y$, add a transition from X to Y labeled a.
5. For each rule of the form $X \rightarrow a$, add a transition from X to # labeled a.
6. For each rule of the form $X \rightarrow \varepsilon$, mark state X as accepting.
7. Mark state # as accepting.

FSM* $\rightarrow$ *Regular grammar: Similarly.

Strings that End with aaaa

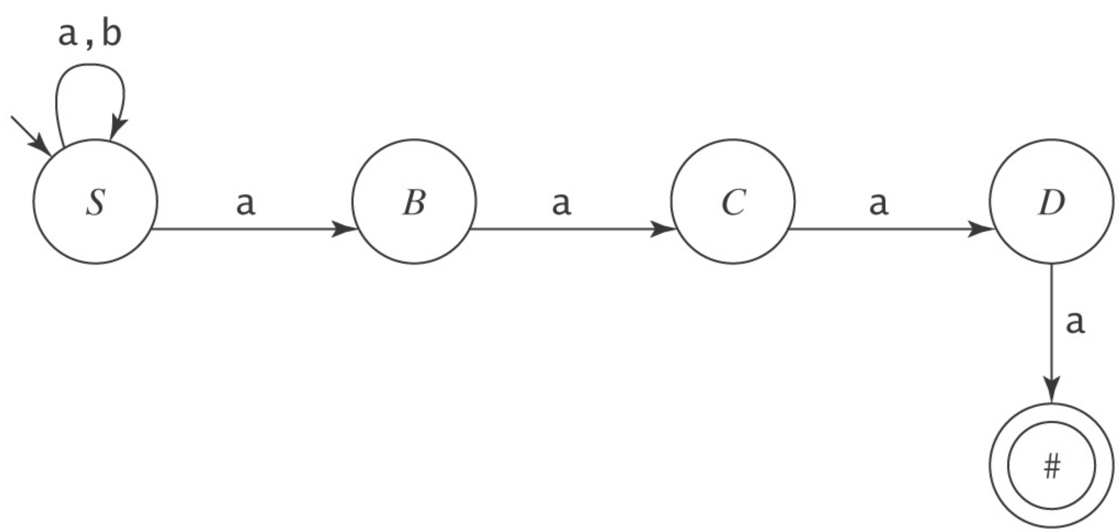
$L = \{w \in \{a, b\}^* : w$ ends with the pattern aaaa$\}$.

$$S \rightarrow aS$$
$$S \rightarrow bS$$
$$S \rightarrow aB$$
$$B \rightarrow aC$$
$$C \rightarrow aD$$
$$D \rightarrow a$$

Strings that End with aaaa

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

$S \rightarrow aS$
$S \rightarrow bS$
$S \rightarrow aB$
$B \rightarrow aC$
$C \rightarrow aD$
$D \rightarrow a$

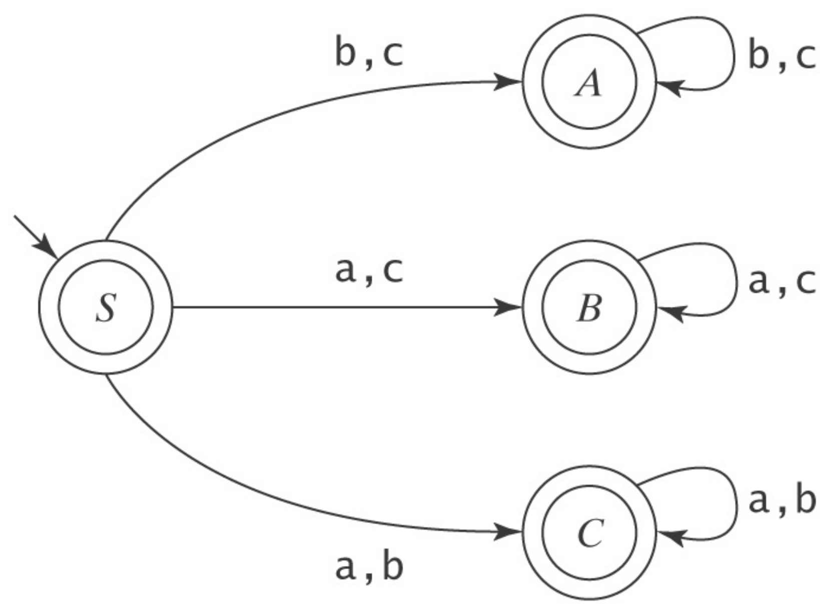


Example 2 – One Character Missing

$S \to \varepsilon$
$S \to aB$
$S \to aC$
$S \to bA$
$S \to bC$
$S \to cA$
$S \to cB$

$A \to bA$
$A \to cA$
$A \to \varepsilon$
$B \to aB$
$B \to cB$
$B \to \varepsilon$

$C \to aC$
$C \to bC$
$C \to \varepsilon$



Conversions