Grammars

Grammars

Grammars express languages

Example: the English language

$$\langle sentence \rangle \rightarrow \langle noun\_phrase \rangle \ \langle predicate \rangle$$

$$\langle noun\_phrase \rangle \rightarrow \langle article \rangle \ \langle noun \rangle$$

$$\langle predicate \rangle \rightarrow \langle verb \rangle$$

$$\langle article \rangle \to a$$
$$\langle article \rangle \to the$$

$$\langle noun \rangle \to cat$$
$$\langle noun \rangle \to dog$$

$$\langle verb \rangle \to runs$$
$$\langle verb \rangle \to walks$$

3

A derivation of "the dog walks":

$$\begin{aligned}
\langle sentence \rangle &\Rightarrow \langle noun_phrase \rangle \ \langle predicate \rangle \\
&\Rightarrow \langle noun_phrase \rangle \ \langle verb \rangle \\
&\Rightarrow \langle article \rangle \ \langle noun \rangle \ \langle verb \rangle \\
&\Rightarrow the \ \ \langle noun \rangle \ \langle verb \rangle \\
&\Rightarrow the \ \ dog \ \ \langle verb \rangle \\
&\Rightarrow the \ \ dog \ \ walks
\end{aligned}$$

4

A derivation of "a cat runs":

$$\begin{aligned}
\langle sentence \rangle &\Rightarrow \langle noun_phrase \rangle \ \langle predicate \rangle \\
&\Rightarrow \langle noun_phrase \rangle \ \langle verb \rangle \\
&\Rightarrow \langle article \rangle \ \langle noun \rangle \ \langle verb \rangle \\
&\Rightarrow a \ \langle noun \rangle \ \langle verb \rangle \\
&\Rightarrow a \ \ cat \ \langle verb \rangle \\
&\Rightarrow a \ \ cat \ \ runs
\end{aligned}$$

5

Language of the grammar:

L = { "a cat runs",
"a cat walks",
"the cat runs",
"the cat walks",
"a dog runs",
"a dog walks",
"the dog runs",
"the dog walks" }

6

Notation

Production Rules

$$\underbrace{\begin{array}{l}\langle noun \rangle \to cat \\ \langle noun \rangle \to dog\end{array}}$$

Variable

Terminal

7

Another Example

Grammar: $S \to aSb$

$S \to \lambda$

Derivation of sentence ab:

$$S \Rightarrow aSb \Rightarrow ab$$

$S \to aSb$ $\qquad$ $S \to \lambda$

8

Grammar: $S \rightarrow aSb$

$S \rightarrow \lambda$

Derivation of sentence $aab b$:

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aabb$$

$S \rightarrow aSb$ $\qquad$ $S \rightarrow \lambda$

9

Other derivations:

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaaSbbb \Rightarrow aaabbb$$

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaaSbbb$$
$$\Rightarrow aaaaSbbbb \Rightarrow aaaabbbb$$

10

Language of the grammar

$$S \to aSb$$
$$S \to \lambda$$

$$L = \{a^n b^n : n \geq 0\}$$

More Notation

Grammar $G = (V, T, S, P)$

$V:$ Set of variables

$T:$ Set of terminal symbols

$S:$ Start variable

$P:$ Set of Production rules

Example

Grammar G : $S \to aSb$

$S \to \lambda$

$$G = (V, T, S, P)$$

$V = \{S\}$ $T = \{a, b\}$

$P = \{S \to aSb,\ S \to \lambda\}$

More Notation

Sentential Form:

A sentence that contains variables and terminals

Example:

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaaSbbb \Rightarrow aaabbb$$

Sentential Forms sentence

We write: $S \overset{*}{\Rightarrow} aaabbb$

Instead of:

$$S \Rightarrow aSb \Rightarrow aaSbb \Rightarrow aaaSbbb \Rightarrow aaabbb$$

15

In general we write: $w_1 \overset{*}{\Rightarrow} w_n$

If: $w_1 \Rightarrow w_2 \Rightarrow w_3 \Rightarrow \cdots \Rightarrow w_n$

16

By default: $w \overset{*}{\Rightarrow} w$

17

Example

Grammar

$S \rightarrow aSb$

$S \rightarrow \lambda$

Derivations

$S \overset{*}{\Rightarrow} \lambda$

$S \overset{*}{\Rightarrow} ab$

$S \overset{*}{\Rightarrow} aabb$

$S \overset{*}{\Rightarrow} aaabbb$

18

Example

Grammar

$$S \rightarrow aSb$$

$$S \rightarrow \lambda$$

Derivations

$$S \overset{*}{\Rightarrow} aaSbb$$

$$aaSbb \overset{*}{\Rightarrow} aaaaaSbbbbb$$

Another Grammar Example

Grammar G:

$$S \rightarrow Ab$$

$$A \rightarrow aAb$$

$$A \rightarrow \lambda$$

Derivations:

$$S \rightarrow Ab \rightarrow b$$

$$S \rightarrow Ab \rightarrow aAbb \rightarrow abb$$

$$S \rightarrow Ab \rightarrow aAbb \rightarrow aaAbbb \rightarrow aabbb$$

More Derivations

$$S \Rightarrow Ab \Rightarrow aAbb \Rightarrow aaAbbb \Rightarrow aaaAbbbb$$
$$\Rightarrow aaaaAbbbbb \Rightarrow aaaabbbbb$$

$$S \overset{*}{\Rightarrow} aaaabbbbb$$

$$S \overset{*}{\Rightarrow} aaaaaabbbbbbb$$

$$S \overset{*}{\Rightarrow} a^n b^n b$$

21

Language of a Grammar

For a grammar G
with start variable S :

$$L(G) = \{w : \quad S \overset{*}{\Rightarrow} w\}$$

String of terminals

22

Example

For grammar G : $S \to Ab$

$$A \to aAb$$

$$A \to \lambda$$

$$L(G) = \{a^n b^n b : \quad n \geq 0\}$$

Since: $S \overset{*}{\Rightarrow} a^n b^n b$

23

A Convenient Notation

$$A \to aAb$$
$$A \to \lambda$$

$$\Rightarrow \quad A \to aAb \mid \lambda$$

$$\langle article \rangle \to a$$
$$\langle article \rangle \to the$$

$$\Rightarrow \quad \langle article \rangle \to a \mid the$$

24

Linear Grammars

25

Linear Grammars

Grammars with
at most one variable at the right side
of a production

Examples:

$$S \to aSb$$
$$S \to \lambda$$

$$S \to Ab$$
$$A \to aAb$$
$$A \to \lambda$$

26

A Non-Linear Grammar

Grammar G :

$$S \to SS$$
$$S \to \lambda$$
$$S \to aSb$$
$$S \to bSa$$

$$L(G) = \{w: \ n_a(w) = n_b(w)\}$$

Number of a in string w

27

Another Linear Grammar

Grammar G :

$$S \to A$$
$$A \to aB \mid \lambda$$
$$B \to Ab$$

$$L(G) = \{a^n b^n : n \ge 0\}$$

28

Right-Linear Grammars

All productions have form: $A \to xB$

or

$A \to x$

Linear Grammars has 2 type

Left and Right

string of terminals

Example: $S \to abS$

$S \to a$

29

Left-Linear Grammars

All productions have form: $A \to Bx$

or

$A \to x$

string of terminals

Example: $S \to Aab$

$A \to Aab \mid B$

$B \to a$

30

Regular Grammars

Regular Grammars

A regular grammar is any
right-linear or left-linear grammar

Examples:

G_1

$$S \rightarrow abS$$
$$S \rightarrow a$$

G_2

$$S \rightarrow Aab$$
$$A \rightarrow Aab \mid B$$
$$B \rightarrow a$$

Observation

Regular grammars generate regular languages

Examples:

G_1

$S \to abS$

$S \to a$

$L(G_1) = (ab)^* a$

G_2

$S \to Aab$

$A \to Aab \mid B$

$B \to a$

$L(G_2) = aab(ab)^*$

Regular Grammars
Generate
Regular Languages

Theorem

$$\left\{ \begin{array}{l} \text{Languages} \\ \text{Generated by} \\ \text{Regular Grammars} \end{array} \right\} = \left\{ \begin{array}{l} \text{Regular} \\ \text{Languages} \end{array} \right\}$$

Theorem - Part 1

$$\left\{ \begin{array}{l} \text{Languages} \\ \text{Generated by} \\ \text{Regular Grammars} \end{array} \right\} \subseteq \left\{ \begin{array}{l} \text{Regular} \\ \text{Languages} \end{array} \right\}$$

Any regular grammar generates
a regular language

Theorem - Part 2

$$\left\{ \begin{array}{l} \text{Languages} \\ \text{Generated by} \\ \text{Regular Grammars} \end{array} \right\} \supseteq \left\{ \begin{array}{l} \text{Regular} \\ \text{Languages} \end{array} \right\}$$

Any regular language is generated by a regular grammar

37

Proof – Part 1

$$\left\{ \begin{array}{l} \text{Languages} \\ \text{Generated by} \\ \text{Regular Grammars} \end{array} \right\} \subseteq \left\{ \begin{array}{l} \text{Regular} \\ \text{Languages} \end{array} \right\}$$

The language $L(G)$ generated by any regular grammar G is regular

38

The case of Right-Linear Grammars

Let G be a right-linear grammar

We will prove: $L(G)$ is regular

Proof idea: We will construct NFA M with $L(M) = L(G)$

39

Grammar G is right-linear

Example:

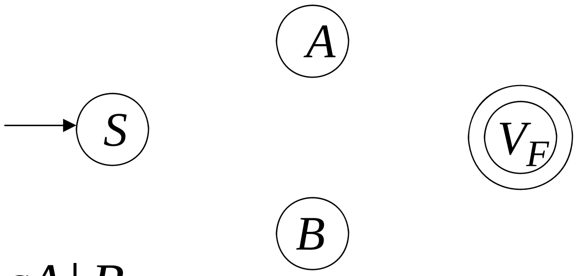
$$S \to aA \mid B$$

$$A \to aa\,B$$

$$B \to b\,B \mid a$$

40

Construct NFA M such that every state is a grammar variable:



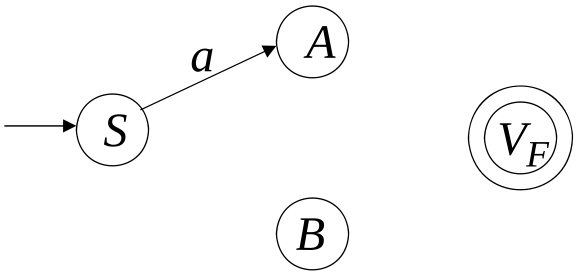
special final state

$S \to aA \mid B$

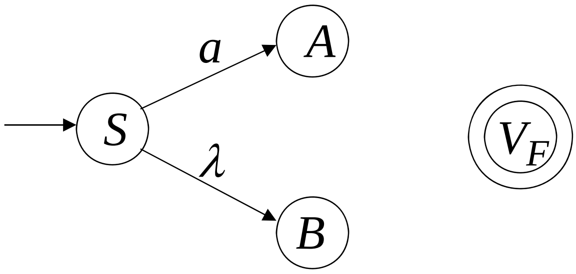
$A \to aa\, B$

$B \to b\, B \mid a$

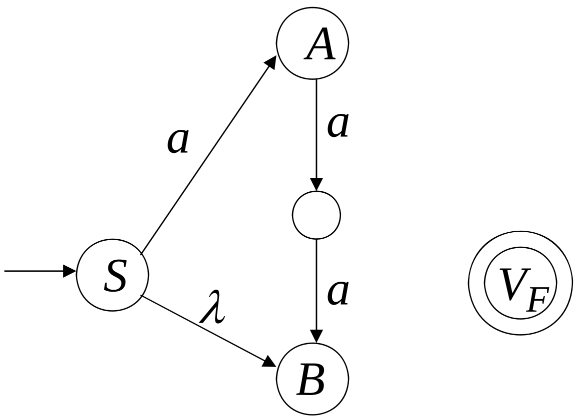
Add edges for each production:



$S \to aA$

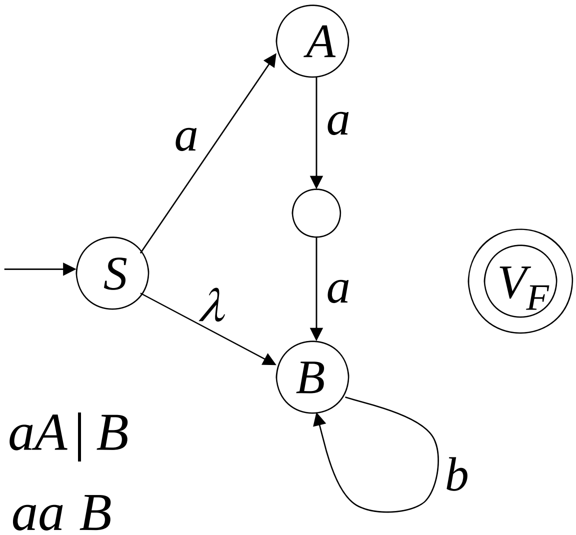


$$S \to aA \mid B$$



$$S \to aA \mid B$$

$$A \to aa\ B$$

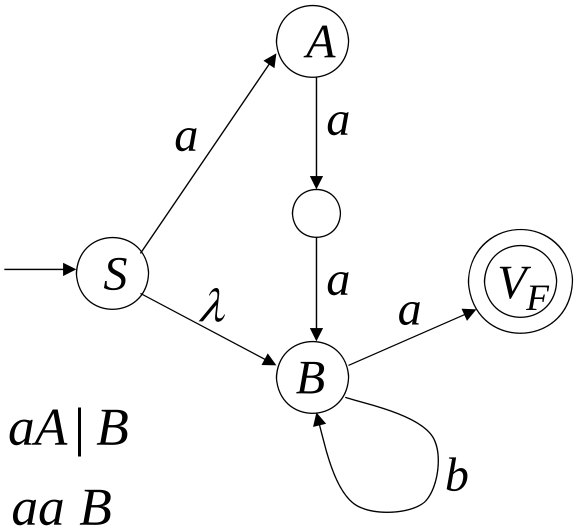


$$S \to aA \mid B$$

$$A \to aa\ B$$

$$B \to bB$$

45

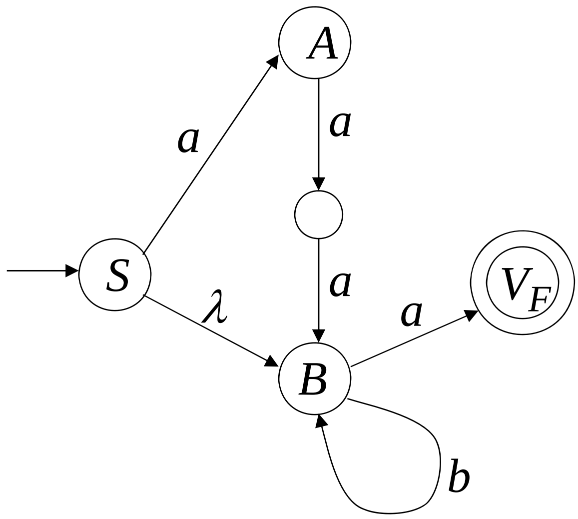


$$S \to aA \mid B$$

$$A \to aa\ B$$

$$B \to bB \mid a$$

46



$$S \Rightarrow aA \Rightarrow aaaB \Rightarrow aaabB \Rightarrow aaaba$$

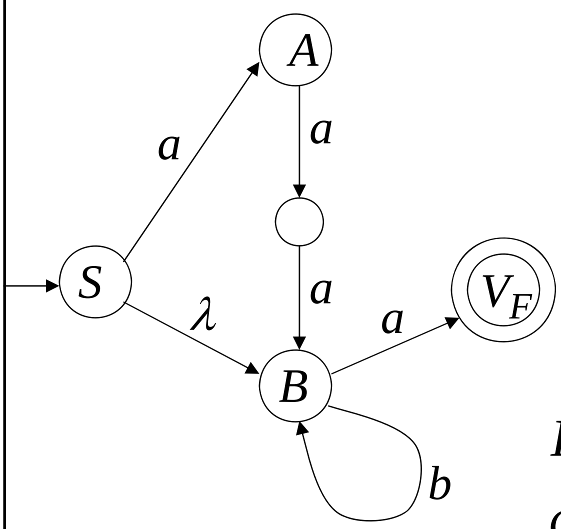
NFA M

Grammar G

$$S \rightarrow aA \mid B$$

$$A \rightarrow aa\ B$$

$$B \rightarrow bB \mid a$$



$$L(M) = L(G) =$$

$$aaab^*a + b^*a$$

In General

A right-linear grammar G

has variables: $V_0, V_1, V_2, \ldots$

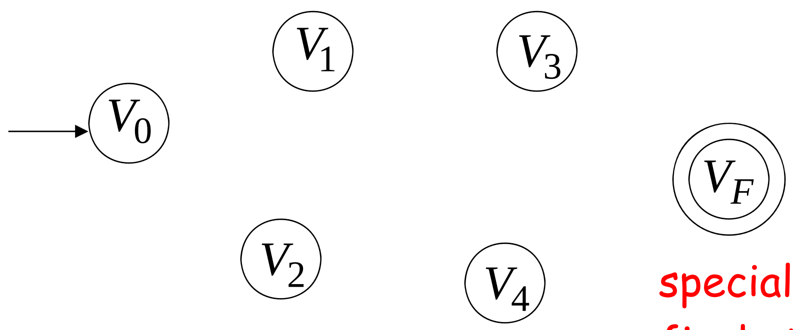
and productions: $V_i \to a_1 a_2 \cdots a_m V_j$

or

$V_i \to a_1 a_2 \cdots a_m$

We construct the NFA M such that:

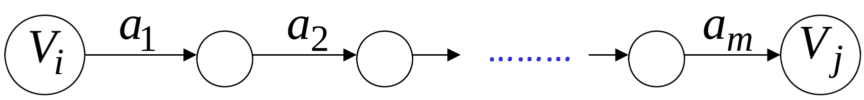
each variable V_i corresponds to a node:



special
final state

For each production: $V_i \to a_1 a_2 \cdots a_m V_j$

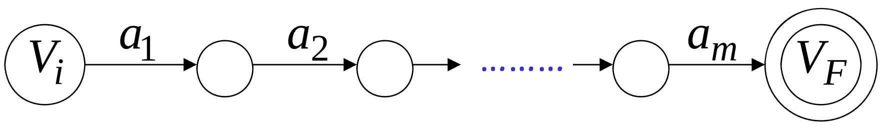
we add transitions and intermediate nodes



51

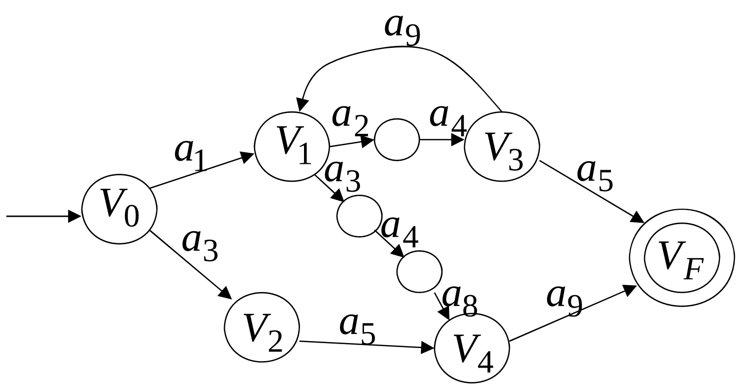
For each production: $V_i \to a_1 a_2 \cdots a_m$

we add transitions and intermediate nodes



52

Resulting NFA M looks like this:



It holds that: $L(G) = L(M)$

The case of Left-Linear Grammars

Let G be a left-linear grammar

We will prove: $L(G)$ is regular

Proof idea:

We will construct a right-linear grammar G' with $L(G) = L(G')^R$

Since G is left-linear grammar
the productions look like:

$$A \to Ba_1a_2 \cdots a_k$$

$$A \to a_1a_2 \cdots a_k$$

55

Construct right-linear grammar G'

Left linear G

$$A \to Ba_1a_2 \cdots a_k$$

$$A \to Bv$$

$$\Downarrow$$

Right linear G'

$$A \to a_k \cdots a_2a_1B$$

$$A \to v^R B$$

56

Construct right-linear grammar G'

Left linear G

$$A \to a_1 a_2 \cdots a_k$$

$$A \to v$$

Right linear G'

$$A \to a_k \cdots a_2 a_1$$

$$A \to v^R$$

57

It is easy to see that: $L(G) = L(G')^R$

Since G' is right-linear, we have:

$L(G')$ $\Rightarrow$ $L(G')^R$ $\Rightarrow$ $L(G)$

Regular Language | Regular Language | Regular Language

58

Proof - Part 2

$$\left\{\begin{array}{l}\text{Languages} \\ \text{Generated by} \\ \text{Regular Grammars}\end{array}\right\} \supseteq \left\{\begin{array}{l}\text{Regular} \\ \text{Languages}\end{array}\right\}$$

Any regular language L is generated by some regular grammar G

Any regular language L is generated by some regular grammar G

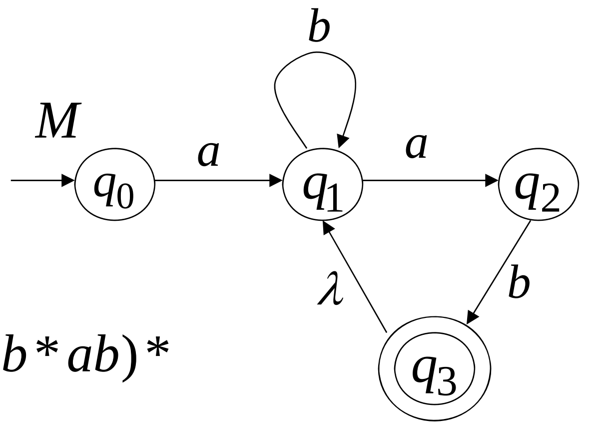
Proof idea:

Let M be the NFA with $L = L(M)$.

Construct from M a regular grammar G such that $L(M) = L(G)$

Since L is regular
there is an NFA M such that $L = L(M)$

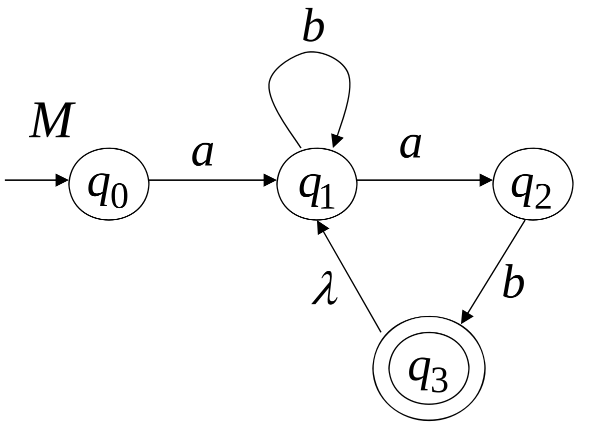
Example:



$L = ab^*ab(b^*ab)^*$
$L = L(M)$

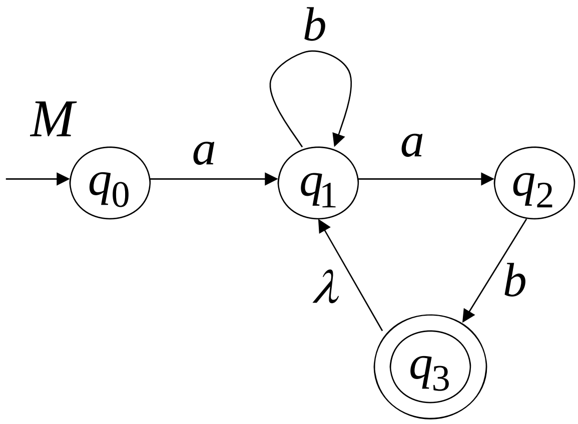
61

Convert M to a right-linear grammar



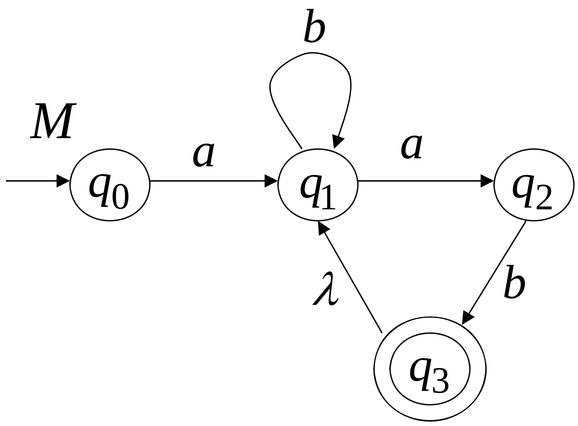
$q_0 \to aq_1$

62



$$q_0 \rightarrow aq_1$$
$$q_1 \rightarrow bq_1$$
$$q_1 \rightarrow aq_2$$

63



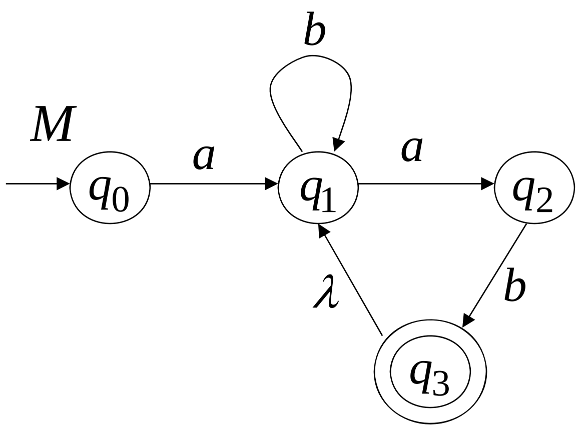
$$q_0 \rightarrow aq_1$$
$$q_1 \rightarrow bq_1$$
$$q_1 \rightarrow aq_2$$
$$q_2 \rightarrow bq_3$$

64

$$L(G) = L(M) = L$$

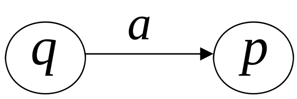
G

$$q_0 \to a q_1$$
$$q_1 \to b q_1$$
$$q_1 \to a q_2$$
$$q_2 \to b q_3$$
$$q_3 \to q_1$$
$$q_3 \to \lambda$$



In General

For any transition:



Add production: $q \to a p$

variable terminal variable

For any final state: q_f

Add production: $q_f \rightarrow \lambda$

67

Since G is right-linear grammar

G is also a regular grammar

with $L(G) = L(M) = L$

68