Chapter 5

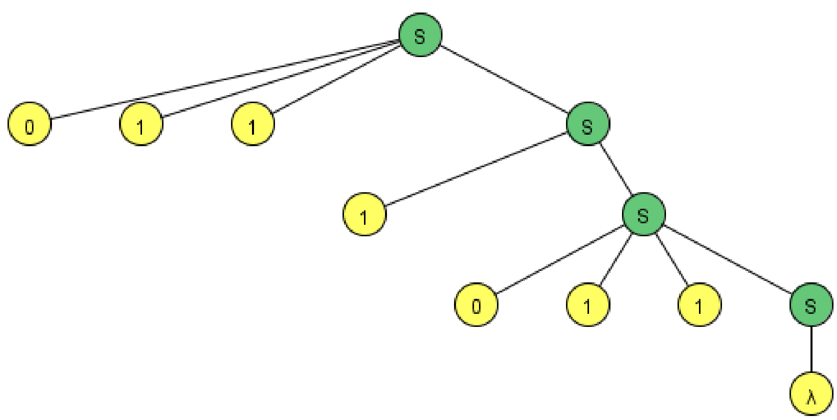
Grammars

A. Construct right-linear or left-linear grammars for the following regular languages:

1. Binary strings in which every 0 is followed by 11. Construct a parse tree for the string 0111011

Solution:

LHS		RHS
S	$\rightarrow$	1S
S	$\rightarrow$	011S
S	$\rightarrow$	λ

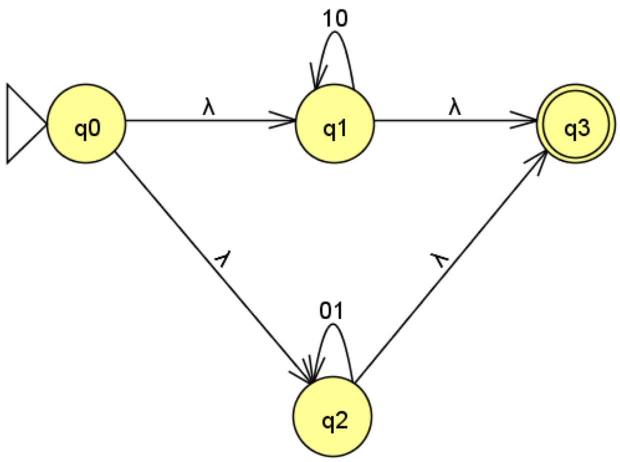


2. Binary strings in which each string has alternating 0 s and 1 s and is of even length.
Also, convert the grammar to an equivalent DFA. What is the minimum number of states in the equivalent DFA?

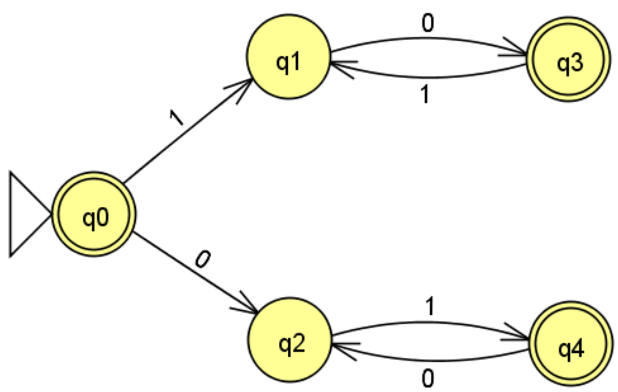
Solution:

LHS		RHS
S	$\rightarrow$	A
S	$\rightarrow$	B
A	$\rightarrow$	01A
B	$\rightarrow$	10B
A	$\rightarrow$	λ
B	$\rightarrow$	λ

Equivalent NFA:



Equivalent DFA:



This is already minimal. The minimum number of states is therefore 5 (since it is incorrect to merge the two branches into a single branch).

3. Binary strings that begin with 11 and end with 11 or begin with 00 and end with 00.

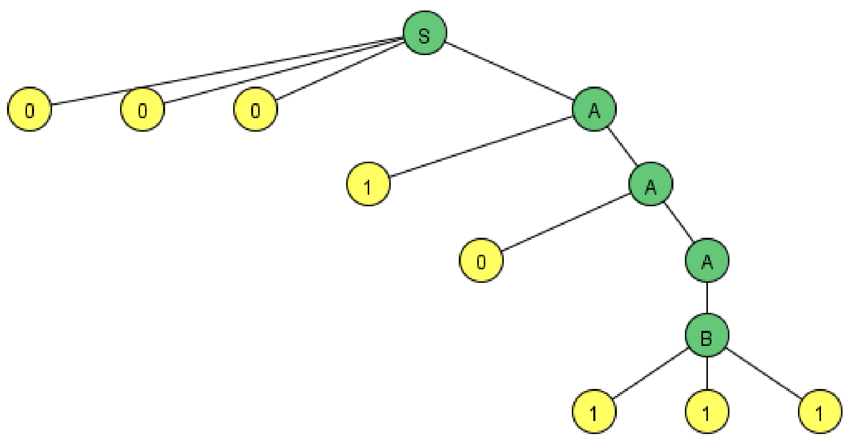
Solution: Assuming that strings in the language are at least 4 symbols long:

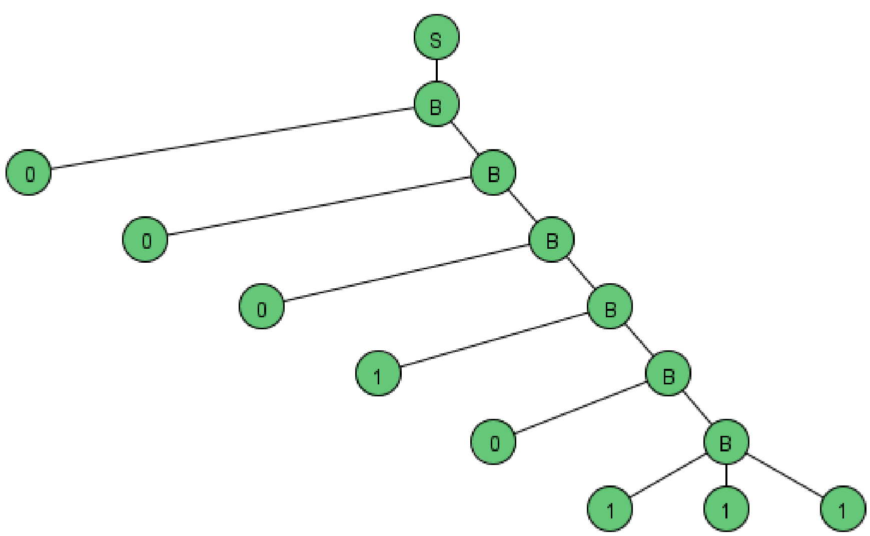
LHS		RHS
S	$\rightarrow$	11A
S	$\rightarrow$	00B
A	$\rightarrow$	0A
A	$\rightarrow$	1A
A	$\rightarrow$	11
B	$\rightarrow$	0B
B	$\rightarrow$	1B
B	$\rightarrow$	00

4. Binary strings starting with 000 or ending with 111 (or both). Show the derivation of 00010111

Solution: Grammar and two possible parse trees for the given string:

LHS		RHS
S	$\rightarrow$	000A
S	$\rightarrow$	B
A	$\rightarrow$	0A
A	$\rightarrow$	1A
A	$\rightarrow$	λ
A	$\rightarrow$	B
B	$\rightarrow$	0B
B	$\rightarrow$	1B
B	$\rightarrow$	111





5. Binary strings in which the sum of the last three digits is even (e.g., 00101011 but not 00101001).

Solution:

LHS		RHS
S	$\rightarrow$	0S
S	$\rightarrow$	1S
S	$\rightarrow$	000
S	$\rightarrow$	011
S	$\rightarrow$	101
S	$\rightarrow$	110

6. Strings over {*a*, *b*, *c*} that contain at least one *a* and at least one *b*.

Solution:

LHS		RHS
S	$\rightarrow$	aA
S	$\rightarrow$	bB
S	$\rightarrow$	cS
A	$\rightarrow$	aA
A	$\rightarrow$	bC
A	$\rightarrow$	cA
B	$\rightarrow$	bB
B	$\rightarrow$	aC
B	$\rightarrow$	cB
C	$\rightarrow$	aC
C	$\rightarrow$	bC
C	$\rightarrow$	cC
C	$\rightarrow$	λ

7. Strings over {*a*, *b*} that contain at least three *a* s or at least two *b* s.

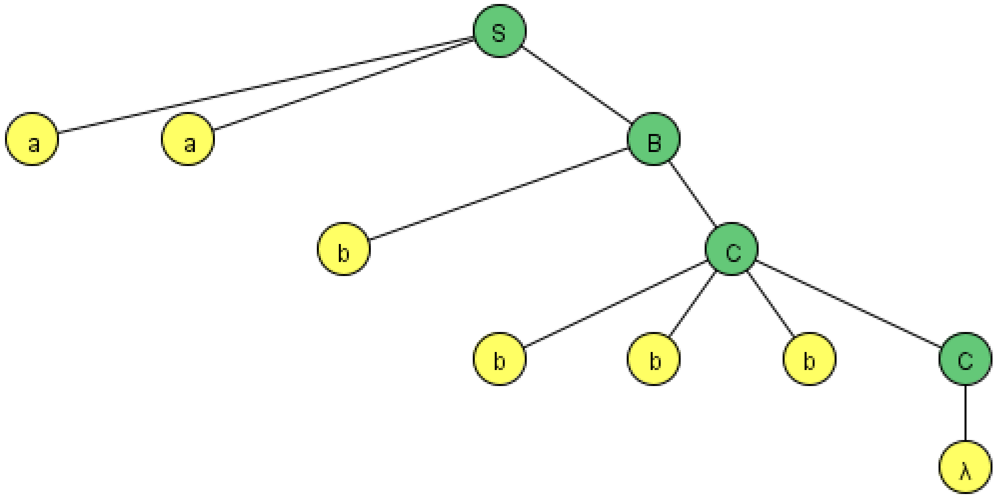
Solution: Please see also the VIDEO SOLUTION to follow the method for constructing the grammar shown below.

LHS		RHS
S	$\longrightarrow$	A
S	$\longrightarrow$	D
A	$\longrightarrow$	aB
B	$\longrightarrow$	aC
C	$\longrightarrow$	a
A	$\longrightarrow$	aA
B	$\longrightarrow$	aB
C	$\longrightarrow$	aC
A	$\longrightarrow$	bA
B	$\longrightarrow$	bB
C	$\longrightarrow$	bC
D	$\longrightarrow$	bE
E	$\longrightarrow$	b
D	$\longrightarrow$	bD
E	$\longrightarrow$	bE
D	$\longrightarrow$	aD
E	$\longrightarrow$	aE

8. Strings over $\{a, b\}$ in which some number of a s is followed by some number of b s with the total length of the string being divisible by 3. Show the parse tree for *aabbbb*.

Solution:

LHS		RHS
S	$\longrightarrow$	aaaS
S	$\longrightarrow$	aA
S	$\longrightarrow$	aaB
S	$\longrightarrow$	C
A	$\longrightarrow$	bbC
B	$\longrightarrow$	bC
C	$\longrightarrow$	bbbC
C	$\longrightarrow$	λ



9. Strings over {a, b} containing at least two a s and ending with an even number of b s.

Solution:

LHS		RHS
S	→	bS
S	→	aA
A	→	aA
A	→	bA
A	→	aB
B	→	bbB
B	→	λ

10. Strings over the alphabet {a, b} of the form $(ab)^n$, e.g., *ababab*.

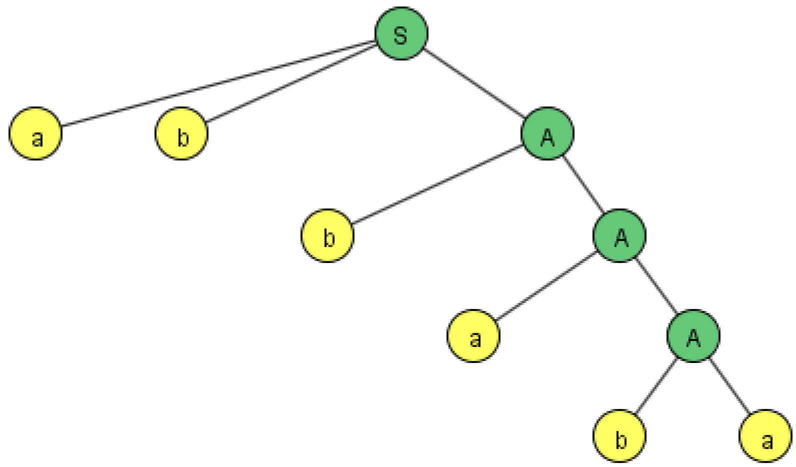
Solution:

LHS		RHS
S	→	abS
S	→	λ

11. Strings of the form *abwba* over the alphabet {a, b} where w is any string over the same alphabet. Show the parse tree for *abbaba*.

Solution:

LHS		RHS
S	→	abA
A	→	aA
A	→	bA
A	→	ba



12. The set of all binary strings that are palindromes of length 4. The alphabet is {*a*, *b*, *c*}.

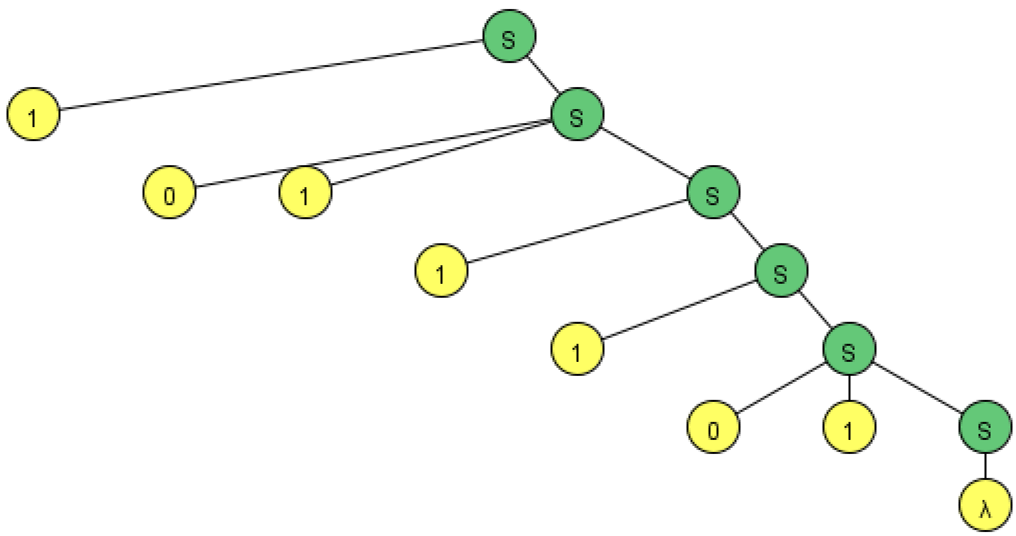
Solution:

LHS		RHS
S	⟶	aA
S	⟶	bB
S	⟶	cC
A	⟶	aD
A	⟶	bE
A	⟶	cF
D	⟶	aa
E	⟶	ba
F	⟶	ca
B	⟶	aG
B	⟶	bH
B	⟶	cI
G	⟶	ab
H	⟶	bb
I	⟶	cb
C	⟶	aJ
C	⟶	bK
C	⟶	cL
J	⟶	ac
K	⟶	bc
L	⟶	cc

13. Binary strings with no consecutive 0 s. Show the derivation of the string w_1 = 1011101 and also why w_2 = 1001 cannot be derived.

Solution:

LHS		RHS
S	$\longrightarrow$	1S
S	$\longrightarrow$	01S
S	$\longrightarrow$	λ
S	$\longrightarrow$	0



String 1001 is not generated by this grammar since it has a 0 that is not followed by a 1.

14. Binary strings in which the first part of each string contains at least four 1 s and the second part contains at least three 0 s.

Solution: Please see also the VIDEO SOLUTION for this exercise.

LHS		RHS
S	$\longrightarrow$	A
A	$\longrightarrow$	1B
B	$\longrightarrow$	1C
C	$\longrightarrow$	1D
D	$\longrightarrow$	1E
E	$\longrightarrow$	0F
F	$\longrightarrow$	0G
G	$\longrightarrow$	0
A	$\longrightarrow$	1A
A	$\longrightarrow$	0A
B	$\longrightarrow$	1B
B	$\longrightarrow$	0B
C	$\longrightarrow$	1C
C	$\longrightarrow$	0C
D	$\longrightarrow$	1D
D	$\longrightarrow$	0D
E	$\longrightarrow$	1E
E	$\longrightarrow$	0E
F	$\longrightarrow$	1F
F	$\longrightarrow$	0F
G	$\longrightarrow$	1G
G	$\longrightarrow$	0G

15. Binary strings with at least two occurrences of at least two consecutive 1 s, the two occurrences not being adjacent (i.e., 011011 is acceptable but 011111 is not).
Solution:

LHS		RHS
S	$\longrightarrow$	0S
S	$\longrightarrow$	1S
S	$\longrightarrow$	11A
A	$\longrightarrow$	1A
A	$\longrightarrow$	0B
B	$\longrightarrow$	1B
B	$\longrightarrow$	0B
B	$\longrightarrow$	11C
C	$\longrightarrow$	0C
C	$\longrightarrow$	1C
C	$\longrightarrow$	λ

16. Strings over $\{a, b\}$ of the form $aa^*(ab + a)^*$.
Solution:

LHS		RHS
S	$\longrightarrow$	aA
A	$\longrightarrow$	aA
A	$\longrightarrow$	λ
A	$\longrightarrow$	abA

17. Strings over {*a*, *b*} where the last two symbols in each string are a reversal of the first two symbols (i.e., last symbol = first symbol and penultimate symbol = second symbol).

Solution:

LHS		RHS
S	$\longrightarrow$	aaA
S	$\longrightarrow$	abB
S	$\longrightarrow$	baC
S	$\longrightarrow$	bbD
A	$\longrightarrow$	aa
B	$\longrightarrow$	ba
C	$\longrightarrow$	ab
D	$\longrightarrow$	bb
A	$\longrightarrow$	aA
A	$\longrightarrow$	bA
B	$\longrightarrow$	aB
B	$\longrightarrow$	bB
C	$\longrightarrow$	aC
C	$\longrightarrow$	bC
D	$\longrightarrow$	aD
D	$\longrightarrow$	bD

18. Student grades in an examination are represented with the letters {*A*, *B*, *C*, *D*, *F*}. A string such as *ABFCAD* indicates the grades obtained by a student in six different subjects. The grammar must generate only those strings that have at most three *D* s and no *F* s.

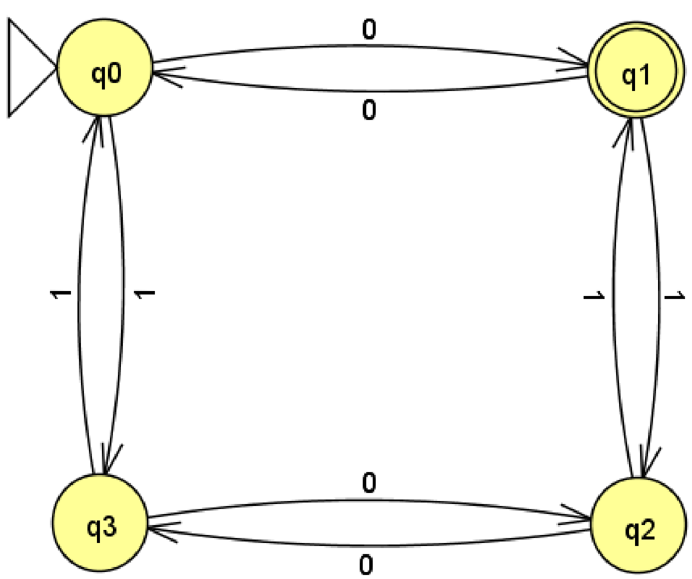
Solution:

LHS		RHS
S	→	AS
S	→	BS
S	→	CS
S	→	DU
U	→	AU
U	→	BU
U	→	CU
U	→	DV
V	→	AV
V	→	BV
V	→	CV
V	→	DW
W	→	AW
W	→	BW
W	→	CW
W	→	λ
S	→	λ
U	→	λ
V	→	λ

B. For the following, first construct an NFA or DFA and then convert it to a regular grammar.

19. Binary strings with an odd number of 0 s and an even number of 1 s.

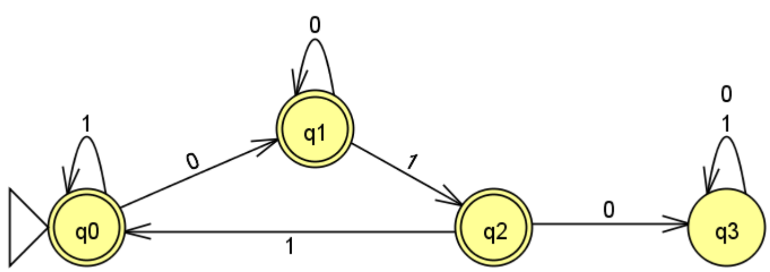
Solution:



LHS		RHS
S	→	0A
S	→	1C
C	→	1S
C	→	0B
B	→	0C
A	→	0S
A	→	1B
B	→	1A
A	→	λ

20. Strings over the binary alphabet that do not contain the substring 010.

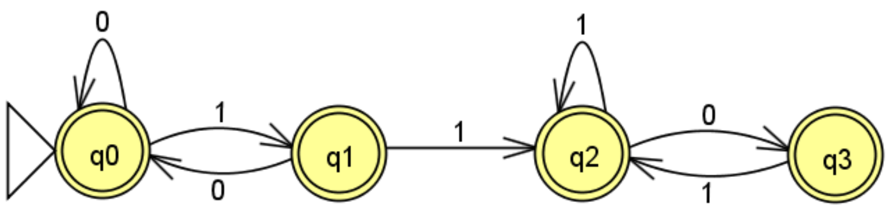
    Solution: Please see also the VIDEO SOLUTION for this exercise, especially to understand what happens in the grammar to the reject state.



LHS		RHS
S	→	0A
S	→	1S
S	→	λ
A	→	λ
A	→	1B
B	→	1S
B	→	λ
C	→	0C
A	→	0A
C	→	1C
B	→	0C

21. Binary strings in which every pair of adjacent 0 s appears before any pair of adjacent 1 s.

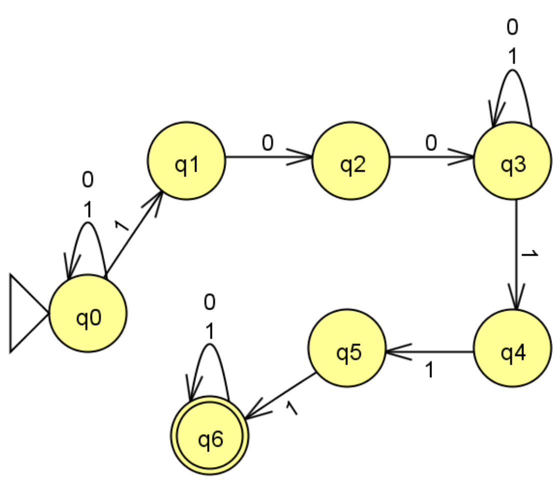
Solution:



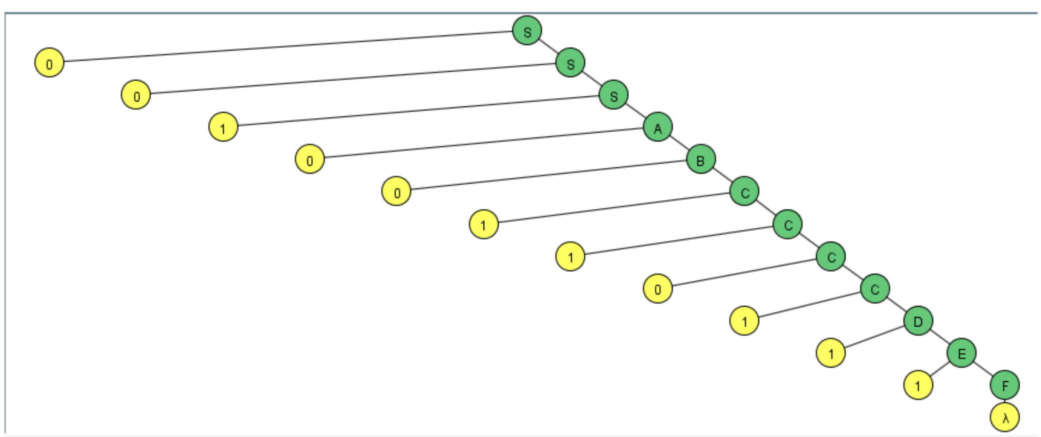
LHS		RHS
S	→	0S
S	→	λ
S	→	1A
A	→	0S
B	→	λ
A	→	λ
C	→	1B
B	→	0C
A	→	1B
C	→	λ
B	→	1B

22. The set of all binary strings having a substring 100 before a substring 111 (both must be present). Show the derivation for the string *w* = 00100110111.

Solution:

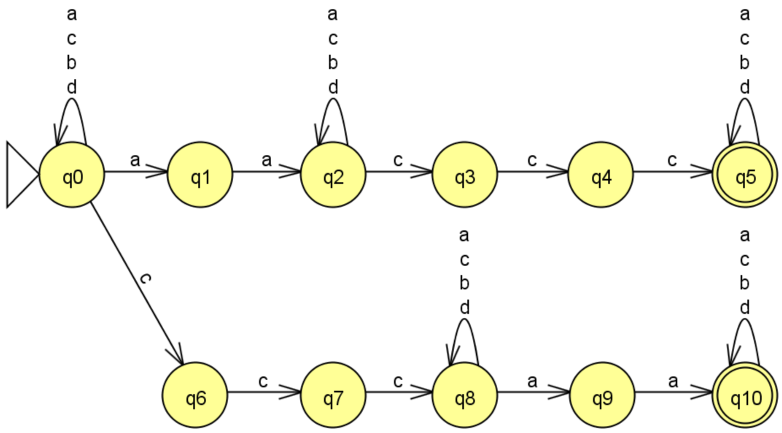


LHS		RHS
S	→	1S
S	→	0S
S	→	1A
B	→	0C
F	→	λ
F	→	0F
C	→	0C
F	→	1F
C	→	1C
C	→	1D
E	→	1F
D	→	1E
A	→	0B



23. A *run* in a string is a substring made up of a single symbol. For example, *aabbbcd* has a run of *a* s of length 2 and a run of *b* s of length 3. The grammar must derive all strings over the alphabet {*a*, *b*, *c*, *d*} with at least one run of *a* s of length 2 or a run of *c* s of length 3.

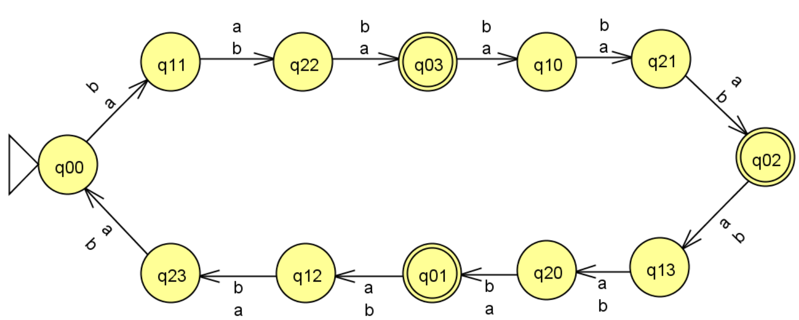
Solution:



LHS		RHS	LHS		RHS
S	$\rightarrow$	aA	H	$\rightarrow$	bH
S	$\rightarrow$	cF	B	$\rightarrow$	bB
S	$\rightarrow$	dS	J	$\rightarrow$	bJ
S	$\rightarrow$	bS	E	$\rightarrow$	cE
S	$\rightarrow$	cS	H	$\rightarrow$	cH
S	$\rightarrow$	aS	B	$\rightarrow$	cB
J	$\rightarrow$	λ	J	$\rightarrow$	cJ
E	$\rightarrow$	λ	B	$\rightarrow$	cC
G	$\rightarrow$	cH	B	$\rightarrow$	aB
C	$\rightarrow$	cD	J	$\rightarrow$	aJ
F	$\rightarrow$	cG	H	$\rightarrow$	aH
J	$\rightarrow$	dJ	E	$\rightarrow$	aE
B	$\rightarrow$	dB	I	$\rightarrow$	aJ
H	$\rightarrow$	dH	H	$\rightarrow$	aI
E	$\rightarrow$	dE	A	$\rightarrow$	aB
E	$\rightarrow$	bE	D	$\rightarrow$	cE

24. Strings over $\{a, b\}$ such that the length of the string is a multiple of 3 but not a multiple of 4.

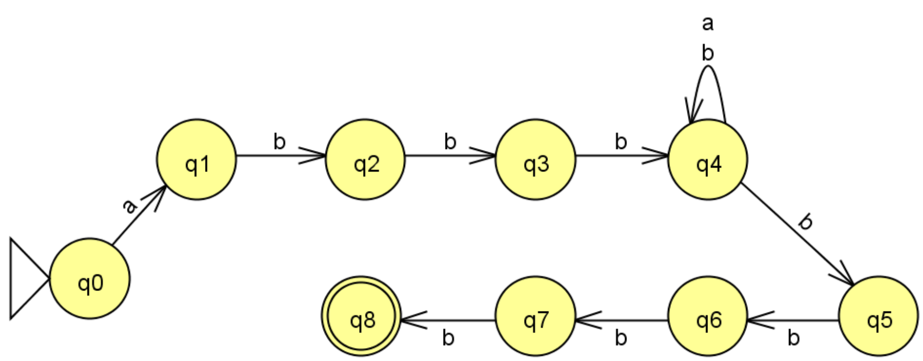
Solution:



LHS		RHS
S	→	aA
S	→	bA
F	→	λ
C	→	λ
I	→	λ
D	→	bE
D	→	aE
G	→	aH
A	→	aB
E	→	bF
E	→	aF
A	→	bB
G	→	bH
F	→	bG
F	→	aG
H	→	aI
H	→	bI
C	→	aD
J	→	bK
J	→	aK
C	→	bD
K	→	bS
K	→	aS
B	→	aC
B	→	bC
I	→	bJ
I	→	aJ

25. Strings of the form ab^3wb^4 where the alphabet is $\{a, b\}$ and w is any string over the same alphabet.

Solution:



LHS		RHS
S	→	aA
H	→	λ
B	→	bC
D	→	bD
C	→	bD
D	→	aD
A	→	bB
F	→	bG
D	→	bE
G	→	bH
E	→	bF

C. Describe the language of the following grammar as concisely as possible.

26. $S \to aS \mid aaS \mid \lambda$

Solution: The language is just a^*, that is, zero or more repetitions of a.

27. $S \to aA \mid \lambda, A \to bS$

Solution: The language is $(ab)^*$, that is, zero or more repetitions of ab.

28. $S \to aA \mid \lambda, A \to aA \mid B, B \to bB \mid \lambda$

Solution: The language is any number of a s followed by any number of b s with the additional constraint that there can't be b s without any a s.

29. $S \to 0S \mid 1S \mid A, A \to 0B, B \to 0C, C \to 0C \mid 1C \mid \lambda$

Solution: Binary strings containing at least one pair of consecutive 0 s.

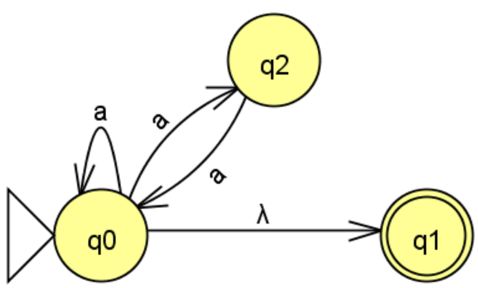
30. $S \to 01S \mid 10S \mid A, A \to 01A \mid 10A \mid \lambda$

Solution: Binary strings of even length with no run of length more than two.

D. Convert the given grammar to an equivalent NFA.

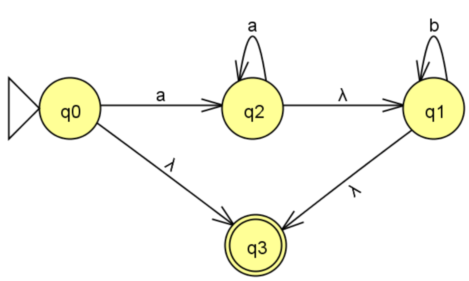
31. The grammar in Exercise 26.

Solution:



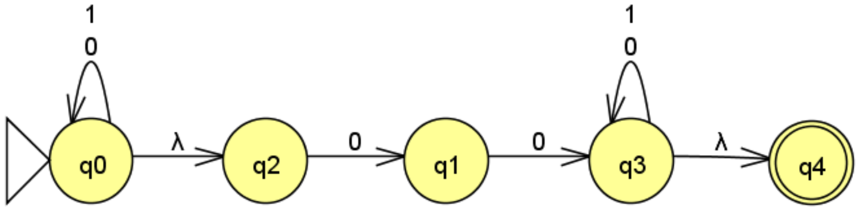
32. The grammar in Exercise 28.

Solution:



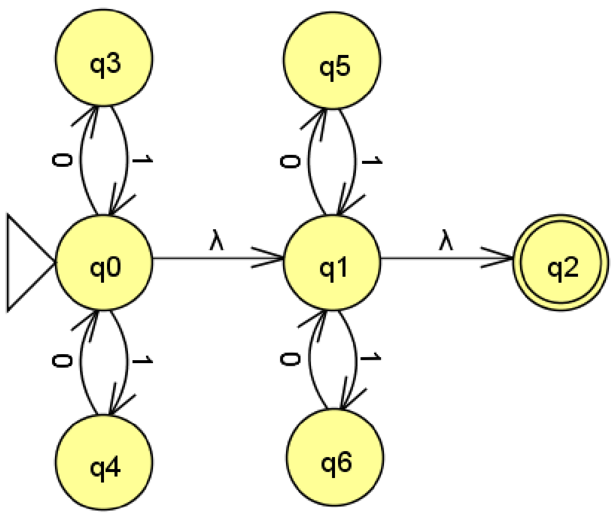
33. The grammar in Exercise 29.

Solution:



34. The grammar in Exercise 30.

Solution:



E. The following grammar is not regular. Convert it to an equivalent regular grammar. What is the language of the grammar?

35. $S \to 0S \mid 1A, A \to A1 \mid \lambda$

Solution: $S \to 0S \mid 1A, A \to 1A \mid \lambda$

The language is any number of 0 s followed by one or more 1 s, that is, 0^n1^m, $n \geq 0$, $m > 0$.

36. $S \to abS \mid baS \mid Sab \mid Sba \mid \lambda$

Solution: $S \to abS \mid baS \mid \lambda$

The language is strings of even length with no run of length greater than two.

F. The following regular grammars are incorrect. Debug and correct them.

37. Binary numbers divisible by 4: $S \to 0S \mid 1S \mid 00S \mid \lambda$

Solution: $S \to 0 \mid 1A, A \to 0A \mid 1A \mid 00$

38. Strings with any numbers of 0, 1, and 2 in that order: $S \to 0S \mid 0A$, $A \to 1A \mid 2B$, $B \to 2B \mid \lambda$

Solution: $S \to 0S \mid A \mid \lambda$, $A \to 1A \mid B \mid \lambda$, $B \to 2B \mid \lambda$

39. Strings with at least two consecutive a s before an occurrence of two consecutive b s:
$S \to bS \mid aaA$, $A \to aA \mid bA \mid bb \mid \lambda$

Solution: The grammar is as shown in the figure:

LHS		RHS
S	$\longrightarrow$	aS
S	$\longrightarrow$	aB
S	$\longrightarrow$	bA
E	$\longrightarrow$	λ
D	$\longrightarrow$	bE
B	$\longrightarrow$	aC
C	$\longrightarrow$	bC
E	$\longrightarrow$	bE
C	$\longrightarrow$	aC
E	$\longrightarrow$	aE
A	$\longrightarrow$	aS
C	$\longrightarrow$	bD

40. Binary strings of even length: $S \to 0S \mid 1S \mid 0A \mid 1A$, $A \to 0S \mid 1S \mid \lambda$

Solution: $S \to 0A \mid 1A \mid \lambda$, $A \to 0S \mid 1S$