

## CLOSURE PROPERTIES OF CFL

Context Free Languages (CFLs) are accepted by pushdown automata. Context free languages can be generated by context free grammars, which have productions (substitution rules) of the form :  $A \rightarrow \rho$  (where  $A \in N$  and  $\rho \in (T \cup N)^*$  and  $N$  is a non-terminal and  $T$  is a terminal)

### Properties of Context Free Languages

**Union :** If  $L_1$  and  $L_2$  are two context free languages, their union  $L_1 \cup L_2$  will also be context free.

For example,

$L_1 = \{ a^n b^n c^m \mid m \geq 0 \text{ and } n \geq 0 \}$  and  $L_2 = \{ a^n b^m c^m \mid n \geq 0 \text{ and } m \geq 0 \}$

$L_3 = L_1 \cup L_2 = \{ a^n b^n c^m \cup a^n b^m c^m \mid n \geq 0, m \geq 0 \}$  is also context free.

$L_1$  says number of a's should be equal to number of b's and  $L_2$  says number of b's should be equal to number of c's. Their union says either of two conditions to be true. So it is also context free language.

So CFL are closed under Union.

**Concatenation :** If  $L_1$  and  $L_2$  are two context free languages, their concatenation  $L_1.L_2$  will also be context free. For example,

$L_1 = \{ a^n b^n \mid n \geq 0 \}$  and  $L_2 = \{ c^m d^m \mid m \geq 0 \}$

$L_3 = L_1.L_2 = \{ a^n b^n c^m d^m \mid m \geq 0 \text{ and } n \geq 0 \}$  is also context free.

$L_1$  says number of a's should be equal to number of b's and  $L_2$  says number of c's should be equal to number of d's. Their concatenation says first number of a's should be equal to number of b's, then number of c's should be equal to number of d's. So, we can create a PDA which will first push for a's, pop for b's, push for c's then pop for d's. So it can be accepted by pushdown automata, hence context free.

So CFL are closed under Concatenation.

**Kleene Closure :** If  $L_1$  is context free, its Kleene closure  $L_1^*$  will also be context free. For example,

$L_1 = \{ a^n b^n \mid n \geq 0 \}$

$L_1^* = \{ a^n b^n \mid n \geq 0 \}^*$  is also context free.

So CFL are closed under Kleen Closure.

**Intersection and complementation :** If  $L_1$  and  $L_2$  are two context free languages, their intersection  $L_1 \cap L_2$  need not be context free. For example,

$L_1 = \{ a^n b^n c^m \mid n \geq 0 \text{ and } m \geq 0 \}$  and  $L_2 = \{ a^m b^n c^n \mid n \geq 0 \text{ and } m \geq 0 \}$

$L_3 = L_1 \cap L_2 = \{ a^n b^n c^n \mid n \geq 0 \}$  need not be context free.

$L_1$  says number of a's should be equal to number of b's and  $L_2$  says number of b's should be equal to number of c's. Their intersection says both conditions need to be true, but pushdown automata can compare only two. So it cannot be accepted by pushdown automata, hence not context free.

Similarly, complementation of context free language  $L_1$  which is  $\Sigma^* - L_1$ , need not be context

[www.binils.com](http://www.binils.com)

## EXAMPLES OF TURING MACHINE

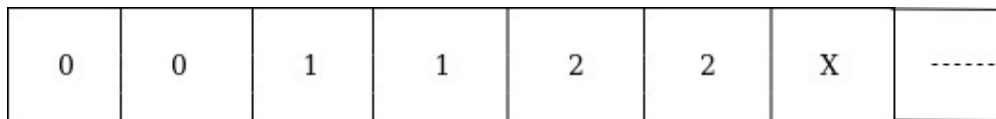
Example 1:

Construct a TM for the language  $L = \{0^n 1^n 2^n\}$  where  $n \geq 1$

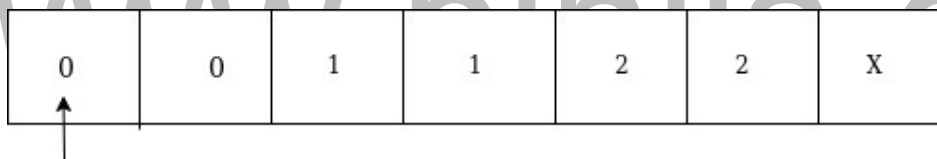
### Solution:

$L = \{0^n 1^n 2^n \mid n \geq 1\}$  represents language where we use only 3 character, i.e., 0, 1 and 2. In this, some number of 0's followed by an equal number of 1's and then followed by an equal number of 2's. Any type of string which falls in this category

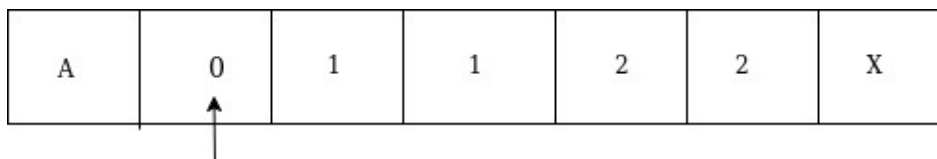
The simulation for 001122 can be shown



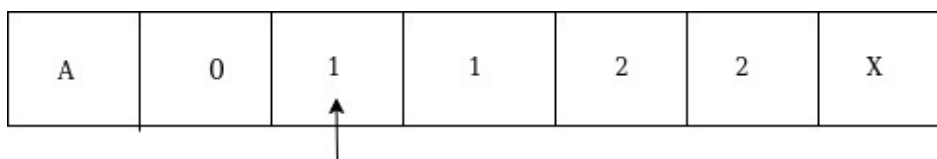
Now, we will see how this Turing machine will work for 001122. Initially, state is  $q_0$  and head points to 0 as:



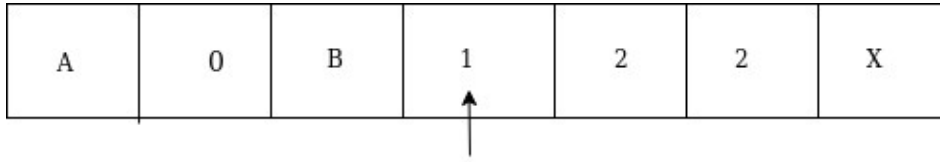
The move will be  $\delta(q_0, 0) = \delta(q_1, A, R)$  which means it will go to state  $q_1$ , replaced 0 by A and head will move to the right as:



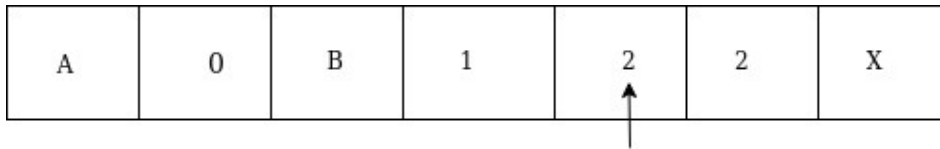
The move will be  $\delta(q_1, 0) = \delta(q_1, 0, R)$  which means it will not change any symbol, remain in the same state and move to the right as:



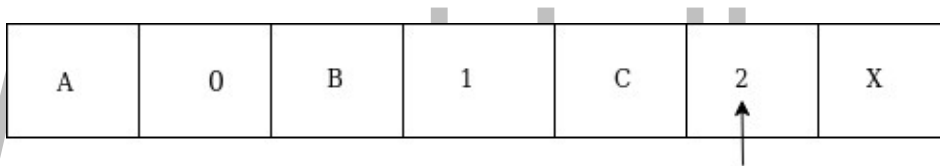
The move will be  $\delta(q_1, 1) = \delta(q_2, B, R)$  which means it will go to state  $q_2$ , replaced 1 by B and head will move to right as:



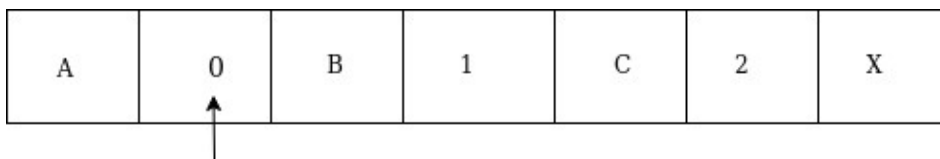
The move will be  $\delta(q_2, 1) = \delta(q_2, 1, R)$  which means it will not change any symbol, remain in the same state and move to right as:



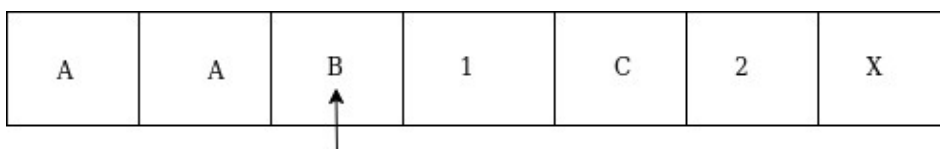
The move will be  $\delta(q_2, 2) = \delta(q_3, C, R)$  which means it will go to state  $q_3$ , replaced 2 by C and head will move to right as:



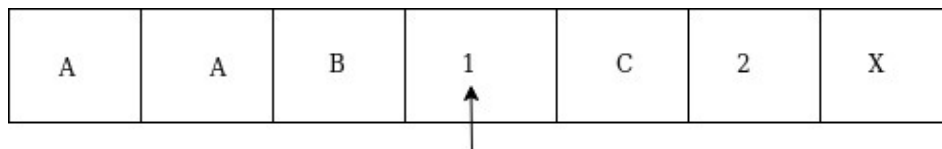
Now move  $\delta(q_3, 2) = \delta(q_3, 2, L)$  and  $\delta(q_3, C) = \delta(q_3, C, L)$  and  $\delta(q_3, 1) = \delta(q_3, 1, L)$  and  $\delta(q_3, B) = \delta(q_3, B, L)$  and  $\delta(q_3, 0) = \delta(q_3, 0, L)$ , and then move  $\delta(q_3, A) = \delta(q_0, A, R)$ , it means will go to state  $q_0$ , replaced A by A and head will m . . . . .



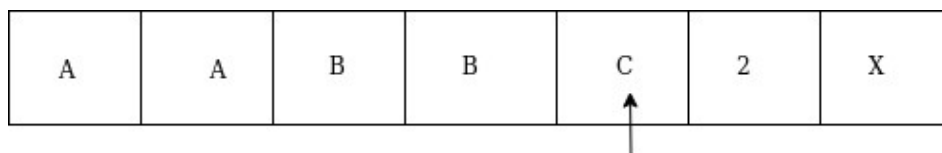
The move will be  $\delta(q_0, 0) = \delta(q_1, A, R)$  which means it will go to state  $q_1$ , replaced 0 by A, and head will move to right as:



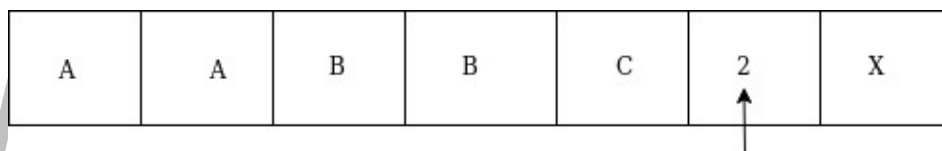
The move will be  $\delta(q_1, B) = \delta(q_1, B, R)$  which means it will not change any symbol, remain in the same state and move to right as:



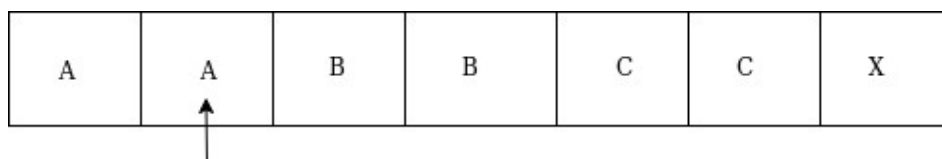
The move will be  $\delta(q_1, 1) = \delta(q_2, B, R)$  which means it will go to state  $q_2$ , replaced 1 by B and head will move to right as:



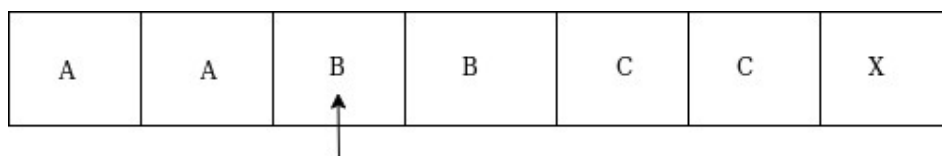
The move will be  $\delta(q_2, C) = \delta(q_2, C, R)$  which means it will not change any symbol, remain in the same state and move to right as:



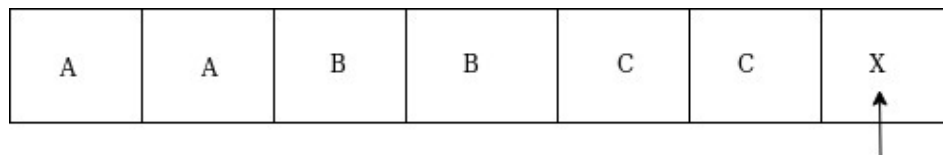
The move will be  $\delta(q_2, 2) = \delta(q_3, C, L)$  which means it will go to state  $q_3$ , replaced 2 by C and head will move to left until we reached A as:



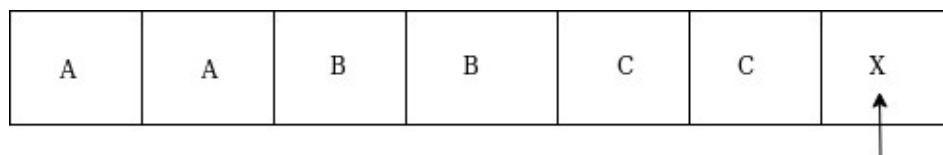
immediately before B is A that means all the 0's are marked by A. So we will move right to ensure that no 1 or 2 is present. The move will be  $\delta(q_3, B) = \delta(q_4, B, R)$  which means it will go to state  $q_4$ , will not change any symbol, and move to right as:



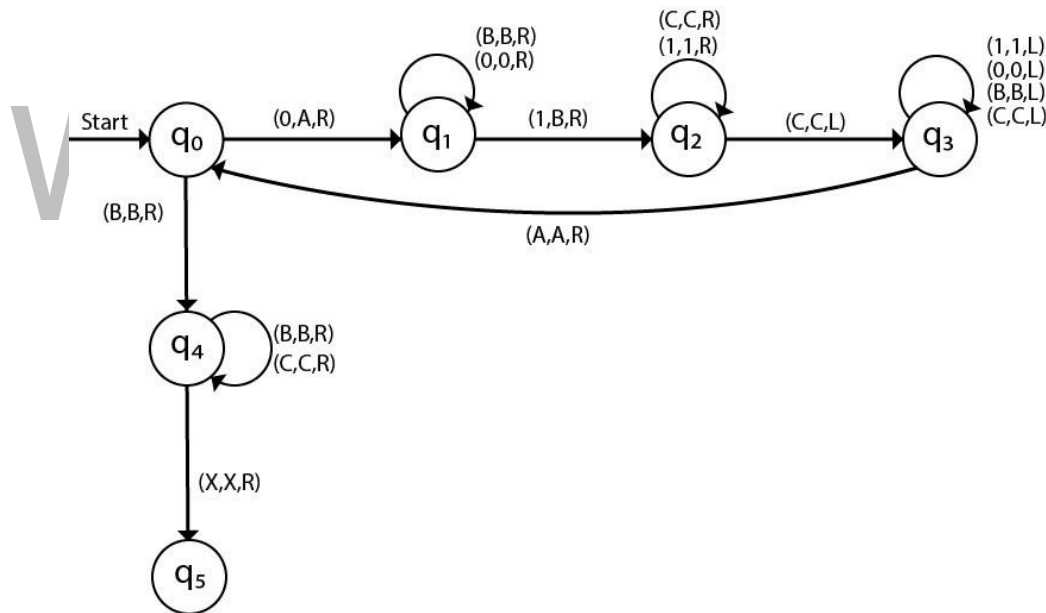
The move will be  $(q_4, B) = \delta(q_4, B, R)$  and  $(q_4, C) = \delta(q_4, C, R)$  which means it will not change any symbol, remain in the same state and move to right as:



The move  $\delta(q_4, X) = (q_5, X, R)$  which means it will go to state  $q_5$  which is the HALT state and HALT state is always an accept state for any TM.



The same TM can be represented by Transition Diagram:



Example 2:

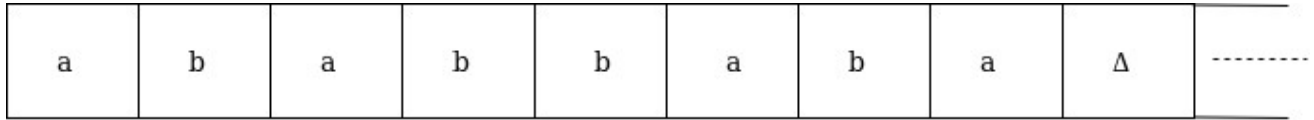
Construct a TM machine for checking the palindrome of the string of even length.

**Solution:**

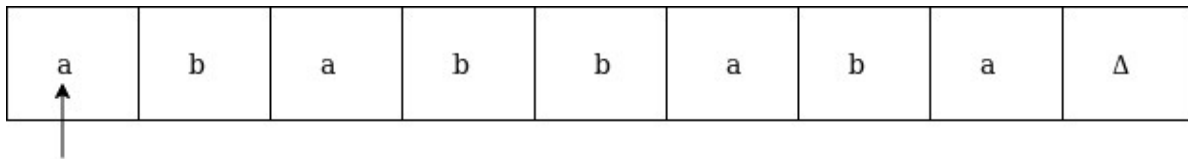
Firstly we read the first symbol from the left and then we compare it with the first symbol from right to check whether it is the same.

Again we compare the second symbol from left with the second symbol from right. We repeat this process for all the symbols. If we found any symbol not matching, we cannot lead the machine to HALT state.

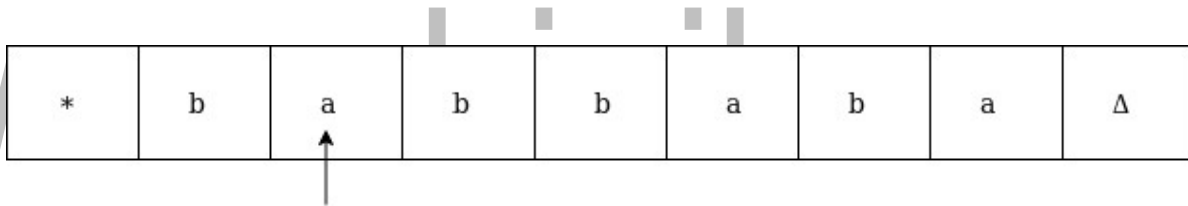
Suppose the string is ababbaba $\Delta$ . The simulation for ababbaba $\Delta$  can be shown as follows:



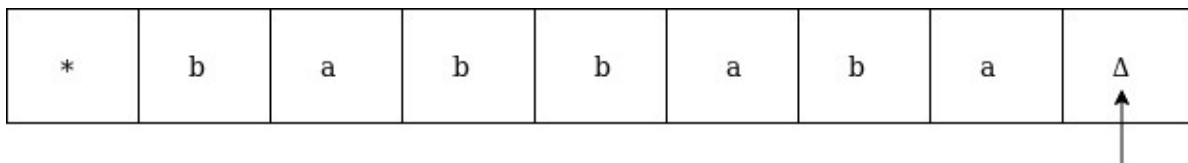
Now, we will see how this Turing machine will work for ababbaba $\Delta$ . Initially, state is  $q_0$  and head points to a as:



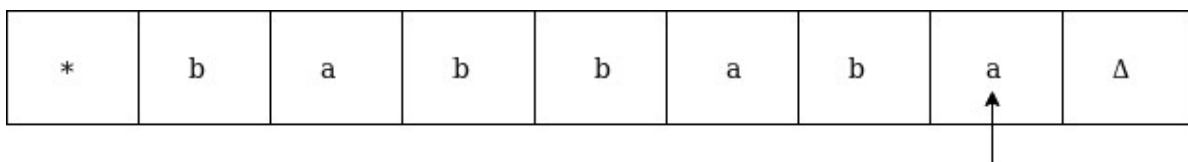
We will mark it by \* and move to right end in search of a as:



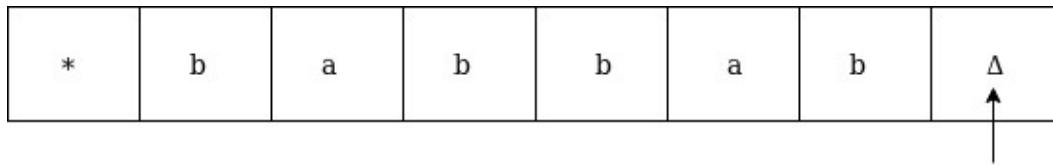
We will move right up to  $\Delta$  as:



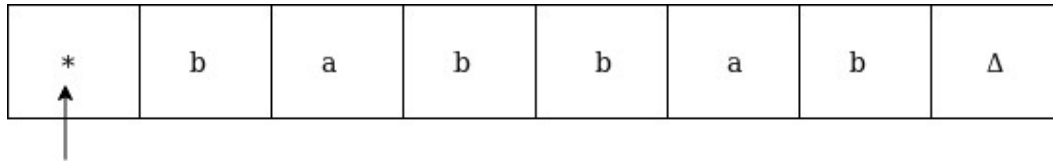
We will move left and check if it is a:



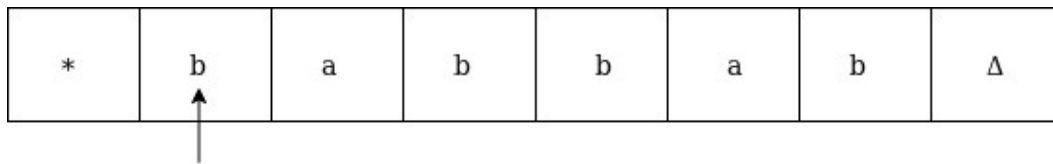
It is 'a' so replace it by  $\Delta$  and move left as:



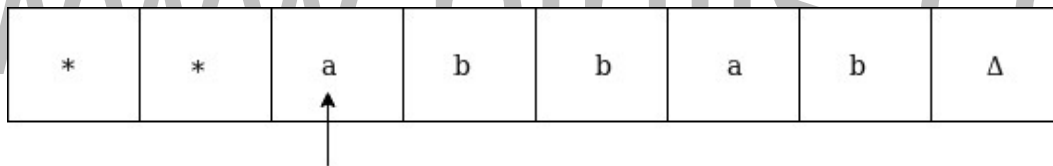
Now move to left up to \* as:



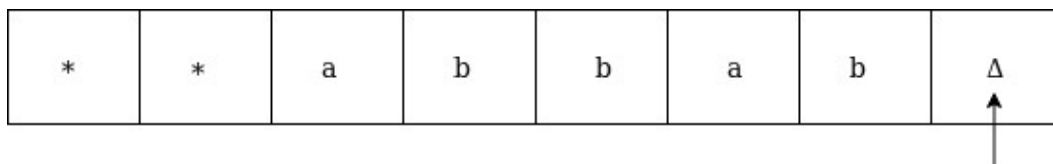
Move right and read it



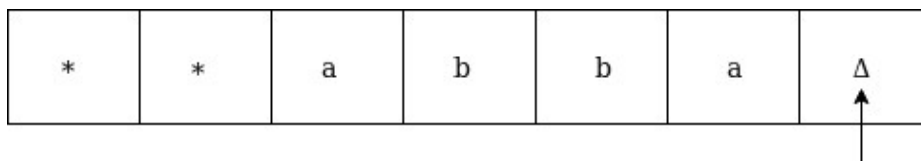
Now convert b by \* and move right as:



Move right up to  $\Delta$  in search of b as:

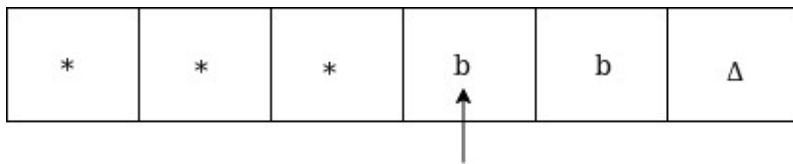
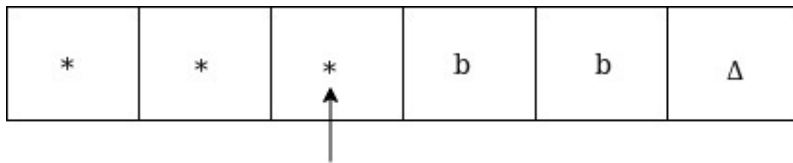
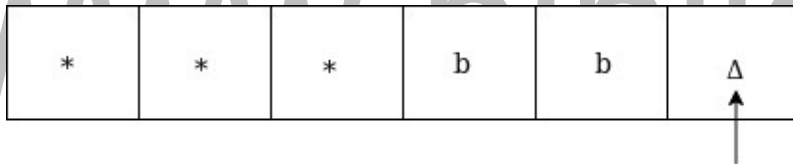
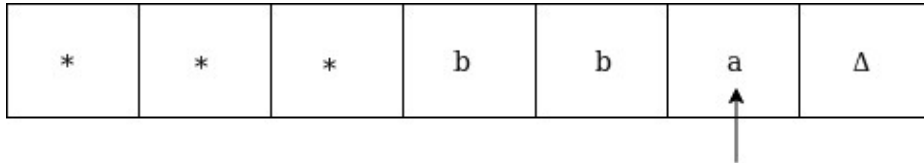
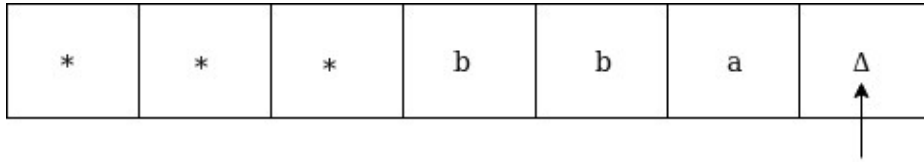
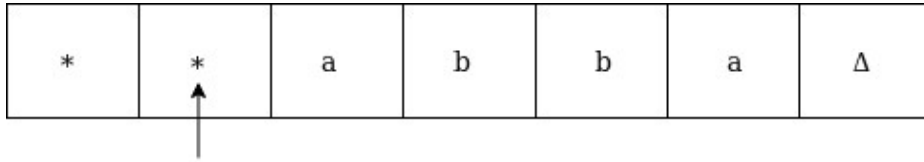


Move left, if the symbol is b then convert it into  $\Delta$  as:

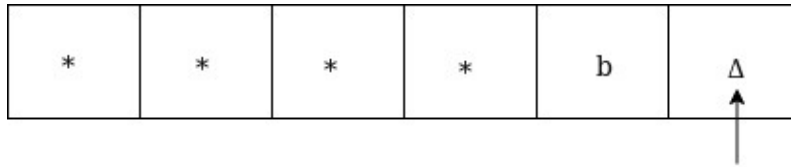


Now move left until \* as:

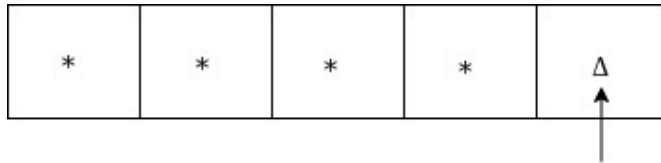




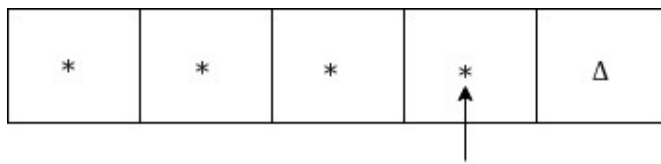
www.binils.com



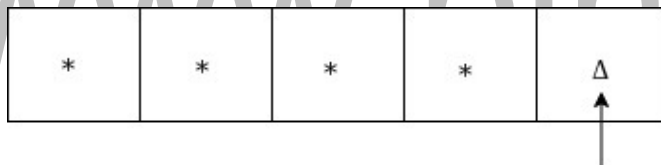
Move left, if the left symbol is b, replace it by Δ as:



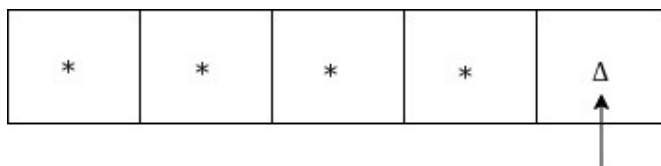
Move left till \*



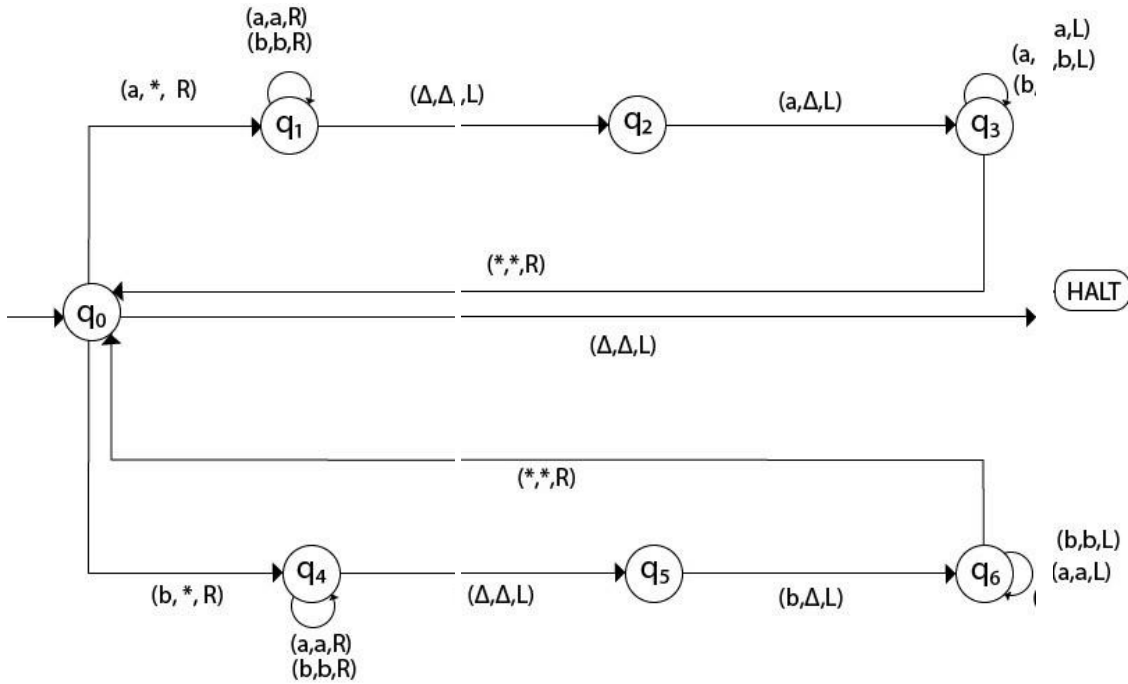
Move right and check whether it is Δ



Go to HALT state



The same TM can be represented by Transition Diagram:



Example 3:

Construct a TM machine for checking the palindrome of the string of odd length.

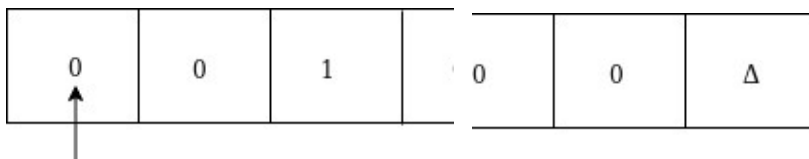
**Solution:**

Firstly we read the first symbol from left and then we compare it with the first symbol from right to check whether it is the same.

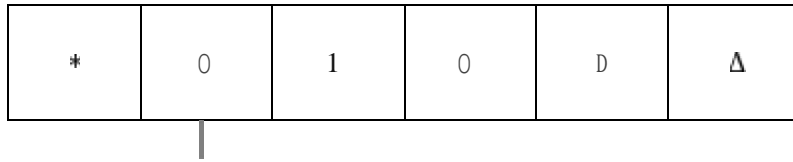
Again we compare the second symbol from left with the second symbol from right. We repeat this process for all the symbols. If we found any symbol not matching, we lead the machine to HALT state.

Suppose the string is 00100Δ. The simulation for 00100Δ can be shown as follows:

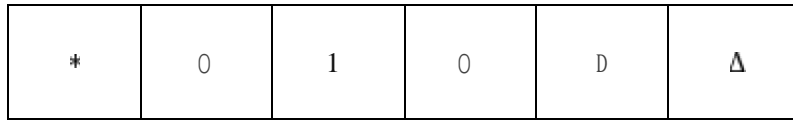
Now, we will see how this Turing machine will work for 00100Δ. Initially, state is  $q_0$  and head points to 0 as:



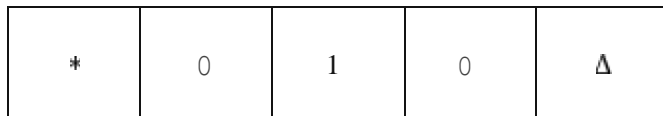
Now replace 0 by \* and move right as:



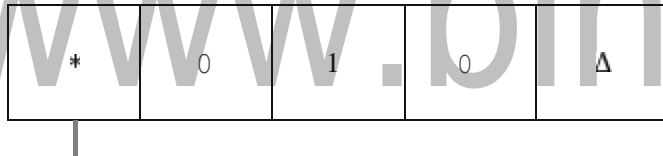
Move right up to  $\Delta$  as:



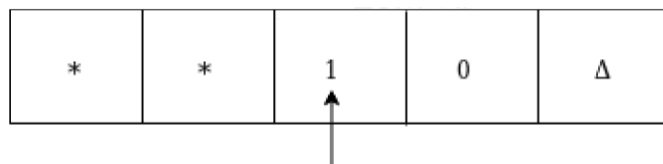
Move left and replace 0 by  $\Delta$  and move left:



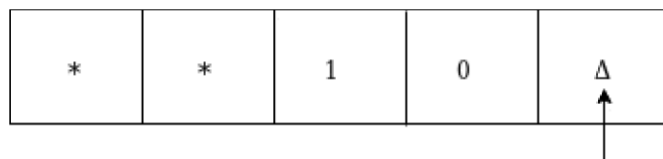
Now move left up to \* as:



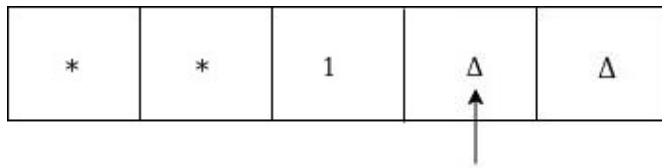
Move right, convert 0 by \* and then move right as:



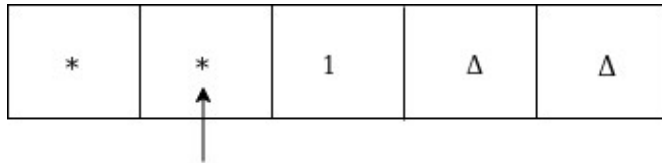
Moved right up to  $\Delta$



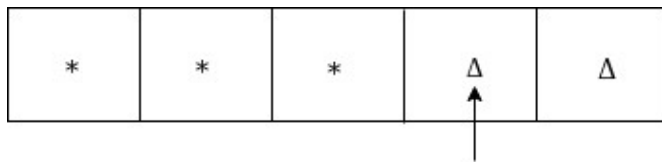
Move left and replace 0 by  $\Delta$  as:



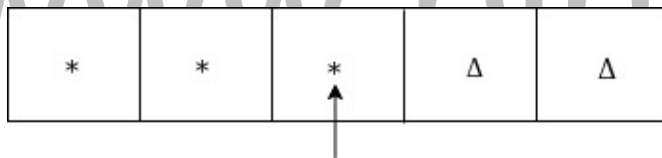
Move left till \* as:



Move right and convert 1 to \* as:

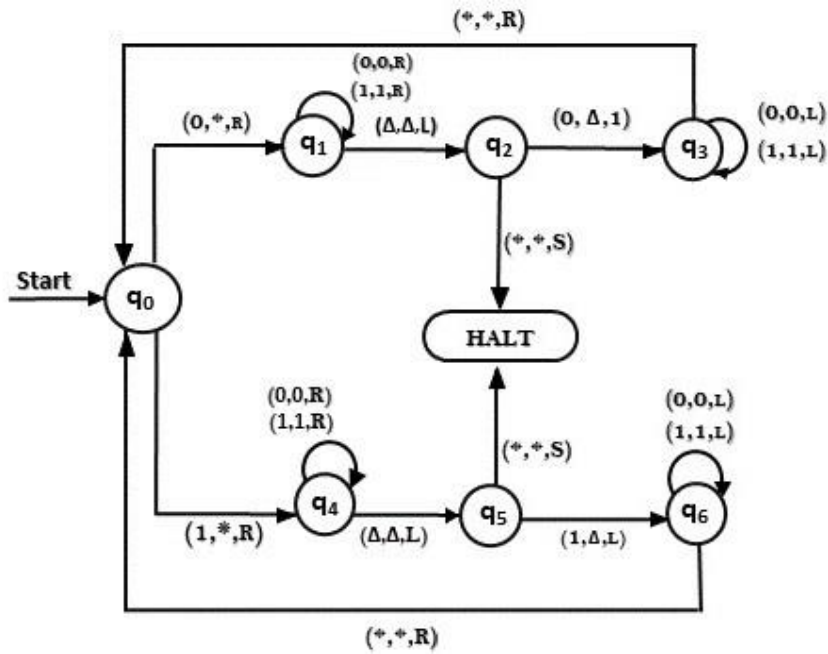


Move left



Since it is \*, goto HALT state.

The same TM can be represented by Transition Diagram:



Example 4:

Construct TM for the addition function for the unary number system.

**Solution:**

The unary number is made up of only one character, i.e. The number 5 can be written in unary number system as 11111. In this TM, we are going to perform the addition of two unary numbers.

**For example**

$$2 + 3$$

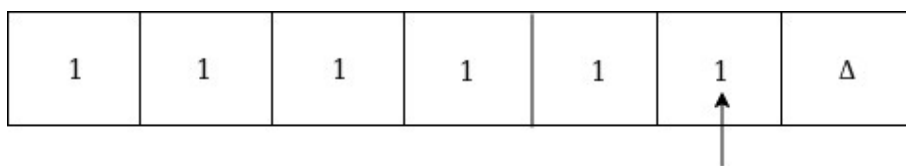
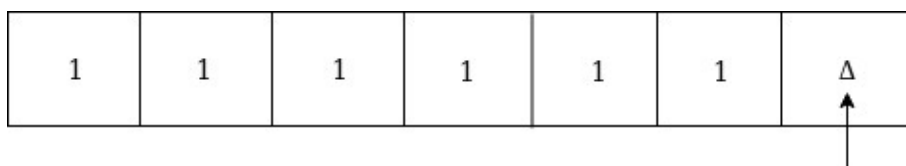
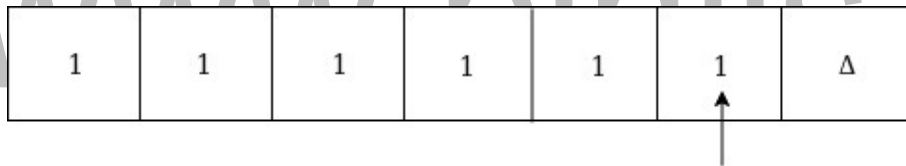
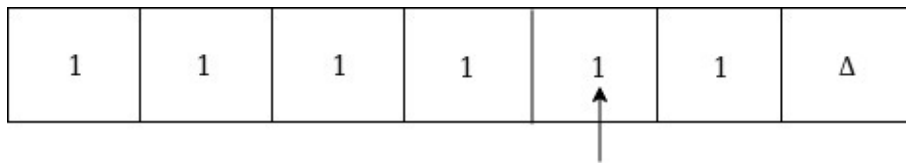
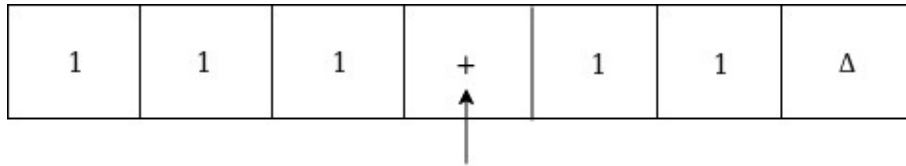
$$\text{i.e. } 11 + 111 = 11111$$

If you observe this process of addition, you will find the resemblance with string concatenation function.

In this case, we simply replace + by 1 and move ahead right for searching end of the string we will convert last 1 to Δ.

**Input:** 3+2

The simulation for 111+11Δ can be shown as below:

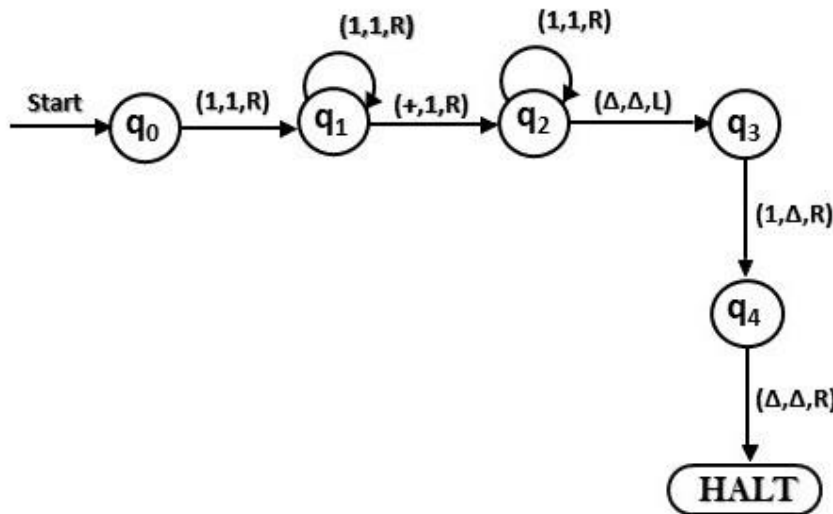


1	1	1	1	1	$\Delta$	$\Delta$
---	---	---	---	---	----------	----------

Thus the tape now consists of the addition of two unary numbers.

The TM will look like as follows:

Here, we are implementing the function of  $f(a + b) = c$ . We assume a and b both are non zero elements.



Example 5:

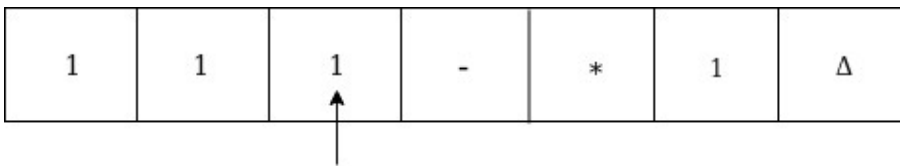
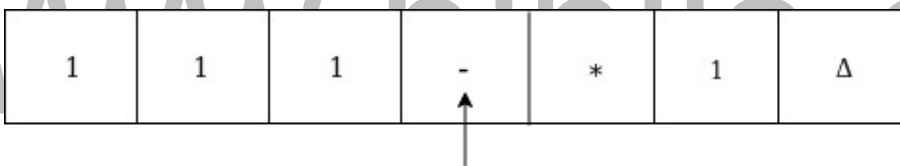
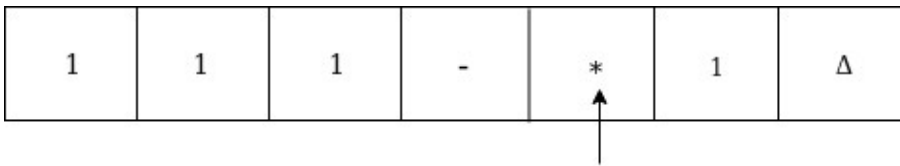
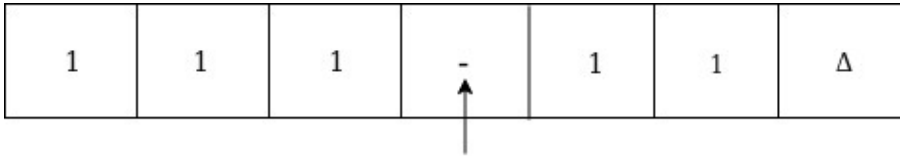
Construct a TM for subtraction of two unary numbers  $f(a-b) = c$  where a is always greater than b.

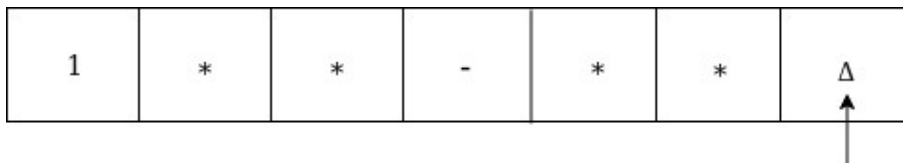
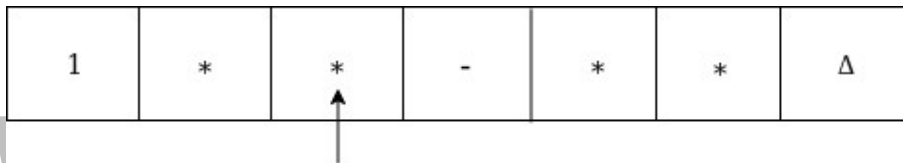
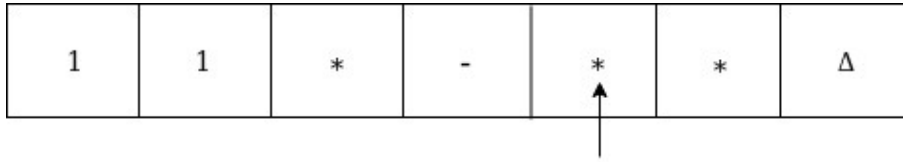
**Solution:** Here we have certain assumptions as the first number is greater than the second one. Let us assume that  $a = 3, b = 2$ , so the input tape will be:

1	1	1	-	1	1	$\Delta$
---	---	---	---	---	---	----------

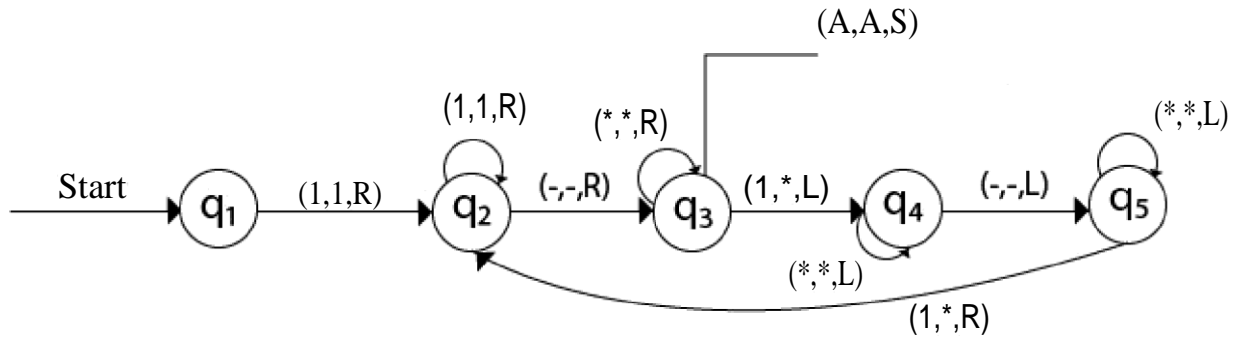
We will move right to - symbol as perform reduction of a number of 1's from the first number. Let us look at the simulation for understanding the logic:







www.binils.com



[www.binils.com](http://www.binils.com)

## NORMAL FORMS FOR CFG

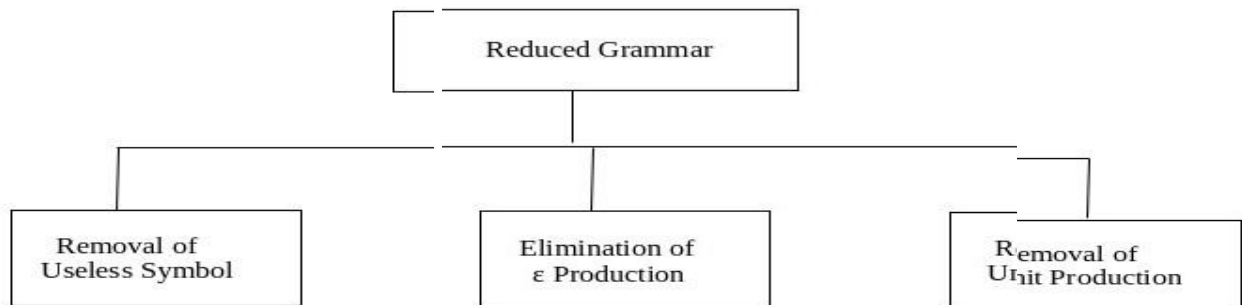
### Simplification of CFG

As we have seen, various languages can efficiently be represented by a context-free grammar. All the grammar are not always optimized that means the grammar may consist of some extra symbols(non-terminal). Having extra symbols, unnecessary increase the length of grammar. Simplification of grammar means reduction of grammar by removing useless symbols. The properties of reduced grammar are given below:

1. Each variable (i.e. non-terminal) and each terminal of  $G$  appears in the derivation of some word in  $L$ .
2. There should not be any production as  $X \rightarrow Y$  where  $X$  and  $Y$  are non-terminal.
3. If  $\epsilon$  is not in the language  $L$  then there need not to be the production  $X \rightarrow \epsilon$ .

Let us study the reduction process in detail.

www.binils.com



#### Removal of Useless Symbols

A symbol can be useless if it does not appear on the right-hand side of the production rule and does not take part in the derivation of any string. That symbol is known as a useless symbol. Similarly, a variable can be useless if it does not take part in the derivation of any string. That variable is known as a useless variable.

For Example:

1.  $T \rightarrow aaB \mid abA \mid aaT$
2.  $A \rightarrow aA$

3.  $B \rightarrow ab \mid b$
4.  $C \rightarrow ad$

In the above example, the variable 'C' will never occur in the derivation of any string, so the production  $C \rightarrow ad$  is useless. So we will eliminate it, and the other productions are written in such a way that variable C can never reach from the starting variable 'T'.

Production  $A \rightarrow aA$  is also useless because there is no way to terminate it. If it never terminates, then it can never produce a string. Hence this production can never take part in any derivation.

To remove this useless production  $A \rightarrow aA$ , we will first find all the variables which will never lead to a terminal string such as variable 'A'. Then we will remove all the productions in which the variable 'B' occurs.

#### Elimination of $\epsilon$ Production

The productions of type  $S \rightarrow \epsilon$  are called  $\epsilon$  productions. These type of productions can only be removed from those grammars that do not generate  $\epsilon$ .

**Step 1:** First find out all nullable non-terminal variable which derives  $\epsilon$ .

**Step 2:** For each production  $A \rightarrow a$ , construct all production  $A \rightarrow x$ , where x is obtained from a by removing one or more non-terminal from step 1.

**Step 3:** Now combine the result of step 2 with the original production and remove  $\epsilon$  productions.

Example:

Remove the production from the following CFG by preserving the meaning of it.

1.  $S \rightarrow XYX$
2.  $X \rightarrow 0X \mid \epsilon$
3.  $Y \rightarrow 1Y \mid \epsilon$

#### **Solution:**

Now, while removing  $\epsilon$  production, we are deleting the rule  $X \rightarrow \epsilon$  and  $Y \rightarrow \epsilon$ . To preserve the meaning of CFG we are actually placing  $\epsilon$  at the right-hand side whenever X and Y have appeared.

Let us take

1.  $S \rightarrow XYX$

If the first X at right-hand side is  $\epsilon$ . Then

1.  $S \rightarrow YX$

Similarly if the last X in R.H.S. =  $\epsilon$ . Then

1.  $S \rightarrow XY$

1.

1. If  $Y = \epsilon$  then

$$S \rightarrow XX$$

1. If Y and X are  $\epsilon$  then,

$$S \rightarrow X$$

1. If both X are replaced by  $\epsilon$

$$S \rightarrow Y$$

1. Now,

$$S \rightarrow XY \mid YX \mid XX \mid X \mid Y$$

1. Now let us consider

2.  $X \rightarrow 0X \mid 0$

$$X \rightarrow 0X$$

Similarly  $Y \rightarrow 1Y \mid 1$

If we place  $\epsilon$  at right-hand side for X then,

Collectively we can rewrite the CFG with removed  $\epsilon$  production as

$$X \rightarrow 0$$

1.  $S \rightarrow XY \mid YX \mid XX \mid X \mid Y$

2.  $X \rightarrow 0X \mid 0$

3.  $Y \rightarrow 1Y \mid 1$

## Removing Unit Productions

The unit productions are the productions in which one non-terminal gives another non-terminal. Use the following steps to remove unit production:

**Step 1:** To remove  $X \rightarrow Y$ , add production  $X \rightarrow a$  to the grammar rule whenever  $Y \rightarrow a$  occurs in the grammar.

**Step 2:** Now delete  $X \rightarrow Y$  from the grammar.

- 1.
- 2.
- 3.
- 4.

[www.binils.com](http://www.binils.com)

- 1.

**Step 3:** Repeat step 1 and step 2 until all unit productions are

For example:

1.  $S \rightarrow 0A \mid 1B \mid$   
 $CA \rightarrow 0S \mid 00$   
 $B \rightarrow 1 \mid A$
1.  $S \rightarrow 0A \mid 1B \mid 01$   
 $C \rightarrow 01$
2.  $A \rightarrow 0S \mid 00$
3.  $B \rightarrow 1 \mid 0S \mid 00$
4.  $C \rightarrow 01$

**Solution:**

Chomsky's Normal Form (CNF) But while removing  $S \rightarrow C$  we have to consider what C gives. So, we can add a rule to S.

CNF stands for Chomsky normal form. A CFG(context free grammar) is in CNF(Chomsky normal form) if all production rules satisfy one of the following conditions:

$$S \rightarrow 0A \mid 1B \mid$$

Similarly,  $B \rightarrow A$  is also a unit production so we can modify

- Start symbol generating  $\epsilon$ . For example,  $A \rightarrow \epsilon$ .
- A non-terminal generating two non-terminals. For example,  $S \rightarrow AB$ .
- A non-terminal generating a terminal. For example,  $S \rightarrow a$ .

For example:

1.  $G1 = \{S \rightarrow AB, S \rightarrow c, A \rightarrow a, B \rightarrow b\}$
2.  $G2 = \{S \rightarrow aA, A \rightarrow a, B \rightarrow c\}$

The production rules of Grammar G1 satisfy the rules specified for CNF, so the grammar G1 is in CNF. However, the production rule of Grammar G2 does not satisfy the rules specified for CNF as  $S \rightarrow aZ$  contains terminal followed by non-terminal. So the grammar G2 is not in CNF.

Steps for converting CFG into CNF

**Step 1:** Eliminate start symbol from the RHS. If the start symbol T is at the right-hand side of any production, create a new production as:

1.  $S1 \rightarrow S$

Where S1 is the new start symbol.

**Step 2:** In the grammar, remove the null, unit and useless productions. You can refer to the Simplification of CFG.

**Step 3:** Eliminate terminals from the RHS of the production if they exist with other non-terminals or terminals. For example, production  $S \rightarrow aA$  can be decomposed as:

1.  $S \rightarrow RA$
2.  $R \rightarrow a$

**Step 4:** Eliminate RHS with more than two non-terminals. For example,  $S \rightarrow ASB$  can be decomposed as:

1.  $S \rightarrow RS$
2.  $R \rightarrow AS$

Example:

Convert the given CFG to CNF. Consider the given grammar G1:



1.  $S \rightarrow a \mid aA \mid B$
2.  $A \rightarrow aBB \mid \epsilon$
3.  $B \rightarrow Aa \mid b$

**Solution:**

**Step 1:** We will create a new production  $S1 \rightarrow S$ , as the start symbol  $S$  appears on the RHS. The grammar will be:

1.  $S1 \rightarrow S$
- $$S \rightarrow a \mid aA \mid B$$
- $$A \rightarrow aBB \mid \epsilon \quad B \rightarrow Aa \mid b$$

**Step 2:** As grammar  $G1$  contains  $A \rightarrow \epsilon$  null production, its removal from the grammar yields:

- $$S1 \rightarrow S$$
- $$S \rightarrow a \mid aA \mid BA \rightarrow aBB$$
- $$B \rightarrow Aa \mid b \mid a$$

Now, as grammar  $G1$  contains Unit production  $S \rightarrow B$ , its removal yield:

- $$S1 \rightarrow S$$
- $$S \rightarrow a \mid aA \mid Aa \mid bA \rightarrow aBB$$
- $$B \rightarrow Aa \mid b \mid a$$

Also remove the unit production  $S1 \rightarrow S$ , its removal from the grammar yields:

1.  $S0 \rightarrow a \mid aA \mid Aa \mid b$
2.  $S \rightarrow a \mid aA \mid Aa \mid b$
3.  $A \rightarrow aBB$
4.  $B \rightarrow Aa \mid b \mid a$

**Step 3:** In the production rule  $S0 \rightarrow aA \mid Aa$ ,  $S \rightarrow aA \mid Aa$ ,  $A \rightarrow aBB$  and  $B \rightarrow Aa$ , terminal  $a$  exists on RHS with non-terminals. So we will replace terminal  $a$  with  $X$ :

1.  $S0 \rightarrow a \mid XA \mid AX \mid b$

2.  $S \rightarrow a \mid XA \mid AX \mid b$
3.  $A \rightarrow XBB$
4.  $B \rightarrow AX \mid b \mid a$
5.  $X \rightarrow a$

**Step 4:** In the production rule  $A \rightarrow XBB$ , RHS has more than two symbols, removing it from grammar yield:

1.  $S_0 \rightarrow a \mid XA \mid AX \mid b$
2.  $S \rightarrow a \mid XA \mid AX \mid b$
3.  $A \rightarrow RB$
4.  $B \rightarrow AX \mid b \mid a$
5.  $X \rightarrow a$
6.  $R \rightarrow XB$

Hence, for the given grammar, this is the required CNF.

#### Greibach Normal Form (GNF)

GNF stands for Greibach normal form. A CFG(context free grammar) is in GNF(Greibach normal form) if all the production rules satisfy one of the following conditions:

- A start symbol generating  $\epsilon$ . For example,  $S \rightarrow \epsilon$ .
- A non-terminal generating a terminal. For example,  $A \rightarrow a$ .
- A non-terminal generating a terminal which is followed by any number of non-terminals. For example,  $S \rightarrow aASB$ .

#### For example:

1.  $G_1 = \{S \rightarrow aAB \mid aB, A \rightarrow aA \mid a, B \rightarrow bB \mid b\}$
2.  $G_2 = \{S \rightarrow aAB \mid aB, A \rightarrow aA \mid \epsilon, B \rightarrow bB \mid \epsilon\}$

The production rules of Grammar  $G_1$  satisfy the rules specified for GNF, so the grammar  $G_1$  is in GNF. However, the production rule of Grammar  $G_2$  does not satisfy the rules specified for GNF as  $A \rightarrow \epsilon$  and  $B \rightarrow \epsilon$  contains  $\epsilon$ (only start symbol can generate  $\epsilon$ ). So the grammar  $G_2$  is not in GNF.

Steps for converting CFG into GNF

**Step 1:** Convert the grammar into CNF.

If the given grammar is not in CNF, convert it into CNF. You can refer the following topic to convert the CFG into CNF: Chomsky normal form

**Step 2:** If the grammar exists left recursion, eliminate it.

If the context free grammar contains left recursion, eliminate it. You can refer the following topic to eliminate left recursion: Left Recursion

**Step 3:** In the grammar, convert the given production rule into GNF form.

1. If any production rule in the grammar is not in GNF form, convert it.

Example:

$$S \rightarrow XB \mid AA$$

$$A \rightarrow a \mid SA \quad B \rightarrow b$$

$$X \rightarrow a$$

**Solution:**

As the given grammar G is already in CNF and there is no left recursion, so we can skip step 1 and step 2 and directly go to step 3.

The production rule  $A \rightarrow SA$  is not in GNF, so we substitute  $S \rightarrow XB \mid AA$  in the production rule  $A$

$\rightarrow SA$  as:

$$S \rightarrow XB \mid AA$$

$$A \rightarrow a \mid XBA \mid AAA \quad B \rightarrow b$$

$$X \rightarrow a$$

The production rule  $S \rightarrow XB$  and  $B \rightarrow XBA$  is not in GNF, so we substitute  $X \rightarrow a$  in the production rule  $S \rightarrow XB$  and  $B \rightarrow XBA$  as:

1.  $S \rightarrow aB \mid AA$
2.  $A \rightarrow a \mid aBA \mid AAA$
3.  $B \rightarrow b$
4.  $X \rightarrow a$

Now we will remove left recursion ( $A \rightarrow AAA$ ), we get:

1.  $S \rightarrow aB \mid AA$
2.  $A \rightarrow aC \mid aBAC$
3.  $C \rightarrow AAC \mid \epsilon$
4.  $B \rightarrow b$
5.  $X \rightarrow a$

Now we will remove null production  $C \rightarrow \epsilon$ , we get:

1.  $S \rightarrow aB \mid AA$  2.  $A \rightarrow aC \mid aBAC \mid a \mid aBA$   $C \rightarrow AAC \mid AA$

$B \rightarrow b$

$X \rightarrow a$

rule  $S \rightarrow AA$  as:

$S \rightarrow aB \mid aCA \mid aBACA \mid aA \mid aBAA$   $A \rightarrow aC \mid aBAC \mid a \mid aBA$

$C \rightarrow AAC$

$C \rightarrow aCA \mid aBACA \mid aA \mid aBAA$   $B \rightarrow b$

$X \rightarrow a$

The production rule  $S \rightarrow AA$  is not in GNF, so we substitute  $A \rightarrow aC \mid aBAC \mid a \mid aBA$  in production

The production rule  $C \rightarrow AAC$  is not in GNF, so we substitute  $A \rightarrow aC \mid aBAC \mid a \mid aBA$  in production

rule  $C \rightarrow AAC$  as:

$S \rightarrow aB \mid aCA \mid aBACA \mid aA \mid aBAA$

$A \rightarrow aC \mid aBAC \mid a \mid aBA$

$C \rightarrow aCAC \mid aBACAC \mid aAC \mid aBAAC$

3.  $C \rightarrow aCA \mid aBACA \mid aA \mid aBAA$

4.  $B \rightarrow b$

5.  $X \rightarrow a$

Hence, this is the GNF form for the grammar G.

## PUMPING LEMMA FOR CFL

If  $L$  is a context-free language, there is a pumping length  $p$  such that any string  $w \in L$  of length  $\geq p$  can be written as  $w = uvxyz$ , where  $vy \neq \epsilon$ ,  $|vxy| \leq p$ , and for all  $i \geq 0$ ,  $uv^ixy^iz \in L$ .

### Applications of Pumping Lemma

Pumping lemma is used to check whether a grammar is context free or not. Let us take an example and show how it is checked.

#### Problem

Find out whether the language  $L = \{x^n y^n z^n \mid n \geq 1\}$  is context free or not.

#### Solution

Let  $L$  is context free. Then,  $L$  must satisfy pumping lemma.

At first, choose a number  $n$  of the pumping lemma. Then, take  $z$  as  $0^n 1^n 2^n$ .

Break  $z$  into  $uvwxy$ , where

$|vwx| \leq n$  and  $vx \neq \epsilon$ .

Hence  $vwx$  cannot involve both 0s and 2s, since the last 0 and the first 2 are at least  $(n+1)$  positions apart. There are two cases –

Case 1 –  $vwx$  has no 2s. Then  $vx$  has only 0s and 1s. Then  $uw^i v^i x^i y^i z^i$ , which would have to be in  $L$ , has  $n$  2s, but fewer than  $n$  0s or 1s.

Case 2 –  $vwx$  has no 0s.

Here contradiction occurs.

Hence,  $L$  is not a context-free language.

### CFL Closure Property

Context-free languages are closed under –

- ▮ Union
- ▮ Concatenation
- ▮ Kleene Star operation

## Union

Let  $L_1$  and  $L_2$  be two context free languages. Then  $L_1 \cup L_2$  is also context free.

### Example

Let  $L_1 = \{ a^n b^n, n > 0 \}$ . Corresponding grammar  $G_1$  will have P:  $S_1 \rightarrow aA^b|ab$

Let  $L_2 = \{ c^m d^m, m \geq 0 \}$ . Corresponding grammar  $G_2$  will have P:  $S_2 \rightarrow cB^d| \epsilon$

Union of  $L_1$  and  $L_2$ ,  $L = L_1 \cup L_2 = \{ a^n b^n \} \cup \{ c^m d^m \}$

The corresponding grammar  $G$  will have the additional production  $S \rightarrow S_1 | S_2$

## Concatenation

If  $L_1$  and  $L_2$  are context free languages, then  $L_1 L_2$  is also context free.

### Example

Union of the languages  $L_1$  and  $L_2$ ,  $L = L_1 L_2 = \{ a^n b^n c^m d^m \}$

The corresponding grammar  $G$  will have the additional production  $S \rightarrow S_1 S_2$

## Kleene Star

If  $L$  is a context free language, then  $L^*$  is also context free.

### Example

Let  $L = \{ a^n b^n, n \geq 0 \}$ . Corresponding grammar  $G$  will have P:  $S \rightarrow aA^b| \epsilon$

Kleene Star  $L_1 = \{ a^n b^n \}^*$

The corresponding grammar  $G_1$  will have additional productions  $S_1 \rightarrow SS_1 | \epsilon$

Context-free languages are not closed under –

- ▮ Intersection – If  $L_1$  and  $L_2$  are context free languages, then  $L_1 \cap L_2$  is not necessarily context free.
- ▮ Intersection with Regular Language – If  $L_1$  is a regular language and  $L_2$  is a context free language, then  $L_1 \cap L_2$  is a context free language.
- ▮ Complement – If  $L_1$  is a context free language, then  $L_1'$  may not be context free.

## TURING MACHINE

Turing machine was invented in 1936 by **Alan Turing**. It is an accepting device which accepts Recursive Enumerable Language generated by type 0 grammar.

There are various features of the Turing machine:

1. It has an external memory which remembers arbitrary long sequence of input.
2. It has unlimited memory capability.
3. The model has a facility by which the input at left or right on the tape can be read easily.
4. The machine can produce a certain output based on its input. Sometimes it may be required that the same input has to be used to generate the output. So in this machine, the distinction between input and output has been removed. Thus a common set of alphabets can be used for the Turing machine.

Formal definition of Turing machine

A Turing machine can be defined as a collection of 7 components:

**Q**: the finite set of states

$\Sigma$ : the finite set of input symbols

**T**: the tape symbol

**q<sub>0</sub>**: the initial state

**F**: a set of final states

**B**: a blank symbol used as a end marker for input

$\delta$ : a transition or mapping function.

The mapping function shows the mapping from states of finite automata and input symbol on the tape to the next states, external symbols and the direction for moving the tape head. This is known as a triple or a program for turing machine.

$$1. (q_0, a) \rightarrow (q_1, A, R)$$

That means in  $q_0$  state, if we read symbol 'a' then it will go to state  $q_1$ , replaced a by X and move ahead right(R stands for right).

**Example:**

Construct TM for the language  $L = \{1^n\}$  where  $n \geq 1$ .

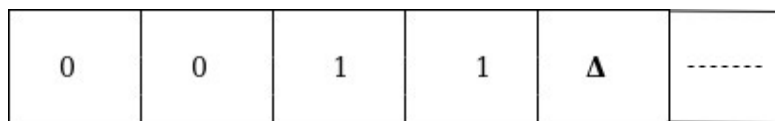
**Solution:**

We have already solved this problem by PDA. In PDA, we have a stack to remember the previous symbol. The main advantage of the Turing machine is we have a tape head which can be moved forward or backward, and the input tape can be scanned.

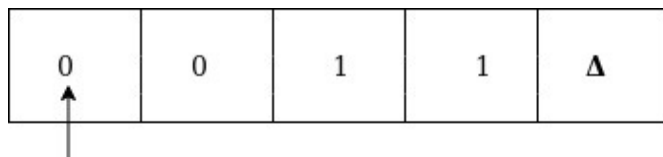
The simple logic which we will apply is read out each '0' mark it by A and then move ahead along with the input tape and find out 1 convert it to B. Now, repeat this process for all a's and b's.

Now we will see how this Turing machine

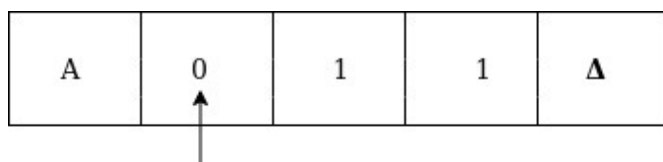
The simulation for 0011 can be shown as below:



Now, we will see how this Turing machine will work for 0011. Initially, state is  $q_0$  and head points to 0 as:

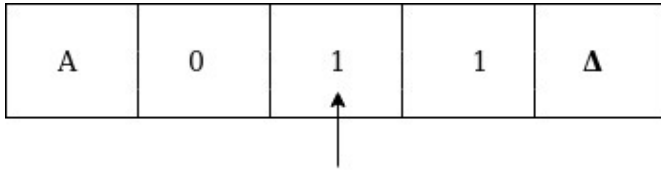


The move will be  $\delta(q_0, 0) = \delta(q_1, A, R)$  which means it will go to state  $q_1$ , replaced 0 by A and head will move to the right as:

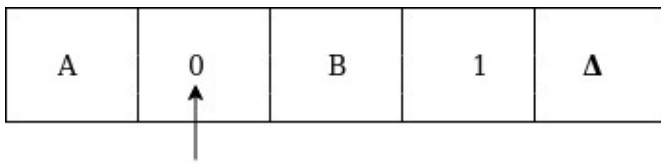




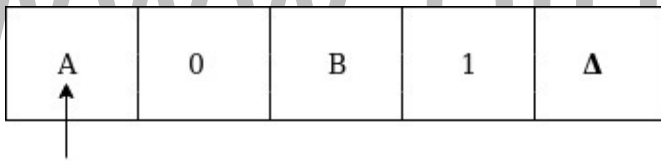
The move will be  $\delta(q_1, 0) = \delta(q_1, 0, R)$  which means it will not change any symbol, remain in the same state and move to the right as:



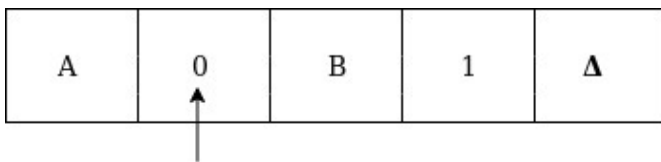
The move will be  $\delta(q_1, 1) = \delta(q_2, B, L)$  which means it will go to state  $q_2$ , replaced 1 by B and head will move to left as:



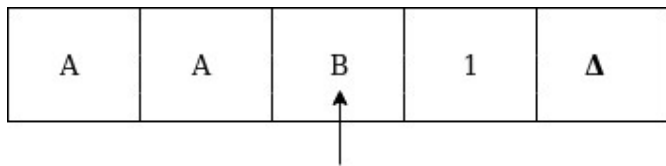
Now move will be  $\delta(q_2, 0) = \delta(q_2, 0, L)$  which means it will not change any symbol, remain in the same state and move to left as:



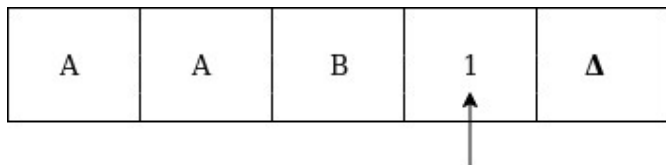
The move will be  $\delta(q_2, A) = \delta(q_0, A, R)$ , it means will go to state  $q_0$ , replaced A by A and head will move to the right as:



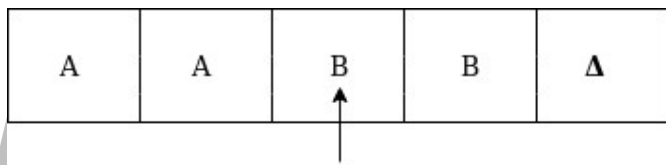
The move will be  $\delta(q_0, 0) = \delta(q_1, A, R)$  which means it will go to state  $q_1$ , replaced 0 by A, and head will move to right as:



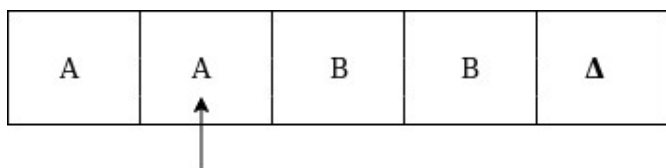
The move will be  $\delta(q_1, B) = \delta(q_1, B, R)$  which means it will not change any symbol, remain in the same state and move to right as:



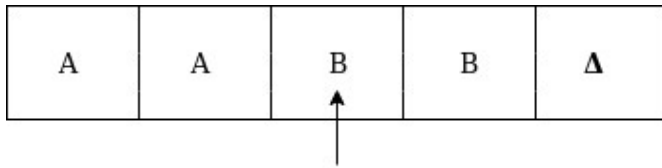
The move will be  $\delta(q_1, 1) = \delta(q_2, B, L)$  which means it will go to state  $q_2$ , replaced 1 by B and head will move to left as:



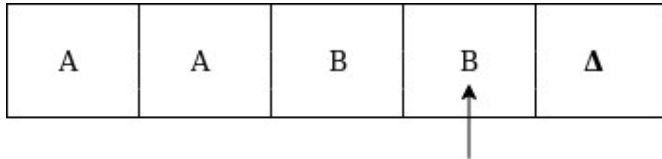
The move  $\delta(q_2, B) = (q_2, B, L)$  which means it will not change any symbol, remain in the same state and move to left as:



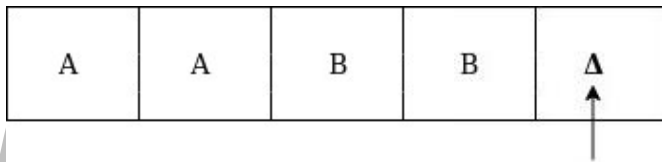
Now immediately before B is A that means all the 0's are marked by A. So we will move right to ensure that no 1 is present. The move will be  $\delta(q_2, A) = (q_0, A, R)$  which means it will go to state  $q_0$ , will not change any symbol, and move to right as:



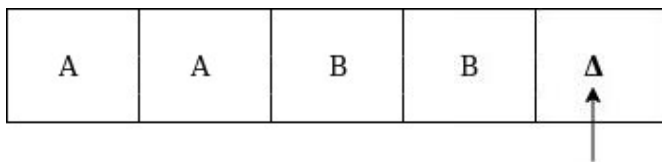
The move  $\delta(q_0, B) = (q_3, B, R)$  which means it will go to state  $q_3$ , will not change any symbol, and move to right as:



The move  $\delta(q_3, B) = (q_3, B, R)$  which means it will not change any symbol, remain in the same state and move to right as:



The move  $\delta(q_3, \Delta) = (q_4, \Delta, R)$  which means it will go to state  $q_4$  which is the HALT state and HALT state is always an accept state for any TM.

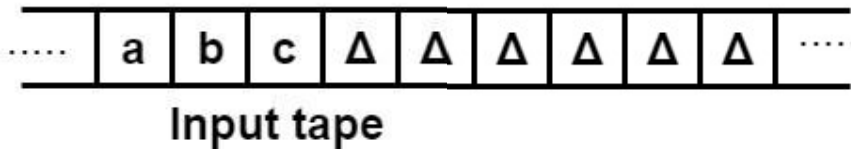


The same TM can be represented by Transition Diagram:

Basic Model of Turing machine

The turning machine can be modelled with the help of the following representation.

1. The input tape is having an infinite number of cells, each cell containing one input symbol and thus the input string can be placed on tape. The empty tape is filled by blank characters.



2. The finite control and the tape head which is responsible for reading the current input symbol. The tape head can move to left to right.

3. A finite set of states through which machine has to undergo.

4. Finite set of symbols called external symbols which are used in building the logic of turing machine.

Language accepted by Turing machine

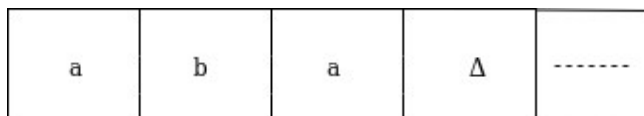
The turing machine accepts all the language even though they are recursively enumerable. Recursive means repeating the same set of rules for any number of times and enumerable means a list of elements. The TM also accepts the computable functions, such as addition, multiplication, subtraction, division, power function, and many more.

**Example:**

Construct a turing machine which accepts the language of aba over  $\Sigma = \{a, b\}$ .

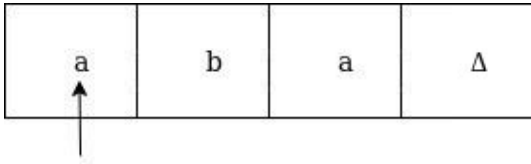
**Solution:**

We will assume that on input tape the string 'aba' is placed like this:

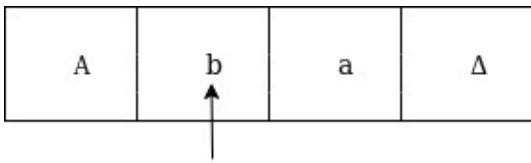


The tape head will read out the sequence up to the  $\Delta$  characters. If the tape head is readout 'aba' string then TM will halt after reading  $\Delta$ .

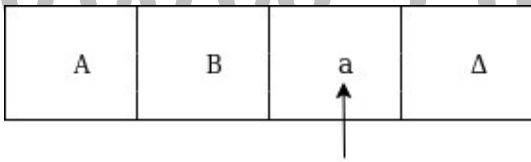
Now, we will see how this turing machine will work for aba. Initially, state is  $q_0$  and head points to a as:



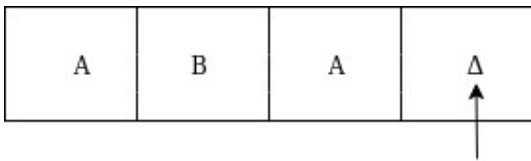
The move will be  $\delta(q_0, a) = \delta(q_1, A, R)$  which means it will go to state  $q_1$ , replaced a by A and head will move to right as:



The move will be  $\delta(q_1, b) = \delta(q_2, B, R)$  which means it will go to state  $q_2$ , replaced b by B and head will move to right as:



The move will be  $\delta(q_2, a) = \delta(q_3, A, R)$  which means it will go to state  $q_3$ , replaced a by A and head will move to right as:



The move  $\delta(q_3, \Delta) = (q_4, \Delta, S)$  which means it will go to state  $q_4$  which is the HALT state and HALT state is always an accept state for any TM.

The same TM can be represented by Transition Table:

States	A	b	$\Delta$
q0	(q1, A, R)	–	–
q1	–	(q2, B, R)	–
q2	(q3, A, R)	–	–
q3	–	–	(q4, $\Delta$ , S)
q4	–	–	–

The same TM can be represented by Transition Diagram:

