

For the other states  $\{A, E, G\}$ , for an input of either 0 or 1, the next state belongs to  $Q_2$ .

This divides  $Q_2$  into two parts:  $\{A, E, G\}$  and  $\{B, D, F, H\}$ . Let us name them as  $Q_3$  and  $Q_4$ .

The divided sets are

$$S_2 : \{\{C\}, \{A, E, G\}, \{B, D, F, H\}\}.$$

Consider the subset of states  $\{A, E, G\}$ . A and E with input 0 and 1 both go to F, B and F, H, respectively, i.e., to the subset  $\{B, D, F, H\}$ . G with input 0 and 1 goes to E, G, i.e., the same subset. Here,  $\{A, E, G\}$  is divided into two subsets:  $\{A, E\}$  and  $\{G\}$ .

The subset  $\{B, D, F, H\}$  can be divided depending on the input and the next state combination. B and H produce the next states C and G for input 0 and 1, respectively.

D and F produce the next states G and C for input 0 and 1, respectively. So, the set  $\{B, D, F, H\}$  is divided into two subsets:  $\{B, H\}$  and  $\{D, F\}$ .

The divided sets are

$$S_3 : \{\{C\}, \{A, E\}, \{G\}, \{B, H\}, \{D, F\}\}.$$

The subsets cannot be divided further making these as the states of minimized DFA. Let us rename the subsets as  $q_0, q_1, q_2, q_3$ , and  $q_4$ . The initial state was A, and so here the initial state is  $\{A, E\}$ , i.e.,  $q_1$ . The final state was C, and so here the final state is  $\{C\}$ , i.e.,  $q_0$ . The tabular representation of minimized DFA is

| Present State       | Next State |         |
|---------------------|------------|---------|
|                     | I/P = 0    | I/P = 1 |
| $\textcircled{q_0}$ | $q_0$      | $q_1$   |
| $\rightarrow q_1$   | $q_4$      | $q_3$   |
| $q_2$               | $q_1$      | $q_2$   |
| $q_3$               | $q_0$      | $q_2$   |
| $q_4$               | $q_2$      | $q_0$   |

### 3.15 Myhill–Nerode Theorem

*John Myhill and Anil Nerode* of the University of Chicago proposed a theorem in 1958 which provides a necessary and sufficient condition for a language to be regular. This theorem can also be used to minimize a DFA. But before going into the details of the theorem statement, we need to know some definitions related to the theorem.

#### 3.15.1 Equivalence Relation

- A relation  $R$  in set  $S$  is reflexive if  $xRx$  for every  $x$  in  $S$ .
- A relation  $R$  in set  $S$  is symmetric if for  $x, y$  in  $S$ ,  $yRx$  whenever  $xRy$ .
- A relation  $R$  in set  $S$  is transitive if for  $x, y$ , and  $z$  in  $S$ ,  $xRz$  whenever  $xRy$  and  $yRz$ .

A relation  $R$  in set  $S$  is called an equivalence relation if it is reflexive, symmetric, and transitive.

### 3.15.1.1 Right Invariant

An equivalence relation  $R$  on strings of symbols from some alphabet  $\Sigma$  is said to be right invariant if for all  $x, y \in \Sigma^*$  with  $xRy$  and all  $w \in \Sigma^*$  we have that  $xwRyw$ . This definition states that an equivalence

relation has the right invariant property if two equivalent strings (x and y) that are in the language still are equivalent if a third string (w) is appended to the right of both of them.

### 3.15.2 Statement of the Myhill–Nerode Theorem

The Myhill–Nerode theorem states that the following three statements are equivalent.

1. The set  $L$ , a subset of  $\Sigma^*$ , is accepted by a *DFA*, i.e.,  $L$  is a regular language.
2. There is a right-invariant equivalence relation  $R$  of finite index such that  $L$  is the union of some of the equivalence classes of  $R$ .
3. Let equivalence relation  $R_L$  be defined as  $xR_Ly$ , if and only if for all  $z$  in  $\Sigma^*$ ,  $xz$  is in  $L$  exactly when  $yz$  is in  $L$  then  $R_L$  is of finite index.

### 3.15.3 Myhill–Nerode Theorem in Minimizing a DFA

**Step I:** Build a two-dimensional matrix labelled by the states of the given DFA at the left and bottom side. The major diagonal and the upper triangular parts are shown as dashes.

**Step II:** One of the three symbols, X, x, or 0 are put in the locations where there is no dash.

1. Mark X at p, q in the lower triangular part such that p is the final state and Q is the non-final state.

2. Make distinguished pair combination of the non-final states. If there are  $n$  number of non-final states, there are  $nC_2$  number of distinguished pairs. Take a pair  $(p, q)$  and find  $(r, s)$ , such that  $r = \delta(p, a)$  and  $s = \delta(q, a)$ . If in the place of  $(r, s)$  there is  $X$  or  $x$ , in the place of  $(p, q)$ , there will be  $x$ .
3. If  $(r, s)$  is neither  $X$  nor  $x$ , then  $(p, q)$  is 0.
4. Repeat (2) and (3) for final states also.

**Step III:** The combination of states where there is 0, they are the states of the minimized machine.

Consider the following examples to get the earlier discussed method.

**Example 3.27** Minimize the following DFA using the Myhill-Nerode theorem.

| <i>PresentState</i> | <i>NextState</i> |           |
|---------------------|------------------|-----------|
|                     | $I/P = a$        | $I/P = b$ |
| $\rightarrow A$     | B                | E         |
| B                   | C                | D         |
| C                   | H                | I         |
| D                   | I                | H         |
| E                   | F                | G         |
| F                   | H                | I         |
| G                   | H                | I         |
| H                   | H                | H         |
| I                   | I                | I         |

Here C, D, F, G are final states.

**Solution:**

**Step I:** Divide the states of the DFA into two subsets: final (F) and non-final ( $Q - F$ ).

$$F = \{C, D, F, G\}, Q - F = \{A, B, E, H, I\}$$

Make a two-dimensional matrix as shown in Fig. 3.50 labelled at the left and bottom by the states of the DFA.

|   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|
| A | - | - | - | - | - | - | - | - | - |
| B |   | - | - | - | - | - | - | - | - |
| C |   |   | - | - | - | - | - | - | - |
| D |   |   |   | - | - | - | - | - | - |
| E |   |   |   |   | - | - | - | - | - |
| F |   |   |   |   |   | - | - | - | - |
| G |   |   |   |   |   |   | - | - | - |
| H |   |   |   |   |   |   |   | - | - |
| I |   |   |   |   |   |   |   |   | - |
|   | A | B | C | D | E | F | G | H | I |

**Fig. 3.50**

**Step II:**

1. The following combinations are the combination of the beginning and final state.

(A, C), (A, D), (A, F), (A, G), (B, C), (B, D), (B, F), (B, G), (E, C), (E, D), (E, F), (E, G), (H, C), (H, D), (H, F), (H, G), (I, C), (I, D), (I, F), (I, G).

Put X in these combinations of states. The modified matrix is given in Fig. 3.51.

|   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|
| A | - | - | - | - | - | - | - | - | - |
| B |   | - | - | - | - | - | - | - | - |
| C | X | X | - | - | - | - | - | - | - |
| D | X | X |   | - | - | - | - | - | - |
| E |   |   | X | X | - | - | - | - | - |
| F | X | X |   |   | X | - | - | - | - |
| G | X | X |   |   | X |   | - | - | - |
| H |   |   | X | X |   | X | X | - | - |
| I |   |   | X | X |   | X | X |   | - |
|   | A | B | C | D | E | F | G | H | I |

**Fig. 3.51**

2. The pair combination of non-final states are (A, B), (A, E), (A, H), (A, I), (B, E), (B, H), (B, I), (E, H), (E, I), and (H, I).

$$r = \delta(A, a) \rightarrow Bs = \delta(B, a) \rightarrow C$$

in the place of (B, C) there is X. So, in the place of (A, B), there will be x.

Similarly,

$(r, s) = \delta((A, E), a) \rightarrow (B, F)$  (there is X). In the place of (A, E), there will be x.

$(r, s) = \delta((A, H), a) \rightarrow (B, H)$  (neither X nor x). In the place of (A, H), there will be 0.

$(r, s) = \delta((A, I), a) \rightarrow (B, I)$  (neither X nor x). In the place of (A, I), there will be 0.

$(r, s) = \delta((B, E), a) \rightarrow (C, F)$  (neither X nor x). In the place of (B, E), there will be 0.

$(r, s) = \delta((B, H), a) \rightarrow (C, H)$  (there is X). In the place of (B, H), there will be x.

$(r, s) = \delta((B, I), a) \rightarrow (C, I)$  (there is X). In the place of (B, I), there will be x.

$(r, s) = \delta((E, H), a) \rightarrow (F, H)$  (there is X). In the place of (E, H), there will be x.

$(r, s) = \delta((E, I), a) \rightarrow (F, I)$  (there is X). In the place of (E, I), there will be x.

$(r, s) = \delta((H, I), a) \rightarrow (H, I)$  (neither X nor x). In the place of (H, I), there will be 0.

**3** The pair combination of final states are (C, D), (C, F), (C, G), (D, F), (D, G), and (F, G).

$(r, s) = \delta((C, D), a) \rightarrow (H, I)$  (neither X nor x). In the place of (C, D), there will be 0.

$(r, s) = \delta((C, F), a) \rightarrow (H, H)$  (there is dash, neither X nor x). In the place of (C, F), there will be 0.

$(r, s) = \delta((C, G), a) \rightarrow (H, H)$  (neither X nor x). In the place of (C, G), there will be 0.

$(r, s) = \delta((D, F), a) \rightarrow (I, H)$  (neither X nor x). In the place of (D, F), there will be 0.

$(r, s) = \delta((D, G), a) \rightarrow (I, H)$  (neither X nor x). In the place of (D, G), there will be 0.



$(r, s) = \delta((F, G), a) \rightarrow (H, H)$  (neither X nor x). In the place of (F, G), there will be 0.

The modified matrix becomes as shown in Fig. 3.52.

## 1 picture

|   |   |   |   |   |   |   |   |   |  |
|---|---|---|---|---|---|---|---|---|--|
| B | x |   |   |   |   |   |   |   |  |
| C | X | X |   |   |   |   |   |   |  |
| D | X | X | 0 |   |   |   |   |   |  |
| E | x | 0 | X | X |   |   |   |   |  |
| F | X | X | 0 | 0 | X |   |   |   |  |
| G | X | X | 0 | 0 | X | 0 |   |   |  |
| H | x | x | X | X | x | X | X |   |  |
| I | x | x | X | X | x | X | X | 0 |  |
|   | A | B | C | D | E | F | G | H |  |

Fig. 3.52

The combination of entries 0 are the states of the modified machine. The states of the minimized machine are  $[A]$ ,  $[B, E]$ ,  $[C, D]$ ,  $[C, F]$ ,  $[C, G]$ ,  $[D, F]$ ,  $[D, G]$ ,  $[F, G]$ , and  $[H, I]$ .

For the minimized machine  $M'$

$Q' = (\{A\}, \{B, E\}, \{C, D, F, G\}, \{H, I\})$ .  $[C, D]$ ,  $[C, F]$ ,  $[C, G]$ ,  $[D, F]$ ,  $[D, G]$ , and  $[F, G]$  are combined to a new state  $[CDEFG]$ .

$\Sigma = \{a, b\}\delta' : \text{(given in the following table)}$