## Problem

**Myhill–Nerode theorem.** Refer to Problem 1.51. Let L be a language and let X be a set of strings. Say that X is ***pairwise distinguishable by*** L if every two distinct strings in X are distinguishable by L. Define the ***index of*** L to be the maximum number of elements in any set that is pairwise distinguishable by L. The index of L may be finite or infinite.

**a.** Show that if L is recognized by a DFA with k states, L has index at most k.

**b.** Show that if the index of L is a finite number k, it is recognized by a DFA with k states.

**c.** Conclude that L is regular iff it has finite index. Moreover, its index is the size of the smallest DFA recognizing it.

## Step-by-step solution

**Step 1** of 4

The definition of **Myhill-Nerode theorem** is as follows:

**Myhill–Nerode theorem:** for any language $L$

• ***Distinguishable by L:*** *x* and *y* are the strings ***distinguishable*** by $L$, for the string *z* in generating of the strings $xz$ or $yz$ is a member of $L$.

• ***Indistinguishable by L:*** *x* and *y* are ***indistinguishable*** by $L$ for the string *z* we have $xz \in L$ every time $yz \in L$. We can write $x \equiv_L y$.

• ***Pair-wise distinguishable by L:*** set of strings contains in $S$, if every two separate strings are distinguishable in $L$.

• ***Index of L:*** It can count as finite or infinite. Language $L$ contains max number of elements which are ***pair-wise distinguishable***.

Comment

**Step 2** of 4

(a) Language $L$ recognized by DFA (Deterministic Finite Automata) as $M$ with number of states is $k$. We have to prove that $L$ has an index at most $k$.

Take a contradiction assumption i.e., $L$ has an index greater than $k$.

If $L$ contains index more than $k$ then $k+1$ strings are at least in any set $S$ which is ***pair wise distinguishable by*** $L$.

**Pigeonhole's** principle:

We will find two distinct strings $x$ and $y$ from $S$, such that the state of DFA $M$ after reading input $x$ is the same as the state of DFA $M$ after reading input $y$.

By applying **Pigeonhole's** principle both $xz$ and $yz$ are not in $L$. This is not satisfying the definition ***Distinguishable by L*** in **Myhill-Nerode theorem**

Hence contradiction occurs. Therefore our assumption that $L$ has index greater than $k$ is wrong. So, $L$ has index at most $k$.

Comment

**Step 3** of 4

(b) Index of Language $L$ contains $k$ finite states i.e., set $S = \{s_1, s_2, \ldots s_k\}$. We have to prove that $L$ recognized by DFA with $k$ states.

• Let $M = (Q, \Sigma, \delta, q_0, F)$ be *DFA* with *k* states that recognizes $L$

• The construction of $M$ is as follows:

o Assume $Q = (q_1, q_2 \ldots q_k)$ is the set of states.

o Transition function is given as: $\delta(q_i, a) = q_j$ if $s_i a$ and $s_j$ are not distinguishable.

o $F = \{q_i \mid s_i \in L\}$ be the setoff

o Start state $q_0$ be the state such that $s_i$ and the empty string $\in$ are not distinguishable by $L$.

• We show that if string *t* and $s_j$ are not distinguishable by $L$, the state of $M$ will be $q_j$ after reading *t* as input.

• By the definition of *F*, $M$ accepts *t* if and only if *t* is in $L$.

• Hence $M$ recognizes $L$.

Comment

**Step 4** of 4

(c) Language $L$ is regular if it contains finite index. Index is size of smallest DFA recognizing it.

(i) <u>if $L$ is regular then $L$ has finite index:</u>

• Let us assume that $L$ is regular.

• $M$ be *DFA* that recognizes $L$.

• Let $k$ be the number of states in $M$.

• Then from part (a), $L$ has index at most $k$

(ii) <u>if $L$ has finite index then $L$ is regular:</u>

• Let us assume that $L$ has finite index $k$

• Then from part (b) we can contract a *DFA* with *k* states recognizing $L$

• We know that “A language is regular if and only if it is recognized by some *DFA*”

• Therefore $L$ is regular language.

Therefore from (i) and (ii) *L* is regular if and only if it has finite index.

• The index $k$ is size of the smallest *DFA* $M$ recognizing it, suppose on the opposing that is not true. From part (a) we could terminate that $L$ has indexed fewer than $k$, which contradicts fact that $L$ has index equal to $k$.

Comment