Myhill-Nerode

Tuesday, September 10, 2019 5:21 PM

The Pumping Lemma is a great tool for proving that a language is not .

However, it is not complete. I.e.

$\exists$ languages which are not regular and can be pumped.

Today we'll see a complete characterization.

Given a set S, possibly infinite, an Equivalence Relation is a set of pairs from S which satisfies 3 properties:

(1) Reflexive $(a,a) \in R \quad \forall a \in S$

(2) Symmetric $(a,b) \in R \Rightarrow (b,a) \in R$

(3) Transitive $(a,b),(b,c) \in R \Rightarrow (a,c) \in R$.

E.g. R = "born in the same city/village".

not an equivalence : "taller than"

"friend of"

In an undirected graph $G=(V,E)$

In an undirected graph $G = (V, E)$
$R(a,b) =$ "path between a and b".
Then all vertices reachable from a form a component, and the equivalence relation R partitions V into connected components.

Given a language L (set of strings)
we define the following relation
$(X,Y) \in R$ if $\forall Z \quad XZ \in L \Leftrightarrow YZ \in L$.
For any suffix Z, $XZ \in L$ iff $YZ \in L$
i.e. there is no suffix Z which makes $XZ \in L$ and $YZ \notin L$, there is no suffix that distinguishes X from Y.

<u>Lemma</u>. R is an equivalence relation.

<u>Thm</u>. [Myhill-Nerode] A language L is regular <u>iff</u> R has a finite number of equivalence classes.

Pf. (i) L is regular $\Rightarrow$ R has a finite number of classes

Pf. (i) L is regular $\Rightarrow$ R has a finite # equivalence classes

Let D be a DFA for L.

Label each string X with the state q that D ends on when given input X.

This partitions all strings.

Now if X, Y have the same label q, then for any Z, XZ and YZ will go to the same state; so both will be accepted or both will not be accepted.

$\Rightarrow$ # equivalence classes of $R \leq$ # states of D.

(ii) R has a finite # equiv. classes $\Rightarrow$ L is regular

We will construct a DFA for L.

states = # equivalence classes.

q_0 is the state for the empty string.

For any q pick x that ends in q.

Σ $\delta(q, a) = q'$

For any (i) pick

and for each $y \in \Sigma$, $\delta(q, y) = q'$

where q' is the state for xy.

states q containing $x \in L$ are accepting states.

$\boxtimes$.

<u>Cor.</u> The minimum #states of a DFA that accepts a regular language L is the number of equivalence classes of R for L.