# TDDD14/TDDD85

Slides for Lecture 6

Myhill-Nerode Relations

Christer Bäckström, 2017

# Myhill-Nerode Relations

Let $\Sigma$ be an alphabet.
Let $L \subseteq \Sigma^*$ be a language
Let $\equiv$ be an equivalence relation on $\Sigma^*$.
Then $\equiv$ is a *Myhill-Nerode relation* for $L$ if it satisfies:

1. It is *right congruent*
i.e. for all $x, y \in \Sigma^*$ and all $a \in \Sigma$, if $x \equiv y$, then $xa \equiv ya$.

2. It *refines* $L$,
i.e. if $x \equiv y$, then $x \in L \Leftrightarrow y \in L$.

3. It is of finite index,
i.e. it has a finite number of equivalence classes.

Note:

$\equiv$ is an equivalence relation on the strings in $\Sigma^*$.
(The relation $\approx$ in a previous lecture was an equivalence relation on the states of a DFA.)

Even if the number of equivalence classes of $\equiv$ is finite, the size of each class need not be finite. (At least one of them must contain an infinite number of strings since $\Sigma^*$ is infinite.)

# Construction $M \mapsto \equiv_M$

Let $M$ be a DFA over an alphabet $\Sigma$ with start state $q_0$ such that all states are accessible from $q_0$.

Define the relation $\equiv_M$ on $\Sigma^*$ such that $x \equiv_M y$ iff $\widehat{\delta}(q_0, x) = \widehat{\delta}(q_0, y)$.

Then $\equiv_M$ is an equivalence relation on $\Sigma^*$.

We will show that $\equiv_M$ is also a Myhill-Nerode relation for $L(M)$.

1. Let $x$ and $y$ be arbitrary strings in $\Sigma^*$ and let $a$ be an arbitrary symbol in $\Sigma$.

Assume $x \equiv_M y$. Then $\hat{\delta}(q_0, x) = \hat{\delta}(q_0, y)$ by definition.

We get

$$\hat{\delta}(q_0, xa) = \delta(\hat{\delta}(q_0, x), a) = \delta(\hat{\delta}(q_0, y), a) = \hat{\delta}(q_0, ya).$$

That is, $\equiv_M$ is right congruent.

2. Let $x$ and $y$ be arbitrary strings in $\Sigma^*$.

Assume $x \equiv_M y$. Then $\widehat{\delta}(q_0, x) = \widehat{\delta}(q_0, y)$.

Obviously, $M$ either accepts both $x$ and $y$
or $M$ rejects both $x$ and $y$,
so $x \in L(M)$ iff $y \in L(M)$.

That is, $\equiv_M$ refines $L(M)$.

3. For each $x \in \Sigma^*$, the equivalence class of $x$ is
$[x] = \{y \in \Sigma^* \mid x \equiv_M y\} = \{y \in \Sigma^* \mid \widehat{\delta}(q_0, y) = \widehat{\delta}(q_0, x)\}$.

Let $Q = \{q_0, \ldots, q_n\}$ be the states of $M$.

For each $i$ ($0 \leq i \leq n$), let $x_i \in \Sigma$ be a string such that $\widehat{\delta}(q_0, x_i) = q_i$. (Such a string must exist since we assume all states are accessible from $q_0$.)

Then, $[x_i] \neq [x_j]$ for all $i \neq j$, i.e. there is an equivalence class $[x_i]$ for each state $q_i \in Q$.

Suppose there is a string $y \in \Sigma^*$ s.t. $[y] \neq [x_i]$ for all $i$ ($0 \leq i \leq n$).
Then $\widehat{\delta}(q_0, y) \neq q_i$ for all $q_i \in Q$.
This is impossible, so $\equiv_M$ must have exactly $n + 1$ equivalence classes.

That is, $\equiv_M$ satisfies conditions 1–3 so it is a Myhill-Nerode relation for $L(M)$.

Since $L(M)$ must be a regular language, it follows that we can define a Myhill-Nerode relation for every regular language.

## Construction $\equiv \mapsto M_\equiv$

Let $\Sigma$ be an alphabet and $L \subseteq \Sigma^*$ a language.
Suppose $\equiv$ is a Myhill-Nerode relation for $L$.

Then $\equiv$ has a finite number of equivalence relations, so we can construct a DFA $M_\equiv = (Q, \Sigma, \delta, q_0, F)$ for $L$ as follows:

- $Q = \{[x] \mid x \in \Sigma^*\}$
- $q_0 = [\varepsilon]$
- $F = \{[x] \mid x \in L\}$
- $\delta([x], a) = [xa]$.

Then $L(M_\equiv) = L$ (see book for proof).

That is, if a language $L$ has a Myhill-Nerode relation, then it must be regular.

# Regular Languages and Myhill-Nerode Relations

The previous two constructions give the following result:

Let $L$ be a language over some alphabet. Then

$L$ is regular iff there is a Myhill-Nerode relation for $L$.

# Automata Isomorphism

Let $M = (Q^M, \Sigma, \delta^M, q_0^M, F^M)$ and $N = (Q^N, \Sigma, \delta^N, q_0^N, F^N)$ be two DFAs.

Then $M$ and $N$ are *isomorphic* if there exists a bijective function $f : Q^M \to Q^N$ such that

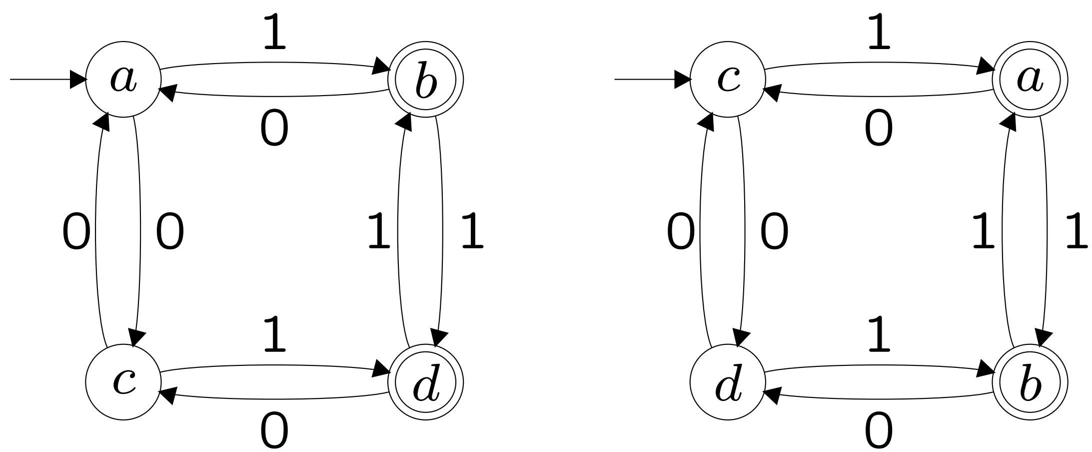
1. $f(q_0^M) = q_0^N$,

2. $f(\delta^M(p, a)) = \delta^N(f(p), a)$ for all $p \in Q^M$ and $a \in \Sigma$,

3. $p \in F^M$ iff $f(p) \in F^N$.

That is we can rename the states of $M$ so it becomes identical to $N$.

The following two DFAs are isomorphic:



# A Closer Analysis of the Constructions

We have shown two constructions:

$M \mapsto \equiv_M$: Given a DFA, construct a Myhill-Nerode relation $\equiv_M$

$\equiv \mapsto M_{\equiv}$: Given a Myhill-Nerode relation, construct a DFA.

These constructions are inverses of each other in the following sense.

Let $L$ be a regular language with a Myhill-Nerode relation $\equiv$.

1. Construct the DFA $M_{\equiv}$ for $\equiv$.

2. Then define the equivalence relation $\equiv_{M_{\equiv}}$.

That is, we do $\equiv \mapsto M_{\equiv} \mapsto \equiv_{M_{\equiv}}$

Then $\equiv$ and $\equiv_{M_{\equiv}}$ are the same relation.

Let $M$ be a DFA with no inaccessible states.

1. Construct the Myhill-Nerode relation $\equiv_M$ for $L(M)$.

2. Then construct the DFA $M_{\equiv_M}$ for $\equiv_M$.

That is, we do $M \mapsto \equiv_M \mapsto \; M_{\equiv_M}$

Then $M$ and $M_{\equiv_M}$ are isomorphic.

# Myhill-Nerodes Theorem

Recall that a relation on $\Sigma^*$ is a subset of $\Sigma^* \times \Sigma^*$, i.e. a set of pairs of strings.

Let $\equiv_1$ and $\equiv_2$ be two equivalence relations on $\Sigma^*$.

Then $\equiv_1$ *refines* $\equiv_2$ if $\equiv_1 \subseteq \equiv_2$
(i.e. if $x \equiv_1 y \Rightarrow x \equiv_2 y$).

We say that $\equiv_1$ is *finer* than $\equiv_2$ and that $\equiv_2$ is *coarser* than $\equiv_1$.

The finest possible relation is $\{(x, x) \mid x \in \Sigma^*\}$.
The coarsest possible relation is $\{(x, y) \mid x, y \in \Sigma^*\}$.

Let $\Sigma$ be an alphabet and let $L \subseteq \Sigma^*$ be a language.
($L$ need not be regular.)

Define the relation $\equiv_L$ such that for all $x, y \in \Sigma^*$,

$$x \equiv_L y \text{ iff for all } z \in \Sigma^* (xz \in L \Leftrightarrow yz \in L)$$

Then $\equiv_L$ is the coarsest possible relation for $L$ that satisfies conditions 1 and 2 for Myhill-Nerode relations.

**Theorem** (Myhill-Nerode): Let $L \subseteq \Sigma^*$ be a language.
Then the following statements are equivalent:

1. $L$ is regular,
2. there exists a Myhill-Nerode relation for $L$,
3. the relation $\equiv_L$ has a finite number of equivalence classes.