Lecture 15

Myhill–Nerode Relations

Two deterministic finite automata

$$M = (Q_M,\ \Sigma,\ \delta_M,\ s_M,\ F_M),$$
$$N = (Q_N,\ \Sigma,\ \delta_N,\ s_N,\ F_N)$$

are said to be *isomorphic* (Greek for "same form") if there is a one-to-one and onto mapping $f : Q_M \to Q_N$ such that

- $f(s_M) = s_N$,
- $f(\delta_M(p,a)) = \delta_N(f(p),a)$ for all $p \in Q_M$, $a \in \Sigma$, and
- $p \in F_M$ iff $f(p) \in F_N$.

That is, they are essentially the same automaton up to renaming of states. It is easily argued that isomorphic automata accept the same set.

In this lecture and the next we will show that if M and N are any two automata with no inaccessible states accepting the same set, then the quotient automata $M/\approx$ and $N/\approx$ obtained by the collapsing algorithm of Lecture 14 are isomorphic. Thus the DFA obtained by the collapsing algorithm is the minimal DFA for the set it accepts, and this automaton is unique up to isomorphism.

We will do this by exploiting a profound and beautiful correspondence between finite automata with input alphabet Σ and certain equivalence relations on Σ^*. We will show that the unique minimal DFA for a regular set R can be defined in a natural way *directly from R*, and that any minimal automaton for R is isomorphic to this automaton.

Myhill–Nerode Relations

Let $R \subseteq \Sigma^*$ be a regular set, and let $M = (Q, \Sigma, \delta, s, F)$ be a DFA for R with no inaccessible states. The automaton M induces an equivalence relation $\equiv_M$ on Σ^* defined by

$$x \equiv_M y \stackrel{\text{def}}{\iff} \widehat{\delta}(s,x) = \widehat{\delta}(s,y).$$

(Don't confuse this relation with the collapsing relation $\approx$ of Lecture 13—that relation was defined on Q, whereas $\equiv_M$ is defined on Σ^*.)

One can easily show that the relation $\equiv_M$ is an equivalence relation; that is, that it is reflexive, symmetric, and transitive. In addition, $\equiv_M$ satisfies a few other useful properties:

(i) It is a *right congruence*: for any $x, y \in \Sigma^*$ and $a \in \Sigma$,

$$x \equiv_M y \Rightarrow xa \equiv_M ya.$$

To see this, assume that $x \equiv_M y$. Then

$$\begin{aligned}\widehat{\delta}(s,xa) &= \delta(\widehat{\delta}(s,x),a)\\ &= \delta(\widehat{\delta}(s,y),a) \quad \text{by assumption}\\ &= \widehat{\delta}(s,ya).\end{aligned}$$

(ii) It *refines* R: for any $x, y \in \Sigma^*$,

$$x \equiv_M y \Rightarrow (x \in R \iff y \in R).$$

This is because $\widehat{\delta}(s,x) = \widehat{\delta}(s,y)$, and this is either an accept or a reject state, so either both x and y are accepted or both are rejected. Another way to say this is that every $\equiv_M$-class has either all its elements in R or none of its elements in R; in other words, R is a union of $\equiv_M$-classes.

(iii) It is of *finite index*; that is, it has only finitely many equivalence classes. This is because there is exactly one equivalence class

$$\{x \in \Sigma^* \mid \widehat{\delta}(s,x) = q\}$$

corresponding to each state q of M.

Let us call an equivalence relation $\equiv$ on Σ^* a *Myhill–Nerode relation for R* if it satisfies properties (i), (ii), and (iii); that is, if it is a right congruence of finite index refining R.

The interesting thing about this definition is that it characterizes exactly the relations on Σ^* that are $\equiv_M$ for some automaton M. In other words, we can reconstruct M from $\equiv_M$ using only the fact that $\equiv_M$ is Myhill–Nerode. To see this, we will show how to construct an automaton $M_\equiv$ for R from any given Myhill–Nerode relation $\equiv$ for R. We will show later that the two constructions

$$M \mapsto \equiv_M,$$
$$\equiv \mapsto M_\equiv$$

are inverses up to isomorphism of automata.

Let $R \subseteq \Sigma^*$, and let $\equiv$ be an arbitrary Myhill–Nerode relation for R. Right now we're not assuming that R is regular, only that the relation $\equiv$ satisfies (i), (ii), and (iii). The $\equiv$-class of the string x is

$$[x] \stackrel{\text{def}}{=} \{y \mid y \equiv x\}.$$

Although there are infinitely many strings, there are only finitely many $\equiv$-classes, by property (iii).

Now define the DFA $M_\equiv = (Q,\ \Sigma,\ \delta,\ s,\ F)$, where

$$\begin{aligned} Q &\stackrel{\text{def}}{=} \{[x] \mid x \in \Sigma^*\}, \\ s &\stackrel{\text{def}}{=} [\epsilon], \\ F &\stackrel{\text{def}}{=} \{[x] \mid x \in R\}, \\ \delta([x], a) &\stackrel{\text{def}}{=} [xa]. \end{aligned}$$

It follows from property (i) of Myhill–Nerode relations that δ is well defined. In other words, we have defined the action of δ on an equivalence class $[x]$ in terms of an element x chosen from that class, and it is conceivable that we could have gotten something different had we chosen another $y \in [x]$ such that $[xa] \neq [ya]$. The property of right congruence says exactly that this cannot happen.

Finally, observe that

$$x \in R \iff [x] \in F. \tag{15.1}$$

The implication ($\Rightarrow$) is from the definition of F, and ($\Leftarrow$) follows from the definition of F and property (ii) of Myhill–Nerode relations.

Now we are ready to prove that $L(M_\equiv) = R$.

Lemma 15.1 $\widehat{\delta}([x], y) = [xy]$.

Proof. Induction on $|y|$.

Basis

$$\widehat{\delta}([x], \epsilon) = [x] = [x\epsilon].$$

Induction step

$$\begin{aligned}\widehat{\delta}([x], ya) &= \delta(\widehat{\delta}([x], y), a) && \text{definition of } \widehat{\delta} \\ &= \delta([xy], a) && \text{induction hypothesis} \\ &= [xya] && \text{definition of } \delta.\end{aligned}$$

□

Theorem 15.2 $L(M_\equiv) = R$.

Proof.

$$\begin{aligned}x \in L(M_\equiv) &\iff \widehat{\delta}([\epsilon], x) \in F && \text{definition of acceptance} \\ &\iff [x] \in F && \text{Lemma 15.1} \\ &\iff x \in R && \text{property (15.1).}\end{aligned}$$

□

$M \mapsto \equiv_M$ and $\equiv \mapsto M_\equiv$ Are Inverses

We have described two natural constructions, one taking a given automaton M for R with no inaccessible states to a corresponding Myhill–Nerode relation $\equiv_M$ for R, and one taking a given Myhill–Nerode relation $\equiv$ for R to a DFA $M_\equiv$ for R. We now wish to show that these two operations are inverses up to isomorphism.

Lemma 15.3 *(i) If $\equiv$ is a Myhill–Nerode relation for R, and if we apply the construction $\equiv \mapsto M_\equiv$ and then apply the construction $M \mapsto \equiv_M$ to the result, the resulting relation $\equiv_{M_\equiv}$ is identical to $\equiv$.*

(ii) If M is a DFA for R with no inaccessible states, and if we apply the construction $M \mapsto \equiv_M$ and then apply the construction $\equiv \mapsto M_\equiv$ to the result, the resulting DFA $M_{\equiv_M}$ is isomorphic to M.

Proof. (i) Let $M_\equiv = (Q, \Sigma, \delta, s, F)$ be the automaton constructed from $\equiv$ as described above. Then for any $x, y \in \Sigma^*$,

$$\begin{aligned}x \equiv_{M_\equiv} y &\iff \widehat{\delta}(s, x) = \widehat{\delta}(s, y) && \text{definition of } \equiv_{M_\equiv} \\ &\iff \widehat{\delta}([\epsilon], x) = \widehat{\delta}([\epsilon], y) && \text{definition of } s \\ &\iff [x] = [y] && \text{Lemma 15.1} \\ &\iff x \equiv y.\end{aligned}$$

(ii) Let $M = (Q, \Sigma, \delta, s, F)$ and let $M_{\equiv_M} = (Q', \Sigma, \delta', s', F')$. Recall from the construction that

$$\begin{aligned}
[x] &= \{y \mid y \equiv_M x\} = \{y \mid \widehat{\delta}(s,y) = \widehat{\delta}(s,x)\}, \\
Q' &= \{[x] \mid x \in \Sigma^*\}, \\
s' &= [\epsilon], \\
F' &= \{[x] \mid x \in R\}, \\
\delta'([x],a) &= [xa].
\end{aligned}$$

We will show that $M_{\equiv_M}$ and M are isomorphic under the map

$$\begin{aligned}
&f : Q' \to Q, \\
&f([x]) = \widehat{\delta}(s,x).
\end{aligned}$$

By the definition of $\equiv_M$, $[x] = [y]$ iff $\widehat{\delta}(s,x) = \widehat{\delta}(s,y)$, so the map f is well defined on $\equiv_M$-classes and is one-to-one. Since M has no inaccessible states, f is onto.

To show that f is an isomorphism of automata, we need to show that f preserves all automata-theoretic structure: the start state, transition function, and final states. That is, we need to show

- $f(s') = s$,
- $f(\delta'([x],a)) = \delta(f([x]),a)$,
- $[x] \in F' \iff f([x]) \in F$.

These are argued as follows:

$$\begin{aligned}
f(s') &= f([\epsilon]) && \text{definition of } s' \\
&= \widehat{\delta}(s,\epsilon) && \text{definition of } f \\
&= s && \text{definition of } \widehat{\delta};
\end{aligned}$$

$$\begin{aligned}
f(\delta'([x],a)) &= f([xa]) && \text{definition of } \delta' \\
&= \widehat{\delta}(s,xa) && \text{definition of } f \\
&= \delta(\widehat{\delta}(s,x),a) && \text{definition of } \widehat{\delta} \\
&= \delta(f([x]),a) && \text{definition of } f;
\end{aligned}$$

$$\begin{aligned}
[x] \in F' &\iff x \in R && \text{definition of } F \text{ and property (ii)} \\
&\iff \widehat{\delta}(s,x) \in F && \text{since } L(M) = R \\
&\iff f([x]) \in F && \text{definition of } f.
\end{aligned}$$

□

We have shown:

Theorem 15.4 *Let Σ be a finite alphabet. Up to isomorphism of automata, there is a one-to-one correspondence between deterministic finite automata over Σ with no inaccessible states accepting R and Myhill–Nerode relations for R on Σ^*.*