Supplementary Lecture D

The Myhill–Nerode Theorem for Term Automata

In the last lecture we generalized DFAs on strings to term automata over a signature Σ and demonstrated that automata-theoretic concepts such as "final states" and "run" were really more general algebraic concepts in disguise. In this lecture we continue to develop this correspondence, leading finally to a fuller understanding of the Myhill–Nerode theorem.

Congruence

First we need to introduce the important algebraic concept of *congruence*. Congruences and homomorphisms go hand-in-hand. Recall from Supplementary Lecture C that a *homomorphism* between two Σ-algebras is a map that preserves all algebraic structure (Definition C.7). Every homomorphism $\sigma : \mathcal{A} \to \mathcal{B}$ induces a certain natural binary relation on $\mathcal{A}$:

$$u \equiv_\sigma v \stackrel{\text{def}}{\iff} \sigma(u) = \sigma(v).$$

The relation $\equiv_\sigma$ is called the *kernel*[1] of σ.

[1] If you have taken algebra, you may have seen the word *kernel* used differently: normal subgroups of groups, ideals of rings, null spaces of linear transformations. These concepts are closely allied and serve the same purpose. The definition of *kernel* as a binary relation is more broadly applicable.

The kernel of any homomorphism defined on $\mathcal{A}$ is an equivalence relation on $\mathcal{A}$ (reflexive, symmetric, transitive). It also respects all algebraic structure in the following sense:

(i) for all n-ary function symbols $f \in \Sigma$, if $u_i \equiv_\sigma v_i$, $1 \le i \le n$, then

$$f^{\mathcal{A}}(u_1, \ldots, u_n) \equiv_\sigma f^{\mathcal{A}}(v_1, \ldots, v_n);$$

(ii) for all n-ary relation symbols $R \in \Sigma$, if $u_i \equiv_\sigma v_i$, $1 \le i \le n$, then

$$R^{\mathcal{A}}(u_1, \ldots, u_n) \iff R^{\mathcal{A}}(v_1, \ldots, v_n).$$

These properties follow immediately from the properties of homomorphisms and the definition of $\equiv_\sigma$.

In general, a *congruence* on $\mathcal{A}$ is any equivalence relation on $\mathcal{A}$ satisfying properties (i) and (ii):

Definition D.1 Let $\mathcal{A}$ be a Σ-algebra with carrier A. A *congruence* on $\mathcal{A}$ is an equivalence relation $\equiv$ on A such that

(i) for all n-ary function symbols $f \in \Sigma$, if $u_i \equiv v_i$, $1 \le i \le n$, then

$$f^{\mathcal{A}}(u_1, \ldots, u_n) \equiv f^{\mathcal{A}}(v_1, \ldots, v_n);$$

(ii) for all n-ary relation symbols $R \in \Sigma$, if $u_i \equiv v_i$, $1 \le i \le n$, then

$$R^{\mathcal{A}}(u_1, \ldots, u_n) \iff R^{\mathcal{A}}(v_1, \ldots, v_n). \qquad \square$$

Thus the kernel of every homomorphism is a congruence. Now, the interesting thing about this definition is that it goes the other way as well: every congruence is the kernel of some homomorphism. In other words, given an arbitrary congruence $\equiv$, we can construct a homomorphism σ such that $\equiv_\sigma$ is $\equiv$. In fact, we can make the homomorphism σ an epimorphism. We will prove this using a general algebraic construction called the *quotient construction*.

We saw an example of the quotient construction in Lecture 13, where we used it to collapse a DFA. We saw it again in Lecture 15, where we constructed an automaton for a set A from a given Myhill–Nerode relation for A. By now you have probably figured out where we are going with this: the collapsing relations $\approx$ of Lecture 13, the Myhill–Nerode relations $\equiv$ of Lecture 15, and the maximal Myhill–Nerode relation $\equiv_R$ of Lecture 16 are all congruences on Σ-algebras! We warned you not to confuse these different kinds of relations, because some were defined on automata and others on strings; but now we can roll them all into a single concept. This is the power and beauty of abstraction.

Lemma D.2 *(i) The kernel of any homomorphism is a congruence.*

(ii) Any congruence is the kernel of an epimorphism.

Proof. For (ii), build a quotient algebra whose elements are the congruence classes $[u]$. There is a unique interpretation of the function and relation symbols in the quotient making the map $u \mapsto [u]$ an epimorphism. We'll leave the details as an exercise (Miscellaneous Exercise 67). $\square$

In fact, if the map $t \mapsto t^{\mathcal{A}}$ is onto (i.e., if each element of $\mathcal{A}$ is named by a term), then the congruences on $\mathcal{A}$ and homomorphic images of $\mathcal{A}$ are in one-to-one correspondence up to isomorphism. This is true of term algebras, since the map $t \mapsto t^{T_\Sigma}$ is the identity. This is completely analogous to Lemma 15.3, which for the case of automata over strings gives a one-to-one correspondence up to isomorphism between the Myhill–Nerode relations for A (i.e., the right congruences of finite index refining A) and the DFAs with no inaccessible states accepting A. Here "no inaccessible states" just means that the map $t \mapsto t^{\mathcal{A}}$ is onto.

The Myhill–Nerode Theorem

Recall from Lecture 16 the statement of the Myhill–Nerode theorem: for a set A of strings over a finite alphabet Σ, the following three conditions are equivalent:

(i) A is regular;

(ii) there exists a Myhill–Nerode relation for A;

(iii) the relation $\equiv_A$ is of finite index, where

$$x \equiv_A y \overset{\text{def}}{\Longleftrightarrow} \forall z \in \Sigma^* \ (xz \in A \Longleftrightarrow yz \in A).$$

Recall that *finite index* means finitely many equivalence classes, and a relation is *Myhill–Nerode for A* if it is a right congruence of finite index refining A.

This theorem generalizes in a natural way to automata over terms. Define a *context* to be a term in $T_{\Sigma \cup \{x\}}$, where x is a new symbol of arity 0. For a context u and ground term $t \in T_\Sigma$, denote by $\mathsf{s}_t^x(u)$ the term in T_Σ obtained by substituting t for all occurrences of x in u. Formally,

$$\begin{aligned} \mathsf{s}_t^x(x) &\overset{\text{def}}{=} t, \\ \mathsf{s}_t^x(ft_1 \cdots t_n) &\overset{\text{def}}{=} f\mathsf{s}_t^x(t_1) \cdots \mathsf{s}_t^x(t_n). \end{aligned}$$

As usual, the last line includes the case of constants:

$$\mathsf{s}_t^x(c) \overset{\text{def}}{=} c.$$

Let Σ be a signature consisting of finitely many function symbols and a single unary relation symbol R. For a given $A \subseteq T_\Sigma$ and ground terms $s, t \in T_\Sigma$, define

$$s \equiv_A t \stackrel{\text{def}}{\iff} \text{for all contexts } u,\ \mathbf{s}_s^x(u) \in A \iff \mathbf{s}_t^x(u) \in A.$$

It is not difficult to argue that $\equiv_A$ is a congruence on $T_\Sigma(A)$ (Miscellaneous Exercise 68).

Theorem D.3 **(Myhill–Nerode theorem for term automata)** *Let $A \subseteq T_\Sigma$. Let $T_\Sigma(A)$ denote the term algebra over Σ in which R is interpreted as the unary relation A. The following statements are equivalent:*

(i) A is regular;

(i′) $T_\Sigma(A)$ has a finite homomorphic image;

(ii) there exists a congruence of finite index on $T_\Sigma(A)$;

(iii) the relation $\equiv_A$ is of finite index.

Proof. We have already observed (Lemma C.11) that to say A is regular (i.e., accepted by a term automaton) is just another way of saying that $T_\Sigma(A)$ has a finite homomorphic image. Thus (i) and (i′) are equivalent.

(i′) $\Rightarrow$ (ii) If $T_\Sigma(A)$ has a finite homomorphic image under epimorphism σ, then the kernel of σ is of finite index, since its congruence classes are in one-to-one correspondence with the elements of the homomorphic image.

(ii) $\Rightarrow$ (iii) Let $\equiv$ be any congruence on $T_\Sigma(A)$. We show that $\equiv$ refines $\equiv_A$; therefore, $\equiv_A$ is of finite index if $\equiv$ is. Suppose $s \equiv t$. It follows by a straightforward inductive argument using Definition D.1(i) that for any context u, $\mathbf{s}_s^x(u) \equiv \mathbf{s}_t^x(u)$; then by Definition D.1(ii), $\mathbf{s}_s^x(u) \in A$ iff $\mathbf{s}_t^x(u) \in A$. Since the context u was arbitrary, $s \equiv_A t$.

(iii) $\Rightarrow$ (i′) Since $\equiv_A$ is a congruence, it is the kernel of an epimorphism obtained by the quotient construction. Since $\equiv_A$ is of finite index, the quotient algebra is finite and therefore a finite homomorphic image of $T_\Sigma(A)$. □

As in Lecture 16, the quotient of $T_\Sigma(A)$ by $\equiv_A$ gives the minimal homomorphic image of $T_\Sigma(A)$; and for any other homomorphic image $\mathcal{B}$, there is a unique homomorphism $\mathcal{B} \to T_\Sigma(A)/\equiv_A$.

Historical Notes

Thatcher and Wright [119] generalized finite automata on strings to finite automata on terms and developed the algebraic connection. The more general version of the Myhill–Nerode theorem (Theorem D.3) is in some sense an inevitable consequence of Myhill and Nerode's work [91, 94] since according to Thatcher and Wright, "conventional finite automata theory goes through for the generalization—and it goes through quite neatly" [119]. The first explicit mention of the equivalence of the term analogs of (i) and (ii) in the statement of the Myhill–Nerode theorem seems to be by Brainerd [13, 14] and Eilenberg and Wright [34], although the latter claim that their paper "contains nothing that is essentially new, except perhaps for a point of view" [34]. A relation on terms analogous to $\equiv_A$ was defined and clause (iii) added explicitly by Arbib and Give'on [5, Definition 2.13], although it is also essentially implicit in work of Brainerd [13, 14].

Good general references are Gécseg and Steinby [42] and Englefriet [35].