COL352 Lecture 10

Contents

1 Recap

Game playing version of pumping lemma.

2 Definitions

Definition 1

$D_1 = (Q_1, \Sigma, \delta_1, q_1, A_1)$ is isomorphic to $D_2 = (Q_2, \Sigma, \delta, q_2, A_2)$ if $\exists$ a bijection $h : Q_1 \to Q_2$ such that:

1. $h(q_1) = q_2$
2. $q \in A \iff h(q) \in A_2$
3. $h(\delta_1(q, a)) = \delta_2(h(q), a)$

Observe that if D_1, D_2 are isomorphic, then $\mathcal{L}(D_1) = \mathcal{L}(D_2)$.

Definition 2

Let $D = (Q, \sigma, \delta, q_0, A)$ be a DFA. Define the relation $\sim_D$ (read "equivalent" wrt D) on Σ^* as

$$x \sim_D y \iff \hat{\delta}(q_0, x) = \hat{\delta}(q_0, y)$$

Definition 3

Let $L \subseteq \Sigma^*$ be <u>any</u> language. Define the relation $=_L$ (read "equivalent" wrt L) on Σ^* as:
$x =_L y \iff \forall z \in \Sigma^*$, either L contains both xz, yz or L doesn't contain any of xz, yz.

3 Content

Consider the language L from the last lecture.

Claim 0.1

L is not regular.

Proof. If L is regular, then $L' = L \cap \mathcal{L}(ab^*c^*)$ is regular. $L' = \{ab^nc^n \mid n \in \mathbb{N} \cup \{0\}\}$ - this is not regular by the pumping lemma. $\square$

Question 1

If L is regular, what is your opponent's strategy?

Answer. $L = \mathcal{L}(D)$ for some DFA $D = (Q, \Sigma, \delta, q_0, A)$. Choose $p = |Q|$. Run D on s which is given in step 2. Find the earliest revisit to some state say q. The first revisit guarantees $|xy| \le p$, the fact that there is a revisit implies $|y| > 0$, and DFA gives that the resulting string is in L.

Now we show a way to characterize the class of regular languages (necessary and sufficient conditions). As a benefit, if L is regular, it gives a systematic way to construct a DFA D for it, and it also gives us the minimum possible number of states among all DFAs that recognize L.

All DFAs that recognize L and have no more states than D are actually isomorphic to D.

Claim 0.2

If $x \sim_D y$ then $\forall z \in \Sigma^*$, we have $xz \sim_D yz$.

Proof. $\hat{\delta}(q_0, xz) = \hat{\delta}(\hat{\delta}(q_0, x), z) = \hat{\delta}(\hat{\delta}(q_0, y), z) = \hat{\delta}(q_0, yz)$. □

Claim 0.3

If $x \sim_D y$ then for all $z \in \Sigma^*$, D either accepts both xz, yz or D rejects xz, yz.

Proof. Follows from the claim above and the fact that the state is either accepting or rejecting. □

Claim 0.4

$\sim_D$ is an equivalence relation.

Proof. Exercise.

We need reflexivity, transitivity and symmetry.

1. Reflexivity: follows from $\hat{\delta}(q_0, x) = \hat{\delta}(q_0, x)$.
2. Symmetry: follows from symmetry of equality.
3. Transitivity: follows from the fact that $\hat{\delta}(q_0, x) = \hat{\delta}(q_0, y) \wedge \hat{\delta}(q_0, y) = \hat{\delta}(q_0, z) \implies \hat{\delta}(q_0, x) = \hat{\delta}(q_0, z)$.

□

Claim 0.5

$=_L$ is an equivalence relation.

Proof. Recall COL202.

1. Reflexivity: follows from the fact that $x = y \implies xz = yz \forall z \in \Sigma^*$.
2. Symmetry: follows from the fact that $xz, yz \in A \equiv yz, xz \in A$.
3. Transitivity: Let $x =_L y$ and $y =_L z$. Consider some $w \in \Sigma^*$. We have two cases here:
   (a) $xw \in L$. In this case, we have $yw \in L$, which implies $zw \in L$, with both implications following from the definition of $=_L$.
   (b) $xw \notin L$. In this case, we have $yw \notin L$, which implies $zw \notin L$, following from the definition of $=_L$ for both implications.

   From here, we see that $xw \in L \iff zw \in L$, so $x =_L z$.

□

Claim 0.6

If $x =_L y$, then $\forall z \in \Sigma^*, xz =_L yz$.

Proof.
Fix z, and consider any string t. Then the string $zt \in \Sigma^*$. Using the fact that $x =_L y$, we have that $x(zt) \in L \iff y(zt) \in L$. Thus, for any $t \in \Sigma^*$, we have $(xz)t \in L \iff (yz)t \in L$, from where the conclusion follows. $\square$

Example 1

Let $\Sigma = \{0, 1\}$, $L = \{x \mid x \text{ is the binary representation of a multiple of } 7\}$

Question 2

How does $=_L$ look like? When is $x =_L y$?

Answer. $x =_L y \iff x \equiv y \pmod 7$.

Proof. Suppose $x \equiv y \pmod 7$. Then $xz = x \cdot 2^{|z|} + z \equiv y \cdot 2^{|z|} + z \equiv yz \pmod 7$. Now suppose $x =_L y$. Suppose $x \not\equiv y \pmod 7$. Then $x \times 2^3 \not\equiv y \times 2^3 \pmod 7$. So $\exists z, |z| \leq 3$ such that $x \times 2^3 + z$ is divisible by 7 but $y \times 2^3 + z$ is not divisible by 7. Considering $x0^{3-|z|}z$ is in L but $y0^{3-|z|}z$ is not in L. So $x \neq_L y$. $\square$

Theorem 1

Let $L \subseteq \Sigma^*$ be a regular language and D be a DFA for L. Then $\forall x, y \in \Sigma^*$, $x \sim_D y \implies x =_L y$. In other words, $\sim_D$ refines $=_L$.

Proof. Follows from the previous claim and definition of $=_L$.

For a bit more detail, consider x, y such that $x \sim_D y$. Then using Claim 0.3, D either accepts both xz, yz or rejects both xz, yz. Hence, we have either both $xz, yz \in L$, or both not in L, from where it follows (by varying z over Σ^*) that $x =_L y$. $\square$

Note 1

This theorem implies that the equivalence classes of $=_L$ are possibly further partitioned by those of $\sim_D$.

Theorem 2

Myhill-Nerode Theorem (John Myhill, Anil Nerode, 1958)

Proof. $\square$