The relation \( x \sim^i_y \equiv (\forall z \exists^* \zeta) \ L(xz) = L(yz) \) is an equivalence relation (for any language \( L \)).

- Reflexive: \( x \sim^i_x \) (immediate)
- Symmetric: \( x \sim^i_y \Leftrightarrow y \sim^i_x \) (immediate "by form")
- Transitive: \( w \sim^i x \land x \sim^i y \Rightarrow w \sim^i y \)

\[\begin{align*}
(\forall z_1)(L(xz_1) = L(yz_1)) & \land (\forall z_2)(L(xz_2) = L(yz_2)) \Rightarrow \\
(\forall z_3)(L(xz_3) = L(yz_3))
\end{align*}\]

Hence, the relation \( \sim^i \) partitions \( \Sigma^* \) into equivalence classes.

**Defn.** A set \( S \) is a system of representatives for an equivalence relation \( \sim \) if it does not contain more than one member from any equiv. class. Complete if it has one.

- Every regular language has a complete system of representatives.

Mylhill–Nerode Theorem: A language \( L \) is regular if and only if \( \sim^i \) partitions \( \Sigma^* \) into finitely many equiv. classes. That is, if all PD sets are finite.

**Proof:** We show \( \Rightarrow \) by contradiction. Now show \( \Leftarrow \)“

Design an automaton \( M = (Q, \Sigma, \delta, s, F) \) where \( Q \) is the set of equivalence classes.

**Note:** if \( x \in L \) and \( x \sim^i y \), then \( y \in L \) (consider \( z = s \)).

This is well-defined because if \( y \) \& any other representative of \( R_x \), then \( R_y = R_x \), so \( S(R_y, c) \) gives the same answer.

Given that \( Q \) is finite, \( M \) is a DFA, so \( L \) is regular.

**Corollary:** If \( k \) \( \in \mathbb{N} \), then no other DFA \( M \) has \( L(M) = L \) can have fewer than \( k \) states, because a complete set of \( k \) representatives is a PD set of size \( k \). Moreover, \( M \) is unique among PDs of \( k \) states. (That is, determined up to

- Every regular language has a unique minimal-size DFA \( M \) s.t. \( L(M) = L \).
Proofs & Nonregularity by MNT: Fill out a "proof script".

Take \( S = \{ z \in \{0, 1\}^* | \exists w \in \{0, 1\}^*, z = 0w1 \} \). (Justify if needed that \( S \) is infinite.)

Let any \( x, y \in S \), \( x \neq y \), be given. Then we can helpfully write \( x = \tilde{x}m \) and \( y = \tilde{y}n \) when non-

Take \( z = D^n \). Then \( L(xz) \neq L(yz) \) because \( xz = \tilde{x}m^n \) gets killed since \( m \) is a multiple of \( n \) but \( \tilde{y}n \) survives. Sino \( x, y \in S \) are arbitrary, \( S \) is PD.

Example: \( L = \) "The spears-and-dragons language"
if you could save any number of stones!

\[ L' = \{ x \in \{0, 1\}^* : \text{there is a suffix w of } x \text{ that makes } xw \text{ balanced} \} \]

If \( (x \in S) \land (x, y \in S, x \neq y, \exists z \in \{0, 1\}^* \land L(xz) \neq L(yz) ) \) then \( L \) is not regular.

---

Which direction, \( \Rightarrow \) or \( \Leftarrow \), did we show in proving that having an infinite PD set makes \( L \) non-regular?

Likewise \( \exists a \in \{0, 1\}^* : \exists a \in \{0, 1\}^* \) is non-regular.

Arbitrary \( S \) is PD \( \Rightarrow \) \( \exists a \in \{0, 1\}^* : \exists a \in \{0, 1\}^* \) is non-regular.

Take \( S = \{ \epsilon \}, \) clearly infinite. Let any \( x, y \in S, x \neq y, \)
be given. Then \( x = \epsilon, y = \epsilon, \) where \( m \neq n \). Note \( z = \epsilon \).

Then \( xz = \epsilon \) is balanced, but \( yz = \epsilon \) is not.

So \( BAL(xz) \neq BAL(yz) \), so \( S \) is an infinite PD set for BAL. \( \Box \)

Note that if \( m > n, \) \( \epsilon^n \) and \( \epsilon^m \) are both in \( L' \).

So we need "wlog \( m \geq n \)" trick to make this work for \( L' \).

In fact, \( L' \) is the same as "spears-and-dragons with \( (-\$) \) = 0."

---

Example: \( L_0 = \{ \text{balanced-paren strings} \} \), called \( BAL \).

Take \( S = \{ \text{strings} \} \), clearly infinite. Let any \( x, y \in S, x \neq y, \)
be given. Then \( x = \text{strings}, y = \text{strings}, \) where \( m \neq n \). Note \( z = \text{strings} \).

Then \( xz = \text{strings} \) is balanced, but \( yz = \text{strings} \) is not.

So \( BAL(xz) \neq BAL(yz) \), so \( S \) is an infinite PD set for BAL. \( \Box \)

Note that if \( m > n, \) \( \text{strings}^n \) and \( \text{strings}^m \) are both in \( L' \).

So we need "wlog \( m \geq n \)" trick to make this work for \( L' \).

In fact, \( L' \) is the same as "spears-and-dragons with \( (-\$) \) = 0."