

2.2: Using Induction to Prove Language Equalities

In this section, we introduce three string induction principles, ways of showing that every $w \in A^*$ has property $P(w)$, where A is some set of symbols.

Typically, A will be an alphabet, i.e., a finite set of symbols. But when we want to prove that all strings have some property, we can let $A = \mathbf{Sym}$, so that $A^* = \mathbf{Str}$.

Each of these principles corresponds to an instance of well-founded induction.

We also look at how different kinds of induction can be used to show that two languages are equal.

Right String Induction

Theorem 2.2.1 (Principle of Right String Induction)

Suppose $A \subseteq \mathbf{Sym}$ and $P(w)$ is a property of a string w .

If

(basis step)

$P(\epsilon)$ and

(inductive step)

for all $a \in A$ and $w \in A^*$, if $(\dagger) P(w)$, then $P(aw)$,

then,

for all $w \in A^*$, $P(w)$.

We refer to the formula (\dagger) as the *inductive hypothesis*.

Left String Induction

Theorem 2.2.2 (Principle of Left String Induction)

Suppose $A \subseteq \mathbf{Sym}$ and $P(w)$ is a property of a string w .

If

(basis step)

$P(\epsilon)$ and

(inductive step)

for all $a \in A$ and $w \in A^*$, if $(\dagger) P(w)$, then $P(wa)$,

then,

for all $w \in A^*$, $P(w)$.

We refer to the formula (\dagger) as the *inductive hypothesis*.

Strong String Induction

Theorem 2.2.3 (Principle of Strong String Induction)

Suppose $A \subseteq \mathbf{Sym}$ and $P(w)$ is a property of a string w .

If

for all $w \in A^*$,

if (\dagger) for all $x \in A^*$, if x is a proper substring of w , then $P(x)$,
then $P(w)$,

then,

for all $w \in A^*$, $P(w)$.

We refer to the formula (\dagger) as the *inductive hypothesis*.

Example: Difference Function

Define $\mathbf{diff} \in \{0, 1\}^* \rightarrow \mathbb{Z}$ by: for all $w \in \{0, 1\}^*$,

$\mathbf{diff} w =$ the number of 1's in w $-$ the number of 0's in w .

Then:

- $\mathbf{diff} \epsilon = 0$;
- $\mathbf{diff} 1 = 1$;
- $\mathbf{diff} 0 = -1$; and
- for all $x, y \in \{0, 1\}^*$, $\mathbf{diff}(xy) = \mathbf{diff} x + \mathbf{diff} y$.

Note that, for all $w \in \{0, 1\}^*$, $\mathbf{diff} w = 0$ iff w has an equal number of 0's and 1's.

Definition of Two Languages

Let X be the least subset of $\{0,1\}^*$ such that:

- (1) $\% \in X$;
- (2) for all $x, y \in X$, $xy \in X$;
- (3) for all $x \in X$, $0x1 \in X$; and
- (4) for all $x \in X$, $1x0 \in X$.

This is an inductive definition.

Let $Y = \{w \in \{0,1\}^* \mid \mathbf{diff} w = 0\}$.

Our goal is to prove that $X = Y$, i.e., that: (the easy direction) every string that can be constructed using X 's rules has an equal number of 0's and 1's; and (the hard direction) that every string of 0's and 1's with an equal number of 0's and 1's can be constructed using X 's rules.

Principle of Induction on X

Proposition Slides-2.2.1 (Principle of Induction on X)

Suppose $P(w)$ is a property of a string w . If

(1)

$$P(\epsilon),$$

(2)

for all $x, y \in X$, if $(\dagger) P(x)$ and $P(y)$, then $P(xy)$,

(3)

for all $x \in X$, if $(\dagger) P(x)$, then $P(0x1)$,

(4)

for all $x \in X$, if $(\dagger) P(x)$, then $P(1x0)$,

then,

for all $w \in X$, $P(w)$.

We refer to (\dagger) as the *inductive hypothesis*.

Easy Direction

Lemma 2.2.11

$X \subseteq Y$.

Proof. We use induction on X to show that, for all $w \in X$, $w \in Y$. There are four steps to show.

- (1) We must show $\% \in Y$. Since $\% \in \{0, 1\}^*$ and $\mathbf{diff} \% = 0$, we have that $\% \in Y$.
- (2) Suppose $x, y \in X$, and assume the inductive hypothesis: $x, y \in Y$. We must show that $xy \in Y$. Since $x, y \in Y$, we have that $xy \in \{0, 1\}^*$ and $\mathbf{diff}(xy) = \mathbf{diff} x + \mathbf{diff} y = 0 + 0 = 0$. Thus $xy \in Y$.

Easy Direction (Cont.)

Proof (cont.).

- (3) Suppose $x \in X$, and assume the inductive hypothesis: $x \in Y$. We must show that $0x1 \in Y$. Since $x \in Y$, we have that $0x1 \in \{0, 1\}^*$ and $\mathbf{diff}(0x1) = \mathbf{diff} 0 + \mathbf{diff} x + \mathbf{diff} 1 = -1 + 0 + 1 = 0$. Thus $0x1 \in Y$.
- (4) Suppose $x \in X$, and assume the inductive hypothesis: $x \in Y$. We must show that $1x0 \in Y$. Since $x \in Y$, we have that $1x0 \in \{0, 1\}^*$ and $\mathbf{diff}(1x0) = \mathbf{diff} 1 + \mathbf{diff} x + \mathbf{diff} 0 = 1 + 0 + -1 = 0$. Thus $1x0 \in Y$.

□

Hard Direction

Lemma 2.2.12

$Y \subseteq X$.

Proof. Since $Y \subseteq \{0,1\}^*$, it will suffice to show that, for all $w \in \{0,1\}^*$,

if $w \in Y$, then $w \in X$.

We proceed by strong string induction. Suppose $w \in \{0,1\}^*$, and assume the inductive hypothesis: for all $x \in \{0,1\}^*$, if x is a proper substring of w , then

if $x \in Y$, then $x \in X$.

We must show that

if $w \in Y$, then $w \in X$.

Suppose $w \in Y$. We must show that $w \in X$. There are three cases to consider.

Hard Direction (Cont.)

Proof (cont.).

- Suppose $w = \%$. Then $w = \% \in X$, by Part (1) of the definition of X .

Hard Direction (Cont.)

Proof (cont.).

- Suppose $w = 0t$ for some $t \in \{0, 1\}^*$. Since $w \in Y$, we have that $-1 + \text{diff } t = \text{diff } 0 + \text{diff } t = \text{diff}(0t) = \text{diff } w = 0$, and thus that $\text{diff } t = 1$.

Let u be the shortest prefix of t such that $\text{diff } u \geq 1$. (Since u is a prefix of itself and $\text{diff } t = 1 \geq 1$, it follows that u is well-defined.) Let $z \in \{0, 1\}^*$ be such that $t = uz$. Clearly, $u \neq \epsilon$, and thus $u = yb$ for some $y \in \{0, 1\}^*$ and $b \in \{0, 1\}$. Hence $t = uz = ybz$. Since y is a shorter prefix of t than u , we have that $\text{diff } y \leq 0$.

Suppose, toward a contradiction, that $b = 0$. Then $\text{diff } y + -1 = \text{diff } y + \text{diff } 0 = \text{diff } y + \text{diff } b = \text{diff}(yb) = \text{diff } u \geq 1$, so that $\text{diff } y \geq 2$. But $\text{diff } y \leq 0$ —contradiction. Hence $b = 1$.

Hard Direction (Cont.)

Proof (cont.).

- (Continuation of second case.)

Summarizing, we have that $u = yb = y1$, $t = uz = y1z$ and $w = 0t = 0y1z$. Since

$\text{diff } y + 1 = \text{diff } y + \text{diff } 1 = \text{diff}(y1) = \text{diff } u \geq 1$, it follows that $\text{diff } y \geq 0$. But $\text{diff } y \leq 0$, and thus $\text{diff } y = 0$. Thus

$y \in Y$. Since $1 + \text{diff } z = 0 + 1 + \text{diff } z =$

$\text{diff } y + \text{diff } 1 + \text{diff } z = \text{diff}(y1z) = \text{diff } t = 1$, it follows that $\text{diff } z = 0$. Thus $z \in Y$.

Because y and z are proper substrings of w , and $y, z \in Y$, the inductive hypothesis tells us that $y, z \in X$. Thus, by Part (3) of the definition of X , we have that $0y1 \in X$. Hence, Part (2) of the definition of X tells us that $w = 0y1z = (0y1)z \in X$.

- Suppose $w = 1t$ for some $t \in \{0, 1\}^*$. This is symmetric to the preceding case.



Language Equality

Proposition 2.2.13

$$X = Y.$$

Proof. Follows immediately from Lemmas 2.2.11 and 2.2.12. \square