Lecture 22: Bijections

Sunday, November 13, 2016  8:03 PM

In the previous lecture we proved two lemmas:

**Lemma 1.** Suppose \( X \rightarrow \rightarrow Y \) is a function.

- \( f \) is injective \( \iff \) \( f \) has a left inverse.

**Lemma 2.** Suppose \( X \rightarrow \rightarrow Y \) is a function.

- \( f \) is surjective \( \iff \) \( f \) has a right inverse.

Using these Lemmas we prove

**Theorem.** Suppose \( X \rightarrow \rightarrow Y \) is a function.

- \( f \) is bijective \( \iff \) \( f \) is invertible. (Lemma 1)

**Proof.** \( f \) is bijective \( \iff \) \( f \) is injective \( \iff \) \( f \) has a left inverse \( \iff \) \( f \) is surjective \( \iff \) \( f \) has a right inverse. (Lemma 2)

\( f \) has a left inverse \( \iff \) \( f \) is invertible.

\( f \) has a right inverse.

**Lemma.** If \( g \) is a left inverse of \( f : X \rightarrow Y \) and \( h \) is a right inverse of \( f \), then \( g = h \).

**Proof.** Consider \( g \circ f \circ h \). We have \( g \circ f \circ h = g \circ (f \circ h) = g \circ I_Y = g \)

and \( g \circ f \circ h = (g \circ f) \circ h = I_X \circ h = h \). So \( g = h \). ■

**Theorem.** Suppose \( X \rightarrow \rightarrow Y \) is a function.

- \( f \) is a bijection \( \iff \) there is a unique \( g : Y \rightarrow X \),

\[ g \circ f = I_X \quad \text{and} \quad f \circ g = I_Y \]

(Such \( Y \rightarrow \rightarrow X \) is called the inverse of \( f \) and it is denoted by \( f^{-1} \).)
Proof. (\(\rightarrow\)) we have to prove two things ① existence of such function ② uniqueness of such function.

① Existence. Since \(f\) is a bijection, by the previous theorem, \(f\) is invertible. So \(f\) has a left inverse \(g\) and a right inverse \(h\). By the above lemma, \(h=g\). So \(g \circ f = I_X\) and \(f \circ g = I_Y\).

② Uniqueness. Suppose both \(g_1, g_2 : Y \to X\) satisfy the above conditions. So \(g_1\) is a left inverse of \(f\) and \(g_2\) is a right inverse of \(f\). Hence, by the above lemma, \(g_1 = g_2\), which shows the uniqueness of such function.

Theorem (a) If \(f\) is a bijection, then \(f\) is the inverse of \(f\).
And so \(f^{-1}\) is a bijection.

(b) If \(X \xrightarrow{f} Y\) and \(Y \xrightarrow{g} Z\) are two bijections, then \(g \circ f: X \to Z\) is a bijection. Moreover,
\[
(g \circ f)^{-1} = f^{-1} \circ g^{-1}.
\]

Proof. (a) By the definition of the inverse function \(f^{-1}\), we have \(f^{-1} \circ f = I_X\) and \(f \circ f^{-1} = I_Y\). Hence \((f^{-1})^{-1} = f\).
And so by the previous theorem \(f^{-1}\) is a bijection.

(b) We show that \((g \circ f) \circ (f^{-1} \circ g^{-1}) = I_Z\)
and \((f^{-1} \circ g^{-1} \circ (g \circ f) = I_X\).
This implies that \(g \circ f\) has an inverse and
\[
(g \circ f)^{-1} = f^{-1} \circ g^{-1}.
\]
Now, by the previous theorem, we can deduce that 
\( g \circ f \) is a bijection.

\[
(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f^{-1} \circ g^{-1}) = g \circ g^{-1} = \mathbb{I}_Z
\]

\[
(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g \circ f) = f^{-1} \circ f = \mathbb{I}_X.
\]

**Definition.** Two sets \( A \) and \( B \) are called **equipotent sets**, and we write \( A \sim B \) if there is a bijection \( f : A \rightarrow B \).

**Lemma.** For any non-empty sets \( A, B, \) and \( C \), we have

1. \( A \sim A \).
2. \( A \sim B \implies B \sim A \).
3. \( A \sim B \implies A \sim C \).
   \[B \sim C \]

**Proof.**

1. \( I_A : A \rightarrow A \) is a bijection.
2. If \( f : A \rightarrow B \) is a bijection, then \( B \rightarrow f \) \( f^{-1} : A \) is a bijection.
3. If \( A \sim B \) and \( B \sim C \) are bijections, then \( A \sim C \) is a bijection.

Based on our intuition of cardinality of finite sets we have:

**Theorem.** Suppose \( A \) and \( B \) are two non-empty finite sets.

Then \( A \sim B \iff |A| = |B| \).

In fact a bit stronger results are true:
**Theorem.** Suppose $X$ and $Y$ are non-empty finite sets and $\mathcal{X} \xrightarrow{f} \mathcal{Y}$ is a function. Then

$$f \text{ is injective } \Rightarrow |X| \leq |Y|.$$ 

The contra-positive form of the above theorem is called the **pigeonhole principle**.

$$|X| > |Y| \Rightarrow \exists x_1, x_2 \in X, x_1 \neq x_2 \land f(x_1) = f(x_2).$$

Alternatively: If there are $n$ pigeons, $m$ pigeonholes and $n > m$, then at least two pigeons share a pigeonhole.

Later we will see some of its applications.