Lecture 1: Set Theory

Tianpei Xie

Nov. 7th., 2022

Contents

1 Fundamental Concepts

1.1 Set Operations and Logics

- **Definition** Given a set X, the collection of all subsets of X, denoted as 2^X, is defined as

$$2^X := \{E : E \subseteq X\}$$

- **Remark** The followings are basic operation on 2^X: For $A, B \in 2^X$,
  1. ***Inclusion***: $A \subseteq B$ if and only if $\forall x \in A$, $x \in B$.
  2. ***Union***: $A \cup B = \{x : x \in A \vee x \in B\}$.
  3. ***Intersection***: $A \cap B = \{x : x \in A \wedge x \in B\}$.
  4. ***Difference***: $A \setminus B = \{x : x \in A \wedge x \notin B\}$.
  5. ***Complement***: $A^c = X \setminus A = \{x : x \in X \wedge x \notin A\}$.
  6. ***Symmetric Difference***: $A \Delta B = (A \setminus B) \cup (B \setminus A) = \{x \in X : x \notin A \vee x \notin B\}$.

  We have ***deMorgan's laws***:

$$\left(\bigcup_{a \in A} U_a\right)^c = \bigcap_{a \in A} U_a^c, \quad \left(\bigcap_{a \in A} U_a\right)^c = \bigcup_{a \in A} U_a^c$$

- **Remark** Note that the following equality is useful:

$$A \Delta B = (A \cup B) \setminus (A \cap B)$$

- The forms of logic statement using "if ... then":
  1. Orignal statement: "*If P then Q*", or "*Q holds **if** P holds*";

$$P \Rightarrow Q$$

  2. <u>***Converse statement***</u>: "*If Q then P*", or "*Q holds **only if** P holds*";

$$Q \Rightarrow P$$

  3. <u>***Contrapositive statement***</u>: "*If not Q then not P*", or "*P not holds **if** Q not holds*";

$$\neg Q \Rightarrow \neg P$$

     The contrapositive and the original statements are ***logically equivalent***.

  If it should happen that both the statement $P \Rightarrow Q$ and its converse $Q \Rightarrow P$ are *true*, we express this fact by the notation

$$P \Leftrightarrow Q$$

  "*P holds **if and only if** Q holds*""

2 Functions

- **Definition** A ***rule of assignment*** is a subset r of the cartesian product $C \times D$ of two sets, having the property that each element of C appears as the first coordinate ***at most one ordered pair belonging to*** r. Thus, a subset r of $C \times D$ is *a rule of assignment* if

$$[(c,d) \in r \text{ and } (c,d') \in r] \Rightarrow [d = d'].$$

Given a rule of assignment r, ***the domain*** of r is defined to be the *subset* of C consisting of *all first coordinates of elements* of r, and ***the image*** set of r is defined as the *subset* of D consisting of *all second coordinates of elements* of r.

A ***function*** f is *a rule of assignment* r, together with a set B that *contains the image set* of r.

- **Definition** $f : X \to Y$ is a ***function*** if for each $x \in X$, there exists a unique $y = f(x) \in Y$. X is called the ***domain*** of f and Y is called the ***codomain or image*** of f. $f(X) = \{y \in Y : y = f(x)\}$ is called the ***range*** of f

The ***pre-image*** of f is defined as

$$f^{-1}(E) = \{x \in X : f(x) \in E\}.$$

- **Definition** If $f : A \to B$ and if A_0 is a subset of A, we define the ***restriction*** of f to A_0 to be the function mapping A_0 into B whose rule is

$$\{(a, f(a)) : \; a \in A_0\}.$$

It is denoted by $f|_{A_0}$, which is read ""*f restricted to A_0*."

- **Remark** The pre-image operation ***commutes*** with ***all basic set operations***:

$$\begin{aligned} A \subseteq B &\Rightarrow f^{-1}(A) \subseteq f^{-1}(B) \\ f^{-1}\left(\bigcup_{\alpha \in A} E_\alpha\right) &= \bigcup_{\alpha \in A} f^{-1}(E_\alpha) \\ f^{-1}\left(\bigcap_{\alpha \in A} E_\alpha\right) &= \bigcap_{\alpha \in A} f^{-1}(E_\alpha) \\ f^{-1}(A \setminus B) &= f^{-1}(A) \setminus f^{-1}(B) \\ f^{-1}(E^c) &= \left(f^{-1}(E)\right)^c \end{aligned}$$

- **Remark** The image operation ***commutes*** with only ***inclusion*** *and* ***union*** *operations*:

$$\begin{aligned} A \subseteq B &\Rightarrow f(A) \subseteq f(B) \\ f\left(\bigcup_{\alpha \in A} E_\alpha\right) &= \bigcup_{\alpha \in A} f(E_\alpha) \end{aligned}$$

For the other operations:

$$f\left(\bigcap_{\alpha\in A} E_\alpha\right) \subseteq \bigcap_{\alpha\in A} f\left(E_\alpha\right)$$
$$f\left(A\setminus B\right) \supseteq f(A)\setminus f(B)$$

- **Definition** A map $f : X \to Y$ is ***surjective, or, onto***, if for every $y \in Y$, there exists a $x \in X$ such that $y = f(x)$. In set theory notation:

  $$f : X \to Y \text{ is surjective } \Leftrightarrow f^{-1}(Y) \subseteq X.$$

  A map $f : X \to Y$ is ***injective, or one-to-one***, if for every $x_1 \neq x_2 \in X$, their map $f(x_1) \neq f(x_2)$, or equivalently, $f(x_1) = f(x_2)$ only if $x_1 = x_2$.

  If a map $f : X \to Y$ is both *surjective* and *injective*, we say f is a ***bijective***, or there exists an ***one-to-one correspondence*** between X and Y. Thus $Y = f(X)$.

- **Remark**

  $$\begin{aligned} f^{-1}(f(B)) &\supseteq B, \quad \forall B \subseteq X \\ f(f^{-1}(E)) &\subseteq E, \quad \forall E \subseteq Y \\ f : X \to Y \text{ is surjective } &\Leftrightarrow f^{-1}(Y) \subseteq X. \\ &\Rightarrow f(f^{-1}(E)) = E. \\ f : X \to Y \text{ is injective } &\Rightarrow f^{-1}(f(B)) = B \\ &\Rightarrow f\left(\bigcap_{\alpha\in A} E_\alpha\right) = \bigcap_{\alpha\in A} f\left(E_\alpha\right) \\ &\Rightarrow f\left(A\setminus B\right) = f(A)\setminus f(B) \end{aligned}$$

- **Proposition 2.1** *The following statements for composite functions are true:*

  1. *If f, g are both injective, then $g \circ f$ is injective.*
  2. *If f, g are both surjective, then $g \circ f$ is surjective.*
  3. *Every* ***injective*** *map $f : X \to Y$ can be writen as $f = \iota \circ f_R$ where $f_R : X \to f(X)$ is a* ***bijective*** *map and ι is the* ***inclusion map****.*
  4. *Every* ***surjective*** *map $f : X \to Y$ can be writen as $f = f_p \circ \pi$ where $\pi : X \to (X/\sim)$ is* ***a quotient map*** *(projection $x \mapsto [x]$) for the equivalent relation $x \sim y \Leftrightarrow f(x) = f(y)$ and $f_p : (X/\sim) \to Y$ is defined as $f_p([x]) = f(x)$* ***constant*** *in each coset $[x]$.*
  5. *If $g \circ f$ is* ***injective****, then f is* ***injective****.*
  6. *If $g \circ f$ is* ***surjective****, then g is* ***surjective****.*

3 Relations

- **Definition** A ***relation*** on a set A is a *subset* R of *the cartesian product* $A \times A$.

If R is a relation on A, we use the notation xRy to mean the same thing as $(x, y) \in R$. We read it "x is in the relation R to y."

- **Remark** ***A rule of assignment*** r for a function $f : A \to A$ is also a *subset* of $A \times A$. But it is a subset of a *very* ***special*** *kind*: namely, one such that ***each element*** of A appears as the ***first coordinate*** of an element of r ***exactly once***. ***Any subset*** *of $A \times A$ is a relation on A.*

3.1 Equivalence Relation

- **Definition** ***An equivalence relation*** on X is a relation R on X such that

  1. (***Reflexivity***): xRx for all $x \in X$;
  2. (***Symmetry***): xRy if and only if yRx for all $x, y \in X$;
  3. (***Transitivity***): xRy and yRz then xRz for all $x, y, z \in X$.

  We usually denote the equivalence relation R as $\sim$.

- **Definition** (***Equivalence Class***)
  The equivalence class of an element x is denoted as $[x] := \{y \in X : xRy\}$.

- **Lemma 3.1** *[Munkres, 2000]*
  Two equivalence classes E and E' are either ***disjoint*** *or* ***equal.***

- **Definition** A ***partition*** of a set A is a collection of ***disjoint*** nonempty subsets of A whose ***union*** is all of A.

- **Remark** The set of equivalence classes provides ***a partition of the set*** X in that every $z \in X$ can must belong to *only one equivalence class* $[x]$. That is $[x] \cap [y] = \emptyset$ if $x \not\sim y$ and $X = \bigcup_{x \in X} [x]$.

- **Definition** *The set of all equivalence classes* of X by $\sim$, denoted $X/\sim := \{[x] : x \in X\}$, is ***the quotient set*** of X by $\sim$. $X = \bigcup_{C \in X/\sim} C$.

- **Remark** Since $x \sim y \Rightarrow y \in [x]$, we see that if $[x] \neq [y]$, then $x \not\sim y$, i.e. representative of different equivalence classes are not in the given relationship.

3.2 Order Relation

- **Definition** A relation C on a set A is called ***an order relation*** (or ***a simple order***, or ***a linear order***) if it has the following properties:

  1. (***Comparability***) For every x and y in A for which $x \neq y$, either xCy or yCx.
  2. (***Nonreflexivity***) For no x in A does the relation xCx hold.
  3. (***Transitivity***) If xCy and yCz, then xCz.

  We denote order relation as $>$ or $<$. We shall use the notation $x \leq y$ to stand for the statement "either $x < y$ or $x = y$"; and we shall use the notation $y > x$ to stand for the statement "$x < y$." We write $x < y < z$ to mean "$x < y$ and $y < z$"

- **Remark** If $x \neq y$, then $x < y$ and $y < x$ cannot hold simultaneously.

- **Definition** (***Order Type***)
  Suppose that A and B are two sets with order relations $<_A$, and $<_B$ respectively. We say that A and B have *the same* ***order type*** if there is a ***bijective*** correspondence between them that ***preserves order***; that is, if there exists a bijective function $f : A \to B$ such that

$$x <_A y \Rightarrow f(x) <_B f(y)$$

- **Definition** (***Dictionary Order Relation***)
  Suppose that A and B are two sets with order relations $\prec_A$ and $\prec_B$ respectively. Define an order relation $\prec$ on $A \times B$ by defining

$$a_1 \times b_1 < a_2 \times b_2$$

  if $a_1 <_A a_2$, ***or*** if $a_1 = a_2$ and $b_1 <_B b_2$. It is called ***the dictionary order relation*** on $A \times B$.

- **Definition** Suppose that A is a set ordered by the relation $<$. Let A_0 be a subset of A. We say that the element b is *the* ***largest element*** *of* A_0 if $b \in A_0$ and $x \leq b$ for every $x \in A_0$.

  Similarly, we say that a is *the* ***smallest element*** of A_0 if $a \in A_0$ and if $a \leq x$ for every $x \in A_0$.

- **Remark** It is easy to see that a set has ***at most one*** largest element and *at most one* smallest element.

- **Definition** (***The Upper Bound and The Supremum of Subset***)
  We say that *the subset* A_0 *of* A *is* ***bounded above*** if there is *an element* b *of* A such that $x \leq b$ for every $x \in A_0$; the element $b \in A$ is called ***an upper bound for*** A_0.

  If *the set of all upper bounds* for A_0 has *a* ***smallest element***, that element is called ***the least upper bound***, or ***the supremum***, of A_0. It is denoted by $\sup A_0$, it may or may not belong to A_0. If it *does*, it is ***the largest element*** of A_0.

- **Definition** (***The Lower Bound and The Infimum of Subset***)
  Similarlly, we say that *the subset* A_0 *of* A *is* ***bounded below*** if there is *an element* a *of* A such that $a \leq x$ for every $x \in A_0$; the element $a \in A$ is called ***a lower bound for*** A_0.

  If *the set of all lower bounds* for A_0 has *a* ***largest element***, that element is called ***the greatest lower bound***, or ***the infimum***, of A_0. It is denoted by $\inf A_0$, it may or may not belong to A_0. If it *does*, it is ***the smallest element*** of A_0.

- **Definition** (***The Least Upper Bound Property*** *and* ***The Greatest Lower Bound Property***)
  An ordered set A is said to have ***the least upper bound property*** if *every nonempty* subset A_0 of A that is *bounded above* has *a least upper bound.*

  Analogously, the set A is said to have ***the greatest lower bound property*** if *every nonempty* subset A_0 of A that is *bounded below* has *a greatest lower bound.*

- **Theorem 3.2** *(**Zorn's Lemma**). [Munkres, 2000]*
  Let A be a set that is ***strictly partially ordered****. If every* ***simply ordered subset*** *of A has an* ***upper bound in*** *A, then A has a* ***maximal element****.*

4 Cartesian Products

- **Definition** (***Indexed Family of Sets***)
  Let $\mathcal{A}$ be a nonempty collection of sets. ***An indexing function*** for $\mathcal{A}$ is a ***surjective*** function f from some set J, called ***the index set***, to $\mathcal{A}$. The *collection* $\mathcal{A}$, together with *the indexing function* f, is called ***an indexed family of sets***. Given $\alpha \in J$, we shall denote the set $f(\alpha)$ by the symbol A_α. And we shall denote the indexed family itself by the symbol

  $$\{A_\alpha\}_{\alpha \in J},$$

  which is read "*the family of all A_α, as a ranges over J.*" Sometimes we write merely $\{A_\alpha\}$, if it is clear what the index set is.

- **Definition** (***Cartesian Product of Indexed Family of Sets***)
  Let m be a positive integer. Given a set X, we define an ***m-tuple of elements*** of X to be a function

  $$x : \{1, \ldots, m\} \to X.$$

  If X is an m-tuple, we often denote the value of x at i by *the symbol* x_i; rather than $x(i)$ and call it ***the i-th coordinate of x***. And we often denote the function x itself by the symbol

  $$(x_1, \ldots, x_m).$$

  Now let $\{A_1, \ldots, A_m\}$ be *a family of sets indexed with the set* $\{1, \ldots, m\}$. Let $X = A_1 \cup \ldots \cup A_m$. We define ***the cartesian product*** *of this indexed family*, denoted by

  $$\prod_{i=1}^{m} A_i \quad \text{or} \quad A_1 \times \ldots \times A_m$$

  to be *the set of all m-tuples* $(x_1, \ldots, x_m)$ of elements of X such that $x_i \in A_i$ for each i.

- **Definition** (***Countable Cartesian Product of Indexed Family of Sets***)
  Given a set X, we define an ***ω-tuple of elements*** of X to be a function

  $$x : \mathbb{Z}_+ \to X;$$

  we also call such a function *a* ***sequence***, or *an* ***infinite sequence***, of elements of X. If x is *an ω-**tuple***, we often denote the value of x at i by x_i rather than $x(i)$, and call it ***the i-th coordinate*** of x. We denote x itself by the symbol

  $$(x_1, x_2, \ldots) \quad \text{or} \quad (x_n)_{n \in \mathbb{Z}_+}$$

  Now let $\{A_1, A_2, \ldots\}$ be a family of sets, *indexed with the positive integers*; let X be the union of the sets in this family. ***The cartesian product*** *of this indexed family of sets*, denoted by

  $$\prod_{i \in \mathbb{Z}_+} A_i \quad \text{or} \quad A_1 \times A_2 \times \ldots,$$

  is defined to be the set of all ω-tuples $(x_1, x_2, \ldots)$ of elements of X such that $x_i \in A_i$ for each i.

5 Countable and Uncountable Sets

- **Definition** See the following definitions
  1. A set is said to be ***countably infinite*** if it admits a ***bijection*** with the set of *positive integers* $f : A \to \mathbb{Z}_+$, and
  2. A set is said to be ***countable*** if it is *finite* or *countably infinite.*
  3. A set that is not countable is said to be ***uncountable***.
- **Proposition 5.1** *Let B be a nonempty set. Then the following are equivalent:*
  1. *B is **countable**.*
  2. *There is a **surjective** function $f : \mathbb{Z}_+ \to B$.*
  3. *There is an **injective** function $g : B \to \mathbb{Z}_+$.*
- **Lemma 5.2** *If C is an infinite subset of $\mathbb{Z}_+$, then C is countably infinite.*

6 The Principle of Recursive Definition

- **Principle 6.1** *(**Principle of Recursive Definition**). [Munkres, 2000]*
  *Let A be a set. Given a **formula** that defines $h(1)$ as a **unique** element of A, and for $i > 1$ defines $h(i)$ **uniquely** as an element of A in terms of the values of h **for positive integers less than** i, this formula determines **a unique function** $h : \mathbb{Z}_+ \to A$.*
- **Theorem 6.2** *(**Principle of Recursive Definition**). [Munkres, 2000]*
  *Let A be a set; let a_0 be an element of A. Suppose ρ is a function that assigns, to **each function** f mapping a **nonempty section** of **the positive integers** into A, an eiement of A. Then there exists a **unique function***

  $$h : \mathbb{Z}_+ \to A$$

  such that

  $$\begin{aligned} h(1) &= a_0, \\ h(i) &= \rho(h|\{1, \ldots, (i-1)\}) \text{ for all } i > 1. \end{aligned} \tag{1}$$

  *The formula (1) is called a **recursion formula** for h. It specifies $h(1)$, and it expresses the value of h at $i > 1$ in terms of the values of h for positive integers less than i. A definition given by such a formula is called a **recursive definition**.*
- **Corollary 6.3** *A subset of a countable set is countable.*
- **Corollary 6.4** *The set $\mathbb{Z}_+ \times \mathbb{Z}_+$ is countably infinite.*
- **Proposition 6.5** *A countable union of countable sets is countable.*
- **Proposition 6.6** *A finite product of countable sets is countable.*
- It is very tempting to assert that *countable products of countable sets should be countable*; but this assertion is in fact ***not true***:

Theorem 6.7 *Let X denote the two element set $\{0,1\}$. Then the set X^{ω} is uncountable.*

- **Theorem 6.8** *Let A be a set. There is* ***no injective map*** *$f : 2^{A} \to A$, and there is* ***no surjective map*** *$g : A \to 2^{A}$.*

- **Proposition 6.9** *Let A be a set. The following statements about A are equivalent:*

  1. *There exists an* ***injective*** *function $f : \mathbb{Z}_{+} \to A$.*
  2. *There exists a* ***bijection*** *of A with a proper subset of itself.*
  3. *A is infinite.*

7 The Axiom of Choice

- **Principle 7.1** ***(Axiom of Choice)****. [Munkres, 2000]*
  Given a collection $\mathscr{A}$ of ***disjoint*** *nonempty sets, there exists a set C consisting of* ***exactly one element from each element of $\mathscr{A}$****; that is, a set C such that C is contained in the* ***union of the elements*** *of $\mathcal{A}$, and for each $A \in \mathscr{A}$, the set $C \cap A$ contains* ***a single element****.*

- **Lemma 7.2** ***(Existence of a Choice Function)****. [Munkres, 2000]*
  Given a collection $\mathscr{B}$ of nonempty sets (not necessarily disjoint), there exists a function

  $$c : \mathscr{B} \to \bigcup_{B \in \mathscr{B}} B$$

  such that $c(B)$ is an element of B, for each $B \in \mathscr{B}$.

  Remark The function c is called ***a choice function*** for the collection $\mathscr{B}$. The difference between this lemma and the axiom of choice is that in this lemma the sets of the collection $\mathscr{B}$ are not required to be disjoint.

- **Remark** The axiom of choice is used when someone construct *an infinite set* using *infinite number of arbitrary choices.*

- **Corollary 7.3** *If $\{A_{\alpha}\}_{\alpha \in J}$ is a* ***disjoint*** *collection of nonempty sets, there is a set $C \subset \bigcup_{\alpha \in J} A_{\alpha}$ such that $C \cap A_{\alpha}$ contains* ***precisely one element*** *for each $\alpha \in J$.*

8 Well-Ordering Theorem and the Maximum Principle

- **Definition** (***Well-Ordered Set***)
  A set A with an order relation $<$ is said to be ***well-ordered*** if *every nonempty subset* of A has a ***smallest element***.

- **Proposition 8.1** ***(Finite Ordered Set is Well-Ordered)*** *[Munkres, 2000]*
  Every nonempty ***finite*** *ordered set has the order type of a section $\{1, \ldots, n\}$ of $\mathbb{Z}_{+}$, so it is* ***well-ordered****.*

- **Theorem 8.2** ***(Well-Ordering Theorem)****. [Munkres, 2000]*
  If A is a set, there ***exists*** *an order relation on A that is a well-ordering.*

- **Remark** The proof of *Well-Ordering Theorem* is based on a construction involving *an infinite number of arbitrary choices*, that is, a construction involving *the choice axiom*.

- **Corollary 8.3** *There exists an uncountable well-ordered set.*

- **Definition** Let X be a well-ordered set. Given $\alpha \in X$, let S_α denote the set

$$S_\alpha = \{x : x \in X \text{ and } x < \alpha\}.$$

  It is called *the* ***section*** *of X by α*.

- **Definition** (***Strict Partial Order***)
  Given a set A, a relation $\prec$ on A is called a ***strict partial order*** on A if it has the following two properties;

  1. (***Nonreflexivity***) The relation $a \prec a$ never holds.
  2. (***Transitivity***) If $a \prec b$ and $b \prec c$, then $a \prec c$.

  Moreover, suppose that we define $a \preceq b$ either $a \prec b$ or $a = b$. Then the relation $\preceq$ is called ***a partial order*** on A.

- **Remark** *The Comparability condition* means ***every two elements are comparable under simple order***. Without this condition, we have partial order $x \prec y$. Consider *the simple ordering* as along ***a chain graph***, while *the partial ordering* is along ***the general graphs***.

- **Theorem 8.4** ***(The Maximum Principle).***
  Let A be a set; let $\prec$ be a ***strict partial order*** *on A. Then there exists a* ***maximal simply ordered subset*** *B of A.*

- **Definition** (***Upper Bound and Maximal Element for Strict Partial Order***)
  Let A be a set and let $\prec$ be a *strict partial order* on A. If B is a subset of A, ***an upper bound*** on B is an element c of A such that for every b in B, either $b = c$ or $b \prec c$.

  A maximal element of A is an element m of A such that for ***no element*** *a of A does the relation $m \prec a$ hold.*

- **Remark** ***The upper bound*** of a set A is *not necessarily in A*, but ***the maximal element*** of A is in A.

- **Theorem 8.5** ***(Zorn's Lemma).*** *[Munkres, 2000]*
  Let A be a set that is ***strictly partially ordered****. If every* ***simply ordered subset*** *of A has an* ***upper bound in*** *A, then A has a* ***maximal element****.*

- **Remark** Note that the inclusion operation $\subset$ defines an order relationship between two sets. One application of Zorn's lemma is on the collection of subsets $\mathscr{A} = \{A_n\}_{n\in J}$ that is partially ordered by $\subset$ operation. For each simply ordered sub-collection $\mathscr{A}_I := \{A_n\}_{n\in I}$, $I \subseteq J$, where $A_i \subset A_{i+1}$ we can see that $A_{\max I}$ is the uppper bound of $\mathscr{A}_I$ in $\mathscr{A}$. Thus there exists ***a maximal subset*** $A_{max} \in \mathscr{A}$ so that $A_n \subset A_{max}$ for all $n \in J$.

References

James R Munkres. *Topology, 2nd*. Prentice Hall, 2000.