18.S097 Introduction to Proofs
IAP 2015
Lecture Notes 1 (1/5/2015)

1. Introduction

The goal for this course is to provide a quick, and hopefully somewhat gentle, introduction to the task of formulating and writing mathematical proofs. We begin by discussing some basic ideas of logic and sets which form the basic ingredients in our mathematical language, and conclude our discussion for the day with a few examples.

2. Propositions and sets

2.1. **Elements of logic and overview of proof techniques.** To formulate some basic language for our discussion, we begin with a light treatment of some ideas of *propositional logic*, which describes how we can manipulate mathematical statements. Suppose that p and q are two mathematical statements, which can each be true or false. We can form new statements out of p and q by several logical operations:

- The statement "p and q" (written as $p \wedge q$) holds whenever both p and q are true.
- The statement "p or q" ($p \vee q$) holds whenever at least one of p, q are true (inclusive or).
- The statement "not(p)" holds whenever p is not true.
- The statement "p implies q" ($p \Rightarrow q$) holds whenever "not(p) or q" is true.

It is clear that the operation of "and" and "or" are *commutative* – the order of p and q does not matter. We immediately see that the implication "p implies q" is equivalent to the *contrapositive*

$$\text{"not}(q) \text{ implies not}(p).''$$

On the other hand, the implication "p implies q" is **not** equivalent to "q implies p" (this second statement is known as the *converse* to the first).

The negation operation combines with "and" and "or" in the following way (De Morgan's laws):

- The statement "not(p and q)" is equivalent to "not(p) or not(q)."
- The statement "not(p or q)" is equivalent to "not(p) and not(q)."

Roughly speaking, we can identify several strategies to prove implications of the form "p implies q":

- Direct proof: Suppose that p holds, and show how to obtain q.
- Proof by contrapositive: Provide a direct proof of $not(q) \Rightarrow not(p)$.
- Proof by contradiction: Suppose that p holds and q fails, and derive a contradiction.
- Proof by induction: Divide the proposition into smaller claims of the form p_n for each positive integer n. Establish the *base case* p_1. Then show that the implication "p_n implies p_{n+1}" holds for every positive integer n.

Common Task 2.1 (Showing p if and only if q). The statement "p if and only if q" holds if both of the implications "p implies q" and "q implies p" are true. To prove this type of statement it is usually best to divide the proof into two parts, one for each of these implications.

Common Task 2.2 (Showing that several statements are equivalent). Another common task is to prove that several statements, say p_1, p_2, p_3 and p_4 are equivalent. Often the most efficient way to do this is by showing a series of implications:

- "p_1 implies p_2,"
- "p_2 implies p_3,"
- "p_3 implies p_4,"
- "p_4 implies p_1."

Remark 2.3. When discovering and writing proofs, it is often good to keep a few points in mind:

- Formulate your arguments in complete (and grammatically consistent) sentences.
- It is usually best to avoid "shorthand" notation such as $\vee$, $\wedge$, $\Rightarrow$, "iff," etc. in presenting your arguments – the idea can almost always be formulated as a complete sentence, and this is beneficial both to avoid mistakes and for the reader's comprehension (an exception might occur in a single step of a proof, for instance when explaining how to manipulate a particularly intricate combination of operations).
- When possible, it is often helpful to find direct proofs (including proof by contrapositive), rather than proofs by contradiction. In many settings, it is possible to "translate" a proof by contradiction into a direct proof – we will see examples of this as we proceed further.

These are not necessarily rigid rules, but merely guidelines in order to help arrive at the most clear (and clearly explained!) proofs.

2.2. **Elementary theory of sets.** We often say that set theory is the "language of modern mathematics." Putting aside the delicate question of giving a precise description of sets (on which we will comment briefly in a remark below), we may informally say that a *set* A is a collection of *elements*, and use the notation

$$x \in A$$

to indicate that the object x is an element belonging to the set A. To be somewhat more specific, we recall the following familiar sets, the existence and basic properties of which we can take for granted in our day-to-day work:

- The set of natural numbers,

  $$\mathbb{N} = \{0, 1, 2, 3, \cdots\}.$$

  There is some inconsistency in notation here – some authors use $\mathbb{N}$ to denote the set $\{1, 2, 3, \cdots\}$. This is not a major issue and the meaning is almost always clear from context or easy to determine.
- The set of integers,

  $$\begin{aligned}\mathbb{Z} &= \{\cdots, -3, -2, -1, 0, 1, 2, 3, \cdots\} \\ &= \mathbb{N} \cup \{-n : n \in \mathbb{N}\}.\end{aligned}$$

- The set of rational numbers,
$$\mathbb{Q} = \left\{ \frac{m}{n} : m, n \in \mathbb{Z},\ n \neq 0 \right\}.$$
- The set of real numbers
$$\mathbb{R} = (-\infty, \infty).$$
Note that the first task in a rigorous analyis course is often to give a thorough discussion of the existence and basic properties of the set of real numbers. We will gloss over this point

A number of basic *set operations* allow us to form new sets out of those that we already have. We begin with some notation: given a set A, one can define
$$B := \{a \in A : \text{ the proposition } p(a) \text{ holds}\},$$
where $p(a)$ is a (suitably well-defined) property of the object a. The set B is then a *subset* of A, written
$$B \subset A.$$

Let X be a set and let $A, B \subset X$ be given. Our elementary set operations then include:

- Union:
$$A \cup B = \{x \in X : x \in A \text{ or } x \in B\}$$
- Intersection:
$$A \cap B = \{x \in X : x \in A \text{ and } x \in B\}$$
- Complement:
$$A^c = \{x \in X : x \text{ does not belong to } A\}$$

Note the analogy with the logical operations of "or," "and," and "not" that we introduced earlier. We also have two additional notions

- Set difference:
$$A \setminus B = A \cap B^c = \{x \in A : x \notin B\}$$
- Symmetric difference:
$$A \Delta B = (A \setminus B) \cup (B \setminus A).$$

Remark 2.4. Some care is required in formulating a "good" notion of set here. We highlight a famous paradox:

- Let $\mathcal{C}$ be the "set" of all sets which do not contain themselves. Does $\mathcal{C}$ belong to $\mathcal{C}$? If so, then $\mathcal{C}$ does not contain itself, giving a contradiction. If not, then $\mathcal{C}$ does contain itself, giving a contradiction once again. This is known as Russell's paradox – roughly speaking, the resolution is to be very careful to avoid self-reference in forming "sets of sets" (for this reason, when we use constructions involving groups of sets, we often prefer to speak of "collections of sets").

From a practical standpoint, these issues do not arise so often and we will not discuss them further in these notes.

2.3. **Quantifiers and negation.** To broaden the range of propositions we can consider, we now consider two "quantifying" operations: given a set X and a collection of propositions $\{p(x) : x \in X\}$, one can form the statements

- "$p(x)$ holds for every $x \in X$," and
- "there exists $x \in X$ such that $p(x)$ holds,"

with the obvious interpretations. The words "for every" and "there exists" can be abbreviated as $\forall$ and $\exists$, respectively, with the provision given above that it is not usually best to use such shorthand in the sentences of a proof; rather, these symbols often serve to compress the expression of the logical relationships when performing manipulations.

Note that when quantifiers are combined in sequence, the order is important: the statement

$$\text{"for all } x \in X \text{ there exists } y \in Y \text{ such that } p(x, y) \text{ holds}''$$

is distinct from

$$\text{"there exists } y \in Y \text{ such that for all } x \in X \text{ such that } p(x, y) \text{ holds.}''$$

In the first statement, y may depend on x, while in the second statement the **same** y must work for all x.

The negation operator interchanges the quantifiers $\forall$ and $\exists$. More precisely:

- the statement "not(for every $x \in X$, $p(x)$ holds)" (that is, not($\forall\, x \in X$, $p(x)$)) is equivalent to

  $$\text{there exists } x \in X \text{ such that not}(p(x)),$$

  that is, $\exists\, x \in X$ s.t. not($p(x)$), and
- the statement "not($\exists\, x \in X$ s.t. $p(x)$)" is equivalent to

  $$\forall\, x \in X, \text{ not}(p(x)).$$

This pair of quantifiers also allows us to give a slightly more general notion of set union and intersection. If A_n is a sequence of subsets of a set X, we can define the union of the sets A_n as

$$\bigcup_{n \in \mathbb{N}} A_n = \{x \in X : \exists n \in \mathbb{N} \text{ s.t. } x \in A_n\},$$

and the intersection as

$$\bigcap_{n \in \mathbb{N}} A_n = \{x \in X : \forall n \in \mathbb{N},\ x \in A_n\}.$$

Common Task 2.5 (Showing that one set is contained in another). Given a set X with two subsets $A, B \subset X$, we often have to show the inclusion

$$A \subset B.$$

The best way to show this is usually to show that for every $a \in A$, one has $a \in B$; that is, we reduce the claim to the implication "$a \in A$ implies $a \in B$." Proofs of this type often begin "Let $a \in A$ be given. ..." and proceed to demonstrate that a belongs to the set B.

Common Task 2.6 (Showing that two sets are equal). With A, B, X as above, the most efficient (and clear) way to show the set equality $A = B$ is to establish the two inclusions $A \subset B$ and $B \subset A$ individually.

2.4. **Functions.** Given two sets A, B, a function $f : A \to B$ is a rule assigning each $x \in A$ to a value $f(x) \in B$. The set A is referred to as the *domain* and the set B is referred to as the "co-domain".

The *image* of A under f is then the set

$$f(A) = \{f(x) : x \in A\},$$

while for every $B_1 \subset B$, the *preimage* of B_1 is the set

$$f^{-1}(B_1) = \{x \in A : f(x) \in B_1\}.$$

Common Task 2.7. Showing that a function is injective (one-to-one). We say that f is injective if for every $b \in f(A)$, the set $f^{-1}(\{b\})$ consists of a single element (it is non-empty by construction). Equivalently, f is injective if the implication "$f(x) = f(y)$ implies $x = y$" holds for all $x, y \in A$.

The usual approach is to establish this latter implication; the argument would typically take the form: "Let $x, y \in A$ be given such that $f(x) = f(y)$. (...) We therefore conclude that $x = y$; since x and y were arbitrary, we conclude that f is injective as desired." (Here "(...)" contains an argument establishing the claim $x = y$).

Common Task 2.8. Showing that a function is surjective (onto). We say that f is surjective if $f(A) = B$. The inclusion $f(A) \subset B$ follows from the definition of f. We are left with showing $B \subset f(A)$. For this, we usually let $b \in B$ be given, and construct $x \in A$ such that $f(x) = b$.

Common Task 2.9. Showing that a function is bijective (one-to-one and onto). We say that f is bijective if it is both injective and surjective. To demonstrate, we usually separate the argument into the proof of each of these claims.

3. Example: working with sets

Example 3.1. Let X, Y be sets, and let $f : X \to Y$ be a given function. Show that

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$$

for all $A, B \subset Y$.

Proof. Let $A, B \subset Y$ be given.

We begin by showing $f^{-1}(A \cup B) \subset f^{-1}(A) \cup f^{-1}(B)$. Let $x \in f^{-1}(A \cup B)$ be given. Then $f(x) \in A \cup B$, so that we have either $f(x) \in A$ or $f(x) \in B$. Suppose first that $f(x) \in A$ holds. Then $x \in f^{-1}(A) \subset f^{-1}(A) \cup f^{-1}(B)$ as desired. Alternatively, suppose $f(x) \in B$. Then $x \in f^{-1}(B) \subset f^{-1}(A) \cup f^{-1}(B)$. Thus, in either case, we have

$$x \in f^{-1}(A) \cup f^{-1}(B).$$

Since $x \in f^{-1}(A \cup B)$ was arbitrary, we have established the desired inclusion.

It remains to show $f^{-1}(A) \cup f^{-1}(B) \subset f^{-1}(A \cup B)$. Let $x \in f^{-1}(A) \cup f^{-1}(B)$ be given. We then have either $x \in f^{-1}(A)$ or $x \in f^{-1}(B)$. Suppose first that

$x \in f^{-1}(A)$; it follows from this that $f(x) \in A \subset A \cup B$, so that $x \in f^{-1}(A \cup B)$. Similarly, if $x \in f^{-1}(B)$, we have $f(x) \in B \subset A \cup B$ and thus $x \in f^{-1}(A \cup B)$ in this case as well. Since we have obtained the desired conclusion in both possible cases, and $x \in f^{-1}(A) \cup f^{-1}(B)$ was arbitrary, we conclude that the desired inclusion holds.

Since we have shown both set inclusions, we have established the desired set equality. □

Remark 3.2. We make two remarks on the above proof

- In the first part of the proof, we came to a situation where we knew that the statement "$f(x) \in A$ or $f(x) \in B$" was true. From here we know that we are in at least one of the following situations: Case 1: $f(x) \in A$ or Case 2: $f(x) \in B$. Note that both conditions may be valid – by showing that the desired conclusion is true in either case, we conclude that it is true regardless of which case we are in (Case 1 only, Case 2 only, or both). A similar situation occurs in the second part of the proof.
- Recall that to show the equality of two sets U and V, it suffices to show $U \subset V$ (every element of U is also an element of V) and $V \subset U$ (there are no other elements of V). We demonstrate these two inclusions independently – note that the proof contains two *different* variables named x: one in the second paragraph of the proof "Let $x \in f^{-1}(A \cup B)$ be given.", and one in the third paragraph "Let $x \in f^{-1}(A) \cup f^{-1}(B)$ be given."