Foundations and Fundamentals of Mathematics
Course material for MAT2362

Pieter Hofstra
Department of Mathematics and Statistics
University of Ottawa

Pieter Hofstra, 2013

Contents

Introduction

In the early stages of our mathematical development, topics are usually introduced in a relatively informal and intuitive manner. Our first acquaintance with calculus consists of intuitive explanations of what functions are, what continuity and differentiability mean, and so on. In this phase, the emphasis is mostly on mastering algorithms or techniques for solving certain types of problems.

After that, however, we need to achieve a more precise and formal understanding. For examples, instead of describing continuity in terms of drawing graphs without having to lift our pencil, we study the mathematical definition in terms of epsilons and deltas. Accordingly, the emphasis shifts towards understanding the intricacies of such mathematical definitions, the ability to work with them, to test them against examples, to relate them to other definitions and concepts, and to combine them to prove theorems.

The purpose of MAT2362/MAT2762 is to help students make the transition to this second phase of mathematical precision. It does so in two ways. First, by introducing some of the essential building blocks of modern mathematics, namely set theory and order theory. Almost all higher mathematics directly or indirectly relies on set theory, or at least is formulated in terms of sets. One of the aims is therefore to make students comfortable with the language of set theory and to illustrate how more sophisticated mathematical concepts can be defined in terms of sets. This is the "Foundations" part of the course.

Second, the course aims to give the student the skills needed to work with abstract definitions, examples and related concepts, and to write good proofs. This is achieved by focusing on the underlying logical structure of the material, paying particular attention to clear formulations, extreme cases, examples and counterexamples, and to proof strategies.

Organization

The notes for this course are divided into 21 lectures. This subdivision was chosen because each lecture presents one idea or coherent family of closely related ideas. Very

roughly, the lectures could correspond to an actual lecture in class, although it should be stressed that some of the lectures in this notes are much more difficult than others.

Each lecture follows the same format. It introduces the phenomenon or question to be studied, then introduces the relevant ideas and develops those. It concludes with a brief summary for easy reference and a series of exercises. Solutions to some but not all exercises can be found in an appendix. The student is advised that the exercises range from elementary (in the sense of testing your understanding of a definition in a small concrete example) to rather advanced or even speculative.

The first three lectures treat the basic principles of mathematical logic, albeit in an informal manner. The purpose of these lectures is to give the student some analytical tools for the rest of the course. Throughout the rest, we will constantly refer to these logical notions by highlighting the logical structure of the mathematical ideas we're discussing. Understanding this structure is the key to being able to work with it.

We have also included an appendix containing some guidelines for writing proofs. I recommend glossing over this appendix at first after you've studied the first three lectures, and then reading it in more detail once you are expected to product some proofs yourself.

Acknowledgements

The larger part of the material for these notes was written in the Fall semesters of 2011 and 2012, when I taught the course. Many students provided very useful feedback, ranging from simple typographical errors to pedagogical flaws. Teaching the course alongside P. Scott resulted in many useful discussions concerning the contents of the course and the exposition of the material. Finally, some of the advice concerning writing proofs originates from a document written for similar purposes by E. Cheng.

LECTURE I

Mathematical Reasoning

We begin by studying some of the basic principles of logic. Logic is often defined as the study of reasoning; mathematical reasoning is therefore considered to be the object of study of mathematical logic. In particular, one of the objectives is to make precise what is meant by a valid argument or proof.

Our goal in the coming lectures is not to give a complete treatment of mathematical logic; rather it is to introduce just enough logical tools for you to understand and analyse the structure of definitions, proofs and counterproofs, and to become familiar with common methods of reasoning in mathematics.

I.1 Propositional logic

There are various different kinds of logic. In this section, we look at a relatively simple but still very useful kind called *propositional logic*.

Proposition:
A *proposition* is a declarative sentence which has a well-defined truth-value.

Thus a proposition is a sentence which makes a claim about a state of affairs (this is what is meant by the adjective *declarative*), and this sentence is either true or false. (We stipulate that there are exactly two truth-values; put plainly: each proposition is

either true or false, and can't be anything in between.) Some examples will make more clear what is a proposition and what is not.

Examples I.1.1.

1. "There are infinitely many prime numbers" is a proposition (which happens to be true).
2. "17 is even" is a proposition (which happens to be false).
3. "Every bachelor is single" is a proposition (which is true).
4. "The cat is on the roof" is a proposition.
5. "Is the cat on the roof?" is not a proposition but a question.
6. "Please open the window!" is not a proposition but a command.
7. "Let p be a prime number" is also not a proposition; it is an imperative.

When in doubt, ask yourself whether it makes sense to say that the sentence in question is true, or that it is false. If that doesn't make sense, it's not a proposition.

Note that some propositions are true (or false) by definition of the words or mathematical notions involved. For example, "17 is even" is false by definition of what it means to be even, and "Every bachelor is single" is true by virtue of the meaning of "bachelor". However, a proposition such as "The cat is on the roof" may be true or false, depending on the specific situation. Note also that it is quite possible that we recognize something to be a proposition without actually knowing its truth value. A famous example is the sentence "There exist infinitely many natural numbers x such that x and $x+2$ are both prime." (This is called the *Twin prime conjecture*.)

I.2 The Calculus of Propositions

The main point of propositional logic is that simple propositions can be combined into more complicated ones using *connectives* such as "and", "or", "if... then", and so on. We introduce special symbols to denote these.

Conjunction

The first connective is the *conjunction* $\wedge$. The intended meaning of $p \wedge q$ is "p *and* q".

Disjunction

The second connective is the *disjunction* $\vee$. The intended meaning of $p \vee q$ is "p *or* q". In the English language, sometimes "or" is intended to be *inclusive* (meaning: p or q or possibly both), and sometimes is it *exclusive* (meaning: p or q but not both). In propositional logic, we adopt the convention that $\vee$ is interpreted in the inclusive sense.

Implication

The third connective is the *implication* $\rightarrow$. The intended meaning of $p \rightarrow q$ is "*if* p *then* q", or "p *implies* q". There are various other English words which are translated by the symbol $\rightarrow$; we refer to Table I.1 for more information.

Bi-implication

The fourth connective is the *bi-implication* $\leftrightarrow$. The intended meaning of $p \leftrightarrow q$ is "p *if, and only if* q". This may be regarded as the conjunction of $p \rightarrow q$ and $q \rightarrow p$.

Negation

The last connective is the *negation* $\neg$. The intended meaning of $\neg p$ is "*not* p", or, "p is false.

We will now describe in a precise manner how propositions are built up using the connectives. Because we don't care about the actual content of the propositions but only about their logical form, we use letters $p, q, r, \ldots$ to denote simple propositions. These are usually called *propositional variables* or *propositional letters*.

Definition I.2.1 (Propositional Calculus)**.** The collection $PROP$ of (formal) propositions is defined by the following clauses:

- Each propositional variable is an element of $PROP$.
- $\bot$ is an element of $PROP$ and $\top$ is an element of $PROP$.
- When α, β are in $PROP$, then so are $(\alpha \wedge \beta), (\alpha \vee \beta), (\alpha \rightarrow \beta), (\alpha \leftrightarrow \beta), \neg\alpha$.

Some remarks are in order. First, this definition is an example of an *inductive definition*: it tells you first what the basic propositions are and then how to build new propositions from old. It is important to note that anything which cannot be obtained according to this recipe is -by definition- not a proposition. (In a later lecture we will study inductive definitions in more detail; for now it suffices to realize that the definition provides you with a method for systematically generating all possible propositions.)

Next, the symbol $\bot$ denotes *falsum* (or *absurdity*), and stands for a proposition which is always false. Similarly, $\top$ denotes *true* (latin: *verum*), and stands for a proposition which is always true. (Having $\top$ in the definition is actually redundant because we could regard it as an abbreviation of $\neg\bot$, but it doesn't hurt.)

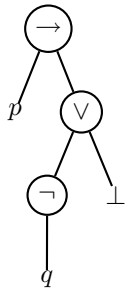
Finally, we use brackets to disambiguate expressions, just as you would do in arithmetic with an expression such as $2 - 3 + 4$. However, there are cases where we omit some of the bracketings when it doesn't create any ambiguity – see the examples below.

Examples I.2.2.

1. $(p \rightarrow \bot)$ is a proposition.
2. $p \rightarrow \bot$ is not a proposition because the brackets are missing. However, we shall adhere to the convention that the outermost brackets may be omitted.

3. $(p \vee q) \to (p \wedge r)$ is a proposition.
4. $p \vee q \wedge r$ is not a proposition (it is ambiguous because it can be bracketed in two different ways).
5. $\neg\neg\neg\neg\neg\top$ is a proposition.
6. $\neg \vee p$ is not a proposition.
7. $p \wedge q \wedge r$ is strictly speaking not a proposition, because there are brackets missing (one should either write $p \wedge (q \wedge r)$ or $(p \wedge q) \wedge r$). However, since we will soon see that both possible bracketings of the expression are logically equivalent, we often omit the brackets. The same principle applies to iterated disjunctions (but *not* to iterated implications!).

A slightly different, more visual, perspective is often useful. Given a proposition, we can represent it in the form of a *tree*. For example, the proposition $p \to (\neg q \vee \bot)$ can be represented by the tree



Thus the *leaves* of the tree are the basic propositions (variables and $\top, \bot$) occurring in the proposition, while the branching points are labelled by the connectives. Each proposition can be represented in this manner, and from the representing tree the proposition can always be recovered. More examples are given in the exercises.

I.3 Translations

We now turn to the art of translating English statements (or arguments) into propositional calculus. We first give a straightforward example to make clear what kind of procedure we have in mind:

Example I.3.1. Consider the sentence:

> If it rains and you don't have an umbrella, you will get wet.

To translate this into propositional logic, note that this is a compound sentence made up of three smaller sentences. We assign a propositional letter to each:

p – It rains.

English	Translation
(both) p and q	$p \wedge q$
p or q	$p \vee q$
p and/or q	$p \vee q$
Either p or q (but not both)	$(p \vee q) \wedge \neg(p \wedge q)$
p implies q	$p \to q$
If p then q	$p \to q$
p only if q	$p \to q$
q if p	$p \to q$
p is a sufficient condition for q	$p \to q$
q is a necessary condition for p	$p \to q$
p unless q	$\neg q \to p$
p if and only if q	$p \leftrightarrow q$
p iff q	$p \leftrightarrow q$
p is necessary and sufficient for q	$p \leftrightarrow q$
p is false	$\neg p$
not p	$\neg p$
It is not the case that p	$\neg p$

Table I.1: Common translations

q – You have an umbrella.

r – You will get wet.

A possible translation is now: $(p \wedge \neg q) \to r$.

More generally, when translating into propositional logic always first identify the basic propositions (those which cannot be further broken down into smaller propositions). Assign propositional letters to each of them, and then find a translation by carefully considering the usage of the English words corresponding to the connectives. In many cases, several different solutions are possible. Translation is not always easy: this is because it is not always obvious which connectives to use and how, and because of the fact that natural language can carry subtle connotations which cannot always be easily translated into the rigorous setting of propositional calculus. Also, English (and other natural languages) was not designed with mathematical precision in mind, and hence some English statements can be ambiguous or imprecise from a logical point of view. Part of the purpose of translating to formal logic is to eliminate such ambiguities and to give a mathematically precise formulation. Table I.1 lists a number of standard constructions and their translations.

Pay special attention to the distinction between "If p then q", "p if q" and "p only if q".

We give several more examples of translations. In each of these, we let

p – Fries are healthy.

q – You put mayonnaise on your fries.

r – Fries are delicious.

Examples I.3.2. Translate the following statements into propositional logic:

(a) If you don't put mayonnaise on your fries then they're not delicious.

(b) Fries are delicious but unhealthy.

(c) Fries are only unhealthy if you put mayonnaise on them.

(d) Fries are delicious, even without mayonnaise.

(e) Unless you put mayonnaise on them, fries are healthy.

(f) It is not true that fries are unhealthy.

(g) The fact that fries are delicious doesn't imply that they are healthy.

(h) Fries are not delicious without mayonnaise but not healthy with mayonnaise.

(i) For fries to be healthy it is necessary but not sufficient that you don't put mayonnaise on them.

Solutions. As a general strategy, first try to identify the *main connective* in the sentence (in the construction tree this would be the connective labelling the root of the tree). Then break down the problem into smaller parts.

(a) $\neg q \to \neg r$. (Note that the main connective is "If ..., then ...".

(b) $r \wedge \neg p$.

(c) $\neg p \to q$.

(d) $r \wedge \neg(\neg q \to \neg r)$. (Literally: fries are delicious and it's not the case that they're not delicious if you don't put mayo on them.) If you read this more as: "Fries are delicious with mayo but also without it" then you could use $(q \to r) \wedge (\neg q \to r)$.

(e) $\neg p \to q$. Alternatively, you could use $\neg q \to p$.

(f) $\neg\neg p$.

(g) $\neg(r \to p)$.

(h) $(\neg q \to \neg r) \wedge (q \to \neg p)$.

(i) $(\neg q \to p) \wedge \neg(p \to \neg q)$.

I.4 Summary

Propositional logic is used to reason about *propositions*, that is, statements which have a definite truth-value. The main feature of propositional logic is that propositions are built up from basic propositions (propositional variables and the two special propositions $\bot$ and $\top$), using the *propositional connectives* $\wedge, \vee, \rightarrow, \leftrightarrow, \neg$.

- Each proposition can be represented via its *construction tree*.
- The propositional connectives have a precise meaning in the calculus; in English, there may be various different words or phrases which are being translated into one connective.

I.5 Exercises

Exercise 1. Which of the following are propositions? Explain why (not).

(a) There's no such thing as a free lunch.

(b) If a lunch is free, then usually there's a catch.

(c) What's the catch with this free lunch?

(d) I'm hungry, let's eat!

(e) $2 + 2 = 5$.

(f) Assume that f is a differentiable function; then f is also continuous.

(g) If f is a differentiable function then it is also continuous.

(h) Find the derivative of f and use this to show that f is continuous.

(i) Propositional logic is useless.

(j) This is not a proposition.

Exercise 2. Which of the following are well-formed propositional formulas? For those which are well-formed, indicate the main connective.

(a) $p \wedge (q \vee r)$

(b) $p \wedge q \wedge r$

(c) $p \neg q$

(d) $p \leftrightarrow q \leftrightarrow r$

(e) $\neg(p \rightarrow r) \wedge \neg s$

(f) $p \wedge (p \wedge (p \wedge (p \wedge p)))$

(g) $((((p \vee p) \vee p) \vee p$

(h) $\bot \leftrightarrow \top$

(i) $\neg\bot \vee \neg$

(j) $\neg\neg\neg\top\top$

Exercise 3. For each of the following propositional formulas, draw the construction tree.

(a) $p \wedge q$

(b) $\bot$

(c) $\neg\neg\neg\top$

(d) $(p \vee q) \wedge (q \rightarrow \neg p)$

(e) $\neg\bot \vee \neg\top$

Exercise 4. In propositional calculus, brackets are needed to disambiguate expressions. When we use construction trees however, we don't see any brackets. Explain why this is not a problem. More precisely: explain why different bracketings of the same expression will necessarily result in different construction trees.

Exercise 5. Consider the following propositions:

p – Politicians are corrupt

q – Politicians keep their promises

r – Voting makes sense

Translate the following sentences into propositional logic:

(a) Voting makes no sense if politicians don't keep their promises.

(b) Voting makes sense only if politicians keep their promises.

(c) Even though some politicians are corrupt, they still keep their promises.

(d) Politicians don't keep their promises, even when they're not corrupt.

(e) Voting only makes sense when politicians are neither corrupt nor break their promises.

(f) If politicians don't keep their promises or if they're corrupt, voting doesn't make sense.

Exercise 6. Translate the following into propositional logic.

(a) I wear gloves if it's very cold.

(b) When I wear gloves I can't cook food.

(c) Either I don't wear gloves and I can cook food or I wear gloves but then I can't cook food.

(d) Is me wearing gloves a necessary condition for it being cold or is it only a sufficient one?

(e) I can cook food unless it's very cold or if I wear gloves.

(f) Not only can I not cook food, it's also very cold.

Exercise 7. Analyse the propositional structure of the statement

> The Twin prime conjecture implies that there exist infinitely many primes, but these two statements are not equivalent.

(You may translate "p is equivalent to q" by $p \leftrightarrow q$, and regard the Twin prime conjecture as a basic proposition.)

LECTURE II

Truth and Validity

Having introduced the language of propositional logic, we now turn to the question of what constitutes valid reasoning. We will explore two techniques for verifying the validity of arguments: the laws of boolean algebra and truth-tables.

II.1 Truth Tables

Recall that the collection of propositions is defined inductively, in the sense that we first specified basic propositions, and then methods for constructing new propositions from old. Because this guarantees that every proposition is built from basic propositions using the connectives, we can systematically break down a proposition into its components. This is particularly useful when we are interested in the *truth* of a proposition, i.e., when we ask whether a proposition is true or false. The following definition is one of many ways of specifying how the truth-value of a complex proposition depends on its constituents. In what follows we write t and f for the two truth-values; alternatively, we use 1 and 0. (Some texts use $\top$ and $\bot$, but we already used those for the two special basic propositions.)

Definition II.1.1 (Truth-assignment)**.** A *truth-assignment* (also called a *valuation*) is a mapping[1] $v : \mathrm{PROP} \to \{\mathsf{t}, \mathsf{f}\}$ with the following properties.

- $v(\top) = \mathsf{t}$
- $v(\bot) = \mathsf{f}$
- $v(\alpha \wedge \beta) = \begin{cases} \mathsf{t} & \text{if } v(\alpha) = \mathsf{t} = v(\beta) \\ \mathsf{f} & \text{otherwise.} \end{cases}$

[1] I assume you already know what a mapping or function is, at least on an informal level. The precise definition will come later.

- $v(\alpha \vee \beta) = \begin{cases} \mathsf{t} & \text{if } v(\alpha) = \mathsf{t} \text{ or } v(\beta) = \mathsf{t} \text{ (or both)} \\ \mathsf{f} & \text{otherwise.} \end{cases}$
- $v(\alpha \to \beta) = \begin{cases} \mathsf{f} & \text{if } v(\alpha) = \mathsf{t} \text{ and } v(\beta) = \mathsf{f} \\ \mathsf{t} & \text{otherwise.} \end{cases}$
- $v(\alpha \leftrightarrow \beta) = \begin{cases} \mathsf{t} & \text{if } v(\alpha) = v(\beta) \\ \mathsf{f} & \text{otherwise.} \end{cases}$
- $v(\neg\alpha) = \begin{cases} \mathsf{f} & \text{if } v(\alpha) = \mathsf{t} \\ \mathsf{t} & \text{if } v(\alpha) = \mathsf{f}. \end{cases}$

Notice that there are no conditions on the truth-values assigned to propositional variables. However, the point of the definition is that once we have specified truth-values to the propositional variables, the clauses above tell you exactly what the truth-values of complex propositions will be.

Note in particular the rule for implication: an implication is false when the antecedent is true but the consequent is false. Hence, *an implication is automatically true when the antecedent is false*, and it is also automatically true when the consequent is true. (Sometimes this is referred to as *vacuous truth*.) For example, the implication

> If the moon is made of blue cheese, then 3+3=5.

is true, because the antecedent "The moon is made of blue cheese" is false. Similarly,

> If fries are healthy, then 2+2=4.

is true because the consequent "$2 + 2 = 4$" is true. (It doesn't matter whether fries are healthy or not.)

Here are some examples of how one calculates with truth-assignments.

Example II.1.2. Suppose that v is a truth-assignment with $v(p) = v(q) = \mathsf{t}$ and $v(r) = \mathsf{f}$. Then

1. $v(p \wedge (q \vee r)) = \mathsf{t}$ because $v(q \vee r) = \mathsf{t} = v(p)$.
2. $v(p \to \neg q) = \mathsf{f}$ because $v(p) = \mathsf{t}$ and $v(\neg q) = \mathsf{f}$.
3. $v(p \to \neg r) = \mathsf{t}$ because $v(\neg r) = 1$.
4. $v(\bot \to (p \wedge r)) = \mathsf{t}$ because $v(\bot) = \mathsf{f}$.

There is another method of carrying out the task of finding the truth-value of a complex proposition in terms of those of its constituents. This method is called *truth-tables*. The idea is that one lists all possible combinations of truth-values that can be assigned to the basic propositions occurring in the complex proposition at hand, and then uses the rules for the connectives to work out for each of these cases what the truth-value of the complex proposition is.

For the proposition $p \wedge q$, for example, we would have the following table:

p	q	$p \wedge q$
f	f	f
f	t	f
t	f	f
t	t	t

Note that we have four rows: each of p, q can have two possible truth-values, so there are $2 \cdot 2 = 4$ possible combinations.

Here are the truth-tables for the other connectives:

p	q	$p \vee q$
f	f	f
f	t	t
t	f	t
t	t	t

p	q	$p \to q$
f	f	t
f	t	t
t	f	f
t	t	t

p	q	$p \leftrightarrow q$
f	f	t
f	t	f
t	f	f
t	t	t

p	$\neg p$
f	t
t	f

We can now use these to form truth-tables of more complicated propositions. For example, the truth-table for $\neg p \leftrightarrow (q \vee \neg p)$ is

p	q	$\neg p$	$q \vee \neg p)$	$\neg p \leftrightarrow (q \vee \neg p)$
f	f	t	t	t
f	t	t	t	t
t	f	f	f	t
t	t	f	t	f

The truth-table for $(p \wedge q) \to (p \vee (\neg r \wedge \neg q))$ is

p	q	r	$\neg q$	$\neg r$	$\neg r \wedge \neg q$	$p \vee (\neg r \wedge \neg q)$	$p \wedge q$	$(p \wedge q) \to (p \vee (\neg r \wedge \neg q))$
f	f	f	t	t	t	t	f	t
f	f	t	t	f	f	f	f	t
f	t	f	f	t	f	f	f	t
f	t	t	f	f	f	f	f	t
t	f	f	t	t	t	t	f	t
t	f	t	t	f	f	t	f	t
t	t	f	f	t	f	t	t	t
t	t	t	f	f	f	t	t	t

This truth-table has a rather special feature: in each row, the truth-value of the formula is t. Such propositions have a special name:

Definition II.1.3 (Tautology)**.** A proposition is called a *tautology* when its truth-value is t in every row of the truth-table of that proposition.

This means that a tautology is true, no matter what truth-values we assign to the propositional letters.

On the other extreme, we have:

Definition II.1.4 (Contradiction)**.** A proposition is a *contradiction* when its truth-value is f in every row of the truth-table of that proposition.

It is clear that a proposition cannot be a tautology and a contradiction at the same time. However, it can be neither:

Definition II.1.5 (Contingency)**.** A proposition is called *contingent* (also: *satisfiable*) when its truth-value is t in at least one of the rows of the truth-table.

Example II.1.6. The formula $(p \vee q) \leftrightarrow (\neg q \wedge \neg p)$ is a contradiction. The truth-table is

p	q	$\neg p$	$\neg q$	$\neg q \wedge \neg p$	$p \vee q$	$(p \vee q) \leftrightarrow (\neg q \wedge \neg p)$
f	f	t	t	t	f	f
f	t	t	f	f	t	f
t	f	f	t	f	t	f
t	t	f	f	f	t	f

from which we see that the formula is false in all rows.

Example II.1.7. Consider the proposition $(p \vee q) \rightarrow (\neg q \wedge \neg p)$. The truth-table is:

p	q	$\neg p$	$\neg q$	$\neg q \wedge \neg p$	$p \vee q$	$(p \vee q) \rightarrow (\neg q \wedge \neg p)$
f	f	t	t	t	f	t
f	t	t	f	f	t	f
t	f	f	t	f	t	f
t	t	f	f	f	t	f

Since the formula is true in at least one row, the formula is contingent.

The last concept we introduce in this section is that of logical equivalence.

Definition II.1.8 (Logical Equivalence)**.** Two propositions p and q are *logically equivalent* when p and q have the same truth-value in each row of the truth-table.

Notation: $p \equiv q$.

Equivalently, $p \equiv q$ when the proposition $p \leftrightarrow q$ is a tautology.

Example II.1.9. The propositions $p \rightarrow \neg q$ and $\neg(p \wedge q)$ are equivalent. Indeed consider the truth table:

p	q	$p \rightarrow \neg q$	$\neg(p \wedge q)$
f	f	t	t
f	t	t	t
t	f	t	t
t	t	f	f

In each row, the two propositions have the same truth-value.

By contrast, the propositions $p \to \neg q$ and $\neg p \vee q$ are not equivalent, as can be seen from the table:

p	q	$p \to \neg q$	$\neg p \vee q$
f	f	t	t
f	t	t	t
t	f	t	f
t	t	f	t

In rows 3 and 4 the propositions have different truth-values.

II.2 Validity of Arguments

There is more to reasoning than just determining the truth or falsity of a single proposition. Most of the time, we wish to know whether a chain of inference steps is correct. First, let us define the general form an argument takes:

Definition II.2.1. An *argument* consists of

- A collection of propositions (usually referred to as *hypotheses* or *assumptions*)
- A proposition called the *conclusion*.

Usually you can tell which part of the argument is the conclusion because it is preceded by a word such as "Therefore". (However, in some cases the order is just the opposite: first the conclusion gets stated, and then the arguments are given, typically preceded by "Because".) Schematically, an argument looks as follows:

$$\begin{array}{c} \text{Assumption 1} \\ \text{Assumption 2} \\ \vdots \\ \text{Assumption n} \\ \hline \text{Conclusion} \end{array}$$

Here is an example of an argument:

> If it rains and you don't have an umbrella you get wet. It rains, and you have an umbrella. Therefore, you don't get wet.

Separating the assumptions from the conclusion, we get:

$$\begin{array}{c} \text{If it rains and you don't have an umbrella you get wet.} \\ \text{It rains and you have an umbrella.} \\ \hline \text{You don't get wet.} \end{array}$$

Is this a valid argument? We must first decide what we mean by that. An argument purports to demonstrate that the conclusion logically follows from the assumptions. It does not matter whether the assumptions are true or not, only whether the conclusion is true whenever the assumptions are. This leads to:

Definition II.2.2 (Validity of an Argument)**.** Consider an argument with assumptions $p_1, \ldots, p_n$ and conclusion q. Then this argument is *valid* when for every truth-assignment with $v(p_1) = \mathsf{t}, \ldots, v(p_n) = \mathsf{t}$ it is also the case that $v(q) = \mathsf{t}$.

This condition can be checked using truth-tables: for the argument to be valid, you must check that in every row in which all of the p_i are true, it is also the case that q is true. On the other hand, if you wish to show that an argument is invalid, you need to find a row in the truth-table in which all the p_i are true, but where q is false.

Let us return to the example argument. In order to analyse it, we translate it into propositional logic first.

p – It rains.

q – You have an umbrella.

r – You get wet.

The translated argument is then

$$\frac{\begin{array}{c}(p \wedge \neg q) \rightarrow r \\ p \wedge q\end{array}}{\neg r}$$

The truth-table is

p	q	r	$(p \wedge \neg q) \rightarrow r$	$p \wedge q$	$\neg r$
f	f	f	t	f	t
f	f	t	t	f	f
f	t	f	t	f	t
f	t	t	t	f	f
t	f	f	f	f	t
t	f	t	t	f	f
t	t	f	t	t	t
t	t	t	t	t	f

There is one row in the table in which both assumptions are true, namely row 8. We must check whether the conclusion is also true in this row. However, it is not. The argument, therefore, is invalid. (If you guessed that the argument was valid, you overlooked the point that you can get wet even when carrying an umbrella - splashing cars for example.)

Proposition II.2.3. *An argument with assumptions $p_1, \ldots, p_n$ and conclusion q is valid if and only if the proposition $p_1 \wedge p_2 \wedge \cdots \wedge p_n \rightarrow q$ is a tautology.*

Proof. This is a consequence of the truth-table definitions of $\wedge$ and $\rightarrow$. □

1.	$p \wedge q$	$\equiv$	$q \wedge p$	commutativity of $\wedge$
2.	$p \wedge (q \wedge r)$	$\equiv$	$(p \wedge q) \wedge r$	associativity of $\wedge$
3.	$p \wedge p$	$\equiv$	p	idempotence of $\wedge$
4.	$p \wedge \top$	$\equiv$	p	$\top$ is a neutral element for $\wedge$
5.	$p \wedge \bot$	$\equiv$	$\bot$	$\bot$ is an absorbing element for $\wedge$
6.	$p \vee q$	$\equiv$	$q \vee p$	commutativity of $\vee$
7.	$p \vee (q \vee r)$	$\equiv$	$(p \vee q) \vee r$	associativity of $\vee$
8.	$p \vee p$	$\equiv$	p	idempotence of $\vee$
9.	$p \vee \top$	$\equiv$	$\top$	$\top$ is an absorbing element for $\vee$
10.	$p \vee \bot$	$\equiv$	p	$\bot$ is a neutral element for $\vee$
11.	$p \wedge (q \vee r)$	$\equiv$	$(p \wedge q) \vee (p \wedge r)$	distributivity of $\wedge$ over $\vee$
12.	$p \vee (q \wedge r)$	$\equiv$	$(p \vee q) \wedge (p \vee r)$	distributivity of $\vee$ over $\wedge$
13.	$\neg\neg p$	$\equiv$	p	Double Negation
14.	$\neg(p \wedge q)$	$\equiv$	$\neg p \vee \neg q$	De Morgan's Law
15.	$\neg(p \vee q)$	$\equiv$	$\neg p \wedge \neg q$	De Morgan's Law
16.	$p \vee \neg p$	$\equiv$	$\top$	Law of Excluded Middle
17.	$\neg\top$	$\equiv$	$\bot$	
18.	$p \to q$	$\equiv$	$\neg p \vee q$	
19.	$p \leftrightarrow q$	$\equiv$	$(p \to q) \wedge (q \to p)$	
20.	$p \leftrightarrow q$	$\equiv$	$(p \wedge q) \vee (\neg p \wedge \neg q)$	

Table II.1: Laws of Propositional Logic

Let me end this section by stressing one feature of arguments which seems to be misunderstood a lot, namely: *If one of the assumptions of an argument is false, then the argument is automatically valid!* (Also: if the conclusion is true then the argument is valid regardless of what the assumptions are.) This is a consequence of the definition of the truth table for implications.

II.3 Laws of Boolean Algebra

Truth-tables are a reliable and straightforward method for determining whether a certain proposition is a tautology, a contradiction or contingent; they can also be used for determining validity of arguments, and for testing logical equivalence of propositions. However, truth-tables are rather time consuming, and in many cases a more algebraic approach is useful.

The algebraic approach is based on the fact that a small number of ubiquitous logical equivalences already suffice to obtain all others. We list these important equivalences in Table II.1.

These equivalences are called the laws of propositional logic, or the laws of Boolean algebra. When we talk about an algebraic proof that two propositions are equivalent, we mean a proof consisting of a sequence of steps of the form in the table. (This means that you're not allowed to use other equivalences!)

Here are some examples.

Examples II.3.1.

1. Prove using the laws of Boolean algebra that $p \wedge \neg p = \bot$.

   Solution. Use the equivalences

$$\begin{array}{rcll} p \wedge \neg p & \equiv & \neg\neg p \wedge \neg p & \text{by 13.} \\ & \equiv & \neg(\neg p \vee p) & \text{by 15.} \\ & \equiv & \neg\top & \text{by 16.} \\ & \equiv & \bot & \text{by 17.} \end{array}$$

2. Prove that $\neg(p \vee q) \to r$ and $(p \vee r) \vee q$ are equivalent.

   Solution. Use the equivalences

$$\begin{array}{rcll} \neg(p \vee q) \to r & \equiv & \neg\neg(p \vee q) \vee r & \text{by 18.} \\ & \equiv & (p \vee q) \vee r & \text{by 13.} \\ & \equiv & p \vee (q \vee r) & \text{by 7.} \\ & \equiv & p \vee (r \vee q) & \text{by 6.} \\ & \equiv & (p \vee r) \vee q & \text{by 7.} \end{array}$$

3. Prove that $(p \wedge q) \wedge (p \wedge r) \equiv (p \wedge q) \wedge r$.

   Solution. Use the equivalences

$$\begin{array}{rcll} (p \wedge q) \wedge (p \wedge r) & \equiv & ((p \wedge q) \wedge p) \wedge r & \text{by 2.} \\ & \equiv & ((q \wedge p) \wedge p) \wedge r & \text{by 1.} \\ & \equiv & (q \wedge (p \wedge p)) \wedge r & \text{by 2.} \\ & \equiv & (q \wedge p) \wedge r & \text{by 3.} \\ & \equiv & (p \wedge q) \wedge r & \text{by 1.} \end{array}$$

Many more examples can be found in the exercises. We end this section with two elementary but important observations concerning implications. Mistakes are often made with these, so make sure you understand them well.

Definition II.3.2 (Converse, contrapositive)**.** Let $p \to q$ be an implicational proposition. Then

- the *contrapositive* of $p \to q$ is the proposition $\neg q \to \neg p$
- the *converse* of $p \to q$ is the proposition $q \to p$.

Using the laws of Boolean Algebra we can show that $p \to q$ is equivalent to its contrapositive $\neg q \to \neg p$:

$$\begin{array}{rcl} \neg q \to \neg p & \equiv & \neg\neg q \vee \neg p \\ & \equiv & q \vee \neg p \\ & \equiv & \neg p \vee q \\ & \equiv & p \to q \end{array}$$

(where we leave it as an exercise to figure out which rules were used). This is often useful in practice: to show that p implies q, you can show instead that $\neg q$ implies $\neg p$. For example: the statements

(i) If a matrix is invertible then it does not have a row of zeros.

(ii) If a matrix has a row of zeros then it is not invertible.

are each other's contrapositive, and hence logically equivalent.

However, $p \to q$ is generally not equivalent to its converse $q \to p$. Indeed, the statement

(iii) If a matrix does not have a row of zeros then it is invertible.

is the converse of (i), but it is false. (Whereas (i) is true.)

Similarly, the *negation* $\neg(p \to q)$ is generally not equivalent to the contrapositive, nor to the converse. (Write out the truth-tables to convince yourself of this fact.)

II.4 Knights and Knaves

The island of Knights and Knaves is inhabited by exactly two kinds of people: Knights, who always speak the truth, and Knaves, who always lie. (No other types of people live there, and there is no way to tell by looking what someone's type is.) This gives rise to puzzling situations, such as this one:

> One day, you arrive on the island, and two people approach you. One of them tells you: "At least one of us is a knave".

Can you figure out what the types of these two inhabitants is? We can reason as follows. First, note that it's not possible for both persons to be knights. If that were so, the first statement would be false, which would imply that the first person would be lying. Hence at least one of them is a knave, meaning that the first statement is true, and hence that the first person is a knight. Therefore the second must be the knave.

Problems like these, although they appear to be far from everyday mathematics, are included here for the following reasons:

- They are good exercise in systematic reasoning.
- They force you to reason using propositional logic, in particular negations.
- Even though it isn't apparent, there are situations in mathematics which are actually rather similar to these puzzles.

II.5 Summary

The key feature of propositional logic is that the truth-value of a complex proposition is completely determined by the truth-values of the propositional letters occurring in

the proposition. In order to determine the truth-value of a complex proposition, there are two methods (which are essentially the same):

- *Truth-assignments*
- *Truth-tables*

We also identified three classes of propositions: tautologies, contradictions and contingencies. Truth-tables can be used to find to which category a given proposition belongs.

The notion of *logical equivalence* is also defined in terms of truth-tables: two formulas are equivalent when their truth-tables are the same.

An *argument* consists of a collection $p_1, \ldots, p_n$ of assumptions and a conclusion q. Such an argument is *valid* when the implication $p_1 \wedge \cdots \wedge p_n \to q$ is a tautology.

Finally, the laws of Boolean algebra can be used to give efficient algebraic proofs of logical equivalencies. When asked to prove an equivalency using these laws, you may only use these laws, nothing else. You must also specify for each step which of the laws you use.

II.6 Exercises

Exercise 8. Construct truth-tables for each of the following propositional formulas. Use the truth-tables to determine whether the formula is a tautology, a contradiction or neither.

(a) $p \to \neg p$

(b) $\neg p \to p$

(c) $p \leftrightarrow \neg p$

(d) $p \to (q \to p)$

(e) $\neg(q \vee p) \to \neg p$

(f) $(p \to (q \to r)) \to ((p \wedge q) \to r)$

(g) $((p \to q) \to r) \to ((p \wedge q) \to r)$

(h) $(p \to \neg q) \leftrightarrow (q \to \neg p)$

(i) $(p \to \neg q) \leftrightarrow (\neg p \to q)$

(j) $\neg(p \to \neg q) \leftrightarrow (\neg p \to q)$

Exercise 9. Suppose that v is a truth-assignment such that $v(p) = \mathsf{t}, v(q) = \mathsf{f}$ and $v(r) = \mathsf{f}$. Find the following:

(a) $v(p \wedge (q \vee r))$

(b) $v(p \to (\neg q \to \neg r))$

(c) $v(\neg p \to \bot)$

(d) $v(p \vee (q \wedge (\neg r \to q)))$

Exercise 10. Suppose that v is a truth-assignment such that $v(p \wedge (\neg q \to r)) = \mathsf{f}$, $v(\neg q \wedge r) = \mathsf{f}$ and $v((q \vee r) \to p) = \mathsf{t}$. What can you conclude about $v(p), v(q)$ and $v(r)$?

Exercise 11. A propositional formula X is built from two propositional variables p and q. The truth-table for X is as follows:

p	q	X
f	f	f
f	t	t
t	f	t
t	t	f

(a) Find such a propositional formula X, using only p, q and propositional connectives.

(b) Find such a propositional formula X, but this time only use $\vee, \neg$.

(c) Find such a propositional formula X, but this time only use $\to, \neg$.

Exercise 12. Find a propositional formula Y using three propositional letters p, q, r such that $v(Y) = \mathsf{f}$ precisely when $v(p) = v(q) = \mathsf{t}$ or when $v(q) = v(r) = \mathsf{t}$. Next, give such a formula but only use the connectives $\to, \neg$.

Exercise 13. Consider the sentence: "If the Twin prime conjecture is not true, then number theory is less interesting than combinatorics." What is the contrapositive of this statement? What is the converse? And what is the negation?

Exercise 14. Consider the following argument.

> If we support the candidate, she will be elected. If the candidate is elected, she will help our cause. Therefore, if we support the candidate, she will help our cause.

Translate the argument into propositional logic and determine whether it is valid.

Exercise 15. Consider the following argument.

> If we support the candidate, she will be elected. If the candidate is elected, she will increase taxes. If taxes don't increase, we can afford to buy a new car. Therefore, if we don't support the candidate, we can afford to buy a new car.

Translate the argument into propositional logic and determine whether it is valid.

Exercise 16. Consider the following argument.

> Only if we don't support the candidate she will not be elected. If the candidate isn't elected or if the economy gets worse, taxes will increase. Therefore taxes won't increase unless we support the candidate.

Translate the argument into propositional logic and determine whether it is valid.

Exercise 17. Consider the propositional formulas $p \to (q \vee r)$ and $(p \to q) \vee (p \to r)$. Are these logically equivalent?

Exercise 18. Prove that $\neg p \equiv p \to \bot$.

Exercise 19. Use the laws of Boolean algebra to show that $(\neg p \to q) \to r$ is equivalent to $(p \to r) \wedge (q \to r)$.

Exercise 20. Using the laws of Boolean algebra, find a proposition which is equivalent to but simpler than (in terms of number of connectives) $p \to (\neg p \vee q)$.

The following exercises are about knights and knaves.

Exercise 21. On the island of knights and knaves, is it possible that someone says "I'm a knave"?

Exercise 22. Suppose now that you meet two people, one of whom says to you: "We're both knaves." What can you conclude about the types of these two people?

Exercise 23. Suppose you meet two people, and that one claims that exactly one of them is a knave. What do you conclude?

Exercise 24. Again you meet two people, but this time one of them says: "Either we're both knights or we're both knaves". The second person then says: "That's not true". What types do these people have?

Exercise 25. Suppose A says, "I am a knave but B isn't." What are A and B?

Exercise 26. If a knave claims that a proposition p is true and also claims that a proposition q is true, does it follow that he also claims that $p \wedge q$ is true? What about the converse?

Exercise 27. You meet four persons (call them A,B,C and D). What can you deduce from their statements?

A: "If I'm a knight, then 2+2=4."

B: "If I'm a knight, then 2+2=5."

C: "If I'm a knave, then 2+2=4."

D: "If I'm a knave, then 2+2=5."

Exercise 28. We have three inhabitants, A, B and C, each of whom is a knight or a knave. Two people are said to be of the same type if they are both knights or both knaves. A and B make the following statements:

A: B is a knave.

B: A and C are of the same type.

What is C?

Exercise 29. Again three people A, B and C. A says "B and C are of the same type." Someone then asks C, "Are A and B of the same type?" What does C answer?

Exercise 30. Someone asks A, "Are you a knight?" He replies, "If I'm a knight, then I'll eat my hat!" Prove that A has to eat his hat.

Exercise 31. Two individuals, X and Y, were being tried for participation in a robbery. A and B were court witnesses, and each of A, B is either a knight or a knave. The witnesses make the following statement:

A: If X is guilty, so is Y.

B: Either X is innocent or Y is guilty.

Are A and B necessarily of the same type? (i.e. either both knights or both knaves.)

LECTURE III

Predicate Logic

Predicate logic allows for a finer analysis of (mathematical) statements than propositional logic. Consider, for example, the following argument which most people would agree is valid (even though they might dispute the truth of the assumptions):

> All lawyers are greedy. Kimberly is a laywer. Therefore, Kimberly is greedy.

If we try to analyse this argument using propositional logic, we don't get very far. We would assign propositional letters to each of the three basic propositions occurring, namely

p - All lawyers are greedy

q - Kimberly is a lawyer

r - Kimberly is greedy

Then, the argument has the form "p and q, therefore r". However, the formula $p \wedge q \rightarrow r$ is not a tautology, and thus we cannot conclude that the argument is valid.

The problem, of course, is that we don't look closely enough at the sentences in the argument. The key aspect of the reasoning has to do with the use of the word "all"; the argument assumes that all lawyers have a certain property, and proceeds to conclude that a certain lawyer then must have that property. Thus in order to judge the merits of the argument, propositional logic is inadequate, and we must ask ourselves how we should reason about properties, individuals and words like "all". This is exactly what predicate logic does.

III.1 The syntax of predicate logic

In propositional logic, we build up formulas from basic propositions $p, q, r, \ldots$ using the propositional connectives $\wedge, \vee, \rightarrow, \neg, \leftrightarrow$. In predicate logic we still have these connectives, but we have a lot more. Specifically, predicate logical formulas are built from the following ingredients:

- *Constants*, or *individuals*, denoted by lower case letters $a, b, c, \ldots$. These are used to refer to specific objects, elements or individuals.
- *Variables*, denoted by $x, y, z, u, v, w, \ldots$. These variables are used to denote arbitrary, unspecified elements, much like in calculus you use an expression like $f(x) = x^2$, where x refers to an arbitrary element in the domain of f.
- *Predicates*, which are used to talk about properties of individuals or relations amongst individuals. Predicates are denoted $P(x), Q(x, y, z), R(x, y)$, etcetera. The number of arguments of the predicate is called the *arity* of the predicate. Thus, e.g. $P(x)$ has arity 1, while $Q(x, y, z)$ has arity 3.
- *Propositional connectives*, which are just as in propositional logic.
- *Quantifiers*, of which there are two: the *existential quantifier*, denoted $\exists$, and the *universal quantifier*, denoted $\forall$. These are used to express that *some* individuals have a certain property, or that *all* individuals have a certain property.

The following definition explains how the above ingredients can be used to form formulas:

- If $P(x_1, \ldots, x_n)$ is an n-ary predicate and each of $t_1, \ldots, t_n$ is either a variable or a constant, then $P(t_1, \ldots, t_n)$ is a formula.
- If ϕ, ψ are formulas, then so are $\phi \wedge \psi, \phi \vee \psi, \phi \rightarrow \psi, \phi \leftrightarrow \psi$ and $\neg\phi$.
- If ϕ is a formula and v is a variable, then $\forall v.\phi$ and $\exists v.\phi$ are formulas.

We give some examples.

Example III.1.1.

1. $P(a)$ is a formula (here, P is a unary predicate and a is a constant).
2. $\forall x.P(x)$ is a formula.
3. $\exists y.R(y, a)$ is a formula.
4. $\forall x \exists y.R(x, y)$ is a formula.
5. $\forall x.(P(x) \rightarrow \exists y.R(y, x))$ is a formula.

We avoid writing formulas in which variables are quantified more than once. For example, $\forall x \exists x.P(x)$ is nonsensical. A formula such as $\forall x.P(a)$, or $\forall x \forall y.P(x)$ is acceptable though. Just as in propositional logic, we use parentheses to disambiguate expressions.

We conclude with some terminology. In a formula $\exists x.\phi(x)$ or $\forall x.\phi(x)$ we say that the occurrences of the variable x in ϕ are *bound* by the quantifier $\exists x$ ($\forall x$ respectively). We also say that the variable x in ϕ is in the *scope* of the quantifier. Variables which are not bound are called *free*. This is similar to calculus, where in an expression such as $\int f(x)dx$ the variable x is bound by dx. Bound variables are merely placeholders and can always be renamed when convenient. For example, if we have a formula $\exists x.P(x)$ and a variable y which does not occur in P then $\exists y.P(y)$ is equivalent to $\exists x.P(x)$.

To see why we require that y should not occur in P, consider the formula $\exists x.x \neq y$. If we rename x to y we would get $\exists y.y \neq y$, which is manifestly false.

Some examples will make the concept of bound and free variables more clear.

Example III.1.2.

1. In $\forall x.(P(x) \to \exists y.R(x,y,z))$, the variables x and y are bound, while z is free.
2. In $\forall x.(P(y) \to \forall y.R(x,y,z))$, the variable x is bound, z is free; the first occurrence of y is free, while the second occurrence is bound.
3. In $\exists x.(P(x) \wedge P(y)) \to \exists z.P(x)$ the first occurrence of x is free, the second occurrence is bound, while y is free.

III.2 Translating into predicate logic

We begin by showing three examples of how predicate logic can be used to express everyday mathematical ideas.

Example III.2.1. 1. The statement

> For all real numbers x, y, there is a positive integer n satisfying $\frac{1}{n} < x$.

is translated as $\forall x \in \mathbb{R} \forall y \in \mathbb{R} \exists n \in \mathbb{N}.(n > 0 \wedge \frac{1}{n} < x)$.

2. The statement

> There exist rational numbers which are larger than their square.

is translated as $\exists x \in \mathbb{Q}.x > x^2$.

3. The statement

> If x is a real number satisfying $x + y > xy$ for all positive real numbers y, then x is negative.

is translated as $\forall x \in \mathbb{R}.(\forall y \in \mathbb{R}.(x + y > xy) \to x < 0)$.

Note how in the last example, it appears at first that the main connective is an implication (because the statement has the form "If..., then...") but that is misleading: the statement is about real numbers x, and can be paraphrased as "For all real numbers x: if x satisfies $x+y > xy$ for all positive y, then x is negative". It is often insightful to see if a complex statement can be reformulated in such a way that the logical structure becomes more clear.

We next turn to the representation in predicate logic of plain (non-mathematical) English sentences. When we wish to translate an English sentence into predicate logic, we have to

- Specify the constants we use.
- Specify the predicates we use.
- Give the actual translation.

For example, if we wish to translate the sentence "John is hungry", we would introduce a constant referring to John and a unary predicate for the property "is hungry".

a - John

$P(x)$ - x is hungry

Then the translation is: $P(a)$.

We can now translate the argument from the beginning of this lecture. Our dictionary is:

k – Kimberly

$L(x)$ – x is a lawyer

$G(x)$ – x is greedy

Then the translation of the argument is:

$$\frac{\begin{array}{c}\forall x.(L(x) \rightarrow G(x)) \\ L(k)\end{array}}{G(k)}$$

Note that we use an implication (and not a conjunction): we want to say that *if* an x is a lawyer, *then* that x is greedy. We don't want to say that all x are laywers and are greedy.

Table III.1 lists a number of standard constructions which you will encounter frequently. In some of these we also use *equality*. It is commonly assumed to be part of the language, and it allows for greater expressivity.

We proceed to give a number of examples. In each case, we make use of the following dictionary:

a - John

English	Translation
All P's are Q's	$\forall x.(P(x) \to Q(x))$
Some P's are Q's	$\exists x.(P(x) \wedge Q(x))$
Not all P's are Q's	$\neg\forall x.(P(x) \to Q(x))$
No P's are Q's	$\neg\exists x.(P(x) \wedge Q(x))$
There is at least one P	$\exists x.P(x)$
There are at least two P's	$\exists x\exists y.(P(x) \wedge P(y) \wedge x \neq y)$
There is at most one P	$\forall x\forall y.(P(x) \wedge P(y) \to x = y)$
There are at most two P's	$\forall x\forall y\forall z.(P(x) \wedge P(y) \wedge P(z) \to x = y \vee x = z \vee y = z)$
There is exactly one P	$\exists x.(P(x) \wedge \forall y.(P(y) \to x = y))$

Table III.1: Common predicate-logical translations

b - Mary

$L(x, y)$ - x loves y

Example III.2.2.

1. John loves Mary — $L(a, b)$
2. John loves Mary but Mary doesn't love John — $L(a, b) \wedge \neg L(b, a)$
3. Everyone loves Mary — $\forall x.L(x, b)$
4. Not everyone loves John — $\neg\forall x.L(x, a)$
5. Everyone loves someone — $\forall x\exists y.L(x, y)$
6. Not everyone is loved by someone — $\neg\forall x\exists y.L(y, x)$
7. Some people don't love themselves — $\exists x.\neg L(x, x)$
8. Everyone except Mary loves John — $\forall x.(\neg x = b \to L(x, a))$
9. John loves at most one person — $\forall x\forall y.(L(a, x) \wedge L(a, y) \to x = y)$
10. John loves at least two persons — $\exists x\exists y.(L(a, x) \wedge L(a, y) \wedge \neg x = y)$
11. John loves exactly one person — $\exists x.(L(a, x) \wedge \forall y.(L(a, y) \to x = y))$
12. If you don't love yourself you can't love others — $\forall x(\neg L(x, x) \to \neg\exists y.L(x, y))$
13. Some people don't love anyone except themselves — $\exists x.(\neg\exists y.(L(x, y) \wedge x \neq y) \wedge L(x, x))$

III.3 Domains and Interpretations

A lot of everyday mathematical statements can be made precise by casting them into predicate logic. As an example, we consider some statements about the ordering relation $\leq$ on the natural numbers.

Example III.3.1. Let $\leq$ denote the usual ordering on $\mathbb{N}$.

1. $\forall x.0 \leq x$ expresses that each x is greater than or equal to 0.
2. $\exists x \forall y.x \leq y$ expresses that there exists an element x with the property that all elements are greater than or equal than x.
3. $\forall x \exists y.y \leq x$ expresses that for each element there an element at least as big as that element.
4. $\forall x \exists y.(y \leq x \wedge x \neq y)$ expresses that for each element there exists a strictly larger element.
5. $\forall x \forall y.(x \leq y \vee y \leq x)$ expresses that every two elements are comparable.

Note that each of the above statements about the ordering $\leq$ on $\mathbb{N}$ are in fact true. A slightly different way of saying the same thing: when we interpret the *symbol* $\leq$ as the ordering relation on $\mathbb{N}$ and the symbol 0 as number 0, then the above statements are true.

However, consider the following statement: $\forall x \exists y.(y < x)$. (Here, we may use $y < x$ as an abbreviation of $y \leq x \wedge y \neq x$.) This expresses that for each element there is an element strictly below it. This is not true of the ordering on $\mathbb{N}$. Similarly, the statement $\forall x \forall y.(x < y \rightarrow \exists z.(x < z \wedge z < y))$ which says that for any two distinct elements there is an element strictly in between the two is false in $\mathbb{N}$.

By contrast, consider now the usual ordering on the rational numbers $\mathbb{Q}$. Then the statement $\forall x.0 \leq x$ is false: 0 is not the smallest rational number. However, the statement $\forall x \forall y.(x < y \rightarrow \exists z.(x < z \wedge z < y))$ is true in $\mathbb{Q}$, because if $x < y$ then the number $\frac{x+y}{2}$ is strictly in between x and y.

The main lesson here is that the same predicate-logical statement may be true in one situation but false in another. Put another way, the truth of the statement depends on the situation in which we interpret it. It is therefore considered good mathematical practice to be explicit about these matters. One common method is to state explicitly which set the variables range over. For example, one could write

$$\forall x \in \mathbb{N} \exists y \in \mathbb{N}.x < y$$

or

$$\forall x \in \mathbb{Q} \forall y \in \mathbb{Q}.(x < y \rightarrow \exists z \in \mathbb{Q}.(x < z \wedge z < y)).$$

It is even possible to mix and match, as in the formula

$$\forall x \in \mathbb{R} \exists y \in \mathbb{Z}.(x \leq y \wedge \forall z \in \mathbb{N}.(x \leq z \rightarrow y \leq z)).$$

The set of elements over which the variables range is usually called the *domain of discourse*. Only omit this information when it is clear from the context or when it doesn't matter what the domain is.

While the above notation is fairly common, it is worth pointing out that there is an alternative: instead of writing $\forall x \in A.P(x)$ (where A is some set) one may write $\forall x(x \in A \to P(x))$. And instead of $\exists x \in A.P(x)$ one may use $\exists x.(x \in A \wedge P(x))$. Note that in the case of the universal quantifier an implication is used ("If x is in A then it satisfies P") while for the existential quantifier one uses conjunction ("There exists an x in A which also satisfies P").

III.4 Reasoning with predicates

Now that we are familiar with the syntax of predicate logic, we will look at the principles of reasoning which govern it. While for propositional logic we gave a complete semantics in the form of truth-tables (complete in the sense that any question about logical equivalence and validity can be decided by the truth-table method), we will not attempt now to achieve something similar for predicate logic. This requires technology which we have not developed yet. Rather, our goal here is to familiarize the reader with the most common manipulations and proof strategies on an informal level, so that these can be used in subsequent lectures.

Of course, since predicate logic is an extension of propositional logic, all laws of propositional logic are still valid. We must therefore describe how reasoning with quantifiers works. The first principle is the following:

> From $\forall x.P(x)$ you may infer $P(a)$ for any a.

This principle is called *instantiation*; if P is true for all x, then it is also true of a particular instance a.

Example III.4.1. Suppose that we already know that the statement

$$\forall x \in \mathbb{N} \exists y \in \mathbb{N}.(x < y \wedge \text{Prime}(y))$$

is true. (Note that it states that there exist arbitrarily large prime numbers.) Then we may deduce that

$$\exists y \in \mathbb{N}.(2012 < y \wedge \text{Prime}(y))$$

is true as well.

Next, suppose that we wish to prove a formula of the form $\forall x.P(x)$.

> To prove $\forall x.P(x)$ take an arbitrary element x and prove $P(x)$.

It is important that in the proof of $P(x)$ you don't assume anything about x, because otherwise it isn't be *arbitrary* any longer!

Example III.4.2. Suppose we wish to prove that $\forall x \in \mathbb{N}.(x > 1 \to x^2 > x)$. We can do this as follows: consider an arbitrary natural number x. Suppose that $x > 0$. Then $x < x^2$ (we don't spell out this part here because it's not what we want to focus on). Thus $x > 0 \to x^2 < x$ holds for arbitrary x. Hence we conclude $\forall x.(x > 1 \to x^2 > x)$.

For the existential quantifier we also have two basic principles.

> To prove $\exists x.P(x)$, show that $P(a)$ for some element a.

For example, since 12 is a number satisfying $12 > 10 \wedge 12^2 < 150$, we may infer $\exists x \in \mathbb{N}.(x > 10 \wedge x^2 < 150)$.

Finally, there is a principle which tells how we can deduce statements when $\exists x.P(x)$ is given.

> Suppose that $\exists x.P(x)$ holds, and that Q is a statement not involving x. If $P(x)$ implies Q for arbitrary x, then Q holds.

This principle corresponds to the following common line of reasoning: if you can prove Q from the assumption $P(x)$, and you don't assume anything about x except for that it satisfies P, then Q already follows from the assumption that there exists an x with $P(x)$.

The above four principles dictate how we can prove statements of the form $\forall x.P(x)$ and $\exists x.P(x)$ and how we draw inferences from these. Using these (and the rules for propositional logic) we can prove various other valid predicate-logical principles. As in propositional logic, we write $\alpha \equiv \beta$ to indicate that α and β are logically equivalent. Also, we say that α is a tautology (contradiction) when $\alpha \equiv \top$ ($\alpha \equiv \bot$).

Example III.4.3. The formula $\forall x.P(x) \to \exists x.P(x)$ is a tautology.

To prove this, assume $\forall x.P(x)$. This implies $P(a)$, where a is arbitrary. Hence we get $\exists x.P(x)$. Therefore $\forall x.P(x) \to \exists x.P(x)$ is true.[1]

Slightly more complicated is the following:

Example III.4.4. The formula $(\forall x.(P(x) \to Q(x)) \wedge \forall x.P(x)) \to \forall x.Q(x)$ is a tautology.

To see why, assume that $\forall x.(P(x) \to Q(x)) \wedge \forall x.P(x)$ holds. To show that $\forall x.Q(x)$ is also true, consider an arbitrary x. From $\forall x.(P(x) \to Q(x))$ we get $P(x) \to Q(x)$.

[1] This only works because we usually assume that we are reasoning about a situation where at least one thing exists.

$\forall x \forall y.P(x,y)$	$\equiv$	$\forall y \forall x.P(x,y)$
$\exists x \exists y.P(x,y)$	$\equiv$	$\exists y \exists x.P(x,y)$
$\neg \forall x.P(x)$	$\equiv$	$\exists x.\neg P(x)$
$\neg \exists x.P(x)$	$\equiv$	$\forall x.\neg P(x)$
$\forall x.(P(x) \wedge Q(x))$	$\equiv$	$\forall x.P(x) \wedge \forall x.Q(x)$
$\exists x.(P(x) \vee Q(x))$	$\equiv$	$\exists x.P(x) \vee \exists x.Q(x)$
$\exists x \forall y.P(x,y) \rightarrow \forall y \exists x.P(x,y)$	$\equiv$	$\top$

Table III.2: Predicate-logical Equivalences

From $\forall x.P(x)$ we get $P(x)$. Thus by Modus Ponens, we deduce $Q(x)$. Since x was arbitrary, we may conclude $\forall x.Q(x)$.

In Table III.2 we list a number of predicate-logical equivalences which occur frequently.

Example III.4.5. The formula $\forall y \exists x.P(x,y) \rightarrow \exists x \forall y.P(x,y)$ is *not* a tautology. For example, when $P(x,y)$ is the predicate $x > y$ and the domain of discourse is the set of integers, then $\forall y \in \mathbb{Z} \exists x \in \mathbb{Z}.x > y$ is certainly true. However, $\exists x \in \mathbb{Z} \forall y \in \mathbb{Z}.x > y$ is manifestly false.

Many more examples of tautologies and equivalences can be found in the exercises.

The rules involving the interplay between quantifiers and negation deserve a close look, as they also tell us how we should negate quantified statements. Especially in formulas with many quantifiers mistakes are easily made, so be extra careful here. For example, consider the statement

For each real number x there exist real numbers y, z such that $x^2 > y$ and $z + y = x + z^2$.

What is the negation of this statement? In predicate-logical language, the statement reads

$$\forall x \in \mathbb{R} \exists y, z \in \mathbb{R}.(x^2 > y \wedge z + y = x + z^2).$$

The negation then is

$$\neg \forall x \in \mathbb{R} \exists y, z \in \mathbb{R}.(x^2 > y \wedge z + y = x + z^2),$$

and using the rules we find that this is equivalent to

$$\exists x \in \mathbb{R} \forall y, z \in \mathbb{R}.\neg(x^2 > y \wedge z + y = x + z^2).$$

This in turn is equivalent (using propositional logic and some facts about the ordering on $\mathbb{R}$) to

$$\exists x \in \mathbb{R} \forall y, z \in \mathbb{R}.(x^2 \leq y \vee z + y \neq x + z^2).$$

III.5 Summary

Predicate logic is a more expressive logic than propositional logic: it analyses the fine structure of propositions by means of predicates and quantifiers.

- $\forall x.P(x)$ means: "for all x, $P(x)$ holds"; the symbol $\forall$ is the *universal quantifier*.
- $\exists x.P(x)$ means: "there exists x for which $P(x)$ holds"; the symbol $\exists$ is the *existential quantifier*.

Most of the time when using predicate logic we have in mind a certain *domain of discourse*; this is a set over which the variables range. The truth of predicate-logical statements depends essentially on this domain; therefore it is often explicitly mentioned in the notation as $\forall x \in A.P(x)$ or $\exists x \in A.P(x)$.

Several concepts from propositional logic can be generalized to predicate logic: logical equivalence, tautology, contradiction.

The negation of $\forall x.P(x)$ is $\neg\forall x.P(x)$, which is equivalent to $\exists x.\neg P(x)$.

The negation of $\exists x.P(x)$ is $\neg\exists x.P(x)$, which is equivalent to $\forall x.\neg P(x)$.

III.6 Exercises

Exercise 32. Write the following mathematical statements in predicate logic:

(a) Some natural numbers are even, and some are prime.

(b) Not all odd numbers are prime.

(c) There exists an even prime number.

(d) Not all prime numbers are odd.

(e) If a number is prime and not equal to 2, then it is odd.

(f) Some numbers are neither prime nor even.

(g) If a number isn't prime then either it isn't odd or it isn't equal to 2.

Exercise 33. For each of the following statements, find out whether it is true or false.

(a) $\forall x \in \mathbb{N}\forall y \in \mathbb{N}.x + y \leq xy$

(b) $\forall x \in \mathbb{N}\forall y \in \mathbb{N}.((x > 0 \wedge y > 0) \rightarrow x + y \leq xy)$

(c) $\forall x \in \mathbb{R}\exists y \in \mathbb{N}.x < y$

(d) $\forall x \in \mathbb{N}.(x > 0 \rightarrow \exists y \in \mathbb{R}.xy = 1)$

(e) $\forall x \in \mathbb{R}.(x > 0 \rightarrow \exists y \in \mathbb{N}.xy = 1)$

(f) $\exists x \in \mathbb{R} \forall y \in \mathbb{R}.(y \leq 0 \vee xy < y)$

(g) $\exists x \in \mathbb{Z}[\forall y \in \mathbb{R}.(xy = y) \wedge \forall z \in \mathbb{R}.((\forall y \in \mathbb{R}.zy = y) \rightarrow x = z)]$

(h) $\exists x \in \mathbb{N} \forall y \in \mathbb{N} \forall z \in \mathbb{N}.(xy > z \vee zx > y)$

Exercise 34. Consider the formula $\forall x \forall y \exists z.(x \neq 0 \rightarrow x \cdot z = y)$.

(a) Work out what the negation of this formula is.

(b) Is the formula true in $\mathbb{N}$ (with the usual interpretation of 0 and $\cdot$)?

(c) Is the formula true in $\mathbb{Q}$ (with the usual interpretation of 0 and $\cdot$)?

Exercise 35. We consider 3×3 matrices with the usual notation

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Among the matrices below, find for which the following statement is true:

$$\forall i \in \{1,2,3\} \exists j \in \{1,2,3\} \left[a_{ij} \leq -5 \vee \exists n \in \mathbb{N}.(a_{ij} = (n+1)^2) \right].$$

$$A = \begin{bmatrix} 1 & 4 & -7 \\ -1 & 2 & 0 \\ 2 & 0 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 1 & -1 \\ 2 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 1 & -10 \\ 0 & 16 & 0 \\ 49 & -12 & 36 \end{bmatrix}$$

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E = \begin{bmatrix} 1 & 0 & -6 \\ -7 & 4 & 25 \\ 0 & 0 & 9 \end{bmatrix} \quad F = \begin{bmatrix} -1 & 0 & -6 \\ 0 & 4 & 2 \\ -1 & 0 & 0 \end{bmatrix}$$

Exercise 36. Walter and Donny are having a disagreement about the statement

$$\exists x \in \mathbb{R} \forall y \in \mathbb{R}.(0 < y \wedge y \leq x \rightarrow y^2 < y).$$

Walter: "This statement is true; take for example $x = -1$. Then for any y, $0 < y \wedge y \leq x$ is false, and hence the whole statement is true."

Donny: "This statement is false; take for example $x = 1$. Then consider $y = 1$; we have $0 < y \leq 1$, but $1^2 = 1$, so the conclusion fails."

Who is wrong?

Exercise 37. Consider again the statement

$$\exists x \in \mathbb{R} \forall y \in \mathbb{R}.((0 < y \wedge y \leq x) \rightarrow y^2 < y).$$

What is the correct negation of this statement?

(a) $\forall y \in \mathbb{R} \exists x \in \mathbb{R}.((0 \geq y \vee x < y) \rightarrow y \leq y^2)$

(b) $\forall x \in \mathbb{R} \exists y \in \mathbb{R}.((0 < y \wedge y \leq x) \wedge y \leq y^2)$

(c) $\forall x \in \mathbb{R} \exists y \in \mathbb{R}.((0 \geq y \vee x < y) \wedge y \leq y^2)$

(d) $\forall x \in \mathbb{R} \exists y \in \mathbb{R}.(y^2 < y \rightarrow (0 < y \wedge y \leq x))$

Exercise 38. Translate to predicate logic:

(a) Not everyone likes John.

(b) Some people only love themselves.

(c) Not everyone is loved by everyone.

(d) At least two people like Mary.

(e) At most one person likes him/herself.

(f) There is exactly one person who likes John, and it is not John himself.

Exercise 39. Translate the following statements into predicate logic.

(a) All dogs chase cats.

(b) There are dogs which don't chase any cats.

(c) Some dogs only chase some cats.

(d) Big dogs don't chase small cats.

(e) Big cats are only chased by big dogs.

(f) There are big cats which don't get chased by any dog.

(g) Not all small dogs chase big cats.

(h) A cat only chases a dog when the cat is big and the dog is small.

Exercise 40. Consider the formulas: $\alpha = \exists x(P(x) \wedge Q(x))$ and $\beta = \exists x.P(x) \wedge \exists x.Q(x)$.

(a) Show that $\alpha \rightarrow \beta$ is a tautology.

(b) Give an example to show that $\beta \rightarrow \alpha$ is not a tautology.

Exercise 41. Consider the formula $\forall x.(\exists y.L(x, y) \vee \forall z.L(z, x))$. Is this a tautology? If yes, prove it. If no, give an example to show why not.

Exercise 42. Which of the following formulas is a tautology? If it is a tautology, explain why. If not, give an example to show why not.

(a) $\forall x \forall y.R(x, y) \rightarrow \forall x.R(x, x)$

(b) $\forall x.P(x) \vee \forall x.\neg P(x)$

(c) $\exists x.P(x) \vee \exists x.\neg P(x)$

(d) $\forall x.(\bot \rightarrow P(x))$

Exercise 43. Find the negations of the following formulas. Write them in a form in which the negation symbols only occur directly in front of a predicate symbol.

(a) $\exists x \exists y.P(x,y)$

(b) $\exists x \exists y.(P(x,y) \vee Q(x,y))$

(c) $\exists x \exists y.(P(x,y) \wedge Q(x,y))$

(d) $\forall x \forall y.P(x,y)$

(e) $\forall x \forall y.(P(x,y) \to Q(x,y)$

(f) $\exists x \forall y.P(x,y)$

(g) $\exists x \forall y.(P(x,y) \to Q(x,y)$

(h) $\forall x \exists y.P(x,y)$

(i) $\forall x.(\exists y.P(x,y) \vee \exists z.Q(x,z))$

(j) $\exists x.(P(x) \to \forall y.P(y))$

(k) $\forall x.(\exists y.P(x,y) \to \forall z.Q(x,y))$

Exercise 44. Investigate the validity of the following argument:

> All men are mortal. All mortals have fear. Therefore, all men have fear.

Exercise 45. Investigate the validity of the following argument:

> Anyone who can climb Mount Everest is strong. If you are strong, you have no fear. The only people who have fear are those which are mortal. Therefore, those who are mortal can't climb Mount Everest.

Exercise 46. Investigate the validity of the following argument:

> John loves everyone except himself. Mary loves everyone who John loves. Therefore, Mary and John love each other.

Exercise 47. Carefully translate the following version of the Twin prime conjecture in predicate logic: "For every natural number x, there exists a strictly larger natural number y such that y and $y+2$ are both prime."

Exercise 48. On the island of Knights and Knaves, you meet two inhabitants. The first says: "All inhabitants of this island are knaves." The second replies: "That's not true!"

What can you conclude about the types of these inhabitants?

Exercise 49. Next, you meet two inhabitants, A and B, who seem to be acting strange. You ask them whether any of them has been drinking. They give you the following answers:

A: All knaves on this island are drunk. I'm perfectly sober though.

B: (Nods.) I'm a knave and I'm quite drunk!

Can you figure out who has been drinking?

LECTURE IV

Sets and Foundations

IV.1 Questioning

Mathematics studies a wide variety of things. Calculus is concerned with differentiation and integration of functions; linear algebra centers around the study of vector spaces and linear mappings; geometry studies shapes and forms; mathematical logic studies proofs and computations, and so on.

While some of the more advanced concepts in a branch of mathematics may seem very abstract and far removed from the basic, elementary notions we are more familiar with, one can try to unravel complex definitions by demanding each part of the definition to be expanded in simpler terms, and then asking for each of these simpler terms to be explained in even simpler terms, and so on. Let's try this for an example of a familiar mathematical structure, namely a (real) vector space.

You: *What is a vector space?*

Linear Algebraist: It is a set X equipped with an addition operation, a zero and scalar multiplication satisfying certain axioms.

You: *What do you mean by "addition operation"?*

Linear Algebraist: An addition operation on a set X is a function from $X \times X$ to X. The axioms say this should be associative and commutative.

You: *A function from $X \times X$ to X? What exactly is $X \times X$?*

Linear Algebraist: That's the cartesian product of the set X with itself. You may take it to be the set of all pairs of elements of X.

You: *Ah. But what do you mean by a pair?*

Linear Algebraist: You don't know what a pair is??

You: *I have an intuitive understanding of what it is, but I want to make sure I have the correct definition.*

Linear Algebraist: Ok then. One way of defining a pair (x, y) is by using a trick by a mathematician called Kuratowski. He defined the pair (x, y) as the set $\{x, \{x, y\}\}$.

You: *That's an interesting approach, although I'm far from clear yet why the set of all such pairs always exists.*

Linear Algebraist: That's just an elementary fact about sets: given any two sets you can form their product. You can even form the product of infinitely many sets if you want. If you can't form a product of sets we might as well give up on mathematics.

You: *I don't think I want to give up on mathematics; but still I'm not sure why the product of two sets should exist. Could you prove it for me?*

Linear Algebraist: Alright then, although it appears as if we won't get to do any interesting linear algebra today! If we use the Kuratowski definition of pair, then an element of $X \times Y$ is simply a special subset of the union $X \cup \mathcal{P}(X \cup Y)$, where $\mathcal{P}$ is the powerset operation.

You: *I see, but that doesn't really address my worries about the set of all such things existing. We now seem to rely on something even more complicated, namely powersets.*

Linear Algebraist: You don't believe that given a set you can form the set of subsets of that set?

You: *It certainly seems very plausible that you can do that, and it would indeed solve this issue. But are you saying that this is not something we can prove, but simply believe in?*

Linear Algebraist: Ultimately we need to take the truth of some self-evident statements for granted, otherwise we can't do anything.

As you can see, insisting on further and further unravelling of definitions and questioning the existence of basic mathematical entities ultimately reveals an acceptance of certain principles of set theory (in the dialogue above this was the existence of powersets). Most mathematicians indeed believe[1] that ultimately all of mathematics rests on set theory; moreover, they take certain axioms about sets to be true and self-evident.

Assuming that sets indeed are of foundational importance in mathematics, several questions arise:

1. What are sets, really?

[1] Many mathematicians simply don't wish to be bothered with questions about foundations. That's a bit like not wanting to think about which political party to vote for because you're not really comfortable with any of them.

2. What principles are sets governed by?
3. How can we correctly reason about sets and the more complex structures we build from them?
4. Can the truth or falsity of any mathematical statement be decided from facts about sets?

Interestingly, mathematicians don't all agree on the answers to these questions. Some axioms about sets, for example those which postulate the existence of certain very large sets, are not universally agreed upon. Some mathematicians reject the existence of powersets, or the notion of an infinite set! What is more, there is disagreement on the validity of certain principles of mathematical reasoning and logic. Ultimately, these things are a matter of philosophical orientation.

However, there is no doubt that sets play a very important role in mathematics and that learning the basics of set theory is a necessity in order to really understand more advanced mathematics. While we may disagree on how important sets are as a *foundation* of mathematics, they are still *fundamental*!

Alternative Foundations

While the majority of mathematicians regard set theory in on form or another as a suitable foundation for mathematics, there are alternative foundational frameworks available. *Category theory* is one such alternative; this is a conceptual approach to mathematics which emphasizes patterns and structures which occur across mathematics. *Type theory* is another, more logic-oriented approach.

IV.2 What are sets?

So far, we argued that sets are important, both from a foundational and from a practical mathematical perspective. However, we didn't attempt to carefully define the notion of a set. This is of course a serious omission: when we claim that large parts of mathematics are built out of sets, we better make the notion of set very precise!

Most people, when pressed to define what a set is, come up with the following:

Definition IV.2.1 (Naive definition of set)**.** A *set* is a collection of things.

We may want to sharpen this a bit, and replace "things" by "mathematical objects". But that still is very suspicious: when we're trying to provide a foundation for mathematics, we certainly shouldn't be assuming that we already know what mathematical objects are. And equally seriously, the word "collection" seems to mean almost the same thing as "set", so aren't we replacing one problematic word with another here?

Yes, we are. However, even if we found a better word or sentence to describe sets, we could again object that we're just replacing words by other words, and we can always keep asking for further clarification. So if we want to continue, we have to settle on something, and definition (IV.2.1) seems good enough for practical purposes. Of course, if we're interested in *foundations* we still have to be more explicit about what we count

as mathematical objects and how we allow ourselves to collect them into sets. For the moment, we sidestep this question and try to get the theory off the ground.

Adopting the above definition, accepting the fact that it is still inherently vague but useful enough for practical purposes is called *naive set theory*. In naive set theory, one simply develops the theory of sets without worrying too much about a formal definition. In the next lecture however, we will explain why this can lead to problems, prompting a more careful approach.

IV.3 The membership relation

The most basic thing which we want to express about a set is that a certain object is, or is not, in the set. When X is a set (we usually use upper-case roman letters to denote sets) and a is some object, we write $a \in X$ for the statement that a is an element of the set X. If we want to express that a is not a member of X, we write $a \notin X$.

Membership relation:
$a \in X$ means: a is an element of the set X.
$a \notin X$ means: a is not an element of X.

We also say that a is a *member* of X, or that a *belongs to* X. The symbol $\in$ is the Greek letter epsilon, and is called the *membership* symbol.[2]

Often we will want to describe a set by explicitly listing all of its elements. In that case, we use the set-bracket notation $\{\ldots\}$. For example, if we wish to introduce the set X whose elements are the numbers $1, 4$ and 22, we could write

$$X = \{1, 4, 22\}.$$

[2] In a later lecture we will explain in what technical sense the membership relation is a relation.

Infinite Regress

An *infinite regress* is a situation where a solution to a problem is never reached because the same problem keeps repeating itself over and over again, ad infinitum. In the case of definitions, this occurs because when we define a concept A in terms of other concepts B,C, we need to define B and C; when we do this (in terms of yet other concepts D,E,F, say), we are now required to define D,E,F as well. This never stops, unless you accept certain *Axioms*.

Combining this with the element-notation from above, we could now express various things about this set, such as

$$1 \in X, \qquad 3 \notin X, \qquad 4 \notin X, \qquad 6 \in X.$$

The first two of these statements are true, while the second two are false.

One important point to note is that the order in which we list the elements of a set does not matter. Thus, we could also have written the set X as $X = \{22, 1, 4\}$ or $X = \{4, 22, 1\}$. Moreover, the multiplicity of elements (meaning: the number of times an element is listed) is also irrelevant. Thus we also have $X = \{1, 4, 4, 4, 22, 22\}$. Saying three times that the number four is an element of X doesn't change anything. In the next lecture we will be more specific about what it means for two sets to be the same.

Let us consider a few common collections of mathematical objects which are used all the time:

Examples IV.3.1.

1. $\mathbb{N} = \{0, 1, 2, \ldots\}$ is the set of *natural numbers*
2. $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ is the set of *integers*
3. $\mathbb{Q}$ is the set of *rational numbers*
4. $\mathbb{R}$ is the set of *real numbers*
5. $\mathbb{C}$ is the set of *complex numbers*
6. The set $\{\mathtt{t}, \mathtt{f}\}$ (often called 2) is the set of *truth values*.

Note that for the first two sets we use the ...-notation to indicate that you have to continue forever. This notation should only be used when it is clear how to continue!

A very useful way of describing sets is called *set-builder notation*. Here's an example:

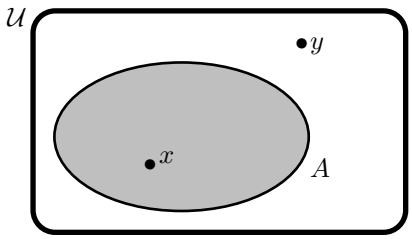
$$P = \{x | x \text{ is prime}\}.$$

Here, we're defining a set called P whose elements are those numbers which are prime. The vertical bar | may be pronounced "such that". Unfortunately, the above definition is ambiguous, since it doesn't specify whether we want to allow negative numbers as well. Better would be $\{x \in \mathbb{N} | x \text{ is prime}\}$ or $\{x \in \mathbb{Z} | x \text{ is prime}\}$. Moral: be explicit when there is potential ambiguity!

IV.4 Venn Diagrams

There also is a useful pictorial tool for visualizing sets and constructions on them. A *Venn diagram* is a visualization of one or more sets and the relations between them. Typically, sets are drawn as ellipses, while elements are drawn as dots. Then if a dot named x is inside an ellipse named A, that is meant to indicate that $x \in A$. If an element y is drawn outside A, we intend to convey that $y \notin A$.

Often, we are working in a specific mathematical situation where there is a particular ambient set in which all things of interest live. For example, if we "do number theory" we might be confining ourselves to the set of integers. In such a situation, where we only consider elements and subsets of a particular given set, we often refer to such a set as the *universe of discourse*, or simply *universe*. It is customary to denote such universe by $\mathcal{U}$. In a Venn Diagram, the outer rectangle (inside of which everything of interest happens) is the universe. See figure IV.4 for an illustration.



A *Venn Diagram* with

- Universe $\mathcal{U}$
- A set A
- An element $x \in A$
- An element $y \notin A$

Figure IV.1: Venn Diagram

It should be stressed that Venn diagrams, while a useful tool for visualization, can never replace precise mathematical reasoning.

IV.5 Summary

Naive set theory is an informal theory of sets which ignores philosophical problems; it is the style of set theory which occurs in everyday mathematics.

- The naive working definition of set is "a collection of mathematical objects".
- Set membership is the most fundamental notion in set theory.
- When listing the elements of a set, the order of the elements and their multiplicity do not matter.
- There are various ways of describing sets: by listing all elements, or by set-builder notation. As an informal visual aid, Venn diagrams can be helpful.

LECTURE V

Formal set theory

The naive version of set theory works perfectly fine for most practical purposes. However, things go wrong when you push it to its limits! This lecture explains what goes wrong and why, and also briefly looks at how formal set theory avoids these problems.

V.1 Russell's Paradox

So far, naive set theory has given us a simple language for talking about which sets contain which elements, and for describing sets using set-builder notation.

It is important to appreciate how powerful and flexible this is. For example, recall our definition of set: a collection of mathematical objects. Now certainly we should consider sets themselves to be mathematical objects. Thus it makes sense to speak of sets whose elements are themselves sets. For example we have the set $\{\mathbb{N}, \mathbb{Q}, \mathbb{R}\}$. As you can see, this gives us a lot of expressive power: the only limit to defining new sets is our imagination.

But now consider the following (defined using set-builder notation):

$$X = \{U | U \notin U\}.$$

To see why this is a problem, we should analyse the set X a bit closer. There are two types of sets: those which are a member of themself and those which aren't. Which type does

Logical Analysis

To make the schematical form of Russell's paradox more explicit, consider a hypothesis A. Suppose now, that with the help of A we can prove two statements:

- $P \to \neg P$
- $\neg P \to P$

where P is some other statement (it doesn't matter which one, but ours was $X \in X$). These two form a contradiction (check the truth-table!). Thus we have shown that $A \to \bot$. This is the same as saying that A is false.

the mysterious set X have? Suppose that X is a member of itself, i.e., that $X \in X$. By definition of the set X, this means that X is not a member of itself (because X consists of *all* sets which are not a member of themselves). So $X \notin X$. This is a contradiction. Next, suppose X is not a member of itself, i.e., $X \notin X$. Then by definition, X belongs to X (again because X consists of *all* sets which are not a member of themselves). So $X \in X$. Again we have a contradiction.

V.2 Formal set theory

Russell's paradox arises because there are no restrictions on how we can define new sets. This led mathematicians to develop forms of set theory which avoid the inconsistency[1] of naive set theory.

The most common approach is to work with a *formal system*; the idea is not to give a direct definition of what a set is, but to axiomatize within a tightly controlled logical system what the essential properties of sets are (and then to agree to call anything with those properties a set). The most common version of formal set theory is called *Zermelo-Fränkel* set theory, or ZFC[2].

Roughly speaking, this system works as follows.

- First, one specifies a *language*. For ZFC and its variants, the language is generated from only one basic symbol, namely $\in$. One can then write *formulas* involving the membership symbol using the propositional connectives and quantifiers. For example, $\forall x \exists y.(y \in x \vee x \in y)$ is such a formula. (Technically, ZFC is a theory in first-order predicate logic.)

- Next, one chooses *axioms*. These axioms tell us what sets can be constructed and how they behave. For example, one of the axioms states that the powerset of an already defined set always exists; another one guarantees the existence of infinite sets; and there is an axiom which specifies what it means for two sets to be equal.

- Finally, one may use the rules of predicate logic to derive theorems about sets.

The fact that one works within a formal system together with a suitable formulation of the axioms, guarantees that problems like Russell's paradox do not occur. Of course, one still wants enough expressive power, and would like to have a method similar to the set-builder notation for constructing new sets from old. This leads to the following axiom, often called *restricted comprehension*. Comprehension is the technical term for the set-builder notation, while the restrictedness refers to the fact that we can only apply it in cases where we select elements from an already defined set by specifying a property they must satisfy.

[1] A theory is called *inconsistent* when a contradiction can be derived from it. Virtually everyone - with the notable exception of a few Australians - agrees that this is a fatal flaw.

[2] The "C" stands for "choice", referring to the axiom of choice which is part of this system.

Axiom of Restricted Comprehension:
Suppose X is a set and $\phi(x)$ is a property elements of X may possess. Then there exists a set

$$A = \{x \in X | \phi(x)\}$$

characterized by $x \in A \Leftrightarrow \phi(x)$.

V.3 Summary

The key lessons from this lecture are:

- Without any restrictions, naive set theory leads to Russell's paradox.
- In order to avoid Russell's paradox and similar problems, one often works with a formal set theory such as ZFC.
- In such a system, Russell's paradox is avoided by using a more careful formulation of the set-builder construction, namely the Axiom of Restricted Comprehension.

V.4 Exercises

Exercise 50. In a certain village there is a male barber who shaves every man who doesn't shave himself. Explain why this is a paradox and how it is related to Russell's paradox.

Exercise 51. Does there exist a set X such that $X \in X$? Why (not)?

Exercise 52. Suppose we define $X = \{U | U \text{ is a set}\}$. That is, X is the set consisting of all sets. Is this a paradox as well?

Exercise 53. Is an infinite regress necessarily a bad thing? Or are there situations where infinite regress is harmless?

Exercise 54. Suppose we write down a number of axioms which we think sets should satisfy. How do we know that the resulting theory is not inconsistent?

LECTURE VI

Subsets and Equality

We will now begin to explore some of the basic theory of sets. We have already seen the membership relation, and different methods for specifying sets. This lecture introduces another fundamental concept, that of the *subset relation*. We will also encounter another axiom of set theory, namely the Axiom of Extensionality, which governs equality between sets.

VI.1 Subsets

We often have occasion to express that one set is contained in another. For example, the natural numbers are contained in the integers. The following definition makes that precise:

Definition VI.1.1 (Subset)**.** Given two sets A, B, we say that A *is a subset of* B when every element of A is also an element of B. In this case we write $A \subseteq B$. In logical notation:

$$A \subseteq B \Leftrightarrow_{def} \forall x(x \in A \to x \in B).$$

How to prove it

To prove $A \subseteq B$: you need to:

- take an arbitrary element $x \in A$ and show that $x \in B$.

To refute $A \subseteq B$ it you must:

- Show that there exists an element $x \in A$ for which $x \notin B$.

If A is not a subset of B, we write $A \not\subseteq B$. Note that this is *not* the same thing as saying that $B \subseteq A$. (See the examples below.)

Examples VI.1.2.

1. We have $\{2, 3, 4\} \subseteq \{2, 3, 4, 5\}$.

LOGICAL ASPECTS

From the definition

$$A \subseteq B \quad \equiv_{def} \quad \forall x(x \in A \to x \in B)$$

We see that the *negation* of $A \subseteq B$ can equivalently be stated as

$$\begin{aligned} A \not\subseteq B \quad &\equiv_{def} \quad \neg\forall x(x \in A \to x \in B) \\ &\equiv \quad \exists x\neg(x \in A \to x \in B) \\ &\equiv \quad \exists x(x \in A \wedge x \notin B). \end{aligned}$$

2. We have $\{2,3,4\} \not\subseteq \{2,3\}$: the first set contains the element 4, but this is not an element of the second set.
3. We have $\mathbb{N} \subseteq \mathbb{Z}$, $\mathbb{Z} \subseteq \mathbb{Q}$, and so on. (This is a consequence of the definition of these sets of numbers.) We also have $\mathbb{Z} \not\subseteq \mathbb{N}$, for the first set has elements which are not in the second, for example -4.
4. We have $\{2,3\} \not\subseteq \{2,4\}$ and $\{2,4\} \not\subseteq \{2,3\}$.

VI.2 WHEN ARE TWO SETS EQUAL?

Consider the following sets: $A = \{4\}, B = \{12,4\}, C = \{4,12\}, D = \{4,12,4\}$. There is a clear sense in which A and B are different: A has one element, while B has two. Are B and C different? The only difference is that the elements are listed in a different order. And the only difference between C and D is that in D the element 4 is mentioned twice. We introduce the following axiom which defines a notion of *equality* between sets: it roughly says that sets are completely determined by their elements.

Axiom of Extensionality:
Two sets are equal precisely when they have the same elements. In logical notation:

$$X = Y \Leftrightarrow \forall z(z \in X \leftrightarrow z \in Y).$$

Thus two sets are distinct when one of them contains an element which is not contained in the other. Since our set B contains the element 12 and A does not, this

Logical Aspects

We have equivalent statements

$$\begin{aligned} X = Y &\equiv_{def} \forall z(z \in X \leftrightarrow z \in Y) \\ &\equiv \forall z((z \in X \to z \in Y) \wedge (z \in Y \to z \in X)) \\ &\equiv \forall z(z \in X \to z \in Y) \wedge \forall z(z \in Y \to z \in X). \end{aligned}$$

Thus the *negation* of $X = Y$ can equivalently be stated as

$$\begin{aligned} X \neq Y &\equiv_{def} \neg\forall z(z \in X \leftrightarrow z \in Y) \\ &\equiv \neg[\forall z(z \in X \to z \in Y) \wedge \forall z(z \in Y \to z \in X)] \\ &\equiv [\neg\forall z(z \in X \to z \in Y)] \vee [\forall z(z \in Y \to z \in X)] \\ &\equiv [\exists z\neg(z \in X \to z \in Y)] \vee [\exists z\neg(z \in Y \to z \in X)] \\ &\equiv [\exists z(z \in X \wedge z \notin Y)] \vee [\exists z(z \in Y \wedge z \notin X)]. \end{aligned}$$

means that $A \neq B$. However, we get $B = C$: each element in B can also be found in C and vice versa. Similarly we find $C = D$.

As with any definition, you should always ask yourself what it takes to prove that a certain example satisfies the definition or to show that it doesn't (see inset).

Exercise 55. Consider the sets $X = \{1, 2, 3\}, Y = \{0, 1, 0, 2, 0, 3\}, Z = \{3, 1, 2, 3\}$. Which of these sets are equal?

Exercise 56. Consider the sets $A = \{a\}, B = \{\{a\}\}$, and $C = \{a, \{a\}\}$. Are any of these sets equal? Which of them are subsets of which?

How to Prove It

In order to prove that $X = Y$ you need to:

- Take an arbitrary element $z \in X$ and prove that $z \in Y$, **AND**
- Take an arbitrary element $z \in Y$ and prove that $z \in X$.

In order to prove that $X \neq Y$ you need to:

- Show that there exists an element z with $z \in X$ but $z \notin Y$, **OR**
- Show that there exists an element z with $z \in Y$ but $z \notin X$

VI.3 Subsets and equality

The notion of subset is closely related to that of equality, as witnessed by the following proposition.

Proposition VI.3.1. *For any two sets A, B we have $A = B$ precisely when both $A \subseteq B$ and $B \subseteq A$.*

Proof. Let A and B be arbitrary sets. Assume first that $A = B$. Then consider $x \in A$. Since $A = B$, also $x \in B$ (by definition of equality of sets). This shows that $A \subseteq B$. Moreover, given $x \in B$ we get $x \in A$ (again since $A = B$) so that $B \subseteq A$.

For the other direction, assume $A \subseteq B$ and $B \subseteq A$. We must show that $A = B$. First consider an element $x \in A$. Since $A \subseteq B$, it follows that $x \in B$ as well. Next, consider an element $x \in B$. Since $B \subseteq A$, we get $x \in A$. Thus A and B have the same elements, whence $A = B$. □

The above proof was spelled out in a bit more detail than you'd normally see. The point is to make clear (a) what the structure of the proof is and (b) how we are using the various definitions involved. As we progress, we will start skipping some of the more familiar details in order to keep the proofs readable. But remember that you should always be able to fill in those details!

VI.4 Summary

We have introduced an axiom defining equality between sets:

- *Axiom of Extensionality*: two sets are equal when they have the same elements.

Moreover, we defined the subset relation and established the following elementary but important fact:

- $A = B$ if and only if $A \subseteq B$ and $B \subseteq A$.

One of the key things to keep in mind is that a definition implicitly (or sometimes explicitly) gives a recipe for proving or disproving that something satisfies the definition.

VI.5 Exercises

Exercise 57. Find all subsets of the set $X = \{a, b, c\}$.

Exercise 58. Find all subsets of $X = \{a, b, c, d\}$.

Exercise 59. Suppose X has n elements. How many subsets does X have?

Exercise 60. Consider the sets

$$A = \{a, b\}, \quad B = \{a, \{a\}, \{b\}\}, \quad C = \{\{a\}, \{b\}, a, b\}, \quad D = \{\{a\}, \{b\}, \{a, b\}\}.$$

Determine which of these sets are elements of which others. Also determine which are contained in which others.

Exercise 61. True or False? Any two sets which both have exactly 6 elements are equal. (If true, give a proof. If false, give a counterexample.)

Exercise 62. Prove that for any set A, we have $A \subseteq A$. Use this to prove that $A = A$.

Exercise 63. Prove that if $A \subseteq B$ and $B \subseteq C$, then also $A \subseteq C$. Use this to prove that if $A = B$ and $B = C$ then also $A = C$.

Exercise 64. Prove that $A = B$ implies $B = A$.

Exercise 65. We introduce the following notation:

$$A \subset B \Leftrightarrow_{def} A \subseteq B \text{ and } A \neq B.$$

Show that $A \subset B$ and $B \subset C$ imply $A \subset C$.

Exercise 66. (This exercise uses the same notation as the previous one.) Is the following true? $A \subset B$ and $B \subseteq C$ implies $A \subset C$. Give a proof or a counterexample.

Exercise 67. Same question, but now for $A \subset B$ and $B \subseteq C$ implies $A \subseteq C$.

Exercise 68. Is it possible for two sets A, B that $A \subset B$ and $B \subset A$ at the same time? Why (not)?

Exercise 69. Is it possible for two sets A, B that $A \subseteq B$ and $A \in B$ at the same time? Why (not)?

Exercise 70. Find sets A, B, C such that $A \in B$, $B \in C$ and also $A \in C$. Is it generally the case that $A \in B$ and $B \in C$ implies $A \in C$? Explain why (not).

LECTURE VII

Existence

In the previous lectures we encountered some well-known sets, such as the set of integers. However, we got ahead of ourselves there: the purpose was to build up mathematics using sets and sets alone, and when we introduced sets such as $\mathbb{Z}$ we were presupposing that we were already given the integers.

So let's back up, and not assume anything about certain mathematical objects already being handed to us. What kind of sets can we then show to exist? Unfortunately, there is no reason why there would be anything at all! This lecture introduces axioms which guarantee the existence of many interesting sets.

VII.1 The empty set

In order to get things off the ground, we have to postulate the existence of at least one set.

Axiom of Existence:
There exists a set $\emptyset$ which has no elements.

It may seem strange to introduce an empty set rather than a more exciting one, and we'll have to see whether this is a useful axiom at all[1]. We'll see that together with the powerset construction defined in the next section we can indeed build many other sets.

[1] It is possible to derive this axiom from the other axioms which we will introduce, but there is only value in doing this if you're interested in obtaining a set of axioms which is somehow minimal.

For now, let's study empty sets. First, we note:

Lemma VII.1.1. *For any set A we have $\emptyset \subseteq A$.*

Proof. We need to show that any element of $\emptyset$ is also an element of A. But there are no elements in $\emptyset$, so there is nothing to check[2]. □

That is, the empty set is a subset of any set.

Proposition VII.1.2. *The empty set is unique. That is, there is exactly one empty set.*

Proof. We already know (because we took it as an axiom) that there exists at least one empty set. So all we need to make sure is that there can't be more than one. The way to do this is to assume that we have two empty sets, and then to prove that they must be equal.

So suppose that $\emptyset$ and $\emptyset'$ are both empty sets. By Lemma VII.1.1, we know that an empty set is contained in any other set. Using this twice, we find $\emptyset \subseteq \emptyset'$ and $\emptyset' \subseteq \emptyset$. But by Proposition VI.3.1, it follows that $\emptyset = \emptyset'$. □

LOGICAL ASPECTS

The statement $\emptyset \subseteq A$ means:

$$\forall x(x \in \emptyset \to x \in A).$$

However, $x \in \emptyset$ is always false, so this becomes

$$\forall x(\bot \to x \in A).$$

An implication $\bot \to p$ is always true, and therefore the whole statement is also true.

VII.2 THE POWERSET AXIOM

We have introduced the empty set, but asking for more isn't exactly unreasonable. The *powerset* construction allows us to form new sets from old: when X is a set, define a new set by

$$\mathcal{P}(X) = \{U | U \subseteq X\}.$$

That is, $\mathcal{P}(X)$ is defined to be the set consisting of all subsets of X. $\mathcal{P}(X)$ is called the *powerset* of X. One of the axioms of set theory now stipulates that this set actually exists[3]:

Powerset Axiom:
Given a set X, there exists a set $\mathcal{P}(X)$ with the property that $U \in \mathcal{P}(X)$ if and only if $U \subseteq X$.

[2] Statements which are true because there are no cases to verify are sometimes called *vacuously true*.

[3] This is one of the more controversial axioms, and there are mathematicians who reject it in this general form; a suitable weakening gives rise to a formulation of set theory called *predicative set theory*.

Some examples will illustrate this (see the first three exercises of the previous lecture if you have trouble understanding these):

Examples VII.2.1.

1. When $X = \{a\}$, we have $\mathcal{P}(X) = \{\emptyset, \{a\}\}$.
2. When $X = \{a, b\}$, we have $\mathcal{P}(X) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.
3. When $X = \{a, b, c\}$, we have
$$\mathcal{P}(X) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$
4. $\mathcal{P}(\emptyset) = \{\emptyset\}$. Indeed, the empty set has exactly one subset, namely itself. This is the only element of $\mathcal{P}(\emptyset)$.
5. $\mathcal{P}(\mathcal{P}(\emptyset)) = \mathcal{P}(\{\emptyset\}) = \{\emptyset, \{\emptyset\}\}$. (Special case of the first example.)

The following lemma is an easy consequence of the definition (check for yourself!):

Lemma VII.2.2. *For any set X, we have*
$$\emptyset \in \mathcal{P}(X) \qquad \text{and} \qquad X \in \mathcal{P}(X).$$

Exercise 71. Suppose a set X has n elements. How many elements does $\mathcal{P}(X)$ have?

Exercise 72. Find $\mathcal{PP}(\{0, 1\})$.

As can be seen from the examples, the powerset axiom gives us genuinely new sets, such as $\{\emptyset\}$ (check for yourself that this set is indeed different from $\emptyset$!).

Now what about a set such as $A = \{\{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}$? This is not a powerset of any other set (check this!), so its existence does not follow directly from the Powerset Axiom. However, the Axiom of Comprehension helps here. First, we do have the set
$$B = \mathcal{PPP}(\emptyset) = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}.$$
We now use comprehension to define B as a subset of A:
$$A = \{U \in B | U \neq \emptyset\}.$$
This proves that A exists!

Exercise 73. Use comprehension to show that $C = \{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}\}$ exists.

VII.3 Properties of powersets

We will now prove a typical result concerning powersets. You should mainly focus on the methods of proof used, as these are good examples of systematically unpacking definitions and using logic to draw the desired conclusions.

Proposition VII.3.1. *For any sets A, B we have $A \subseteq B$ if and only if $\mathcal{P}(A) \subseteq \mathcal{P}(B)$.*

Proof. Assume A, B are arbitrary sets. First, we prove $A \subseteq B \Rightarrow \mathcal{P}(A) \subseteq \mathcal{P}(B)$. Assume $A \subseteq B$. This means: for all x, if $x \in A$ then also $x \in B$.

We must show $\mathcal{P}(A) \subseteq \mathcal{P}(B)$, meaning: for each U: if $U \in \mathcal{P}(A)$ then also $U \in \mathcal{P}(B)$. By definition of powerset, this is equivalent to: for each U: if $U \subseteq A$ then also $U \subseteq B$.

So assume that we have U with $U \subseteq A$. To show $U \subseteq B$, consider an element $x \in U$. Since $U \subseteq A$, it follows that $x \in A$. By the hypothesis $A \subseteq B$, it now follows that $x \in B$ as well. Thus $x \in U$ implies $x \in B$, so $U \subseteq B$, as needed. This shows that $U \subseteq A$ implies $U \subseteq B$.

For the other direction, assume $\mathcal{P}(A) \subseteq \mathcal{P}(B)$, or equivalently that for any U, $U \subseteq A$ implies $U \subseteq B$. In particular, this holds for $U = A$; as $A \subseteq A$ we get $A \subseteq B$ as desired. □

The first part of this result is often expressed by saying that $\mathcal{P}$ is a *monotone* operation. (A general notion of monotonicity will be discussed in Lecture XX.)

Note that the second part involves a little trick: in order to use the hypothesis that $\forall U(U \subseteq A \to U \subseteq B)$, we need to pick a suitable U. This is why proving things is not always an automatic enterprise: sometimes clever choices need to be made.

The first part of the proof is more straightforward, but one has to be systematic. In order to highlight both the logical structure of the proof and the "dynamics", i.e., the order in which things unfold, Figure VII.3 gives the structure of the proof in tabular format.

VII.4 Summary

This lecture introduced two new axioms:

- *Axiom of Existence*: an empty set exists
- *Powerset Axiom*: the powerset of any set exists

Using these axioms, we can build up a stockpile of sets. Using comprehension, we can also show various other sets to exist (as subsets of powersets). This usually involves coming up with a description which singles out the elements of the subset which is being defined.

We also established the uniqueness of the empty set, and proved that the powerset operation is *monotone*, meaning that $A \subseteq B$ implies $\mathcal{P}(A) \subseteq \mathcal{P}(B)$.

Given/Assumed	To Prove	Auxiliary	Justification
$--$	$A \subseteq B \to \mathcal{P}(A) \subseteq \mathcal{P}(B)$		
$A \subseteq B$	$\mathcal{P}(A) \subseteq \mathcal{P}(B)$		proof strat. for $\to$
$\forall x(x \in A \to x \in B)$	$\forall U(U \in \mathcal{P}(A) \to U \in \mathcal{P}(B))$		def. of $\subseteq$
$\forall x(x \in A \to x \in B)$	$\forall U(U \subseteq A \to U \subseteq B)$		def. of $\mathcal{P}$
$\forall x(x \in A \to x \in B)$ $U -$ arbitrary	$U \subseteq A \to U \subseteq B$		proof strat. for $\forall$
$\forall x(x \in A \to x \in B)$ $U -$ arbitrary $U \subseteq A$	$U \subseteq B$		proof strat. for $\to$
$\forall x(x \in A \to x \in B)$ $U -$ arbitrary $\forall x(x \in U \to x \in A)$	$\forall x(x \in U \to x \in B)$		def. of $\subseteq$
$\forall x(x \in A \to x \in B)$ $U -$ arbitrary $\forall x(x \in U \to x \in A)$ $x -$ arbitrary	$x \in U \to x \in B$		proof strat. for $\forall$
$\forall x(x \in A \to x \in B)$ $U -$ arbitrary $\forall x(x \in U \to x \in A)$ $x -$ arbitrary $x \in U$	$x \in B$		proof strat. for $\to$
$\boxed{\forall x(x \in A \to x \in B)}$ $U -$ arbitrary $\boxed{\forall x(x \in U \to x \in A)}$ $x -$ arbitrary $x \in U$	$x \in B$	$\boxed{x \in A \to x \in B}$ $\boxed{x \in U \to x \in A}$	instantiation
$\forall x(x \in A \to x \in B)$ $U -$ arbitrary $\forall x(x \in U \to x \in A)$ $x -$ arbitrary $\boxed{x \in U}$	$x \in B$	$x \in A \to x \in B$ $\boxed{x \in U \to x \in A}$ $\boxed{x \in A}$	Modus Ponens
$\forall x(x \in A \to x \in B)$ $U -$ arbitrary $\forall x(x \in U \to x \in A)$ $x -$ arbitrary $x \in U$	$\boxed{x \in B}$	$\boxed{x \in A \to x \in B}$ $x \in U \to x \in A$ $\boxed{x \in A}$	Modus Ponens

Figure VII.1: A proof that $A \subseteq B$ implies $\mathcal{P}(A) \subseteq \mathcal{P}(B)$

VII.5 Exercises

Exercise 74. For each of the following sets, describe all elements in that set.

(a) $\mathcal{P}(\mathcal{P}(\emptyset))$

(b) $\mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset)))$

(c) $\{x \in \mathcal{P}(\mathcal{P}(\emptyset)) | x \subseteq \mathcal{P}(\emptyset)\}$

(d) $\{x \in \mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) | \emptyset \in x\}$

(e) $\{x \in \mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) | \{\emptyset\} \in x\}$

(f) $\{x \in \mathcal{P}(\mathcal{P}(\mathcal{P}(\emptyset))) | \mathcal{P}(\emptyset) \subseteq x\}$

Exercise 75. Which of the following statements are true?

(a) $\emptyset \in \mathcal{P}(\emptyset)$

(b) $\emptyset \subseteq \mathcal{P}(\emptyset)$

(c) $\mathcal{P}(\emptyset) \in \mathcal{P}\mathcal{P}(\emptyset)$

(d) $\mathcal{P}(\emptyset) \in \mathcal{P}\mathcal{P}\mathcal{P}(\emptyset)$

(e) $\mathcal{P}(\emptyset) \subseteq \mathcal{P}\mathcal{P}\mathcal{P}(\emptyset)$

(f) $\{X\} \in \mathcal{P}(X)$

(g) $\{X\} \subseteq \mathcal{P}(X)$

(h) $\{X\} \in \mathcal{P}(\{X\})$

(i) $\{X\} \subseteq \mathcal{P}(\{X\})$

Exercise 76. Show that $A = B$ implies $\mathcal{P}(A) = \mathcal{P}(B)$. Does the converse hold?

Exercise 77. Is the set $\{a, \{a\}, \{\{a\}\}, \{\{\{a\}\}\}\}$ a powerset?

Exercise 78. Can there exist a set X with $\mathcal{P}(X) \subseteq X$? If yes, give an example. If no, explain why not.

Exercise 79. Prove that the existence of an empty set can also be derived from the powerset axiom together with comprehension.

Exercise 80. Can there exist sets A, B, C with $A \subseteq B \subseteq C, A \in B \in C$ and $A \in C$? Give an example or prove that this is impossible.

Exercise 81. Can there exist sets A, B with $\{A\} \subseteq \{B\}$ and $B \in A$? Give an example or prove that this is impossible.

LECTURE VIII

Operations

This lecture introduces the so-called *Boolean operations* on sets[1]. Not only do these occur a lot in practice and help us defining new sets, they are also closely connected with logical reasoning, as we shall see.

VIII.1 Four operations

The simplest of the operations we introduce is that of *(binary) intersection*: given sets A, B we define

$$A \cap B =_{def} \{x | x \in A \wedge x \in B\}.$$

Next, the *(binary) union* of two sets A, B is defined to be

$$A \cup B =_{def} \{x | x \in A \vee x \in B\}.$$

Here, the word *or* is an "inclusive or", i.e. a disjunction.

The operation of *complementation* only makes sense when we work relative to a universe $\mathcal{U}$. Then we may define, for a set A:

$$A^c =_{def} \{x \in \mathcal{U} | x \notin A\}.$$

Finally, we define the *difference* of two sets A and B as

$$A - B =_{def} \{x | x \in A \wedge x \notin B\}.$$

[1] Named after George Boole (1815–1864), author of the landmark text *The Laws of Thought*, which develops the principles which govern both propositional logic and the operations on sets we discuss here.

Figure VIII.1 illustrates all four operations using Venn diagrams. In each of the four cases, the doubly shaded region is the one being defined.

As you can see, these operations are defined in terms of the logical connectives $\wedge, \vee, \neg$. This means that if we wish to prove things about intersections, unions, differences and complements, we will have to use propositional logic.

Before moving on to examples, we introduce one more important definition:

Definition VIII.1.1 (Disjointness)**.** Two sets A and B are called *disjoint* when $A \cap B = \emptyset$.

In plain English: A and B are disjoint precisely when they have no elements in common.

VIII.2 Examples

We will illustrate the Boolean operations on sets by means of a couple of examples. The examples in this section are presented relatively informally to give you the general idea of how these operations work; in the next section we will look at some detailed proofs.

Let's begin with a very concrete example. Suppose we have the following sets:

$$A = \{a, b, c\}, \qquad B = \{b, c, d, e\}, \qquad C = \{c, e, f, g\}.$$

Then we have

- $A \cap B = \{b, c\}$
- $A \cap C = \{c\}$
- $B \cap C = \{c, e\}$
- $A \cup B = \{a, b, c, d, e\}$
- $A \cup C = \{a, b, c, e, f, g\}$
- $B \cup C = \{b, c, d, e, f, g\}$
- $A - B = \{a\}$
- $A - C = \{a, b\}$
- $B - C = \{b, d\}$
- $A \cap B \cap C = \{c\}$
- $A \cup B \cup C = \{a, b, c, d, e, f, g\}$.

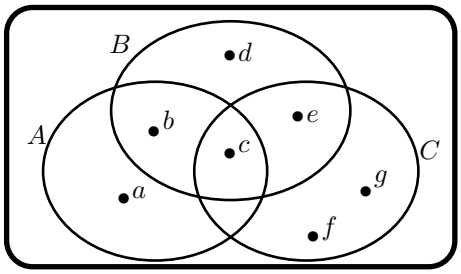
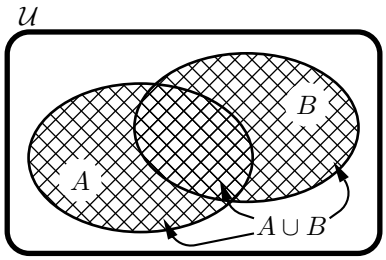
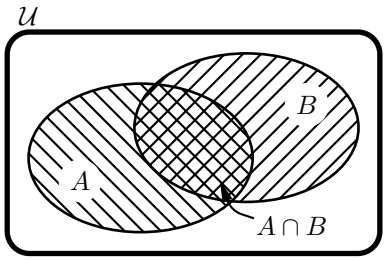


Figure VIII.2: Three sets



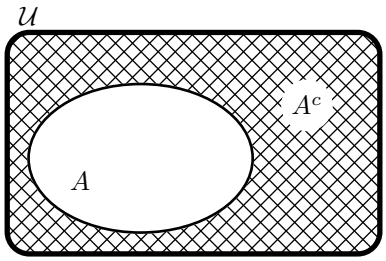
Union

Definition: $A \cup B = \{x | x \in A \vee x \in B\}$
English: The elements in A **or** in B (or both)
Diagram: Region covered by A and B together



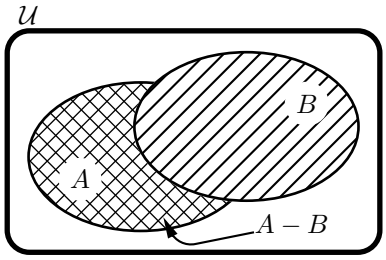
Intersection

Definition: $A \cap B = \{x | x \in A \wedge x \in B\}$
English: The elements both in A **and** in B
Diagram: Overlap between A and B



Complement

Definition: $A^c = \{x \in \mathcal{U} | a \notin A\}$
English: The elements which are **not** in A
Diagram: Everything outside A



Difference

Definition: $A - B = \{x | x \in A \wedge x \notin B\}$
English: Those elements in A **but not** in B
Diagram: A minus the overlap between A and B

Figure VIII.1: Union, Intersection, Complement and Difference.

For a somewhat more involved question based on the same example, we try to determine the set $((A \cap B) - C) \cup (C - (A \cup B))$. You can do this in steps:

1. First, $A \cap B = \{b, c\}$, so $(A \cap B) - C = \{b, c\} - \{c, e, f, g\} = \{b\}$.
2. Next, $A \cup B = \{a, b, c, d, e\}$, so $C - (A \cup B) = \{c, e, f, g\} - \{a, b, c, d, e\} = \{f, g\}$.
3. Finally, the set we're looking for is

$$((A \cap B) - C) \cup (C - (A \cup B)) = \{b\} \cup \{f, g\} = \{b, f, g\}.$$

You may wish to verify this by determining the corresponding region in the Venn diagram.

Here's a different example: suppose we are working in the universe $\mathcal{U} = \mathbb{N}$, the set of natural numbers. Define

$$X = \{x \in \mathbb{N} | x \text{ is prime}\}, \qquad Y = \{x \in \mathbb{N} | x < 8\}, \qquad Z = \{x \in \mathbb{N} | x \text{ is odd}\}.$$

Then we have the following calculations:

1. $X \cap Y \cap Z = \{x \in \mathbb{N} | x \text{ is prime and is odd and } x < 8\} = \{3, 5, 7\}$.
2. $(X^c - Z) \cap Y = \{x \in \mathbb{N} | x \text{ is not prime and } x \text{ is not odd and } x < 8\} = \{0, 4, 6\}$
3. $$\begin{aligned} ((Y - X^c) \cup (X \cup Z)^c) - Y &= (\{2, 3, 5, 7\} \cup \{1, 2, 3, 5, 7, 9, 11, 13, \ldots\}^c) - Y \\ &= \{0, 4, 6, 8, 10, 12, \ldots\} - Y \\ &= \{8, 10, 12, \ldots\}. \end{aligned}$$

VIII.3 Boolean operations and propositional logic

As remarked above, the definitions of the Boolean operations involve the propositional connectives. Suppose that we use the operations to form, starting with some sets $A_1, \ldots, A_n$, a new set A. (For example, A could be something like $A = A_2 \cup (A_5^c - A_3)$.) Now we ask what it means for an element x to satisfy $x \in A$. When we unpack this using the definitions, we get some complicated expression built from statements of the form $x \in A_i$ using the connectives $\wedge, \vee, \neg$ (or their English counterparts). In the example, we'd have

$$x \in A \Leftrightarrow x \in A_2 \vee (\neg x \in A_5 \wedge \neg x \in A_3).$$

To make this a bit more readable, we introduce proposition letters α_i (intended interpretation: $x \in A_i$). Then the statement $x \in A$ becomes

$$\alpha_2 \vee (\neg \alpha_5 \wedge \neg \alpha_3).$$

Let A, B be sets built from sets $A_1, \ldots, A_n$ are sets using Boolean operations.
Step 1. Introduce propositional letters $\alpha_1, \ldots, \alpha_n$.
Step 2. Define a translation τ by

$$\begin{array}{rcl} \tau(A_i) & = & \alpha_i \\ \tau(X \cap Y) & = & \tau(X) \wedge \tau(Y) \\ \tau(X \cup Y) & = & \tau(X) \vee \tau(Y) \\ \tau(X^c) & = & \neg\tau(X) \\ \tau(X - Y) & = & \tau(X) \wedge \neg\tau(Y) \end{array}$$

Step 3. To verify whether $A \subseteq B$, check whether $\tau(A)$ logically implies $\tau(B)$; equivalently, check whether $\tau(A) \to \tau(B)$ is a tautology. To verify whether $A = B$, check whether $\tau(A)$ and $\tau(B)$ are logically equivalent; equivalently, check whether $\tau(A) \leftrightarrow \tau(B)$ is a tautology.

Figure VIII.3: Translating sets to propositions

Next, suppose we have another set B, also built from the sets $A_1, \ldots, A_n$. (For concreteness, suppose say that $B = (A_3 - A_2)^c \cup (A_5^c \cap A_2)$.) Then we could also translate the statement $x \in B$ as

$$\neg(\alpha_3 \wedge \neg\alpha_2) \vee (\neg\alpha_5 \wedge \alpha_2).$$

Finally, suppose we were wondering whether $A \subseteq B$. That is, we want to know whether $x \in A$ implies $x \in B$. This now reduces to verifying whether the implication

$$[\alpha_2 \vee (\neg\alpha_5 \wedge \neg\alpha_3)] \to [\neg(\alpha_3 \wedge \neg\alpha_2) \vee (\neg\alpha_5 \wedge \alpha_2)]$$

is a tautology, something we can check using logical manipulations or truth-tables.

Figure VIII.3 summarizes the procedure.

As an example, we prove the statement $A - (B \cup C) = (A - B) \cap (A - C)$ in two ways; one informally, and one using logical notation.

Proof 1. Consider first an arbitrary $x \in A - (B \cup C)$. By def. of difference this means $x \in A$ and $x \notin B \cup C$. If x is not an element of $B \cup C$ then we can't have $x \in B$ and we also can't have $x \in C$. So $x \in A$, $x \notin B$ and $x \notin C$. But then $x \in A - B$ and $x \in A - C$, so that $x \in (A - B) \cap (A - C)$.

For the other direction, let $x \in (A - B) - (A - C)$. Then $x \in (A - B)$ and $x \in (A - C)$. The first means that $x \in A$ and $x \notin B$. The second means $x \in A$ and $x \notin C$. From $x \notin B$ and $x \notin C$ we may infer $x \notin B \cup C$ (for if $x \in B \cup C$ we would have $x \in B$ or $x \in C$, neither of which is possible). Hence $x \in A$ and $x \notin B \cup C$, i.e., $x \in A - (B \cup C)$. □.

Proof 2. Translated to propositional logic, we see that we must prove the logical equivalence

$$\alpha \wedge \neg(\beta \vee \gamma) \equiv (\alpha \wedge \neg\beta) \wedge (\alpha \wedge \neg\gamma).$$

But we have

$$\begin{array}{rcll} \alpha \wedge \neg(\beta \vee \gamma) & \equiv & \alpha \wedge (\neg\beta \wedge \neg\gamma) & \text{by De Morgan} \\ & \equiv & (\alpha \wedge \alpha) \wedge (\neg\beta \wedge \neg\gamma) & \text{by idempotence of } \wedge \\ & \equiv & (\alpha \wedge \neg\beta) \wedge (\alpha \wedge \neg\gamma) & \text{by associativity and commutativity of } \wedge \end{array}$$

□

Note that in the first proof we're systematically unpacking the definitions of $\cap, \cup$ and $-$ so that we obtain statements in English involving "and", "or" and "not". Then we're using propositional logic informally to manipulate these statements, and then repack them in terms of Boolean operations again. The second proof bypasses all this translating and is hence much more efficient. However, you should understand both types of proof!

VIII.4 Summary

The *Boolean operations* of *intersection, union, complement* and *difference* allow us to form new combinations of sets.

In reasoning about such combinations, we may

- Translate statements into English and use informal logic, or
- Translate statements into formal propositional logic. We can then use logical methods such as truth-tables to establish results.

VIII.5 Exercises

Exercise 82. Let $A = \{0, 1, 2\}$, $B = \{2, 3, 4, 5\}$, $C = \{1, 3, 5\}$, and suppose that $\mathcal{U} = \mathbb{N}$. Find the following sets:

(a) $A \cap B$

(b) $A \cup B$

(c) $A \cap B \cap C$

(d) $A - B$

(e) $A^c - B^c$

(f) $(A - B)^c$

(g) $(A^c \cap C^c)^c$

Exercise 83. Give a simple description of each of the following sets: (the universal set is $\mathbb{R}$)

(a) $\mathbb{Q}^c - \{x \in \mathbb{R} | x > 0\}$

(b) $(\mathbb{Q} \cap \{x \in \mathbb{R} | 0 < x < 1\})^c - \{x \in \mathbb{R} | e^x \leq 1\}$

(c) $[0,1] - \{x \in \mathbb{R} | x^3 < 1\}$ (here, $[0,1]$ denotes the closed unit interval)

Exercise 84. Show that for any set A, we have:

(a) $A \cap \emptyset = \emptyset$

(b) $A \cup \emptyset = A$

(c) $A \cup A = A$

(d) $A \cap A = A$

(e) $A - A = \emptyset$

Exercise 85. Show that $A \subseteq B$ if and only if $A \cap B = A$

Exercise 86. Show that $A \subseteq B$ if and only if $A \cup B = B$

Exercise 87. Prove the following identities for any sets A, B, C;

(a) $(A \cap B) \cap C = A \cap (B \cap C)$

(b) $A \cap B = B \cap A$

(c) $(A \cup B) \cup C = A \cup (B \cup C)$

(d) $A \cup B = B \cup A$

Exercise 88. True or false? $A - B = B - A$. If true, give a proof. If false, give a counterexample.

Exercise 89. True or false? $A - (B - C) = (A - B) - C$. If true, give a proof. If false, give a counterexample.

Exercise 90. Prove: $(A - B) \cap (B - A) = \emptyset$

Exercise 91. Define the *symmetric difference* of A and B to be

$$A \Delta B =_{def} (A - B) \cup (B - A).$$

(a) Show that $x \in A \Delta B$ if and only if $x \in A \leftrightarrow x \notin B$ is true.

(b) Show that $A \Delta A = \emptyset$.

(c) Show that $(A \Delta B) = (A \cup B) - (A \cap B)$.

(d) Show that $A \Delta B = B \Delta A$.

(e) Prove or disprove: $(A \Delta B) \Delta C = A \Delta (B \Delta C)$.

Exercise 92. Prove or disprove: $\mathcal{P}(A \cap B) = \mathcal{P}(A) \cap \mathcal{P}(B)$

Exercise 93. Prove or disprove: $\mathcal{P}(A \cup B) = \mathcal{P}(A) \cup \mathcal{P}(B)$

Exercise 94. Prove or disprove: $\mathcal{P}(A - B) = \mathcal{P}(A) - \mathcal{P}(B)$

LECTURE IX

Products and Sums

The Boolean operations we introduced in the last lecture helped us construct a variety of new sets. This lecture introduces two more ways of building new sets. These operations are called *Cartesian product* and *disjoint sum* (or *coproduct*).

IX.1 Products and pairs

Let A, B be sets. Informally, we wish to define a new set $A \times B$ whose elements are pairs (a, b), where $a \in A$ and $b \in B$. There are two questions: first, what are pairs (a, b)? (This is a genuine issue because we wish to build everything from things we have already.) And second, how do we know that all such pairs form a set?

In order to answer the first question, let's see what we expect of pairs. First, we expect that any elements $a \in A, b \in B$ can be combined to form a pair (a, b). Second, we want that given a pair (a, b), we can retrieve a and b. In particular, given two pairs (a, b) and (a', b'), we want that these two pairs are equal precisely when $a = a'$ and $b = b'$. Thus a pair should be completely determined by its two coordinates.

Can we, given element a, b as above, create a set (a, b) meeting these desiderata? Your first guess might be to use $\{a, b\}$. But there is a problem: suppose A and B are actually the same set. Since $\{a, b\} = \{b, a\}$, we can't tell apart the pairs (a, b) and (b, a). Since these are different pairs we want to represent them by different sets.

One trick, due to Kuratowski, is to use the set $\{a, \{a, b\}\}$ instead. Note that we can never have $a = \{a, b\}$, because that would give $a \in a$. Thus the set $\{a, \{a, b\}\}$ has two distinct elements. Moreover, the sets $\{a, \{a, b\}\}$ and $\{b, \{a, b\}\}$ are distinct when a and b are. Thus the criteria are met, and we can set $(a, b) = \{a, \{a, b\}\}$.

Now we define

Definition IX.1.1 (Cartesian product)**.** Let A, B be sets. Then the *Cartesian product* of A and B is the set

$$A \times B = \{(a, b) | a \in A, b \in B\}.$$

A few remarks: first, the trick used above is not the only possible one. What really matters is that pairs, no matter how you define them, behave the way you want them to. Second, we should ask ourselves whether $A \times B$, as defined, exists! Well, note first that $\{a, \{a, b\}\}$ is a subset of $A \cup \mathcal{P}(A \cup B)$. Thus the set $A \times B$ of all such pairs is a subset of $\mathcal{P}(A \cup \mathcal{P}(A \cup B))$. The latter exists, so we may obtain $A \times B$ via comprehension.[1]

Example IX.1.2. Suppose $A = \{a, b\}$ and $B = \{b, c, d\}$. Then

$$A \times B = \{(a, b), (a, c), (a, d), (b, b), (b, c), (b, d)\}.$$

When we want to work out what this means using the Kuratowski definition of pair (normally we don't want to do that, but suppose that we care), we get

$$A \times B = \{\{a, \{a, b\}\}, \{a, \{a, c\}\}, \{a, \{a, d\}\}, \{b, \{b\}\}, \{b, \{b, c\}\}, \{b, \{b, d\}\}\}.$$

The product operation $\times$ behaves a bit like multiplication of numbers. However, not literally so: we don't have $A \times (B \times C) = (A \times B) \times C$ in general, nor do we have $A \times B = B \times A$. (Try to find the smallest possible counterexamples!) However, a slightly modified form of associativity and commutativity still holds – we will revisit this in Lecture XII.

IX.2 Coproducts

We now turn to the "dual" operation, namely sums (also called coproducts). Given two sets A and B, we wish to combine them in such a way that the resulting set has both the elements of A and the elements of B, and in such a way that we can still tell which are which. Taking the union $A \cup B$ is not quite accurate, because when A and B are not disjoint (meaning: $A \cap B \neq \emptyset$), we are losing track of where the elements in the overlap came from.

The solution is to first make sure that A and B are disjoint by giving the elements a unique identifying label. For example, we can give elements from A the label 0 and elements from B the label 1. (The actual labels don't matter, as long as they are distinct. We may let $0 = \emptyset$ and $1 = \{\emptyset\}$ for example.)

Definition IX.2.1. Given sets A, B define

$$A + B = A \times \{0\} \cup B \times \{1\}.$$

Here are some concrete examples:

Examples IX.2.2.

[1] However, there are mathematicians (especially those who have doubts about powersets) who simply introduce an axiom stating that product sets exist.

1. Let $A = \{a, b\}$ and $B = \{b, c, d\}$. Then
$$A + B = \{(a, 0), (b, 0), (b, 1), (c, 1), (d, 1)\}.$$
Note how the element b is represented twice in the sum: once with label 0 and once with label 1.
2. Even when A and B are disjoint, the coproduct is not the same as the union. For example, when $A = \{a, b\}$ and $B = \{c, d\}$, we have $A+B = \{(a, 0), (b, 0), (c, 1), (d, 1)\}$. So we always make sure to label everything, even when it is not necessary.
3. $\mathbb{N} + \emptyset = \{(x, 0)|x \in \mathbb{N}\}$.
4. $\{\emptyset\} + \{\emptyset\} = \{(\emptyset, 0), (\emptyset, 1)\}$.

Again, there is a sense in which + behaves like addition of numbers, but this will be made more precise in Lecture XII.

IX.3 Summary

Two new constructions were defined:

- *Cartesian product* of sets $A \times B = \{(a, b)|a \in A, b \in B\}$.
- *Disjoint union* of sets $A + B = A \times \{0\} \cup B \times \{1\}$.

The first is defined using the *Kuratowski definition of pairs*: $(a, b) = \{a, \{a, b\}\}$. The second is defined by first making the sets disjoint and then taking their union.

IX.4 Exercises

Exercise 95. Compute $\{\emptyset\} \times \{\emptyset\}$.

Exercise 96. Is the set $\{\{\emptyset, \{\emptyset\}\}, \{\{\emptyset\}, \{\{\emptyset\}\}\}\}$ a product of two other sets?

Exercise 97. For $A = \{a, b, \{a\}\}, B = \{\{a, b\}, b\}$, find $A \times B$.

Exercise 98. Give examples (as economical as possible) to show that generally $A \neq A \times A$, that $A \times B \neq B \times A$ and that $A \times (B \times C) \neq (A \times B) \times C$. Also find the exception to these statements.

Exercise 99. Compute $\{\emptyset, \{\emptyset\}\} + \{\emptyset\}$.

Exercise 100. Do we have $A + B = B + A$? $A + (B + C) = (A + B) + C$? $A + \emptyset = A$?

Exercise 101. Compute $A \times (B + C)$, where $A = \{a, b\}, B = \{a, c\}, C = \{b, c\}$.

Exercise 102. Prove that if $A \subseteq A'$ and $B \subseteq B'$ then also $A \times B \subseteq A' \times B'$

Exercise 103. Prove that if $A \subseteq A'$ and $B \subseteq B'$ then also $A + B \subseteq A' + B'$

Exercise 104. Investigate the validity of the following:

(a) $A \times (B \cup C) = (A \times B) \cup (A \times C)$

(b) $A \times (B \cap C) = (A \times B) \cap (A \times C)$

(c) $A + (B \cup C) = (A + B) \cup (A + C)$

(d) $A + (B \cap C) = (A + B) \cap (A + C)$

(e) $(A + B) \cup C = (A \cup C) + (B \cup C)$

(f) $(A \times B) \cap C = (A \cap C) \times (B \cap C)$.

Exercise 105. Can it ever happen that $A + B = A \times B$? Find all possible examples, and prove that your list is exhaustive.

LECTURE X

Relations

In this lecture we start the study of *relations* between sets. Our first concern is to define relations, to get a feel for what kind of relations exist, and to see how they may be combined.

X.1 Basic definition

Suppose we are given two sets A and B.

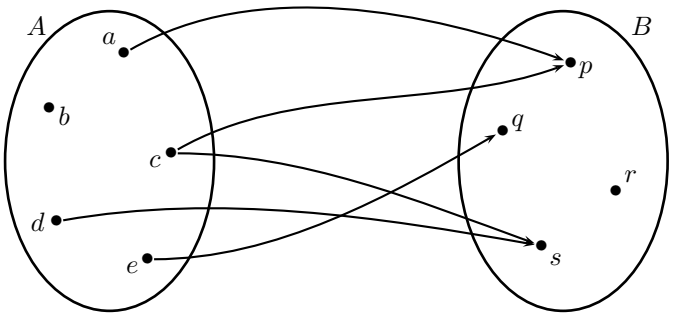
Definition X.1.1. A *(binary*[1]*) relation* from a set A to a set B is a subset R of $A \times B$.

There is a variety of notations in use in connection to relations; when $(a, b) \in R$, one often writes $R(a, b)$ or aRb. One says that "a is R-related to b" (or simply that a is related to b when the relation R is understood). To indicate that $(a, b) \notin R$, we also write $\neg R(a, b)$ or $\neg aRb$. Also, some texts use a special arrow like $\nrightarrow$ to indicate relations: in that case, $R : A \nrightarrow B$ means: R is a relation from A to B.

Note two things: first, that a relation is a very general thing: *any* subset of $A \times B$ is a relation from A to B, no matter how strange it looks. Thus there are generally many different relations from one set to another.

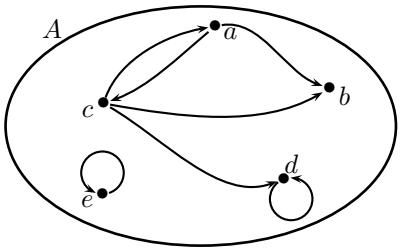
Second, a relation from A to B is not the same thing as a relation from B to A. Relations from A to B and relations from B to A are related (in a sense we will explore in more detail later) but they are manifestly different and should be carefully distinguished. Of course, nothing prevents considering the special case when $A = B$; in

[1] More generally, one may define n-ary relations for any $n \in \mathbb{N}$; a ternary relation is a subset of $A_1 \times A_2 \times A_3$, and so on. A special case is $n = 1$; a unary relation is then simply a subset. You may want to think about what a nullary relation should be!



A relation
$R \subseteq A \times B$ with
$R = \{(a,p),(c,p),$
$(c,s),(d,s),(e,q)\}$

Figure X.1: Example of a relation R from A to B



A relation $R \subseteq A \times A$ with
$R = \{(a,b),(a,c),(c,a),(c,b),$
$(c,d),(d,d),(e,e)\}$

Figure X.2: Example of a relation R from A to A

that case a relation $R \subseteq A \times A$ is called a (binary) relation on A, or a relation from A to itself.

Relations may also be rendered pictorially. See Figure X.1 for an example, and Figure X.2 for an example of a relation from a set to itself. Note the directionality of the connections, which is necessary to remove ambiguity.

Note that there is nothing preventing a particular element being related to several other elements, including itself. Note also that some elements may not be related to anything. Indeed, we have the following extreme examples:

Example X.1.2. Given sets A, B, the empty set (which is a subset of $A \times B$) is a relation from A to B. It is called the *empty relation*. On the other extreme, $A \times B$ itself is a subset of $A \times B$, and hence a relation from A to B. This is called the *maximal relation*.

No element of A is related to any of the elements of B in the empty relation, while for the maximal relation *every* element of A relates to *every* element of B.

In the special circumstance where $A = B = \emptyset$, the two extremes coincide: after all, in that case $A \times B = \emptyset$, and the empty set has exactly one subset, namely itself.

Example X.1.3. For any set A, consider the relation

$$\Delta_A \subseteq A \times A; \qquad \Delta_A = \{(x,x) | x \in A\}.$$

The relation Δ_A goes by many different names: it is often called the *identity relation* on A; it is also called the *diagonal relation* on A.

The following example is a special instance of an *ordering* on a set; it will be explored in greater detail in a later lecture, but for now it is important to recognize that the standard ordering relations you are familiar with are all examples of relations.

Example X.1.4. Let $\leq$ denote the usual ordering on $\mathbb{N}$. Then let consider the subset

$$\{(n, m) \in \mathbb{N} \times \mathbb{N} | n \leq m\} \subseteq \mathbb{N} \times \mathbb{N}.$$

This is a relation from $\mathbb{N}$ to itself.

For a variant, consider the subset

$$\{(n, m) \in \mathbb{N} \times \mathbb{N} | n < m\} \subseteq \mathbb{N} \times \mathbb{N}.$$

This is also a relation from $\mathbb{N}$ to itself.

Similarly, the standard orderings on the integers, rationals and reals can be regarded as relations.

The following example is of a different nature:

Example X.1.5. Define

$$R \subseteq \mathbb{Z} \times \mathbb{Z}; \qquad nRm \Leftrightarrow_{def} n \text{ divides } m.$$

This is a relation on $\mathbb{Z}$, called the *divisibility relation.*

One more:

Example X.1.6. Let X be a set, and consider $A = \mathcal{P}(X)$. Then

$$\{(U, V) \in \mathcal{P}(X) \times \mathcal{P}(X) | U \subseteq V\}$$

is the *subset relation* on the powerset of X.

> How to prove it
>
> To prove that two relations R and S are equal, you have to:
>
> - Take arbitrary x, y with xRy and show that xSy, **AND**
> - Take arbitrary x, y with xSy and show that xRy.

X.2 The calculus of relations

Since relations are defined as subsets, we can do anything with relations which we can do with subsets. In particular, we may compare relations:

Definition X.2.1. Let A, B be sets and let R, S be two relations from $A \to B$. Then we say that R is a *smaller* relation than S (or that S is a *larger* relation than R) when $R \subseteq S$.

We may also express this as: R is smaller than S when aRb implies aSb for all a, b. For example, consider the relations $\leq$ and $<$ on $\mathbb{N}$. Then $<$ is smaller than $\leq$, because $a < b$ implies $a \leq b$.

Exercise 106. Verify that the empty relation from A to B is smaller than any other relation from A to B, and that the maximal relation is larger than any other relation.

HOW TO PROVE IT

To prove that one relation R is smaller than S (i.e., that $R \subseteq S$) you have to:

- Take arbitrary x, y with xRy and show that xSy.

We may also form intersections, unions, complements and differences of relations; see the exercises for some examples and applications.

Earlier, we stressed that a relation from A to B is not the same thing as a relation from B to A. However, there is a construction which allows us to change one into the other:

Definition X.2.2 (Converse of a relation)**.** Let R be a relation from A to B, i.e., $R \subseteq A \times B$. Define a relation R° from B to A by

$$bR^\circ a \Leftrightarrow_{def} aRb.$$

Equivalently, $R^\circ = \{(b, a) \in B \times A | (a, b) \in R\}$. The relation R° is called the *converse* of R (also: the *opposite* of R).

In the pictorial representation of a relation, all you have to do to get R° from R is reverse the orientation of the arrows.

For example, consider the ordering $\leq$ on $\mathbb{N}$. Then the converse of this relation, denoted $\leq^\circ$, is simply the relation $\geq$. We could write this as $\leq^\circ = \geq$, but of course that's unreadable. (Nevertheless, it is technically correct, because both sides of the equation are subsets of $\mathbb{N} \times \mathbb{N}$, and hence the equation expresses that two relations on $\mathbb{N}$ are equal.) Similarly, the converse of $<$ is $>$.

The following is a straightforward consequence of the definition.

Lemma X.2.3. *For a relation $R \subseteq A \times B$, we have $(R^\circ)^\circ = R$.*

Exercise 107. Verify this statement.

Next, we consider a way of combining two relations into a new one: *composition*.

Definition X.2.4 (Composition of Relations)**.** Let $R \subseteq A \times B$ and $S \subseteq B \times C$ be relations. Define the *composite* relation $S \circ R \subseteq A \times C$ to be

$$S \circ R =_{def} \{(a, c) | \exists b \in B. aRb \wedge bSc\}.$$

In words: in the composite relation $S \circ R$, an element $a \in A$ is related to an element $c \in C$ when there exists an element $b \in B$ for which aRb and bSc.

In Figure X.3 we illustrate this pictorially: note that you can simply trace the arrows to see which elements from A relate to which in C.

Here are a couple of concrete examples where we calculate a composite relation:

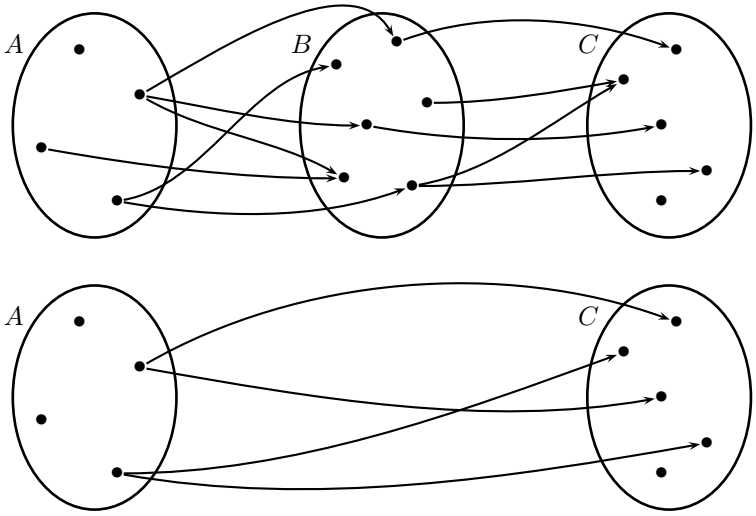


Figure X.3: Two relations and their composite

Examples X.2.5.

1. Let $A = \{a, b\}, B = \{c, d, e\}, C = \{f, g, h\}$ and consider the relations $R = \{(a, c), (a, d), (b, d), (b, e)\}$ and $S = \{(c, f), (d, f)\}$. Then $S \circ R$ is the relation $S \circ R = \{(a, f), (b, f)\}$.
2. Let $X = \{x, y, z\}$, and let $R = \{(x, y), (y, x), (x, z)\}$. Then we get $R \circ R = \{(x, x), (y, y), (y, z)\}$.
3. Consider the standard ordering relation $\leq$ on $\mathbb{N}$. Then $\leq \circ \leq$ is the same relation as $\leq$: on the one hand, suppose $a(\leq \circ \leq)c$: that is, there is a $b \in \mathbb{N}$ with $a \leq b$ and $b \leq c$. Then it follows that $a \leq c$ (this is the *transitivity of the ordering relation*). On the other hand, if $a \leq c$ then there exists $b \in \mathbb{N}$ with $a \leq b$ and $b \leq c$ (take for example $b = a$), so that $a(\leq \circ \leq)c$.
4. With the same notation as in the previous example, we find that $\leq \circ \leq^\circ$ is the maximal relation on $\mathbb{N}$. (Check this for yourself!)

More examples can be found in the exercises at the end of this lecture.

X.3 Properties

In this section we establish some useful results about the converse and composition of relations. We spell out a few proofs in detail, but leave most results as exercises.

Proposition X.3.1. *The operations $(-)^\circ$ and $- \circ -$ on relations satisfy the following properties:*

1. *$R \subseteq R'$ implies $R^\circ \subseteq (R')^\circ$.*

2. $T \circ (S \circ R) = (T \circ S) \circ R$
3. $\Delta_B \circ R = R = R \circ \Delta_A$ *(where* $R \subseteq A \times B$*)*
4. $(S \circ R)^\circ = R^\circ \circ S^\circ$
5. $R \subseteq R'$ *and* $S \subseteq S'$ *implies* $S \circ R \subseteq S' \circ R'$.

Proof. For part 1, suppose that $R \subseteq R'$. We need to show $R^\circ \subseteq (R')^\circ$. Thus, take x, y with $xR^\circ y$; that means (by definition of converse relation) that yRx. Since $R \subseteq R'$, it follows that $yR'x$ as well. Now again by definition of converse, we get $x(R')^\circ y$, as needed.

For part 3, we show $\Delta_B \circ R = R$. First, suppose that x, y are such that xRy. Then by definition of Δ_B we have $y\Delta_B y$. Hence $x(\Delta_B \circ R)y$ as well. Conversely, if $x(\Delta_B \circ R)y$, then by definition of composition there is an element z with xRz and $z\Delta_B y$. But by definition of Δ_B, we get $z = y$, so that xRy follows.

Finally, we prove one half of part 4. Suppose first that x, z are such that $z(S \circ R)^\circ x$. By definition of converse, this means $x(S \circ R)z$; by definition of composition there is an element y with xRy and yRz. Using the definition of converse twice, this gives $yR^\circ x$ and $zS^\circ y$. This shows that $z(R^\circ \circ S^\circ)x$. □

The first part of the proposition states that the operation $(-)^\circ$ is monotone. The second states that composition is associative, and the third that the identity relation is a neutral element for composition (i.e., that composition with an identity relation doesn't do anything). The fourth describes how composition and converse interact: the converse of a composite is the composite of the converses (you may want to draw a picture). Finally, part 5 states that composition is monotone.

Exercise 108. Complete the above proofs.

X.4 Summary

This lecture introduced the notion of a *relation* R from A to B as a subset of the Cartesian product $A \times B$. Three special relations are often considered:

- The empty relation $R = \emptyset$
- The maximal relation $R = A \times B$
- The diagonal relation, or identity relation $\Delta_A = \{(a, a) | a \in A\} \subseteq A \times A$.

Two fundamental operations on relations are:

- Composition
- Converse

The *calculus of relations* (of which the most important aspects are listed in Proposition X.3.1) describes how these two operations behave.

X.5 Exercises

Exercise 109. Let $A = \{a, b, c, d\}$ and $B = \{p, q, r\}$. Let $R \subseteq A \times A$ be the relation $R = \{(a, b), (c, d), (a, d)\}$, let $S \subseteq A \times B$ be the relation $S = \{(a, p), (b, p), (c, p), (d, p), (d, r), (a, q)\}$. Find the following relations:

(a) R°

(b) $R \circ R$

(c) $R \circ R^\circ$

(d) $R^\circ \circ R$

(e) $S \circ R$

(f) $S \circ R^\circ$

(g) $S \circ S^\circ$

(h) $S^\circ \circ S$

(i) $S \circ R \circ S^\circ$

Exercise 110. Let $C = \{0, 1\}$. Find all relations R on C for which $R = R^\circ$. Also find all relations for which $R = R \circ R$.

Exercise 111. Prove that if either R or S is the empty relation, then $S \circ R$ is empty as well. Also show that if R is empty then so is R°.

Exercise 112. Show that when both R and S are maximal, then so is $S \circ R$. Also show that if R is maximal then so is R°.

Exercise 113. Can there be relations R, S such that neither of R, S are empty but $S \circ R$ is? Give an example or a proof that this can't happen.

Exercise 114. Can there be relations R, S such that neither of R, S are maximal but $S \circ R$ is? Give an example or a proof that this can't happen.

Exercise 115. Given a relation $R \subseteq A \times B$, we may consider its complement

$$R^c = \{(a, b) | (a, b) \notin R\}.$$

Investigate how the complement behaves w.r.t. the other operations. For example, see whether $(S \circ R)^c = (S^c \circ R^c)$, and whether $(R^\circ)^c = (R^c)^\circ$.

Exercise 116. Consider the relation $<$ on $\mathbb{N}$. What is $< \circ <$? What is $< \circ <^\circ$? And $<^\circ \circ <$?

LECTURE XI

FUNCTIONS

As we have seen, relations are very general. There are various special classes of relations which we shall study in more depth: functions, equivalence relations and orderings. This lecture is devoted to the beginnings of the study of functions between sets.

XI.1 DEFINITION

Our goal in this section is to give a definition of function in terms of relations. This is perhaps a bit unusual, given that most of us learned what a function was before we heard of relations. However, in line with our overall aim of seeing how mathematics can ultimately be reduced to set theory, we don't wish to take the concept of a function as a primitive notion, but define it in terms of sets.

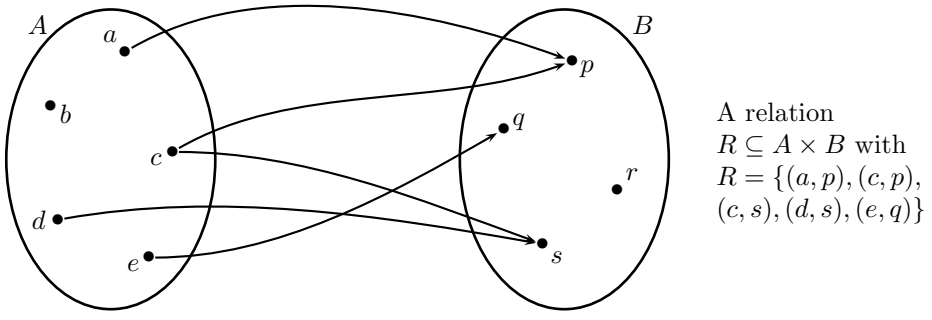


Figure XI.1: Example of a relation R from A to B

Consider again a typical example of a relation as in Figure XI.1. There are some ways in which this may remind us of a function: the relation assigns to elements in the set A on the left elements in the set B on the right[1]. But there are two ways in which it is not a function: first, not every element of A is assigned an element of B (functions have to assign something to every element, and R does not assign anything to b); second, some elements are assigned more than one element (functions have to be non-ambiguous, and assign exactly one value, while R assigns two values to c).

This leads to the following definition:

Definition XI.1.1 (Function)**.** Let A, B be sets and $R \subseteq A \times B$ a relation.

1. R is called *total* when for each $x \in A$ there is at least one $y \in B$ with xRy.
2. R is called *single-valued* when for each $x \in A$ there is at most one $y \in B$ with xRy.
3. When R is both total and single valued, R is called a *function(al relation)*.

HOW TO PROVE IT

To prove $R \subseteq A \times B$ is total you have to:

- Take an arbitrary $x \in A$ and show that there exists $y \in B$ with xRy

To prove $R \subseteq A \times B$ is not total you have to:

- Show that there exists $x \in A$ for which there exists no $y \in B$ with xRy

It is common to denote functional relations by f, g, h, and so on. Also, to indicate the f is a function from A to B we use the standard notation $f : A \to B$. Moreover, when $f : A \to B$ is a function and $x \in A$, we write $f(x) = y$, where y is the (necessarily unique!) element of B with $(x, y) \in f$.

Here's an illustration: let $A = \{a, b, c, d\}$ and let $B = \{p, q, r, s\}$. Consider the relations

$$\begin{aligned} R &= \{(a, q), (b, q), (b, s), (c, p), (d, p)\} \\ S &= \{(a, q), (c, r), (d, p)\} \\ T &= \{(a, q), (b, q), (c, r), (d, p)\} \end{aligned}$$

The relation R is total but not single-valued (as the element b is related to two distinct elements); the relation S is single-valued but not total (as the element b is not related to anything); the relation T is a function.

HOW TO PROVE IT

To prove $R \subseteq A \times B$ is single-valued you have to:

- Show that *if* xRy and xRy' *then* $y = y'$

To prove $R \subseteq A \times B$ is not single-valued you have to:

- Show that there exist $x \in A, y, y' \in B$ with xRy, xRy' and $y \neq y'$.

Here are some extreme examples:

Example XI.1.2. Let A, B be sets. Then the empty relation from A to B is single-valued. Why? This is another example of something which is vacuously true: we need to check that if xRy and xRy' then $y = y'$. But the antecedent is always false, so the implication is always true.

[1] Presumably, the "definition" of function you learned was something like: a function $f : A \to B$ is an assignment of elements of B to elements of A.

Example XI.1.3. Let A, B be a sets. Then the empty relation from A to B is total precisely when $A = \emptyset$.

In one direction, suppose $A = \emptyset$. Then the empty relation R is total: we need to check that if $x \in A$ then there is $y \in B$ with xRy. But $x \in A$ is false, so the implication is true.

Conversely, suppose the empty relation is total. Then we know that for all $x \in A$ there exists $y \in B$ with xRy. But if R is empty, the statement xRy is always false. Thus the consequent of the implication is false, and hence the antecedent is also false.

As an exercise, you may want to find the smallest[2] possible examples of the following:

- A relation which is total but not single-valued.
- A relation which is single-valued but not total.
- A relation which is neither single-valued nor total.

Finally, let us see how common functions can be viewed as relations. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$. Then the *graph* of f is the relation

$$Gr(f) = \{(x, x^2) | x \in \mathbb{R}\}.$$

This is a relation from $\mathbb{R}$ to itself. Generally, given a function $f : A \to B$, define the graph of f to be the relation

$$Gr(f) = \{(x, f(x)) | x \in A\} \subseteq A \times B.$$

Thus what we normally (in calculus, say) call the graph of a function really is what we here call the function itself!

The following important exercise states that the properties of totality, single-valuedness and functionality are closed under composition.

Exercise 117. Let $R \subseteq A \times B$ and $S \subseteq B \times C$ be relations. Prove:

(i) If R and S are total, then so is $S \circ R$.

(ii) If R and S are single-valued then so is $S \circ R$.

(iii) If R and S are functional, then so is $S \circ R$.

(iv) Conclude that the composite of functional relations indeed gives the usual composition of functions $(g \circ f)(x) = g(f(x))$.

XI.2 Functions and products

In calculus, it is common to consider not only functions $\mathbb{R} \to \mathbb{R}$, but more generally functions of several variables $\mathbb{R}^n \to \mathbb{R}$, or even vector-valued functions $\mathbb{R}^n \to \mathbb{R}^m$. In

[2] Finding the smallest or simplest possible example of something is an excellent exercise in mastering definitions and the connections amongst them.

this section we briefly explain what functions between Cartesian products of sets look like and how we can work with them.

First of all, let us consider a function $f : A \to B \times C$. Such a function takes as input elements from A, and gives as output elements from $B \times C$. The latter are pairs, so given $a \in A$, we get $f(a) \in B \times C$, with $f(a) = (f_B(a), f_C(a))$. Thus we see that f actually harbours two functions: $f_B : A \to B$ and $f_B : A \to C$ which have been paired together. Moreover, these two functions together completely determine f.

This proves:

Lemma XI.2.1. *Given functions $p : A \to B$ and $q : A \to C$, there exists a unique function $\langle p, q\rangle : A \to B \times C$ with $\langle p, q\rangle(a) = (p(a), q(a))$. Every function $A \to B \times C$ uniquely arises in this way.*

To put this succinctly: a function into a product is a pair of functions. A slightly more rigorous formulation of this statement can be found in Lecture XII. See also the exercises.

The above lemma explains what functions *into* product sets look like. What about functions *out of* them? Generally, there is not much we can say. A function $f : A \times B \to C$ simply sends pairs $(a, b) \in A \times B$ to elements $f(a, b) \in C$. However, there is one special situation which often occurs in practice. Given functions $p : A \to C$ and $g : B \to D$, we can combine these to form a new function $f \times g : A \times B \to C \times D$, defined by

$$(f \times g)(a, b) = (f(a), g(b)) \in C \times D.$$

It is left as an exercise for you to check that this is indeed a function. Note that we're slightly abusing notation: $f \times g$ is a name for this new function, but is not the same as the cartesian product of the functional relations f and g.

XI.3 Special functions

This section introduces three special classes of functions which are prevalent throughout all of mathematics.

Definition XI.3.1 (Injective, surjective, bijective)**.** Let $f : A \to B$ be a function.

1. f is called *injective* when for all $x, x' \in A$, $f(x) = f(x')$ implies $x = x'$. In logical notation: $\forall x, x' \in A.(f(x) = f(x') \to x = x')$.
2. f is called *surjective* when for all y in B there exists at least one $x \in A$ with $f(x) = y$. In logical notation: $\forall y \in B \exists x \in A.f(x) = y$.
3. f is called *bijective* when it is both injective and surjective.

Often, a bijective function is also called a *bijective correspondence*, or a *1-1 correspondence*, because it relates each element of the first set to a unique element of the second, and vice versa.[3]

[3] And if you want to display your understanding of foundations and show off your mathematical sophistication you should call them *isomorphisms*, because that's what they really are.

As usual, we begin with a few examples to illustrate these concepts.

Examples XI.3.2.

1. Let $A = \{a, b, c\}$ and $B = \{p, q, r, s\}$. Consider the functions

$$\begin{array}{ll} f(a) = p & g(a) = p \\ f(b) = p & g(b) = r \\ f(c) = q & g(c) = s \end{array}$$

   Then f is not injective, because we have $f(a) = f(b)$ but $a \neq b$. Also, f is not surjective, because there is no $x \in A$ with $f(x) = s$. The function g is injective, but not surjective.

2. Let $A = \{a, b, c, d\}$ and $B = \{p, q, r\}$. Consider the functions

$$\begin{array}{ll} f(a) = p & g(a) = p \\ f(b) = p & g(b) = r \\ f(c) = q & g(c) = r \\ f(d) = r & g(d) = r \end{array}$$

   Then f is surjective but not injective, and g is neither injective nor surjective.

3. Let $A = \{a, b, c, d\}$ and $B = \{p, q, r, s\}$. Consider the function

$$f(a) = p, f(b) = s, f(c) = r, f(d) = q.$$

   Then f is both injective and surjective, hence bijective.

HOW TO PROVE IT

To prove that $f : A \to B$ is injective you must

- Take arbitrary $x, x' \in A$, assume $f(x) = f(x')$ and show that $x = x'$.

To prove that $f : A \to B$ is not injective you must

- Find $x, x' \in A$ such that $f(x) = f(x')$ but $x \neq x'$.

To prove that $f : A \to B$ is surjective you must

- Take an arbitrary $y \in B$ and show that there exists an $x \in X$ with $f(x) = y$.

To prove that $f : A \to B$ is not surjective you must

- Find an element $y \in B$ such that $f(x) \neq y$ for all x.

Example XI.3.3. Recall that on any set A there is a relation Δ_A, the identity relation. This relation is functional, and is in fact bijective. A more common notation for Δ_A, when we regard it as a function, is $1_A : A \to A$. In the notation for functions, we have $1_A(x) = x$, and we call 1_A the *identity function on A*.

Here are some examples involving functions from calculus:

Examples XI.3.4.

1. The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is neither injective nor surjective. It is not injective[4] because, for example $f(-2) = 4 = f(2)$, but $2 \neq -2$. It is not surjective because, for example, there is no $x \in \mathbb{R}$ with $f(x) = -1$.

[4] In calculus, you may have seen the *horizontal line test* for injectivity. Verify that this indeed is a visual way of determining our concept of injectivity.

2. The function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = 2x - 5$ is injective: if $g(x) = g(x')$ then $2x - 5 = 2x' - 5$, giving $2x = 2x'$ and hence $x = x'$. It is also surjective: given $y \in \mathbb{R}$, to find x with $g(x) = y$, we solve $2x - 5 = y$ to get $x = \frac{y+5}{2}$. Hence g is bijective.

3. Consider the tangent function $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$. (Note that we restrict the domain to a single period of the function.) Then tan is bijective.

In the special case where $f : A \to A$ is a bijection from A to itself, we sometimes call f a *permutation* of A, because such a function permutes the elements of A.

We conclude with a useful fact about bijective functions.

Lemma XI.3.5. *Let $f : A \to B$ be a function. Then f is bijective if and only if there exists a function $g : B \to A$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$.*

Proof. Suppose first that f is a bijective function. Then define a function $g : B \to A$ as follows:

$$g(y) = \text{ the unique element } x \in A \text{ with } f(x) = y.$$

This is well-defined because f is a bijection: by surjectivity such x exists, and by injectivity it is unique. (Alternatively, you may let $g(y) = x \Leftrightarrow f(x) = y$.) Now

$$g(f(x)) = \text{ the unique element } x \in A \text{ with } f(x) = f(x).$$

But clearly that gives $g(f(x)) = x$ (again because f is injective). Hence $g \circ f = 1_A$. Also, we have

$$f(g(y)) = f(x), \text{ where } x \in A \text{ is the unique element with } f(x) = y.$$

Hence $f(g(y)) = y$, so that $f \circ g = 1_B$.

Conversely, suppose there exists g with $g \circ f = 1_A$ and $f \circ g = 1_B$. To show that f is injective, suppose we have x, x' with $f(x) = f(x')$. Then:

$$x = g(f(x)) = g(f(x')) = x'.$$

To show that f is surjective, consider $y \in B$. Then for $x = g(y)$ we have $f(x) = f(g(y)) = y$. □

A function g with the properties $g \circ f = 1_A$ and $f \circ g = 1_B$ is called an *inverse* for f. It is easily seen that if g and g' are both inverses for f, then they must be the same:

$$g(y) = g(f(g'(y)) = g'(y)$$

where the first step uses $f(g'(y)) = y$ and the second uses $g(f(x)) = x$. Hence we are justified in speaking of *the* inverse of f, and giving it the special notation f^{-1}.

XI.4 Sets of functions

When we consider two fixed sets A and B, we can consider the collection of all relations from A to B. Let's denote this set by $\mathsf{Rel}(A, B)$. Thus by definition

$$\mathsf{Rel}(A, B) = \{R | R \text{ is a relation from } A \text{ to } B\} = \mathcal{P}(A \times B).$$

In the previous sections, we learned that functions are special kinds of relations. Thus, still for fixed A and B, we may also consider the set of all functions from A to B. We denote this set by $\mathsf{Fun}(A, B)$. By definition:

$$\mathsf{Fun}(A, B) = \{R \in \mathsf{Rel}(A, B) | R \text{ is functional } \}.$$

A slightly prettier way would be to write $\mathsf{Fun}(A, B) = \{f | f : A \to B\}$. We will also use the "exponentiation" notation B^A for this set. This will be justified later by the observation that sets of functions behave a bit like exponentials.

We can now play a game: I write down a set (involving functions and perhaps also products) and you give me an element in it.[5]

Let's begin simple: A^A. This is the set of all functions from A to A. (Remember, I haven't told you anything about A itself!) Can you find an element of this set? Yes: amongst all the functions from A to A there is the identity function: $Id_A \in A^A$. Even though we don't know anything about A, we know that this function exists.

Now, let's look at $A^{A\times A}$, the set of functions from $A \times A$ to A. Can you find me an element of this set? To do so, you must define a function $A \times A \to A$. In fact, there are two solutions to the problem: the first is the function

$$\pi_1 : A \times A \to A; \qquad \pi_A(a, b) = a.$$

The second solution is

$$\pi_2 : A \times A \to A; \qquad \pi_A(a, b) = b.$$

These functions are sometimes called the *projection functions* . See also the exercises of this chapter.

Next, consider $(A^A)^A$. An element of this set is a function from A to A^A. Can you give me such a function? In fact, this problem is a bit like the previous one, and there are two solutions again. For the first, consider the function $\eta : A \to A^A$ which sends $a \in A$ to the function $c_a \in A^A$ defined by $c_a(x) = a$. For the second, take the function $\iota : A \to A^A$ which sends a to the identity function.

Last one: find me an element of $\mathsf{Fun}(A^A \times A, A)$. That is, find me a function from $A^A \times A$ to A. Again, there are exactly two possible answers. The first is the function $\pi : A^A \times A \to A$ defined by $\pi(f, a) = a$. (This is again a projection function.) The second is more interesting:

$$\epsilon : A^A \times A \to A; \qquad \epsilon(f, a) = f(a).$$

This is called the *evaluation function*: it takes as input a function and an element in the domain of that function, and it outputs the value $f(a)$ of the function on that element.

We will revisit these matters in the next chapter. For now, note that there are games you can't win: for example, suppose I gave you $\mathsf{Fun}(A^A, A)$. You're supposed to define a function from A^A to A, but what could you possibly do? Indeed, you see you won't be able to win unless you already know of a specific element in A. So if $A = \emptyset$, you're doomed.

Exercise 118. Can you win the game $\mathsf{Fun}(A^A, A^A)$? In how many ways can you win?

[5] This game-theoretic behaviour of sets built in this way can actually be made very precise, and is due to James Dolan.

XI.5 Summary

Functional relations (functions) were defined as relations satisfying two special properties:

- Totality
- Single-valuedness

The usual notion of composition of functions then is recovered as the relational composition of functional relations.

Given functions $f : A \to B$, $g : A \to C$, we get a unique function $\langle f, g\rangle : A \to B \times C$ with $\langle f, g\rangle(a) = (f(a), g(a))$. Every function into a product set is uniquely of this form. Moreover, given functions $f : A \to C, g : B \to D$, we define $f \times g : A \times B \to C \times D$ to be $(f \times g)(a, b) = (f(a), g(b))$.

Amongst functions, there are three special classes of interest:

- Injections (also called: *one-to-one*)
- Surjections (also called: *onto*)
- Bijections (also called: *bijective correspondences*)

Every bijective function $f : A \to B$ has a unique *inverse*, that is, a function $g : B \to A$ for which $g \circ f = 1_A$ and $f \circ g = 1_B$.

We write $\mathsf{Rel}(A, B)$ for the set of relations from A to B, and $\mathsf{Fun}(A, B)$ or B^A for the set of functions from A to B.

XI.6 Exercises

Exercise 119. Given a relation $R \subseteq A \times B$.

(i) If R is total, does it follow that R° is also total?

(ii) If R is single-valued, does it follow that R° is also single-valued?

(iii) If R is functional, does it follow that R° is functional?

Exercise 120. Consider the following relation from $\mathbb{R}$ to itself:

$$R = \{(x^2, x) | x \in \mathbb{R}\}.$$

Is this relation functional?

Exercise 121. Consider the relation $S = \{(x, x^3) | x \in \mathbb{R}\}$ from $\mathbb{R}$ to itself. Is this relation functional? What about S°?

Exercise 122. Find all functions from $\{0,1\}$ to $\{0,1\}$. (That is, describe all elements of the set $\mathsf{Fun}(\{0,1\},\{0,1\})$.)

Exercise 123. Show that if $f : A \to \emptyset$ is a function, then $A = \emptyset$.

Exercise 124. Show that any function $f : \emptyset \to B$ is injective.

Exercise 125. Show that if B has exactly one element, then for any set A there exists exactly one function $f : A \to B$.

Exercise 126. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^3$. Is this function injective? Surjective? Bijective?

Exercise 127. Same question, but now for the absolute value function

$$|x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

Exercise 128. Consider the affine function $f(x) = ax + b$ (as a function from $\mathbb{R}$ to $\mathbb{R}$). For which values of a, b is this function injective? Surjective? Bijective?

Exercise 129. Show that if $f : A \to B$ and $g : B \to C$ are both surjective, so is $g \circ f$. (We express this by saying that surjective functions are *closed under composition*.)

Exercise 130. Show that if $f : A \to B$ and $g : B \to C$ are both injective, so is $g \circ f$. (We express this by saying that injective functions are *closed under composition*.)

Exercise 131. Show that if $f : A \to B$ and $g : B \to C$ are both bijective, so is $g \circ f$. (We express this by saying that bijective functions are *closed under composition*.)

Exercise 132. Can it happen that neither f nor g is injective but that $g \circ f$ is? Same question for surjective and bijective.

Exercise 133. Show that if f is bijective with inverse f^{-1}, then f^{-1} is also bijective, and $(f^{-1})^{-1} = f$.

Exercise 134. Show that for any set A, the identity function 1_A on A is bijective. What is its inverse?

Exercise 135. Show that if f and g are bijective, then $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Exercise 136. Find all permutations of the set $A = \{a, b, c\}$.

Exercise 137. Same question, but now for the set $B = \{a, b, c, d\}$.

Exercise 138. Suppose A has n elements. How many permutations from A to A do you think there are? (Note: this is really a combinatorics question.)

Exercise 139. Find all injective functions from the set $A = \{a, b\}$ to the set $B = \{p, q, r, s\}$.

Exercise 140. Find all surjective functions from the set $A = \{a, b, c, d\}$ to the set $B = \{p, q\}$.

Exercise 141. Is it possible for a function $f : X \to X$ to be injective but not surjective? Can it be surjective but not injective? Give examples.

Exercise 142. Let A, B be sets. Show that there is a function $\pi_A : A \times B \to A$ which sends (x, y) to x, and similarly $\pi_B : A \times B \to B$, sending (x, y) to y. These are called the *product projections*.

Exercise 143. Let $f : A \to B$ and $g : A \to C$ be functions. Show that there exists a function $\langle f, g\rangle : A \to B \times C$ given by $\langle f, g\rangle(x) = (f(x), g(x))$. Show also that $\pi_B \circ \langle f, g\rangle = f$ and $\pi_C \circ \langle f, g\rangle = g$.

Exercise 144. Show that for $f : A \to B$ and $g : C \to D$ there exists a function $f \times g : A \times C \to B \times D$, which sends (x, y) to $(f(x), g(y))$.

Exercise 145. Suppose A has n elements, and B has m elements. How many elements do you think B^A has?

LECTURE XII

Bijective Correspondences

It often happens that two sets look very different, but that they are in fact closely related. In fact, many profound results in mathematics state that two seemingly different looking sets are really "the same". Bijective correspondences are used to make such statements precise. To say that there is a bijective correspondence between two sets A and B simply means that there exists a bijective function $f : A \to B$.

In this lecture we take a look at some non-trivial bijective correspondences. Before we do this, we introduce one piece of notation which is commonly used in this context: given sets A, B, write

$A \cong B \Leftrightarrow_{def}$ there exists a bijection $f : A \to B$.

The intuition is that sets which are in bijective correspondence are the same set *up to a renaming of their elements*. In particular, the two sets carry the same amount of information, since we can go back and forth between them without losing any information.

> How to prove it
>
> Proving that $A \cong B$ can be much more difficult than simply proving that a given function $f : A \to B$ is a bijection. The reason is that finding a candidate bijection may require insight and/or creativity.
> Usually the best thing is to look carefully at what the elements of A and those of B look like, and then try to see how they relate.

From the exercises in the previous lecture we have the following facts about bijective correspondences:

Proposition XII.0.1. *The relation $\cong$ satisfies the following properties:*

1. $A \cong A$
2. $A \cong B$ *implies* $B \cong A$
3. $A \cong B$ *and* $B \cong C$ *implies* $A \cong C$

Keep in mind that there may be many different bijections from A to B.

Warning: the material in this chapter is on a higher level of abstraction than what we've seen so far.

XII.1 Products

In an earlier lecture, we proved that there is exactly one empty set. But what about sets with one element? These are certainly not unique in the sense of extensionality: $\{\emptyset\}, \{\{\emptyset\}\}$ and $\{\mathbb{N}\}$ are all sets with one element, but they are different since they have different elements. However, they are in bijective correspondence, and it is in this sense in which one-element sets are unique.

Lemma XII.1.1. *Suppose that A and B are both sets with exactly one element. Then $A \cong B$.*

Proof. Denote by a the unique element of A, and by b the unique element of B. There's not much choice now: there is only one function $f : A \to B$, namely the one defined by $f(a) = b$. Conversely, there is only one function $g : B \to A$, namely $g(b) = a$. Clearly these are inverses of each other. (See the exercises for a slicker proof.) □

Often, this is expressed by saying that one-element sets are unique *up to isomorphism*[1] It is customary to write 1 for one-element sets when we don't care about the particular element.

Proposition XII.1.2. *Let A be any set. Then $A \times 1 \cong A \cong 1 \times A$.*

Proof. We prove $A \times 1 \cong A$, leaving the rest to you. For convenience, we write $*$ for the element of 1. Then note that

$$A \times 1 = \{(a, *) | a \in A\}.$$

Thus, $A \times 1$ almost[2] is the same as A, only its elements have a "tag". Thus the bijection between $A \times 1$ and A should simply "remove the tag".

To turn this into a concrete proof, set

$$\pi_A : A \times 1 \to A; \qquad f(a, *) = a.$$

Then π_A is a bijection; its inverse is $\pi^{-1}(a) = (a, *)$. □

Note that this may not be the only bijection from $A \times 1$ to A. (If σ is a permutation of A, then $\sigma \circ f$ is also a bijection $A \times 1 \to A$.) However, the bijection f given is *canonical*.

Next, recall that we observed earlier that $A \times B$ is generally not the same set as $B \times A$. But surely they are closely related! Indeed:

[1] "Iso" is Greek for "same", while "morphè" means "form". Thus any two one-element sets have the same form.

[2] Often you'll hear mathematicians say things like: "Morally, these two sets are the same." What they usually mean is that they might not technically be identical, but they behave the same for the purposes at hand.

Proposition XII.1.3. *For any sets A, B we have $A \times B \cong B \times A$.*

Proof. Again it is fairly obvious what the bijection is going to be:

$$\tau : A \times B \to B \times A; \qquad \tau(x, y) = (y, x).$$

You should now check the following claims:

(i) τ is indeed a well-defined function.

(ii) τ is injective.

(iii) τ is surjective.

Alternatively, you can directly give an inverse $\sigma : B \times A \to A \times B$ for τ. □

For a last correspondence involving products, suppose that $f : A \to B$ and $g : C \to D$ are functions. As discussed in lecture IX, we get a function

$$f \times g : A \times C \to B \times D; \qquad (f \times g)(a, c) = (f(a), g(c)).$$

We wish to show that if f and g are both bijective, then so is $f \times g$. This will prove:

Proposition XII.1.4. *If $A \cong B$ and $C \cong D$ then $A \times C \cong B \times D$.*

Proof. We show that $f \times g$ is injective and surjective.

Injectivity: suppose we have two elements (x, y) and (x', y') of $A \times C$ with $(f \times g)(x, y) = (f \times g)(x', y')$. By definition of $f \times g$, this means that $(f(x), g(y)) = (f(x'), g(y'))$. This in turn gives $f(x) = f(x')$ and $g(y) = g(y')$ (why?). Since f and g are both injective, this forces $x = x'$ and $y = y'$. Hence also $(x, y) = (x', y')$, as needed.

Surjectivity: consider an element $(u, v) \in B \times D$. Since f is surjective, there exists $x \in A$ with $f(x) = u$. Since g is surjective, there exists $y \in C$ with $g(y) = v$. Hence $(f \times g)(x, y) = (f(x), g(y)) = (u, v)$ as needed. □

Exercise 146. Give an alternative proof by showing that $f^{-1} \times g^{-1}$ is the inverse of $f \times g$.

Exercise 147. Show that for any sets A, B, C we have $A \times (B \times C) \cong (A \times B) \times C$.

Finally, an example which is of a slightly different nature. We'll show (slightly informally, because these matters will be revisited in a later chapter), that $\mathbb{N} \cong P$, where P is the set of positive prime numbers. This will use the fact (Euclid's Theorem) that there exist arbitrarily large prime numbers.

We need to set up a bijective function $\mathbb{N} \to P$. The idea is to send n to the n-th prime number. This make sense, because we can enumerate all prime numbers in increasing order: $p_0, p_1, p_2, \ldots$. We may assume that there is no repetition in this list. Because of Euclid's Theorem, this list never stops. Now the function $n \mapsto p_n$ is the desired bijection.

XII.2 Characteristic functions

In this section we discuss a slightly more sophisticated example of a bijective correspondence. The starting point is the set $\{0,1\}$, which sometimes goes by the name 2 (because it has two elements). Sometimes the elements of this set are denoted $\top, \bot$, or t, f, for *true* and *false*. We'll stick to $\{0,1\}$.

As a warm-up example, consider the set $A = \{a,b,c\}$. We're going to look at functions $A \to \{0,1\}$. That is, we're going to study the set $\mathsf{Fun}(A, \{0,1\})$. Since A is quite small, we can actually list all such functions (we'll give them random names for now):

$$\begin{array}{lll} f(a) = f(b) = f(c) = 0 & g(a) = g(b) = 0, g(c) = 1 & h(a) = h(c) = 0, h(b) = 1 \\ i(a) = 0, i(b) = i(c) = 1 & j(a) = 1, j(b) = j(c) = 0 & k(1) = k(1) = 0, k(b) = 0 \\ l(a) = l(b) = 1, l(c) = 0 & m(a) = m(b) = m(c) = 1 & \end{array}$$

Thus there are 8 functions $A \to \{0,1\}$ in total, and hence $\mathsf{Fun}(A, \{0,1\})$ has 8 elements.

Now these functions can be interpreted in a slightly different way. Since the elements $\{0,1\}$ play the role of truth values, a function $p : A \to \{0,1\}$ may be thought of as deciding for each element of A whether it is sent to true or to false. Then the set of all those elements which are sent to true form a subset of A. For example, the function i sends b and c to 1, and a to 0, so the subset of elements which are sent to true is $\{b,c\}$.

We can do this for each function $A \to \{0,1\}$: each such function determines a subset of A. This gives

$$\begin{array}{ll} f \rightsquigarrow \emptyset & j \rightsquigarrow \{a\} \\ g \rightsquigarrow \{c\} & k \rightsquigarrow \{a,c\} \\ h \rightsquigarrow \{b\} & l \rightsquigarrow \{a,b\} \\ i \rightsquigarrow \{b,c\} & m \rightsquigarrow \{a,b,c\} \end{array}$$

This sets up a function $\mathsf{Fun}(A, \{0,1\}) \to \mathcal{P}(A)$. Note that not only does each function $A \to \{0,1\}$ determine a subset of A, we get *every* subset of A in this way. And what is more, distinct functions give distinct subsets. Thus, we have a bijective correspondence

$$\mathsf{Fun}(A, \{0,1\}) \cong \mathcal{P}(A).$$

We have discussed this example in terms of how functions $A \to \{0,1\}$ give rise to subsets of A. But the other direction is equally important. How do subsets of A give rise to functions? Suppose we have a subset $U \subseteq A$. We want to represent this as a function $A \to \{0,1\}$. Thinking of 1 as true and 0 as false again, we want this function to output "true" when an element is indeed in U, and "false" when it isn't. This function is called the *characteristic function*[3] of the subset U. Formally:

$$\chi_U(x) = \begin{cases} 1 & \text{if } x \in U \\ 0 & \text{if } x \notin U \end{cases}.$$

For example, if $U = \{a,c\}$ then $\chi_U(a) = \chi_U(c) = 1, \chi_U(b) = 0$.

[3] Some texts call this the *indicator function* instead.

Representing subsets via characteristic functions is not as esoteric as one might think: computers only know 0's and 1's, so when you want to encode a subset in a computer it is precisely the characteristic function which you use.

We now generalize the above example. The idea of the proof is exactly the same as in the case we looked at, but we'll write out a detailed proof. The main difficulty is not in getting the general idea here (after all, I already told you what that was) but in being very systematic about what has to be done; since we're now working with functions between sets of sets and sets of functions it is easy to lose track.

Proposition XII.2.1. *Let X be any set. Then there is a bijective correspondence $\mathcal{P}(X) \cong \mathsf{Fun}(X, \{0,1\})$.*

Proof. The goal is to construct a bijective function from $\mathcal{P}(X)$ to $\mathsf{Fun}(X, \{0,1\})$. Let's call this function (which we yet have to specify) ϕ. Thus for every subset U of X, we must define an element $\phi(U) \in \mathsf{Fun}(X, \{0,1\})$. That is, $\phi(U)$ must be a function $X \to \{0,1\}$. We define, given $U \subseteq X$, $\phi(U)$ to be the characteristic function of U:

$$\phi(U) = \chi_U : X \to \{0,1\}; \qquad \chi_U(x) = \begin{cases} 1 & \text{if } x \in U \\ 0 & \text{if } x \notin U \end{cases}.$$

This defines $\phi : \mathcal{P}(X) \to \mathsf{Fun}(X, \{0,1\})$. It remains to be shown that this is a bijection. We will exhibit the inverse of ϕ. (Alternatively you could try to show that ϕ is injective and surjective, but you'll need roughly the same ideas.)

So let us try to define $\phi^{-1} : \mathsf{Fun}(X, \{0,1\}) \to \mathcal{P}(X)$ by

$$\phi^{-1}(f) = \{x \in X | f(x) = 1\}.$$

That is, ϕ^{-1} sends an element $f : X \to \{0,1\}$ to the set of those x for which $f(x) = 1$. (Note that ϕ^{-1} takes as input a function and gives as output a subset.)

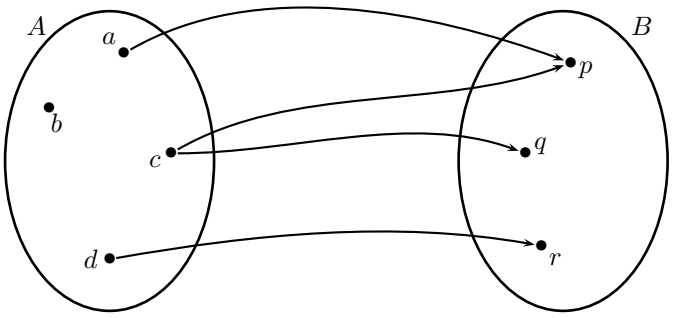
Finally, we must show that ϕ^{-1} is indeed inverse to ϕ. To show that $\phi^{-1} \circ \phi = 1_{\mathcal{P}(X)}$, consider $U \in \mathcal{P}(X)$ and compute

$$\phi^{-1}\phi(U) = \phi^{-1}(\chi_U) = \{x \in X | \chi_U(x) = 1\} = U.$$

And to show that $\phi \circ \phi^{-1} = 1_{\mathsf{Fun}(X,\{0,1\})}$, take $f : X \to \{0,1\}$; then we wish to prove that $\phi(\phi^{-1}(f)) = f$. Since these are functions, we show that they are equal by showing that they agree on all inputs. So consider an arbitrary input $x \in X$. Then

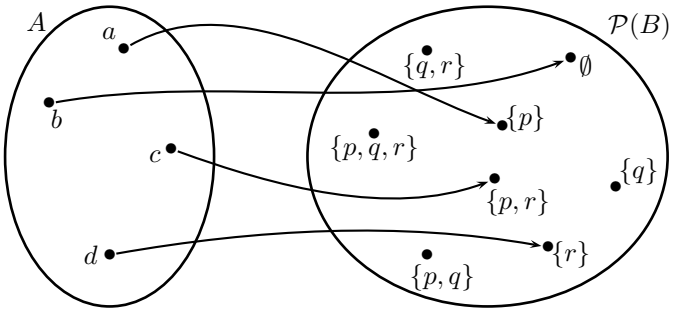
$$\begin{aligned} \phi(\phi^{-1}(f))(x) &= \phi(\{x \in X | f(x) = 1\})(x) \\ &= \chi_{\{x \in X | f(x) = 1\}}(x) \\ &= \begin{cases} 1 & \text{if } f(x) = 1 \\ 0 & \text{if } f(x) = 0 \end{cases} \\ &= f(x) \end{aligned}$$

□



A relation
$R \subseteq A \times B \ldots$

Figure XII.1: Relation R from A to B



$\ldots$ and the
corresponding
function

Figure XII.2: Function from A to $\mathcal{P}(B)$

XII.3 Relations

The last example of a bijective correspondence (for now) is that between relations from A to B and functions $A \to \mathcal{P}(B)$. To give the idea behind this, consider once more the relation in Figure XII.1. How could we reinterpret this picture as a function from A to $\mathcal{P}(B)$? Such a function should associate to each element of A a subset of B. What should the subset associated to c be? The relation R related c to the elements p and s. Thus it makes sense to associate to c the subset $\{p, s\}$. Generally, to an element of A we associate the set of those elements of B which are related to it via R. Calling the sought-after function $r : A \to \mathcal{P}(B)$, we thus get

$$r(a) = \{p\}, r(b) = \emptyset, r(c) = \{p, s\}, r(d) = \{s\}, r(e) = \{q\}.$$

Figure XII.2 depicts this corresponding function.

Thus relations can be viewed as "set-valued" functions. And now for the general statement:

Proposition XII.3.1. *For any sets A, B there is a bijective correspondence*

$$\mathsf{Rel}(A, B) \cong \mathsf{Fun}(A, \mathcal{P}(B)).$$

Recalling that the set of relations from A to B is $\mathcal{P}(A \times B)$, and the set of functions from A to $\mathcal{P}(B)$ is $\mathcal{P}(B)^A$, we may also write this as

$$\mathcal{P}(A \times B) \cong \mathcal{P}(B)^A.$$

Proof. We will define a bijective function $\phi : \mathsf{Rel}(A, B) \to \mathsf{Fun}(A, \mathcal{P}(B))$. Consider an element $R \in \mathsf{Rel}(A, B)$, i.e., a relation $R \subseteq A \times B$. We must define $\phi(R)$, which is to be an element of $\mathcal{P}(B)^A$, that is, a function $A \to \mathcal{P}(B)$. Define this function as follows:

$$\phi(R) : A \to \mathcal{P}(B); \qquad \phi(R)(x) = \{y \in B | xRy\}.$$

Thus $\phi(R)$ is the function which sends x to the set of all elements related to it.

We now define a function (optimistically called ϕ^{-1}) in the other direction. This function ϕ^{-1} takes as input an element of $\mathsf{Fun}(A, \mathcal{P}(B))$ (that is, a function $A \to \mathcal{P}(B)$) and gives as output a relation from A to B. Given $r : A \to \mathcal{P}(B)$, define

$$\phi^{-1}(r) \subseteq A \times B; \qquad \phi^{-1}(r) = \{(x, y) | y \in r(x)\}.$$

The verification that ϕ^{-1} is indeed inverse to ϕ is left as an exercise. □

Exercise 148. Complete the above proof.

XII.4 Summary

This lecture revolved around several examples of *bijective correspondences* between sets.

- Two sets are in bijective correspondence when there exists a bijection from one to the other (and hence also in the other direction).
- When two sets are in bijective correspondence, we often think of their elements as containing the same information, but described in a different manner.
- When $A \cong B$, this means we can go back and forth between A and B without losing any information.

Bijective correspondences occur all over mathematics. We limited ourselves to:

- Bijective correspondences in the context of Cartesian products of sets;
- The correspondence between subsets and functions into $\{0, 1\}$;
- The correspondence between relations and functions into powersets.

When trying to establish that $A \cong B$, the following strategy is usually recommended:

- First write out carefully what a typical element of the set A looks like; similarly, spell out what elements of B amount to.
- Ask yourself what information is contained in an element of A, and how that information can be used to specify an element of B.
- Turn this into a definition of a function $f : A \to B$.

- Prove that f is a bijection.

If you're having difficulties seeing how to turn an element of A into an element of B, the best thing is often to consider a concrete example and then to try to generalize that.

XII.5 Exercises

Exercise 149. List all possible bijective functions from $A = \{a, b, c\}$ to $B = \{0, 1, 2\}$.

Exercise 150. True or False? (Give a proof or a counterexample.) If a function $f : A \to A$ is bijective, then it is the identity function.

Exercise 151. Consider the set $B = \{\text{John, Mary, Kimberly}\}$, and the subset $B = \{Mary, Kimberly\}$. Work out the characteristic function of this subset.

Exercise 152. Consider the set $A = \{0, 1, 2, 3\}$, as well as the following functions from A to $\{0, 1\}$.

$$f(0) = f(2) = 0, f(1) = f(3) = 1 \quad g(0) = 0, g(1) = g(2) = g(3) = 1$$

Find the subsets of A corresponding to these functions.

Exercise 153. Consider the set of natural numbers $\mathbb{N}$, and the subset $P \subseteq \mathbb{N}$ of prime numbers. Describe the characteristic function of this subset.

Exercise 154. Consider the function $\max : \{0, 1\} \times \{0, 1\} \to \{0, 1\}$. What is the subset of $\{0, 1\}$ corresponding to this function?

Exercise 155. Let U, V be subsets of A with characteristic functions χ_U and χ_V respectively. Show that $\chi_{U \cap V}$ is the function

$$\chi_{U \cap V}(x) = \chi_U(x) \cdot \chi_V(x)$$

(i.e., multiply the two characteristic functions). Find similar formulas for $\chi_{U \cup V}, \chi_{U^c}$ and $\chi_{U - V}$.

Exercise 156. Consider the set $B = \{\text{John, Mary, Kimberly}\}$. Suppose John loves Mary and Kimberly, Mary loves John, and Kimberly only loves herself. Work out the function $B \to \mathcal{P}(B)$ corresponding to this relation.

Exercise 157. Under the correspondence between relations $R \subseteq A \times A$ and functions $r : A \to \mathcal{P}(A)$, which function corresponds to the empty relation? To the maximal relation? And to the identity relation?

Exercise 158. Suppose a relation $R \subseteq A \times B$ is total. Show that the corresponding function $r : A \to \mathcal{P}(B)$ is not surjective.

Exercise 159. For every set X there is a function $\eta_X : X \to \mathcal{P}(X)$ defined by

$$\eta_X(x) = \{x\}.$$

To which relation $X \to X$ does this function correspond?

Exercise 160. Prove the following bijective correspondences. (We write 0 for the empty set.)

1. $A + 0 \cong A \cong 0 + A$
2. $A + B \cong B + A$
3. $A + (B + C) \cong (A + B) + C$
4. $A \times 0 \cong 0 \cong 0 \times A$

Exercise 161. Prove the *distributive law* $A \times (B + C) \cong (A \times B) + (A \times C)$.

Exercise 162. Recall that Y^X is the set of all functions $Y \to X$. Prove:

(a) $A^1 \cong A$

(b) $1^A \cong 1$

(c) $A^0 \cong 1$

(d) $(A \times B)^C \cong A^C \times B^C$

(e) $A^{B+C} \cong A^B \times A^C$

(f) $A^{B \times C} \cong (A^B)^C$

Exercise 163. Do we have $0^A \cong 0$?

Exercise 164. Recall that $\mathsf{Rel}(A, B)$ denotes the set of relations from A to B. Prove: $\mathsf{Rel}(A, B) \cong \mathsf{Rel}(B, A)$. Hint: Lemma X.2.3.

Exercise 165. Prove that $\mathsf{Rel}(A \times B, C) \cong \mathsf{Rel}(A, B \times C)$.

LECTURE XIII

EQUIVALENCE RELATIONS

We have studied one particular class of relations, namely functions. This lecture is dedicated to a different but equally important class, namely *equivalence relations*.

XIII.1 DEFINITION

We begin by introducing several properties relations may have. We stress that these only make sense for relations from a set to itself.

Definition XIII.1.1. Let A be a set and $R \subseteq A \times A$ a relation on A. Then R is called:

1. *reflexive* when $\forall x \in A.xRx$
2. *irreflexive* when $\forall x \in A.\neg xRx$
3. *symmetric* when $\forall x, y \in A.[xRy \to yRx]$
4. *anti-symmetric* when $\forall x, y \in A.[xRy \wedge yRx \to x = y]$[1]
5. *transitive* when $\forall x, y, z \in A.[xRy \wedge yRz \to xRz]$.

For an example, consider the relations R, S, V and W in Figure XIII.1.

First, let's see which of these relations satisfy reflexivity. Pictorially, this means checking whether each element has a loop attached to it. Relation R on A certainly meets this criterion, but the others don't.

[1] Sometimes, one encounters the stronger definition $\forall x, y \in A.\neg(xRy \wedge yRx)$; this excludes the possibility that xRx. We will use the weaker version.

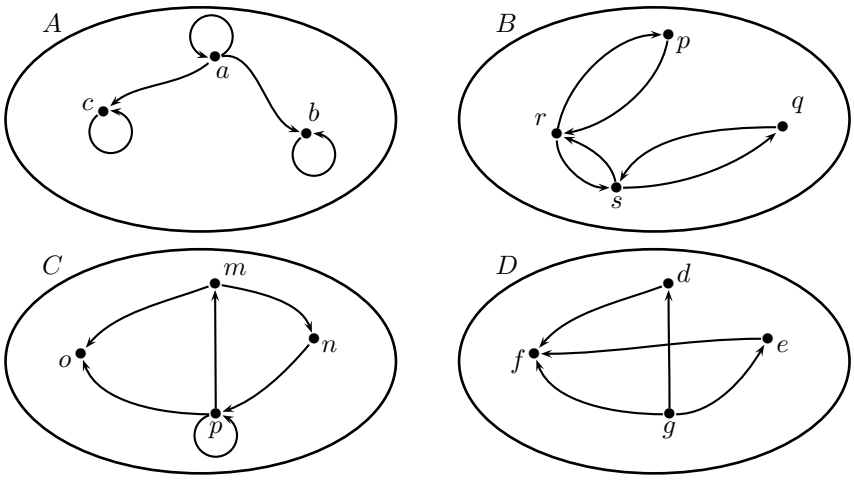


Figure XIII.1: Relations $R \subseteq A \times A, S \subseteq B \times B, V \subseteq C \times C, W \subseteq D \times D$

To test irreflexivity, no element may have a loop attached to it. Relations S and W have this property, while R and V don't. Note that relation V is neither reflexive nor irreflexive, because some elements have a loop, others don't.

Next, symmetry: this means that *if* there is an arrow from one element to another, *then* there must also be an arrow back. The only symmetric relation here is S.

For anti-symmetry we test the opposite: *if* there is an arrow from one element to another, *then* there cannot be an arrow back (unless the two elements are actually the same). Thus S is not anti-symmetric because sSq and qSs but not $s = q$. All other relations here are anti-symmetric.

Finally, transitivity: we must check that *if* there exists an arrow from x to y and one from y to z *then* we must also have an arrow from x to z. Relation R meets this criterion: every two-step path can also be done in one step. Relation S does not: for example, we can get from p to r and from r to p, but not from p to p. Also, V is not transitive: we can get from p to m and from m to n, but not from p to n. Finally, W is transitive. The results are summarized in Table XIII.1.

Here are some examples from everyday life:

Examples XIII.1.2.

Relation	**Reflexive**	**Irreflexive**	**Symmetric**	**Antisymmetric**	**Transitive**
$R \subseteq A \times A$	Yes	No	No	Yes	Yes
$S \subseteq B \times B$	No	Yes	Yes	No	No
$V \subseteq C \times C$	No	No	No	Yes	No
$W \subseteq D \times D$	No	Yes	No	Yes	Yes

Table XIII.1: Properties of R, S, V, W

1. The ordering relation $\leq$ on $\mathbb{N}$ is reflexive: for each $x \in \mathbb{N}$ we have $x \leq x$. It is not irreflexive, because, for example, we have $3 \leq 3$. It is not symmetric: we have, for example, $2 \leq 6$ but not $6 \leq 2$. It is anti-symmetric: if $x \leq y$ and $y \leq x$ then $x = y$. And it is transitive: if $x \leq y$ and $y \leq z$ then also $x \leq z$.
2. The strict ordering $<$ on $\mathbb{N}$ is not reflexive, is irreflexive, not symmetric, anti-symmetric, and transitive. Make sure you understand why it is anti-symmetric!
3. The subset relation on $\mathcal{P}(X)$ is reflexive, not irreflexive, not symmetric, anti-symmetric and transitive (check this).

The difficult part (according to most students!) is that some relations can *vacuously* satisfy some or even all of these properties. For example, consider the set $A = \{a, b\}$ and the relation $R = \{(a, b)\}$. This relation is vacuously transitive: there are no elements x, y, z with xRy and yRz, so the antecedent of the implication in the definition is always false.

Here's the important exercise you should do: for every pair of properties, find: (a) a relation satisfying both properties, (b) a relation satisfying neither, (c) a relation satisfying the first but not the second property and (d) a relation satisfying the second but not the first property. Find the smallest possible examples! You can't claim to understand the definitions if you haven't successfully completed this exercise.

Logical Aspects

Keep in mind that a sentence such as

$$\forall xyz.(xRy \wedge yRz \rightarrow xRz)$$

is true in case for each x, y, z we have

if xRy and yRz *then* xRz.

Thus you need to check this implication for every possible choice of x, y and z, including choices where $x = y$, $y = z$, or $x = z$. For each choice, keep in mind that if the antecedent is false, the implication is true.

XIII.2 Equivalence relations

The previous definitions were just a preparatory step. The real concept of interest in this lecture is the following:

Definition XIII.2.1 (Equivalence Relation)**.** A relation $R \subseteq A \times A$ is an *equivalence relation* when it is reflexive, symmetric and transitive.

Often equivalence relations are denoted by symbols such as $\sim$. We discuss a number of examples to illustrate. For each of them, you should check the details.

Examples XIII.2.2.

1. When $A = \{*\}$ (that is, A has one element), there is only one equivalence relation, namely $\{(*, *)\}$. (The other relation on A, namely the empty one, is not reflexive.)
2. When A is any set, the identity relation $\Delta_A = \{(x, x) | x \in A\}$ is an equivalence relation. It is the smallest equivalence relation on A, in the sense that when R is any equivalence relation on A we have $\Delta_A \subseteq R$. (It is also the smallest reflexive relation on A.)

3. When A is any set, the maximal relation $A \times A$ is an equivalence relation. It is the largest equivalence relation on A.

4. Here's an example of an equivalence relation on $\mathbb{R}$: set

$$x \sim y \Leftrightarrow |x| = |y|.$$

That is, declare x and y equivalent if they have the same absolute value.

5. Another equivalence relation on $\mathbb{R}$: set

$$x \sim y \Leftrightarrow |x - y| \in \mathbb{N}.$$

That is, x and y are equivalent when the distance between them is an integer.

6. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Define

$$x \sim y \Leftrightarrow f(x) = f(y).$$

(Example 4 was a special case with $f(x) = |x|$.)

7. Here's an equivalence relation on the set $\mathbb{R} \times \mathbb{R}$: set

$$(u, v) \sim (x, y) \Leftrightarrow u^2 + v^2 = x^2 + y^2.$$

That is, two points in the plane are equivalent when they have the same distance to the origin.

8. Finally, we discuss an equivalence relation on $\mathbb{Z}$: First fix a natural number $n > 0$. Then define

$$x \sim_n y \Leftrightarrow x - y \text{ is divisible by } n.$$

This is reflexive: for $x \in \mathbb{N}$, we have $x - x = 0$, which is certainly divisible by n. It is also symmetric: if $x - y$ is divisible by n, then so is $y - x = -(x - y)$. And finally, it is transitive: if $x - y = kn$ and $y - z = ln$ for some integers k, l, then

$$x - z = x - y + y - z = kn + ln = (k + l)n,$$

so $x - z$ is also divisible by n.

Note that we have one such equivalence relation for each choice of n.

XIII.3 Equivalence classes

When $\sim$ is an equivalence relation on a set A and $x \sim y$, we say that x and y are *equivalent*. Often, we want to identify equivalent elements, because for the particular purposes we have in mind equivalent elements are to be regarded as equal. For a very practical example, consider a game of cards. If you get dealt a hand you don't care about the order of the cards: you regard two hands as the same when they contain the same cards. Thus there is an equivalence relation on hands: two hands are equivalent if you can get one from the other by permuting the cards.

The idea of "regarding equivalent elements as equal" is made formal in the following way. First, fix an element $x \in A$. Given that element, we can ask for all elements which are equivalent to x. Denote the set of these elements by $[x]_\sim$, or simply by $[x]$ when the equivalence relation is understood[2] . Thus:

$$[x]_\sim =_{def} \{y \in A | x \sim y\}.$$

The set $[x]_\sim$ (which by definition is a subset of A) is called the *equivalence class* of x.

To continue the card example, suppose that in a certain game you get dealt three cards. Then A would be the set of all possible ways of dealing three cards (where $AcKdQh$ would be a different element from $KdQhAc$, say); given a particular hand x, for example $AcKdQh$, the equivalence class would be

$$[AcKdQh] = \{AcKdQh, AcQhKd, KdAcQh, KdQhAc, QhAcKd, QhKdAc\}.$$

Thus the equivalence class consists of all possible permutations of the hand.

Here's another example: let $A = \{a, b, c, d, e, f\}$, and consider the equivalence relation

$$\{(a,a), (b,b), (c,c), (d,d), (e,e), (f,f), (a,b), (b,a), (c,d), (d,c), (c,e), (e,c), (d,e), (e,d)\}.$$

Then we have

$$[a] = [b] = \{a, b\}, \qquad [c] = [d] = [e] = \{c, d, e\}, \qquad [f] = \{f\}.$$

Several more elaborate examples will be discussed in the following section.

XIII.4 The canonical quotient map

Now that we have defined equivalence classes, we consider them all at once:

Definition XIII.4.1. Let $\sim$ be an equivalence relation on a set A. Then the *quotient* of A by $\sim$ is the set

$$A/\sim =_{def} \{[x] | x \in A\}.$$

In words: $A/\sim$ is the set of all equivalence classes. Note that this is a subset of $\mathcal{P}(A)$, because the $[x]$ are subsets of A.

How are the sets A and $A/\sim$ related to each other? To each element $x \in A$ we can associate the element $[x] \in A/\sim$. This is in fact a function

$$\pi : A \to A/\sim; \qquad \pi(x) = [x].$$

This function is called the *canonical quotient map* associated to the equivalence relation $\sim$. This terminology is justified by the fact that π is a surjective function: each $[x] \in$

[2] In certain contexts other notations are common; in modular arithmetic for example, the equivalence class of an integer n modulo p is denoted $\overline{n}$.

$A/\sim$ is of the form $\pi(x)$. It is this quotient function which makes precise the sense in which we identify equivalent elements of A.

In the remainder of this section, we revisit the examples of equivalence relations and explain what the quotient map amounts to in these cases.

Examples XIII.4.2.

1. When $A = \{*\}$ there is nothing to identify: $[*] = \{*\}$, and $A/\sim = \{[*]\} = \{\{*\}\}$, again a 1-element set. Informally, when you have a set with just one element there is nothing to identify, so the quotient map is a bijection.

2. When A is any set and $\Delta_A = \{(x, x)|x \in A\}$ is the identity relation, we have $[x] = \{x\}$. Then $A/\sim = \{\{x\}|x \in A\}$, an the quotient map $\pi(x) = \{x\}$ is not only surjective, but also injective.

3. When A is any set and $\sim$ is the maximal relation $A \times A$, we have $[x] = A$ for any $x \in A$. Thus $A/\sim = \{A\}$, and the quotient map is $\pi(x) = A$ for all x.

4. When $\sim$ is the equivalence relation on $\mathbb{R}$ given by

$$x \sim y \Leftrightarrow |x| = |y|,$$

   we get $[x] = \{x, -x\}$. Thus each class (except for that of 0) has precisely two elements.

5. For the equivalence relation on $\mathbb{R}$ given by

$$x \sim y \Leftrightarrow |x - y| \in \mathbb{N}.$$

   we find that $[x] = \{x + n|n \in \mathbb{Z}\}$. Thus each equivalence class has infinitely many elements.

6. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Define

$$x \sim y \Leftrightarrow f(x) = f(y).$$

   Then $[x] = \{y|f(x) = f(y)\}$, that is, the class of x is the set of all elements which are sent by f to the same value as x. We will have a closer look at this example in the next lecture, where we will also explore the quotient map.

7. For the equivalence relation on the set $\mathbb{R} \times \mathbb{R}$ given by

$$(u, v) \sim (x, y) \Leftrightarrow u^2 + v^2 = x^2 + y^2$$

   the class of a point (x, y) is the set of points which lie on the same circle around the origin as (x, y). Thus $(\mathbb{R} \times \mathbb{R})/\sim$ is the set of circles around the origin.

8. Finally, consider the equivalence relation on $\mathbb{Z}$ given by

$$x \sim_n y \Leftrightarrow x - y \text{ is divisible by } n.$$

   Then $[x] = \{x + kn|k \in \mathbb{Z}\}$. There are n equivalence classes: $[0], [1], \ldots, [n - 1]$. (Check for yourself that these classes are all different, and that $[0] = [n]$, so that there aren't any other classes.) Working with these classes is called *arithmetic modulo n*.

XIII.5 Summary

We introduced *equivalence relations* as relations which are reflexive, symmetric and transitive. Given an equivalence relation $\sim$ on A, the following constructions are the focus of attention:

- The *equivalence classes*, defined by $[x]_\sim = \{y | y \sim x\} \subseteq A$, and
- The *canonical quotient map* from A to the set $A/\sim = \{[x]_\sim | x \in A\}$.

The quotient map, which is always surjective, allows us to identify equivalent elements.

XIII.6 Exercises

Exercise 166. Let A be a non-empty set, and R a single-valued relation on A. Can a single valued relation be reflexive? Irreflexive? Symmetric? Anti-symmetric? Transitive? Can it be an equivalence relation? In each case, give as simple an example as possible, or explain why this is not possible.

Exercise 167. Same question, but now for total relations.

Exercise 168. Same question, but now for functional relations.

Exercise 169. Let A be the set of all people. What properties does the relation "is a brother of" have? What about "is an ancestor of", and "is a sibling of"?

Exercise 170. Let $R \subseteq A \times A$ be a relation on A. Prove:

(a) R is reflexive if and only if $\Delta_A \subseteq R$.

(b) R is irreflexive if and only if $\Delta_A \cap R = \emptyset$.

(c) R is symmetric if and only if $R^\circ \subseteq R$ if and only if $R \subseteq R^\circ$ if and only if $R = R^\circ$.

(d) R is anti-symmetric if and only if $R \cap R^\circ \subseteq \Delta_A$.

(e) R is transitive if and only if $R \circ R \subseteq R$.

Exercise 171. Do transitive relations necessarily satisfy $R \circ R = R$? Give a proof or a counterexample. What about equivalence relations?

Exercise 172. Let $A = \{a, b, c, d, e\}$ and consider the relation

$$R = \Delta_A \cup \{(a, e), (e, a), (c, b), (b, c)\}.$$

Verify that this is an equivalence relation. Describe the equivalence classes. How many distinct equivalence classes are there? Describe the quotient map.

Exercise 173. Same question, but now for the relation

$$S = R \cup \{(a,b),(b,a),(a,c),(c,a),(c,e),(e,c),(b,e),(e,b)\}.$$

Exercise 174. Suppose I'm thinking of a set and an equivalence relation on it. I tell you that there is precisely one equivalence class. What can you conclude about the equivalence relation?

Exercise 175. Consider the following relation on $A = \mathbb{Z} \times \mathbb{Z}$:

$$(x,y) \sim (u,v) \Leftrightarrow x + y = u + v.$$

Prove that this is an equivalence relation. Describe the equivalence classes.

Exercise 176. Let $p : \mathbb{N} \to \{0,1\}$ be the characteristic function of the even numbers, i.e., $p(x) = 1$ if x is even and $p(x) = 0$ if x is odd. Set

$$x \sim y \Leftrightarrow p(x) = p(y).$$

Prove that $\sim$ is an equivalence relation. Describe the equivalence classes. What is the set $\mathbb{N}/\sim$?

Exercise 177. Let $\mathsf{sng} : \mathbb{Z} \to \{1,0,-1\}$ be the sign function, i.e. the function which outputs 1 if the input is positive, 0 if the input is 0 and -1 if the input is negative. Define an equivalence relation on $\mathbb{Z}$ by $x \sim y \Leftrightarrow \mathsf{sgn}(x) = \mathsf{sgn}(y)$.

Prove that this is an equivalence relation. Describe the equivalence classes. What is the quotient $\mathbb{Z}/\sim$?

Exercise 178. Suppose R and S are equivalence relations on a set A.

1. Prove that $R \cap S$ is also an equivalence relation.
2. Prove that $R \cup S$ is not always an equivalence relation.
3. Is $R \circ S$ an equivalence relation?
4. Suppose that $R \circ S = S \circ R$. Prove that $R \circ S$ is an equivalence relation.

Exercise 179. Suppose that R and S are both equivalence relations on A with $R \subseteq S$. Prove that there exists a unique function $\phi : A/R \to A/S$ such that the composite $A \to A/R \to A/S$ equals the canonical quotient $A \to A/S$.

Exercise 180. Consider the set $\mathsf{Fun}(\mathbb{N},\mathbb{N})$ of functions from the set of natural numbers to itself. Define a relation R on this set by

$$fRg \Longleftrightarrow_{\text{def}} f(x) \neq g(x) \text{ for at most one } x \in \mathbb{N}.$$

That is, call two functions equivalent when they differ on at most one argument. Is this an equivalence relation?

LECTURE XIV

PARTITIONS

We have learned that an equivalence relation $\sim$ on a set A gives rise to equivalence classes $[x] = \{y | y \sim x\}$. Moreover, these equivalence classes are subsets of A, and are the elements of the quotient set $A/\sim$. In this lecture we take a closer look at sets of the form $A/\sim$. The main result we prove is that equivalence relations on a set A are really "the same thing" as partitions of A.

XIV.1 PARTITIONINGS

We begin with a small illustrative example. Consider the set A on the left in Figure XIV.1 and the equivalence relation $\sim$ on it.

As is clear pictorially, the elements of A which are in the same equivalence class are connected via arrows; those who are not equivalent (i.e., have distinct equivalence

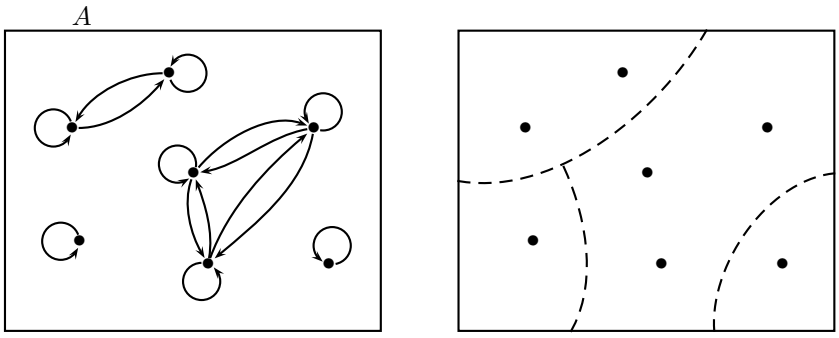


Figure XIV.1: An equivalence relation and the corresponding partitioning

classes) are not connected. On the right, we have indicated how the equivalence classes partition the set into four subsets. Note how both pictures contain the same information, but represent it in a different manner.

The collection of four subsets in the picture on the right has three important features: first, none of the subsets is empty. Second, the subsets don't overlap, i.e., the subsets are *pairwise disjoint*. And third, each element of A is in precisely one of the subsets: we express this by saying that the subsets jointly *cover* the set A. We now take these properties as a definition:

Definition XIV.1.1. Let A be a set and let $\mathcal{U}$ be a collection of subsets of A. (Thus $\mathcal{U} \subseteq \mathcal{P}(A)$.) Then $\mathcal{U}$ is a *partitioning* of A when

- Each $U \in \mathcal{U}$ is nonempty
- For each $U, V \in \mathcal{U}$: $U \neq V$ implies $U \cap V = \emptyset$ (the elements of $\mathcal{U}$ are pairwise disjoint)
- For all $x \in A$ there is a $U \in \mathcal{U}$ with $x \in U$ (the elements of $\mathcal{U}$ cover A).

Intuitively, a partitioning of A divides up A into "compartments". We give several examples of partitionings first; we return to the connection with equivalence relations in the next section.

Examples XIV.1.2.

1. Let A be any set. Then the collection $\mathcal{U} = \{\{x\} | x \in A\}$ is a partitioning of A. This is the *finest* partitioning of A.
2. On the other extreme, there is a partitioning $\mathcal{U} = \{A\}$ of A. This is the *coarsest* partitioning of A.
3. The collection $\{\text{even}, \text{odd}\}$ is a partitioning of $\mathbb{Z}$.
4. The collection of sets $\{C_r | r \geq 0\}$, where $C_r = \{(x, y) | x^2 + y^2 = r\}$ is a partitioning of $\mathbb{R}^2$. (It partitions the plane into concentric circles.)

XIV.2 From equivalence relation to partition

The example given at the beginning of this lecture illustrates how an equivalence relation gives rise to a partitioning. In fact, we had already seen this, because the partitioning is none other than the set $A/\sim$, the quotient of A by the equivalence relation. This is the content of the following proposition:

Proposition XIV.2.1. *Let $\sim$ be an equivalence relation on a set A. Then $A/\sim$ is a partitioning of A.*

Proof. We need to show that $\{[x]_\sim | x \in A\}$ is a partitioning of A.

First, note that we have $x \in [x]$ (because $\sim$ is reflexive). Therefore each $[x]$ is nonempty.

Second, suppose we have $[x], [y]$ with $[x] \neq [y]$. Then $x \not\sim y$, because if $x \sim y$ we get $[x] = [y]$. Now if we had $z \in [x] \cap [y]$ we had $z \sim x$ and $z \sim y$. But since $\sim$ is symmetric and transitive this would force $x \sim y$. Contradiction. Hence $[x] \cap [y] = \emptyset$.

Finally, the sets $[x]$ cover A, since for each $x \in A$ we have $x \in [x]$. □

XIV.3 From partition to equivalence relation

Now start with a partitioning $\mathcal{U}$ of a set A. We want to turn this into an equivalence relation on A. What should that relation be? The partitioning tells us: two elements of A should be equivalent when they are in the same partition. Thus we try:

$$x \sim_{\mathcal{U}} y \Leftrightarrow \exists U \in \mathcal{U}. x \in U \wedge y \in U.$$

We call this the equivalence relation *induced* by the partitioning $\mathcal{U}$. Of course, we must prove that this is actually an equivalence relation.

Proposition XIV.3.1. *When $\mathcal{U}$ is a partitioning of A, then the induced relation $\sim_{\mathcal{U}}$ on A is an equivalence relation.*

Proof. First, we show that $\sim_{\mathcal{U}}$ is reflexive. So consider $x \in A$. To show $x \sim_{\mathcal{U}} x$, we must find $U \in \mathcal{U}$ with $x \in U$. But since $\mathcal{U}$ covers A, such U is guaranteed to exist.

Next, we must show that $\sim_{\mathcal{U}}$ is symmetric. But that is trivial: if there is $U \in \mathcal{U}$ with $x \in U$ and $y \in U$, then for that same U we have $y \in U$ and $x \in U$.

Finally, to see that $\sim_{\mathcal{U}}$ is transitive, suppose we have $x \sim_{\mathcal{U}} y$ and $y \sim_{\mathcal{U}} z$. Then there is $U \in \mathcal{U}$ with $x, y \in U$, and there is $V \in \mathcal{U}$ with $y, z \in V$. Therefore $z \in U \cap V$. Since the elements of $\mathcal{U}$ are pairwise disjoint, it follows that $U = V$. Thus $x, z \in U$, and hence $x \sim_{\mathcal{U}} z$. □

To wrap up the discussions so far: each equivalence relation on A gives a partitioning of A, and conversely each partitioning of A gives an equivalence relation on A. We are now in a position to prove the following important result:

Theorem XIV.3.2. *There is a bijective correspondence between the set of equivalence relations on A and the set of partitionings of A.*

Before giving the proof, let us introduce notation: we will write $\mathbf{EqRel}(A)$ for the set of equivalence relations on A; $\mathbf{Part}(A)$ will denote the set of partitionings of A.

Proof. We must show that $\mathbf{EqRel}(A) \cong \mathbf{Part}(A)$. Thus we must produce a bijection between these sets. We already have defined, in the previous two sections, the functions we have in mind:

$$\Phi : \mathbf{EqRel}(A) \to \mathbf{Part}(A); \qquad \Phi(R) = \{[x]_R | x \in A\}$$

and

$$\Psi : \mathbf{Part}(A) \to \mathbf{EqRel}(A); \qquad \Psi(\mathcal{U}) = \sim_{\mathcal{U}} .$$

We are thus left with the task of showing that these functions are inverse to each other.

First, to show that $\Phi \circ \Psi = 1_{\mathbf{Part}(A)}$, consider a partitioning $\mathcal{U}$ of A. Then $\Psi(\mathcal{U})$ is the relation $\sim_{\mathcal{U}}$, so $\Phi\Psi(\mathcal{U})$ is the partitioning associated to $\sim_{\mathcal{U}}$. Then

$$[x]_{\sim_{\mathcal{U}}} = \{y | y \sim_{\mathcal{U}} x\} = \{y | \exists U \in \mathcal{U}. y, x \in U\} = U.$$

Thus each class $[x]_{\sim_{\mathcal{U}}}$ is a member of $\mathcal{U}$. Conversely given $U \in \mathcal{U}$, pick $x \in U$; then the same reasoning gives $U = [x]_{\sim_{\mathcal{U}}}$. Hence $\mathcal{U} = \{[x]_{\sim_{\mathcal{U}}} | x \in A\} = \Phi\Psi(\mathcal{U})$.

Second, to show that $\Psi \circ \Phi = 1_{\mathbf{EqRel}(A)}$, consider an equivalence relation $\sim$ on A. Then the equivalence relation $\Psi\Phi(\sim)$ has

$$(x, y) \in \Psi\Phi(\sim) \quad \Leftrightarrow \quad \exists U \in \Phi(\sim).(x, y \in U) \quad \Leftrightarrow \quad \exists [z] \in A/\sim .(x, y \in [z]) \quad \Leftrightarrow \quad x \sim y$$

The last step follows again from the fact that $x \in [z], y \in [z]$ implies $x \sim z$ and $y \sim z$. □

XIV.4 Summary

We have introduced *partitionings* as collections of nonempty subsets which are pairwise disjoint and jointly covering. The relationship between partitionings and equivalence relations then has two aspects:

- Each equivalence relation $\sim$ on A gives a partitioning on A, namely $A/\sim$;
- Each partitioning $\mathcal{U}$ of A gives an equivalence relation on A, namely $x \sim_{\mathcal{U}} y \Leftrightarrow \exists U \in \mathcal{U}.(x, y \in U)$.

These two constructions are mutually inverse, in the sense that they define a bijective correspondence between the set of equivalence relations on A and the set of partitionings of A.

XIV.5 Exercises

Exercise 181. Let $A = \{a, b, c, d\}$. Find all possible partitionings of A.

Exercise 182. For each of the partitionings of the set $A = \{a, b, c, d\}$, find the corresponding equivalence relation on A.

Exercise 183. Show that the collection $\{[n, n+1) | n \in \mathbb{Z}\}$ is a partitioning of $\mathbb{R}$.

Exercise 184. Is the collection $\{(-n, n) | n \in \mathbb{N}\}$ a partitioning of $\mathbb{R}$?

No, because these sets are not pairwise disjoint, e.g., $(-1, 1) \cap (-2, 2) = (-1, 1) \neq \emptyset$.

Exercise 185. Show that for any set A and any subset $U \subset A$, the set $\{U, U^c\}$ is a partitioning of A.

Exercise 186. Suppose $f : \mathbb{R} \to \mathbb{R}$ is a function, and we define

$$x \sim y \Leftrightarrow f(x) = f(y).$$

Find a geometric description in terms of the graph of f of the partitioning associated to this equivalence relation.

Exercise 187. Find the partitioning associated to the equivalence relation on $\mathbb{R}$ given by

$$x \sim y \Leftrightarrow |x - y| \in \mathbb{N}.$$

Exercise 188. What is $\mathsf{Part}(\emptyset)$?

Exercise 189. Let S be the set of strings of 0's and 1's of length 4. That is,

$$S = \{abcd | a, b, c, d \in \{0, 1\}\}.$$

(a) List all elements of S

(b) Consider the relation on S defined by $s \sim t$ iff s is a permutation of t. Show that this is an equivalence relation.

(c) Describe all equivalence classes of $\sim$ and the associated partitioning of S.

LECTURE XV

FAMILIES

Often, we are not just interested in one specific set, but a whole collection of them at once. In this case it is advantageous to formulate this explicitly using *families of sets*. This lecture introduces the basic ideas surrounding families of sets, as well as connections with the concepts we have studied in previous lectures.

XV.1 INDEXINGS

We begin with a common example of a family of sets. Consider the closed interval $[a, a+1] = \{x \in \mathbb{R} | a \leq x \leq a+1\} \subseteq \mathbb{R}$, where $a \in \mathbb{Z}$ is an integer. Since we can consider different values for a, we in fact have many sets, one for each $a \in \mathbb{Z}$. We express this by saying that $[a, a+1]_{a \in \mathbb{Z}}$ is a *family of sets indexed by* $a \in \mathbb{Z}$, or simply by $\mathbb{Z}$.

Generally, a family of sets is given by two pieces of data: first, a set I (called the *indexing set*); second, for each $i \in I$, a set A_i. Often, the notation $\mathcal{A} = (A_i)_{i \in I}$ is used to denote such a family. It is important to note that we don't require that these sets A_i are all distinct. Thus a family of sets is not the same thing as a set of sets!

A useful intuition is that a family of sets is like a filing system: the elements of the indexing set I are thought of as labels, and then associated to each label i we have a set A_i "stored under that label".

Here are some more examples:

Examples XV.1.1.

1. Consider the set $I = \{1, 2, 3\}$ and the sets $A_1 = \{a, b\}, A_2 = \{b, c, d\}, A_3 = \{b, c\}$. In this case, the family has three members. Note that in this case these happen to be distinct, but they may overlap. A pictorial representation is given in Figure XV.1.

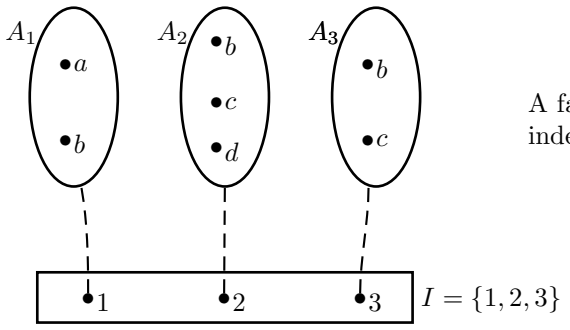


Figure XV.1: Example of a family of sets

2. In our terminology, a family of sets indexed by $I = \{1, 2, \ldots, n\}$ consists of n sets. In particular for $n = 1$, the family has just 1 member, while for $n = 0$, the family is empty. (Which is not the same as the 1 member family consisting of the empty set!)

3. A *bit string* is a finite sequence of 0's and 1's. For examples, 101101000 is a bit string, as is 0010. Each bit string has a *length*, namely the number of bits in it. Write B_n for the set of bit strings of length n. This is a family of sets: $(B_n)_{n\in\mathbb{N}}$.

4. When a is a positive real number, we may consider the circle around the origin with radius a;
$$C_a = \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 = a^2\}.$$
Since we have such a circle for each a, we have in fact a family $(C_a)_{a\in\mathbb{R}^+}$.

5. When I is a set and X is another set, we can consider the family $(X_i)_{i\in I}$, where simply $X_i = X$. Thus all members of this family are identical to X (and hence to each other as well). Such families are sometimes called *constant*.

6. One more extreme example: when I is any set, we have the family of elements of I: $\{i\}_{i\in I}$.

Remark XV.1.2. Technically, we have been a bit sloppy here. How does it follow from the axioms of set theory that families of sets indeed exist? The answer is to be a bit more careful about introducing them. When I is an indexing set and A is another set, then we can consider a function $I \to \mathcal{P}(A)$. This amounts to an I-indexed family (of subsets of A). Now to be guaranteed that a general I-indexed family $(A_i)_{i\in I}$ exists (in the sense that its existence follows from the axioms of set theory) we merely need to make sure that each of the sets A_i can be regarded as a subset of a set A (which has been shown to exist already).

XV.2 Unions and Intersections

Earlier we considered the Boolean operations on sets, in particular union $A \cup B$ and intersection $A \cap B$. We also hinted at the fact that more generally one has n-ary unions

and intersections $A_1 \cup \cdots \cup A_n$ and $A_1 \cap \cdots \cap A_n$. We may now generalize from these finite operations to operations on arbitrary families of sets.

Definition XV.2.1. Let $\mathcal{A} = (A_i)_{i \in I}$ be a family of sets. Then the *union* of $\mathcal{A}$ is the set

$$\bigcup \mathcal{A} = \{x | \exists i \in I . x \in A_i\}$$

and the *intersection* of $\mathcal{A}$ is the set

$$\bigcap \mathcal{A} = \{x | \forall i \in I . x \in A_i\}.$$

In words, an element x is in the union of $\mathcal{A}$ precisely when it is in at least one of the members A_i. An element x is in the intersection of $\mathcal{A}$ when it is in all of the members A_i. Note that while $\mathcal{A}$ is a *family* of sets, both $\bigcup \mathcal{A}$ and $\bigcap \mathcal{A}$ are ordinary sets.

A slight variant on the notation is $\bigcup_{i \in I} A_i$, $\bigcap_{i \in I} A_i$.

Examples XV.2.2.

1. Consider the set $I = \{1, 2, 3\}$ and the sets $A_1 = \{a, b\}, A_2 = \{b, c, d\}, A_3 = \{b, c\}$. Then $\bigcup_{i \in I} A_i = \{a, b, c, d\}$ and $\bigcap_{i \in I} A_i = \{b\}$.
2. Consider the family $(A_n)_{n \in \mathbb{N}}$ of subsets of the real line given by $A_n = [-n, n]$. Then $\bigcup_{n \in \mathbb{N}} A_n = \mathbb{R}$ and $\bigcap_{n \in \mathbb{N}} A_n = \{0\}$.
3. When $X_i = X$ is a constant family over I, we have $\bigcup_{i \in I} X_i = X = \bigcap_{i \in I} X_i$.

The following lemma captures some properties of unions and intersections which will be used in the next section.

Lemma XV.2.3. *Let $(A_i)_{i \in I}$ be a family of sets.*

- $\bigcap_{i \in I} A_i \subseteq A_i$ *for all* $i \in I$.
- *When B is a set with $B \subseteq A_i$ for all i, then $B \subseteq \bigcap_{i \in I} A_i$.*
- $A_i \subseteq \bigcup_{i \in I} A_i$ *for all* $i \in I$.
- *When B is a set with $A_i \subseteq B$ for all $i \in I$, then $\bigcup_{i \in I} A_i \subseteq B$.*

Exercise 190. Prove these statements.

Unions and intersections satisfy algebraic properties common to their finite counterparts. Most are what you would expect, but there are also a few differences. One important fact is that there is still a version of the distributive law:

Lemma XV.2.4. *When (A_i) is a family of sets and B is a set, we have*

$$B \cap \bigcup_{i \in I} A_i = \bigcup_{i \in I} (B \cap A_i).$$

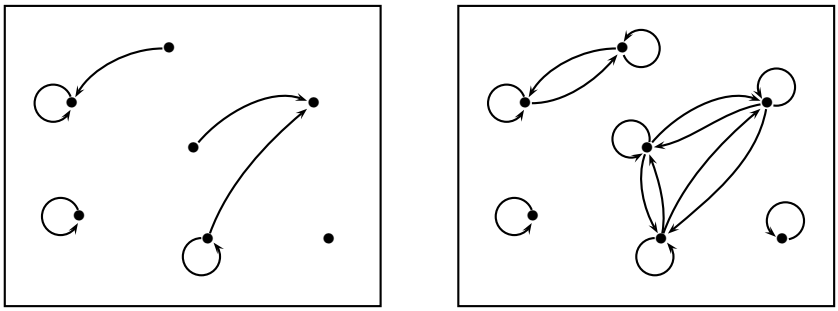


Figure XV.2: A relation and the equivalence relation generated from it

Proof. We'll do a proof using predicate logic:

$$\begin{aligned} x \in B \cap \bigcup_{i\in I} A_i &\leftrightarrow x \in B \wedge x \in \bigcup_{i\in I} A_i \\ &\leftrightarrow x \in B \wedge \exists i. x \in A_i \\ &\leftrightarrow \exists i. x \in B \wedge x \in A_i \\ &\leftrightarrow x \in \bigcup_{i\in I} (B \cap A_i) \end{aligned}$$

where the third step uses the predicate-logical equivalence $\exists i.(B \wedge A(i)) \equiv B \wedge \exists i.A(i)$ (which is allowed when i does not occur in B). □

XV.3 Closure

We give an application of unions and intersections of families to the theory of equivalence relations. As you know, being an equivalence relation is rather special; most relations are certainly not equivalence relations. Suppose we have a relation R on A which is not an equivalence relation. Can we somehow force the issue and modify R so that it becomes one? (Preferably in a way that leaves as much of R intact as possible.)

What we have in mind is illustrated in Figure XV.2: on the left we have a relation R which fails to be an equivalence relation (it is not reflexive, not symmetric and not transitive), and on the right we have "fixed" R by adding all the arrows which are needed.

How did we do it? Step 1: to make the relation reflexive, add loops to every element (which don't have one yet). Step 2: to make it symmetric, add an arrow from y to x whenever there is an arrow from x to y. Step 3: To make it transitive, add an arrow from x to z whenever there are arrows from x to y and from y to z.

A complication is that adding arrows can actually break transitivity and symmetry. Therefore, you can't simply first make the relation symmetric and then make the result transitive, because the result of that needn't be symmetric anymore! Also, in step 3 you may not end up with a transitive relation because adding arrows creates new problems.

But if that happens, you simply keep repeating step 2 and step 3 until the result is both symmetric and transitive (fortunately reflexivity never breaks).

How do we know that this process stops eventually? Well, in the example this is because there are only finitely many elements, so there are only finitely many arrows we could add. However, generally there is no guarantee that this will ever stop. For example, consider the set $\mathbb{N}$, and the relation $R = \{(x, x+1)|x \in \mathbb{N}\}$. If we try to repeat step 3, we keep adding arrows: first we add $(x, x+2)$, then $(x, x+3)$, and so on. This never stops.

Here's where families come in. First, notice that if all we cared about is finding an equivalence relation containing R then the answer would be simple: just take the maximal relation. However, we want to do better than that: we want to find the *least* equivalence relation containing R. This is worth a definition:

Definition XV.3.1. When R is a relation on A, the equivalence relation *generated* by R is a relation $\overline{R}$ on A such that

- $\overline{R}$ is an equivalence relation
- $R \subseteq \overline{R}$
- Whenever S is an equivalence relation containing R, we have $\overline{R} \subseteq S$.

This makes precise the sense in which we aim for a minimal solution; it excludes lazy solutions such as the maximal equivalence relation.

Proposition XV.3.2. *For every relation R on A, there exists a (necessarily unique) least equivalence relation generated by R.*

Proof. First, consider the family $\mathcal{A}$ of all possible equivalence relations on A which contain R. (This is indeed a family of sets, because every equivalence relation is a subset of $A \times A$.)

Now consider the intersection of this family:

$$\overline{R} =_{def} \bigcap \mathcal{A} = \bigcap \{S|S \text{ is an eq. rel with } R \subseteq S\}.$$

Notice that this family is nonempty, because it always contains the maximal equivalence relation $A \times A$. (So the lazy solution is good for something after all.)

Next, we claim that this $\overline{R}$ is the equivalence relation generated by R. We must first show that $\overline{R}$ is indeed an equivalence relation.

- Reflexivity: given $x \in A$, we know that $(x, x) \in S$ for all $S \in \mathcal{A}$, because all $S \in \mathcal{A}$ are reflexive. Thus $(x, x) \in \overline{R}$ as well.
- Symmetry: suppose $(x, y) \in \overline{R}$. Then $(x, y) \in S$ for all $S \in \mathcal{A}$. Hence (because all $S \in \mathcal{A}$ are symmetric) also $(y, x) \in S$ for all $S \in \mathcal{A}$. Thus $(y, x) \in \mathcal{R}$ as well.
- Transitivity: suppose $(x, y) \in \overline{R}$ and $(y, z) \in \overline{R}$. Then $(x, y) \in S$ and $(y, z) \in S$ for all $S \in \mathcal{A}$. This gives $(x, z) \in S$ for all $S \in \mathcal{A}$, hence also $(x, z) \in \overline{R}$.

Second, we must show $R \subseteq \overline{R}$. But that follows from lemma XV.2.3.

And third, we must show that if S is any equivalence relation containing R, then $\overline{R} \subseteq S$. But if S is indeed an equivalence relation containing R, then $S \in \mathcal{A}$. Hence (again by lemma XV.2.3) we get $\overline{R} \subseteq S$. □

XV.4 Summary

The main reason for studying *families of sets* is that often we want to consider many sets at the same time, and that many naturally occurring constructions result in sets which are indexed by some other set. When $\mathcal{A} = (A_i)_{i \in I}$ is a family of sets, we call I the *indexing set*. There are two important operations:

- Union: $\bigcup_{i \in I} A_i = \{x | \exists i \in I . x \in A_i\}$
- Intersection: $\bigcap_{i \in I} A_i = \{x | \forall i \in I . x \in A_i\}$.

Intersections can be used to show that for every relation R on A there is a *least* equivalence relation *generated by* R. This is done by taking the intersection of the family of equivalence relations containing R.

XV.5 Exercises

Exercise 191. Consider the family of sets A_i, $i \in \{1, 2, 3\}$ given by

$$A_1 = \{1, 3, 5, 7, \ldots\}, \quad A_2 = \{n \in \mathbb{N} | n \text{ divides } 99\}, \quad A_3 = \{p \in \mathbb{N} | p \text{ is prime}\}.$$

Find $\bigcap_i A_i$ and $\bigcup_i A_i$.

Exercise 192. What are the union and the intersection of the family $(-a, a)_{a \in \mathbb{N}}$? (Here, (a, b) stands for the open real interval.)

Exercise 193. Same question, but now for the family $(a, a + 1)_{a \in \mathbb{Z}}$.

Exercise 194. Consider the family of open intervals $(-1/n, 1/n)_{n \in \mathbb{N} - \{0\}}$. What is the union of this family? What is the intersection?

Exercise 195. Consider the family of half-open intervals $[0, \log_2 n)_{n \in \mathbb{N} - \{0\}}$. What is the union? And what is the intersection?

Exercise 196. Consider the family of relations $(R_c)_{n \in \mathbb{N}}$ on $\mathbb{Z}$, defined by $aR_cb \Leftrightarrow a + c \leq b$. Describe a few of the members R_c of this family. Also find the union and the intersection.

Exercise 197. Consider the relation $R \subseteq \mathbb{Z} \times \mathbb{Z}$ defined by $R(x, y) \Leftrightarrow y = x + 2012$. Find the smallest equivalence relation containing R.

Exercise 198. What is the smallest equivalence relation on $\mathbb{R}^2$ which contains the relation $(x, y) \sim (u, v) \Leftrightarrow x^2 + y^2 \leq u^2 + v^2$?

Exercise 199. Find the smallest equivalence relation on $\mathbb{R}$ containing the relation $x \sim y \Leftrightarrow x = -y$.

Exercise 200. A *doubly indexed family* of sets consists of sets I, J together with an $I \times J$-indexed family of sets $(A_{i,j})_{i \in I, j \in J}$. Show that for such families we have

$$\bigcup_{i \in I} \bigcup_{j \in J} A_{i,j} = \bigcup_{j \in J} \bigcup_{i \in I} A_{i,j}.$$

Prove a similar formula for the intersection.

Exercise 201. Consider the sets $I = J = \mathbb{N} - \{0\}$ and the doubly indexed family $A_{i,j} = \{x \in \mathbb{R} | \frac{1}{i} < x < j^2\}$. Find

$$\bigcup_{i \in I} \bigcap_{j \in J} A_{i,j}, \qquad \text{and} \qquad \bigcap_{i \in I} \bigcup_{j \in j} A_{i,j}.$$

LECTURE XVI

FIBRES

This lecture studies a particularly prominent way of obtaining families of sets, namely by means of functions.

XVI.1 DIRECT AND INVERSE IMAGE

Before we explain what fibres are, we first consider some more general phenomena about functions. Fix a function $f : A \to B$. As you know, f relates elements of A to elements of B. But can we also relate subsets of A to subsets of B using f?

Suppose $U \subseteq A$ is a subset of A. We can apply f to each of the elements of U; this will give a set of elements of B. (See Figure XVI.1 for an illustration.)

Definition XVI.1.1 (Direct Image)**.** Given $f : A \to B$ and a subset $U \subseteq A$, the *direct*

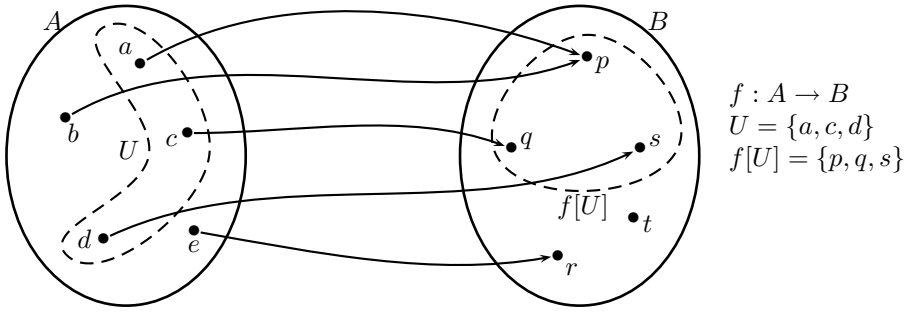


Figure XVI.1: Direct image of U under f

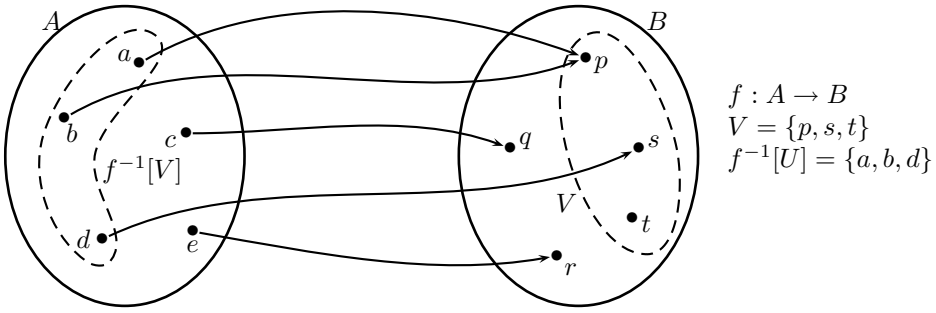


Figure XVI.2: Inverse image of V under f

image of U under f is

$$f[U] =_{def} \{f(x) | x \in U\} \subseteq B.$$

For example, when $f : \mathbb{R} \to \mathbb{R}$ is $f(x) = x^2 + 1$ and U is the closed interval $[-2, 1]$, we have $f[U] = [1, 5]$.

We can also go in the other direction: given a subset V of B, we may consider all elements which are mapped by f to elements in V. This is illustrated in Figure XVI.2

Definition XVI.1.2 (Inverse Image)**.** Given $f : A \to B$ and $V \subseteq B$, the *inverse image* of V under f is

$$f^{-1}[V] = \{x \in A | f(x) \in V\}.$$

For example, when $f : \mathbb{R} \to \mathbb{R}$ is $f(x) = x^2 + 1$ and $V = \{0, 1\}$, we have $f^{-1}[V] = \{0\}$ (because only $f(0) = 1$ and for no $x \in \mathbb{R}$ we have $f(x) = 0$). When W is the closed interval $[-2, 3]$, we have $f^{-1}[W] = [-\sqrt{2}, \sqrt{2}]$.

Warning: The notation $f^{-1}[V]$ is *not* meant to imply or even suggest that f has an inverse! The constructions of direct and inverse image make sense for any function, not just for bijections.

XVI.2 Fibres of a function

The concept of inverse image of a subset under a function has an important special case:

Definition XVI.2.1 (Fibre)**.** Let $f : A \to B$ be a function and $b \in B$. Then the *fibre* of f over b is

$$f^{-1}(b) = f^{-1}[\{b\}] = \{x \in A | f(x) = b\}.$$

This is illustrated in Figure XVI.3. Note that the fibre of an element may be empty. Note also that two elements $x, x' \in A$ are in the same fibre precisely when $f(x) = f(x')$. That means that "being in the same fibre of f" is an equivalence relation on A.[1]

[1] In fact, we have already seen that any function $f : A \to B$ gives an equivalence relation on A defined by $x \sim x' \Leftrightarrow f(x) = f(x')$.

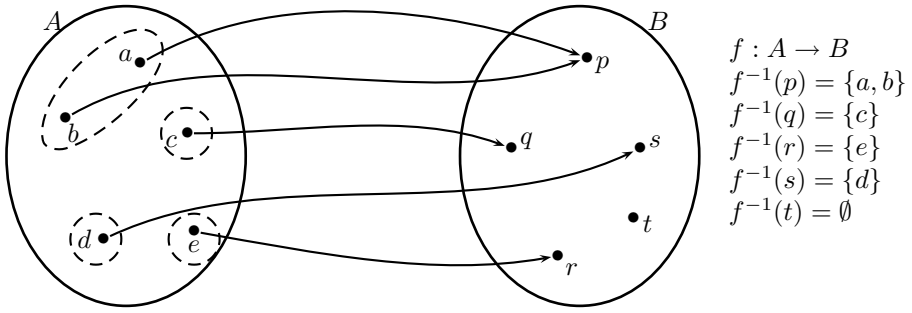


Figure XVI.3: The fibres of f.

The following exercise gives some insight into how properties of f can be translated into statements about the fibres of f.

Exercise 202. Let $f : A \to B$ be a function. Then

1. f is injective if and only if every fibre has at most one element.
2. f is surjective if and only if every fibre has at least one element.
3. f is bijective if and only if every fibre has exactly one element.

We can organize things in a slightly different manner. For each element $y \in B$, we have a set $f^{-1}(b)$. Therefore, we have a family of sets $(f^{-1}(y))_{y \in B}$. The main observation is now the following:

Proposition XVI.2.2. *When $f : A \to B$ is surjective, the family $(f^{-1}(y))_{y \in B}$ is a partitioning of A.*

Proof. First, by the preceding exercise, surjective functions have non-empty fibres. Second, given two fibres $f^{-1}(y)$ and $f^{-1}(y')$, if $z \in f^{-1}(y)$ and $z \in f^{-1}(y')$ then $f(z) = y$ and $f(z) = y'$, whence $y = y'$. Therefore the fibres are pairwise disjoint. Third, given $x \in A$ we have $x \in f^{-1}(f(x))$, so the fibres cover A. □

XVI.3 Families as Fibres

In the previous section we started with a function and showed how this information could be regarded as a family of sets, namely the fibres of the function. In this section we go the opposite route: starting with a family of sets, we wish to show that a family of sets can be regarded as the fibres of a function.

To start, let $(A_i)_{i \in I}$ be a family of sets. We would like to repackage this data in the form of a function $f : A \to B$ for suitable sets A, B. The general picture of a family of sets (Figure XV.1) suggests that $B = I$ might be a good choice. What could A be? We want to collect all the sets A_i into one big set A. Could it be $\bigcup_{i \in I}(A_i)$? The problem is

that the sets A_i may overlap, or that some of them may be empty. In that case forming the union loses information.

The trick is to take the *disjoint union* of the A_i instead. This is done (just as in the special case of a disjoint union of two sets $X + Y$) by attaching tags to the elements to remember which set they came from. This leads to:

Definition XVI.3.1 (Disjoint union of a family)**.** Let $(A_i)_{i\in I}$ be a family of sets. Then the *disjoint union* of this family is

$$\coprod_{i\in I} A_i =_{def} \{(x,i)|x \in A_i\}.$$

Now when we get an element (x, i) from this disjoint union we can tell immediately that it must have come from the set A_i.

Example XVI.3.2. (Cf. Figure XV.1.) The disjoint union of the family with $A_1 = \{a, b\}, A_2 = \{b, c, d\}, A_3 = \{b, c\}$ is the set

$$\coprod_{i\in\{1,2,3\}} A_i = \{(a,1),(b,1),(b,2),(c,2),(d,2),(b,3),(c,3)\}.$$

We now have our sets $A = \coprod_{i\in I}$ and $B = I$, so all we need is a function from $A \to I$. But this is now obvious: we set $f(x, i) = i$, i.e., we send an element to the label of the set it came from.

Now what are the fibres of this newly defined f? Over an element $i \in I$, we have $f(x, j) = i \Leftrightarrow i = j$, so $f^{-1}(i) = \{(x, i)|x \in A_i\}$. This is not literally the same set as A_i, because we have attached tags. However, there is an obvious bijection between the two:

$$\phi : A_i \to f^{-1}(i); \qquad \phi(x) = (x, i).$$

This shows:

Proposition XVI.3.3. *Every family $(A_i)_{i\in I}$ gives rise to a function f for which $f^{-1}(i) \cong A_i$ for each i.*

I leave it as an exercise for you to think about how this construction relates to the one in the previous section.

XVI.4 Summary

To each function $f : A \to B$ we associate two operations:

- *Direct image*: this sends a subset $U \subseteq A$ to $f[U] = \{f(x)|x \in A\}$
- *Inverse image*: this sends $V \subseteq B$ to $f^{-1}[V] = \{x \in A|f(x) \in V\}$.

As a special case, when $V = \{y\}$, we get the *fibre* of f over y, i.e.,

- $f^{-1}(y) = \{x \in A | f(x) = y\}$.

The important facts are:

- The fibres $(f^{-1}(y))_{y \in B}$ form a family of sets indexed by B.
- When f is a surjective function, this family is a partitioning of A (if not, the non-empty fibres form a partitioning).
- Given any family $(A_i)_{i \in I}$, there is a function $f : \coprod_{i \in I} A_i \to I$ with $f^{-1}(i) \cong A_i$.

The last statement says that *up to isomorphism*, each family of sets is the family of fibres of some function.

XVI.5 Exercises

Exercise 203. Let $A = \{0, 1, 2, 3, 4\}, B = \{5, 6, 7, 8\}$ and let f be defined by $f(0) = 7, f(1) = f(4) = 6, f(2) = f(3) = 8$. Find the direct image under f of the sets $U_0 = \emptyset, U_1 = \{0\}, U_2 = \{0, 1\}, U_3 = \{0, 1, 2\}, U_4 = \{1, 4\}, U_5 = \{0, 1, 2, 3, 4\}$. Also find the inverse image under f of the sets $V_0 = \emptyset, V_1 = \{5\}, V_2 = \{5, 6\}, V_3 = \{6, 7\}, V_4 = \{7, 8\}, V_5 = \{6, 7, 8\}$.

Exercise 204. Consider the function $f(x) = x^2$ from the reals to the reals. Describe the fibres of this function. Also work out the inverse image of the set $(-2, 1]$. Idem for $[-1, 2)$.

Exercise 205. For a function $f : A \to B$ show that $U \subseteq U'$ implies $f[U] \subseteq f[U']$.

Exercise 206. For a function $f : A \to B$ show that $V \subseteq V'$ implies $f^{-1}[V] \subseteq f^{-1}[V']$.

Exercise 207. Give an example of a function $f : \mathbb{R} \to \mathbb{R}$ for which all fibres are infinite.

Exercise 208. Consider the function $f : \mathbb{Z} \to \mathbb{Z}$ defined by $f(x) = x + 2$. Find all subsets $U\mathbb{Z}$ for which $f[U] = U$.

Exercise 209. Suppose $f : A \to B$ is a constant function (in the sense that $f(x) = f(y)$ for all $x, y \in A$). Prove that f has at most one non-empty fibre. Is the converse true?

Exercise 210. Let $f : A \to B$ be a function. Prove or disprove: $f[U^c] = f[U]^c$.

Exercise 211. Let $f : A \to B$ be a function. Prove or disprove: $f[U \cap U'] = f[U] \cap f[U']$.

Exercise 212. Let $f : A \to B$ be a function. Prove or disprove: $f[U \cup U'] = f[U] \cup f[U']$.

Exercise 213. Let $f : A \to B$ be a function. Prove or disprove: $f^{-1}[V^c] = f[V]^c$.

Exercise 214. Let $f : A \to B$ be a function. Prove or disprove: $f^{-1}[V \cap V'] = f^{-1}[V] \cap f^{-1}[V']$.

Exercise 215. Let $f : A \to B$ be a function. Prove or disprove: $f^{-1}[V \cup V'] = f^{-1}[V] \cup f^{-1}[V']$.

Exercise 216. Consider the function $f : X \to Y$ with $X = \{a, b, c\}, Y = \{p, q\}$ and $f(a) = f(b) = p, f(c) = q$. How many fibres are there? What are they? What is the associated partition of X?

Exercise 217. Describe the fibres of the identity function on a set A.

Exercise 218. Recall that for sets A, B there is a *projection function* $\pi_A : A \times B \to A$ defined by $\pi_A(a, b) = a$. Prove that for any element $a \in A$, we have $\pi_A^{-1}(a) \cong B$, i.e., that all fibres of π_A are isomorphic to B.

Exercise 219. Consider the set $I = \{1, 2, 3\}$ and the sets $A_1 = \{a, b\}, A_2 = \{b, c, d\}, A_3 = \{c\}$. Work out what the set $A = \{(a, i) | a \in A_i\}$ is in this case, and what the associated function $f : A \to I$ is. How many fibres does this function have? What are the fibres?

Exercise 220. Consider the $\mathbb{Z}$-indexed set $\{\{n - 1, n, n + 1\} | n \in \mathbb{Z}\}$. (Thus, each member has three elements.) Work out what the corresponding function is. What are the fibres of this function?

Exercise 221. Find a function $f : \mathbb{N} \to \mathbb{N}$ such that for each $n \in \mathbb{N}$, there exists a fibre of f with exactly n elements.

Exercise 222. Consider the function $\pi : \mathbb{Z} \to \mathbb{Z}/3$, the quotient by the equivalence relation $x \sim_3 y \Leftrightarrow x - y$ is divisible by 3. How many fibres does this function have? What are the fibres?

Exercise 223. Let $f : A \to B$ and $g : B \to C$ be functions. Show that for any $c \in C$, we have $(g \circ f)^{-1}(c) = \bigcup_{g(x)=c} f^{-1}(x)$.

LECTURE XVII

The Axiom of Choice

In many situations in mathematics we know that a certain set is non-empty, but at the same time it may be difficult to actually get our hands on a concrete element of that set. For example, consider the set A of solutions to the polynomial equation $x^5 - 4x^4 - 2x^3 + x^2 - 8$. We know that this set is non-empty, but finding an element is rather tricky.

In some cases, there are obvious solutions to this "problem". For example, suppose $A = \mathbb{N}$. Surely you can pick a natural number. There is in fact a *canonical* choice: simply pick 0, the first number. In other cases you might need be a bit more creative. If A is the set of irrational numbers between 5 and 6 there is no obvious choice. But you could, for example, pick the number $\pi + 2$. In general, however, we may lack a concrete description of the elements of the set A, making it difficult to come up with a choice.

Still, in many cases, you'll see phrases such as: "The set A is non-empty; let x be an element of A ...". Typically, the text will then prove some statement which depends on the existence of such x, but without having exhibited a concrete instance. What's worse, in many cases the proof will use not just one, but infinitely many such choices. This of course raises the question whether we can safely assume that such choices can always be made, or whether this casts doubt on these kinds of proof.

And this is exactly what the Axiom of Choice is about: can we always choose elements from non-empty sets, regardless of what we know about these sets or how they are presented to us?

XVII.1 Choice functions

In fact, we may formulate a more general problem. Given a set A, is it possible, for each non-empty subset U of A, to pick an element of U? A solution to this problem is

called a *choice function* for the set A. Technically, it is a function

$$s : \mathcal{P}_+(A) \to A$$

subject to the condition that $s(U) \in U$ for all U. Here, $\mathcal{P}_+(A)$ denotes the set of non-empty subsets of A. The condition merely ensures that the element we pick for U is actually an element of U.

To illustrate the concept of a choice function, consider the set $A = \{a, b, c\}$. Then

$$\mathcal{P}_+(A) = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$

To give a choice function for A, we need to choose, for each of the non-empty subsets $U \in \mathcal{P}_+(A)$, an element of that subset. Now for the subset $\{a\}$, there is not much to choose, since there is only one element to choose from. So we have to set $s(\{a\}) = a$. Similarly, we need $s(\{b\}) = b, s(\{c\}) = c$. Next, what could $s(\{a, b\})$ be? Here we need to choose either a or b. Similarly, $s(\{b, c\})$ can be equal to either b or to c. And so on. Thus, an example of a choice function s of A would be

$$\begin{array}{lll} s(\{a\}) = a, & s(\{b\}) = b, & s(\{c\}) = c, \\ s(\{a, b\}) = a, & s(\{a, c\}) = c, & s(\{b, c\}) = c, \\ s(\{a, b, c\}) = c. & & \end{array}$$

As an exercise you may want to try and find other choice funcions for the set A (there are 24 of them in total).

Exercise 224. Try to find choice functions for the sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$. (Hint: don't waste too much time on the last one!)

We can now state a version of the Axiom of Choice:

Axiom of Choice (first formulation):
Every non-empty set has at least one choice function.

The impact of this axiom seems a bit underwhelming when considering small sets, because in that case we can always "pick by hand". However, for sets such as the real numbers the Axiom tells us that we can achieve something which is rather amazing: pick an element from each non-empty set of real numbers, no matter how complicated that set may be.

Before we turn to different aspects and formulations of this powerful principle, we point out that of all the axioms of formal set theory, the Axiom of Choice is by far the most controversial one. While the other axioms seem to capture aspects of naive set theory which are intuitively plausible, the Axiom of Choice postulates the existence of functions for which nobody has any intuition. Indeed, when David Hilbert used a

version of the axiom in the proof of his famous Basis Theorem (1888), his work was criticised by a colleague as belonging to theology rather than mathematics.[1]

So, should we accept the Axiom of Choice? This is largely a philosophical problem, but one thing is clear: many of the theorems in mathematics we cherish rely on it in an essential manner. For example, you need it to prove that every vector space has a basis, or that every field has an algebraic closure.[2]

On the other hand, the axiom allows us to prove several counterintuitive results as well, such as the famous Banach-Tarksi Paradox, which states that you can take a sphere, break it up into pieces, and then rearrange the pieces so that you get two spheres of the same size as the original one.

While most mathematicians are content to accept the Axiom of Choice as part of the foundations of set theory, several opt to omit it, or adopt only a weakened version, such as the Axiom of Countable Choice, which only postulates the existence of choice functions for countable sets.

XVII.2 Choice and Families of Sets

As defined in the previous section, a choice function has the form $\mathcal{P}_+(A) \to A$; it picks, for each non-empty $U \subseteq A$, an element $a \in U$. Now $\mathcal{P}_+(A)$ is in particular a family of sets (recall that every set of sets may be regarded as a family of sets). This leads to a more general formulation of choice functions for arbitrary families of sets.

Consider a family of sets $(A_i)_{i\in I}$, such that each A_i is non-empty (otherwise we can't choose anything from it). A choice function for this family is, intuitively, a function s which gives us, for each $i \in I$, an element of A_i. In symbols: $s(i) \in A_i$. The only non-obvious thing is what the codomain of f should be: it should contain elements of each of the A_i. It suffices to consider the set $\bigcup_{i\in I} A_i$, and consider $s : I \to \bigcup_{i\in I} A_i$. That way, it is possible for $s(i)$ to be an element of A_i, for every i. Thus:

Definition XVII.2.1. A *choice function* for a family of sets $(A_i)_{i\in I}$ is a function

$$s : I \to \bigcup_{i\in I} A_i$$

such that $s(i) \in A_i$ for all $i \in I$.

And we have a corresponding version of the Axiom of Choice:

[1] However, the same colleague later admitted that while the axiom required a certain leap of faith, there were undeniable advantages to importing a bit of religious content into mathematics.

[2] Books have been written about the multitude of mathematical results depending on the Axiom of Choice (many of them actually being equivalent to it). See for example the two-part book by Rubin & Rubin titled "Equivalents of the Axiom of Choice ".

Axiom of Choice (second formulation):
Every family of non-empty sets has at least one choice function.

As an example, consider again the family of sets given by $I = \{1, 2, 3\}$ and $A_1 = \{a, b\}, A_2 = \{b, c, d\}, A_3 = \{c\}$. Then we have $\bigcup A_i = \{a, b, c, d\}$. A choice function for this family should thus be a function $s : \{1, 2, 3\} \to \{a, b, c, d\}$ such that $s(1) \in \{a, b\}, s(2) \in \{b, c, d\}$ and $s(3) \in \{c\}$. Note that this forces $s(3) = c$, but that there are still various options for $s(1)$ and $s(2)$.

Exercise 225. Construct all possible choice functions for the example above.

XVII.3 Sections

We now prepare for our third formulation of AC. Recall that whenever we have an I-indexed family of sets A_i, there exists a function $f : A \to I$ whose fibres are exactly the sets A_i. The idea is then, to formulate the Axiom of Choice in terms of the function f instead of in terms of the family $(A_i)_{i \in I}$.

Definition XVII.3.1. Let $f : A \to I$ be a function. A *section* of f is a function $s : I \to A$ such that $fs = 1_I$.

Clearly not all functions have sections. For example, if we take $A = \{0\}$ and $I = \{0, 1\}$ and $f(0) = 0$, then there exists a function $s : I \to A$ (sending both $0, 1$ to 0) but this gives $fs(1) = 0$, so it fails the requirement $fs = 1_I$. What goes wrong in this example is that for f to have a section, it is necessary for f to be surjective.

Exercise 226. Prove that if f has a section s, then f is surjective and s is injective.

Thus it only makes sense to ask for sections of surjective functions.

Axiom of Choice (third formulation):
Every surjective function has a section.

As an illustration of why sections are related to choice, consider the surjection $f : \{a, b, c, d\} \to \{1, 2\}$ defined by $f(a) = f(d) = 1, f(b) = f(c) = 2$. To specify a section s of f, we need to give $s(1)$ and $s(2)$. But it is required that $fs(1) = 1, fs(2) = 2$. This means that for $s(1) \in \{a, d\}$ and $s(2) \in \{b, c\}$. This gives four possible sections. Note that a choice of section amounts exactly to a choice of an element in each fibre of f, i.e. a choice function for the family of fibres of f.

Exercise 227. Find a couple of sections of the canonical surjection $\pi : \mathbb{Z} \to \mathbb{Z}/3$.

Recall that given sets A, B we can define their cartesian product $A \times B = \{(a, b)|a \in A, b \in B\}$. We now show how the concept of a section can be used to define the product of infinite families of sets as well. We will make use of the fact that any family $(A_i)_{i\in I}$ of sets gives rise to a projection function $f : \coprod_{i\in I} A_i \to I$.

Definition XVII.3.2. Let $(A_i)_{i\in I}$ be a family of sets. Then the *cartesian product* of $(A_i)_{i\in I}$ is the set

$$\prod_{i\in I} A_i = \{s : I \to \coprod_{i\in I} A_i | fs = 1_I\}.$$

Thus the product of the sets A_i is the set of sections of the projection $f : \coprod_{i\in I} \to I$. Concretely, such a section s selects an element from A_i for each $i \in I$. As such, giving a section of f is the same thing as giving a choice function for the family (A_i). This gives rise to yet another formulation of the Axiom of Choice:

Axiom of Choice (fourth formulation):
If $(A_i)_{i\in I}$ is a family of non-empty sets, then the product $\prod_{i\in I} A_i$ is non-empty.

XVII.4 All formulations are equivalent

We will now prove that all four formulations of the Axiom of Choice are equivalent.

First, let us look at the first formulation, which says that every set has a choice function. We want to show that it implies the second formulation, which says that every family of sets has a choice function. Thus we assume that the first formulation holds, consider a family of sets $(A_i)_{i\in I}$, and show that the latter has a choice function. Such a function should be of the form

$$s : I \to \bigcup_{i\in I} A_i$$

and satisfy $s(i) \in A_i$ for all $i \in I$. Now to use the assumption that the first formulation holds, we need to find a set A to which we can apply AC. But note that all sets A_i (the ones we're supposed to pick elements from) are actually subsets of the big union $\bigcup_{i\in I} A_i$. This suggests setting $A = \bigcup_{i\in I} A_i$. Then we have $A_i \in \mathcal{P}_+(A)$. Let's see how this works out: we have, by assumption, a choice function

$$v : \mathcal{P}_+(A) \to A$$

satisfying $v(U) \in U$ for all $U \subseteq A$. In particular, this gives that $v(A_i) \in A_i$. This is nice, but we want a function starting at I, not at $\mathcal{P}_+(A)$. If we could find a function

from $I \to \mathcal{P}_+(A)$ then we could compose the two, giving the function $I \to A$ we want. However, there is an obvious candidate: take $r(i) = A_i$. Then we get

$$vr(i) = v(A_i) \in A_i$$

so that the composite vr is a choice function for the family (A_i).

Next, let us see how the second formulation relates to the third. So assume that every family of non-empty sets has a choice function, and suppose that we are given a surjection $f : X \to I$. We need to find a section of f, i.e. a function $s : I \to X$ with $fs = 1_I$. To use the assumption, we need to come up with a family of sets. But we know that functions with codomain I can be regarded, via their fibres, as I-indexed families of sets. So consider the I-indexed family $\{f^{-1}(i) | i \in I\}$. Each member is non-empty, because f was assumed to be surjective. Thus there exists a choice function

$$v : I \to \bigcup_{i \in I} f^{-1}(i)$$

satisfying $v(i) \in f^{-1}(i)$ for all $i \in I$. This function is the section we're looking for: it does pick, for each $i \in I$, an element of the fibre $f^{-1}(i)$ (which is what a section is supposed to do as well). The only problem is that we need a function $s : I \to X$ and that we have a function $v : I \to \bigcup_{i \in I} f^{-1}(I)$. But we have shown earlier that $X = \bigcup_{i \in I} f^{-1}(i)$, so we're done.

We proceed by showing that the third formulation implies the first. Thus we assume that every surjection has a section, and we want to show that each set has a choice function. So consider a set A. We would like to define a surjective function, so that a section of that function makes the choice we need, i.e., the choice of an element from each subset of A. Thus we need to cook up a surjection, somehow using A and its subsets. Here's one which works: let

$$X = \{(a, U) | a \in U \subseteq A\}; \qquad f : X \to \mathcal{P}_+(A); \qquad f(a, U) = U$$

By assumption, this surjection has a section $s : \mathcal{P}_+(A) \to X$. By definition of section, it satisfies $fs(U) = U$, meaning that $s(U)$ has to be of the form (a, U) with $a \in U$. Thus effectively, s picks an element $a \in U$ for each U. This is what we want, except for that we really need a function $\mathcal{P}_+(A) \to A$, and not one $\mathcal{P}_+(A) \to X$. However, we can again fix that by composing with a suitable function $X \to A$, in this case $g(a, U) = a$. Then we have that $gs(U) \in U$, as needed.

Finally, it is clear that the third and fourth formulations are equivalent: given a surjection $f : A \to I$ we may regard it as an I-indexed family of sets; hence an element of the product of this family is simply a section of f.

Thus we have proved:

Theorem XVII.4.1. *All four formulations of the Axiom of Choice are equivalent.*

XVII.5 Summary

In this lecture we studied the *Axiom of Choice*, which states, informally, that whenever we have a collection of non-empty sets, then we can choose an element from each

of these sets. The following notions are used in the various formulations:

- A *choice function* for a set A is a function $c : \mathcal{P}_+(A) \to A$ for which $c(U) \in U$ for all $U \subseteq A$.
- A *choice function* for a family $(A_i)_{i\in I}$ of sets is a function $c : I \to \bigcup_{i\in I} A_i$ for which $c(i) \in A_i$ for all $i \in I$.
- A *section* of a function $f : A \to B$ is a function $s : B \to A$ for which $fs = 1_B$.
- The *product* of a family of sets $(A_i)_{i\in I}$ is the set of sections of the projection $\coprod_{i\in I} A_i \to I$.

The four formulations of the Axiom of Choice are then:

1. Every set has a choice function.
2. Every family of sets has a choice function.
3. Every surjection has a section.
4. The product of non-empty sets is non-empty.

The main result is that all of these formulations are equivalent.

XVII.6 Exercises

Exercise 228. Find all choice functions for the set $\{0, 1\}$.

Exercise 229. Try to find choice functions for the following sets:

(a) $\emptyset$

(b) $\{\emptyset\}$

(c) $\{\mathbb{N}\}$

(d) $\mathbb{N}$

(e) $\{x \in \mathbb{Z} | x \text{ is prime } \}$

(f) $\{x \in \mathbb{Q} | 0 < x < 1\}$

Exercise 230. Suppose a set A has n elements. How many choice functions does A have? (Note: this requires a bit of combinatorics.)

Exercise 231. Consider the function $f : \{a, b, c, d, e\} \to \{p, q, r\}$ given by $f(a) = f(c) = q, f(b) = f(d) = r, f(e) = p$. Find all possible sections of f.

Exercise 232. Let $f : A \to B$ be a bijective function. How many possible sections can f have?

Exercise 233. Consider the family of sets indexed by $I = \{1, 2, 3\}$ given by $A_1 = \{a, b, c\}, A_2 = \{b, c, d\}, A_3 = \{c, e\}$. Find three different choice functions for this family. How many choice functions are there in total?

Exercise 234. Construct choice functions for the $\mathbb{Z}$-indexed set $\{\{n-1, n, n+1\} | n \in \mathbb{Z}\}$.

Exercise 235. Consider a set I and the I-indexed family $A_i = \{i\}$. Find all possible choice functions for this family.

Exercise 236. Consider a set I and the I-indexed family $A_i = I$. Find all possible choice functions for this family.

Exercise 237. Suppose that $R \subseteq A \times B$ is a total (but not necessarily single-valued) relation. Show that R contains a function. (Hint: use AC.)

Exercise 238. Consider the relation $<$ on $\mathbb{R}$. Find a function contained in this relation. Did you need AC?

Exercise 239. Consider the function $f : \mathbb{R} \to [-1, 1]$ given by $f(x) = \sin(x)$. Find at least two different sections of f.

Exercise 240. Suppose that $f : A \to B$ and $g : B \to C$ are surjections, and that s is a section of f and t is a section of g. Show that st is a section of gf.

Exercise 241. Consider the family of sets $(A_i)_{i \in I}$ where $I = \{1, 2, 3, 4\}$,

$$A_1 = \{a, b, c\}, A_2 = \{p, q\}, A_3 = \{c\}, A_4 = \{p, a\}.$$

Describe explicitly the domain of the corresponding surjective function $e : X \to I$, as well as the surjective function itself. Describe also the function $X \to \bigcup_{i \in I} A_i$. Finally, create an explicit choice function for $(A_i)_{i \in I}$, and explain how this choice function corresponds to a section of e.

Exercise 242. Consider the equivalence relation on $\mathbb{R}$ defined by $x \sim y \Leftrightarrow x - y \in \mathbb{Z}$. Find a couple of examples of sections of the quotient map $\pi : \mathbb{R} \to \mathbb{R}/\sim$.

Exercise 243. Consider the family of sets $\{(n - 1, n + 1) \subseteq \mathbb{R} | n \in \mathbb{Z}\}$. Find a couple of examples of choice functions for this family. Construct the associated surjection and show how your choice functions give rise to sections of that surjection.

Exercise 244. Let $X = \{U \subseteq \mathbb{R} | U \text{ is finite }\}$. Can you find a choice function for X?

LECTURE XVIII

Cardinality

What do the sets $A = \{0, 1, 2\}, B = \{\text{John, Mary, Kimberly}\}$ and $C = \{\mathbb{N}, \mathbb{Z}, \mathbb{R}\}$ in common? They all have three elements. This means that they are in bijective correspondence: $A \cong B \cong C$. Generally, two finite sets X, Y which have the same number of elements are in bijective correspondence. But what can we say about two infinite sets? What does "same number of elements" mean? Do $\mathbb{N}$ and $\mathbb{Z}$ have the same number of elements? And what about $\mathbb{Q}$ and $\mathbb{R}$?

This chapter answers these questions.

XVIII.1 Size

As indicated, two sets may be thought of as "having the same size" when they are in bijective correspondence. This has its own terminology:

Definition XVIII.1.1 (Cardinality)**.** Two sets X and Y *have the same cardinality* when $X \cong Y$. In this case X and Y are also said to be *equipotent*.

Recall that the relation $\cong$ has the following properties:

- $A \cong A$ for every set A
- $A \cong B$ implies $B \cong A$
- $A \cong B$ and $B \cong C$ implies $A \cong C$

Therefore, the notion of "having the same cardinality" may be regarded as an equivalence relation on the collection[1] of all sets. The equivalence class of a set A is denoted

[1] As we know, there is no such thing as the sets of all sets!

$|A|$; the elements of this class are all the sets which are in bijective correspondence with A.

Remark XVIII.1.2 (which you may skip). A natural question is whether the class $|A|$ has some canonical representative in it. In other words, is there, for every possible "size" a set may have, a canonical set with that size? The answer is yes; but it requires a bit of theory. The canonical representatives are called *cardinal numbers*, and generalize the natural numbers (which are used to count the number of elements of finite sets). You can read more about this in any textbook on set theory, for example Halmos' book *Naive Set Theory*.

XVIII.2 Examples

We now establish a number of positive results, meaning that we will show for various sets that they have the same cardinality.

Examples XVIII.2.1.

1. Let $A = \mathbb{N}$ and $B = \mathbb{N} - \{0\}$. We claim that $|A| = |B|$. Define $\phi : A \to B$ by $\phi(n) = n + 1$. Then clearly ϕ is injective and surjective, hence bijective. This proves $\mathbb{N} \cong \mathbb{N} - \{0\}$.
2. Let $A = \mathbb{N}$ and $B = \mathbb{N} + \mathbb{N}$. (Thus B consists of two disjoint copies of the natural numbers.) Then $|A| = |B|$. Let $\psi : B \to A$ be $\psi(x, 0) = 2x$ and $\psi(y, 1) = 2y + 1$. Then ψ is bijective, so $\mathbb{N} \cong \mathbb{N} + \mathbb{N}$.
3. $|\mathbb{N}| = |\mathbb{Z}|$: define $f : \mathbb{Z} \to \mathbb{N}$ by $f(x) = 2x$ if $x \geq 0$ and $f(x) = -2x - 1$ if $x < 0$. Then f is bijective.

These examples show in particular that a set can be in bijective correspondence with a proper subset of itself. Clearly for finite sets this is impossible. Thus it may be taken as a *definition* of what it means for a set to be infinite:

Definition XVIII.2.2 (Infinity)**.** A set X is *infinite* if there exists a subset $Y \subseteq X$ with $Y \neq X$ but $|X| = |Y|$. A set is *finite* if it is not infinite.

In fact, it suffices to require that there is a proper subset Y of X and a surjective function $Y \to X$.

It follows from the definition that finite sets have the following property (which you can verify by hand in small examples):

Lemma XVIII.2.3. *Let X be a finite set. Then for a function $f : X \to X$ the following are equivalent:*

- *f is injective*
- *f is surjective*
- *f is bijective*

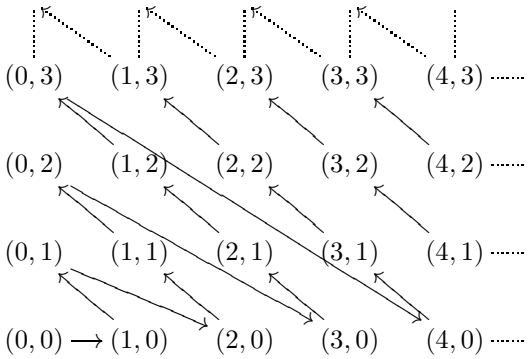


Figure XVIII.1: A coding of the plane

Proof. It suffices to show that injectivity implies surjectivity and vice versa. Suppose f is injective. Then we have $X \cong f[X]$ because f is injective and each $y \in f[X]$ is in the image of f. But $f[X]$ is a subset of X, and hence must be equal to X if X is finite. Therefore $f[X] = X$, meaning that f is surjective.

Conversely, if f is surjective then suppose $f(x) = f(x')$, but $x \neq x'$. Then $X - \{x'\}$ is a proper subset of X, and there exists a surjection $X - \{x'\} \to X$ (which is simply f restricted to this subset). But this can't be if X is finite. □

Sometimes we wish to express that one set is larger than another; the following definition makes that precise:

Definition XVIII.2.4. For sets X, Y, write $|X| \leq |Y|$ if there exists an injection from X to Y. Write $|X| < |Y|$ if $|X| \leq |Y|$ but $|X| \neq |Y|$.

For example, it trivially follows that if $X \subseteq Y$, then $|X| \leq |Y|$. Also, if $f : Y \to X$ is a surjection it follows that $|X| \leq |Y|$. (Choose a section!)

The following example is slightly more involved:

Example XVIII.2.5. The set $\mathbb{N}$ is equipotent to $\mathbb{N} \times \mathbb{N}$. (That is: $|\mathbb{N}| = |\mathbb{N} \times \mathbb{N}|$.)

To prove this, we need to give a bijective function from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$. The picture in Figure XVIII.1 illustrates the idea behind this bijection.

As you can see, we are systematically enumerating the antidiagonals $y + x = c$. It is clear that this sets up a bijection $p : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ (p for pairing).

There are also two "decoding" functions $p_0, p_1 : \mathbb{N} \to \mathbb{N}$, which have the following properties:

$$p_0(p(n, m)) = n, \qquad p_1(p(n, m)) = m, \qquad p(p_0(x), p_1(x)) = x.$$

Thus, the function $p^{-1}(x) = (p_0(x), p_1(x))$ is the inverse to p. The above equations are often called the *pairing equations*.

The function p actually satisfies

$$p(x, y) = \frac{1}{2}(x + y)(x + y + 1) + y.$$

In fact, it turns out that (assuming the Axiom of Choice) *every* infinite set X satisfies $|X| = |X \times X|$. In fact, this is one of the many statements equivalent to AC. For a proof of this statement we refer to standard textbooks on set theory.

A consequence of the previous is the following:

Example XVIII.2.6. We have $|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$.

The first equality has already been proved above. We clearly have $|\mathbb{Z}| \leq |\mathbb{Q}|$, so the only remaining question is whether $|\mathbb{Q}| \leq |\mathbb{Z}|$. To construct the witnessing injection, recall that an element q of $\mathbb{Q}$ may be represented as $\frac{a}{b}$, where $a \in \mathbb{Z}, b \in \mathbb{N}$ and where a, b are relatively prime. Call this the *minimal representation* of q, and note that it is unique. Now consider

$$f : \mathbb{Q} \to \mathbb{Z} \times \mathbb{N}; \qquad f(q) = (a, b) \text{ where } q = \frac{a}{b} \text{ is the min. repr. of } q.$$

This is functional because the minimal representation is unique; it is injective because distinct rational numbers have distinct representations. This concludes the proof.

One more important example: for any pair of real numbers a, b with $a < b$, we have $|\mathbb{R}| = |(a, b)|$. I'll show this for the interval $(-\pi/2, \pi/2)$ and leave it as an exercise to show that any two open bounded intervals (a, b) are of the same cardinality. But the function $\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$ is bijective (with inverse $\tan$). This proves the claim. In the exercises you will also show that all intervals $[a, b]$ have the same cardinality as $\mathbb{R}$.

We have introduced the notation $|A| \leq |B|$, but we haven't quite justified it. One important question is the following: if $|A| \leq |B|$ and $|B| \leq |A|$, does it follow that $|A| = |B|$? Unpacking the definition, this becomes: if there are injective functions $f : A \to B$ and $g : B \to A$, is there a bijection $A \cong B$?

Theorem XVIII.2.7 (Cantor-Bernstein)**.** *If $|A| \leq |B|$ and $|B| \leq |A|$, then $|A| = |B|$.*

We will not give a proof here; it is possible to derive the result from the Axiom of Choice, but the latter is not necessary. For an elegant "back-and-forth" proof the reader is referred to the supplementary readings.

XVIII.3 Cantor's Diagonal Argument

In the previous section we considered various sets with the same cardinality. In this section we introduce a technique for showing that two sets have different cardinality.

Definition XVIII.3.1. A set A is *countable* when $|A| \leq |\mathbb{N}|$. If A is not countable, we call A *uncountable*.

Note that according to this definition, finite sets are countable. Infinite sets which are countable are also called *countably infinite*. Intuitively, if a set is countable, then it is possible to "enumerate its elements": we may write $A = \{a_0, a_1, a_2, \ldots\}$. Technically,

this means that we choose an injection $s : A \to \mathbb{N}$, and then choose a function $e : \mathbb{N} \to A$ of which s is a section. The function e - which must be surjective! - is thought of as enumerating the elements of A (possibly with repetitions).

Our goal is to show: $\mathbb{R}$ is uncountable. Clearly it suffices to prove that there exists a subset of $\mathbb{R}$ which already is uncountable. We will do this for the open interval $(0, 1)$. The proof of this is due to Georg Cantor, the father of set theory; his technique is called a *diagonal argument*.

Theorem XVIII.3.2. *The set* $(0, 1)$ *is uncountable.*

Proof. Suppose, toward a contradiction, that $(0, 1)$ were countable. That means that we can find an enumeration $e : \mathbb{N} \to (0, 1)$. This function allows us to list the elements of $(0, 1)$ one by one. For example, the first couple of elements could look something like this:

$$\begin{array}{lcccccccccccc}
e(0): & \mathbf{3} & 4 & 1 & 0 & 0 & 0 & 2 & 9 & 9 & 2 & 5 & \cdots \\
e(1): & 2 & \mathbf{5} & 5 & 5 & 5 & 3 & 2 & 9 & 7 & 1 & 0 & \cdots \\
e(2): & 2 & 4 & \mathbf{5} & 3 & 5 & 0 & 0 & 4 & 4 & 4 & 7 & \cdots \\
e(3): & 5 & 2 & 0 & \mathbf{7} & 1 & 6 & 1 & 9 & 7 & 1 & 7 & \cdots \\
e(4): & 1 & 2 & 0 & 8 & \mathbf{3} & 0 & 3 & 3 & 8 & 8 & 5 & \cdots \\
e(5): & 2 & 5 & 4 & 2 & 3 & \mathbf{6} & 2 & 9 & 7 & 5 & 8 & \cdots \\
e(6): & 8 & 8 & 8 & 9 & 9 & 7 & \mathbf{0} & 8 & 1 & 0 & 1 & \cdots \\
e(7): & 3 & 3 & 2 & 1 & 1 & 0 & 7 & \mathbf{3} & 6 & 9 & 0 & \cdots \\
e(8): & 3 & 5 & 2 & 6 & 8 & 0 & 0 & 2 & \mathbf{3} & 3 & 0 & \cdots \\
e(9): & 8 & 5 & 1 & 7 & 4 & 0 & 0 & 6 & 6 & \mathbf{0} & 2 & \cdots \\
\vdots &&&&&&&&&&&&
\end{array}$$

This means that the first number in the list is the real number with decimal expansion starting with $0.34100029925\ldots$, and so on. The thing to keep in mind is that by hypothesis *every* element of $(0, 1)$ occurs somewhere in this list.

We will derive a contradiction by finding a real number in $(0, 1)$ which cannot possibly be in the list. The way to construct this number is as follows. Focus on the diagonal of the list, i.e., the i-th decimal place of the i-th element in the list. These are the bolded digits in diagram.

Now consider the real number which has the following decimal expansion:

$$c = 0.4668471441\ldots.$$

In words, take the entries on the diagonal and add one to each. (If an entry is 9, make it 0.) Claim: $c \neq e(i)$ for any $i \in \mathbb{N}$. Proof: suppose $c = e(i)$. The i-th digit of c is different from the i-th digit of $e(i)$ (because we made it so). Contradiction. □

The same technique (or variations thereon) can be used to establish various other results. For example:

Theorem XVIII.3.3. *The set* $\mathcal{P}(\mathbb{N})$ *is uncountable.*

Proof. Suppose that we have an enumeration $e : \mathbb{N} \to \mathcal{P}(\mathbb{N})$. A subset U of $\mathbb{N}$ may be represented via its characteristic function $\chi_U : \mathbb{N} \to \{0, 1\}$. We can view χ_U as

a sequence of 0s and 1s: $\chi_U(0), \chi_U(1), \chi_U(2), \chi_U(3)$, and so on. The first couple of elements in the enumeration then look something like this:

$$\begin{array}{lcccccccccccc}
e(0): & \mathbf{1} & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & \cdots \\
e(1): & 1 & \mathbf{0} & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & \cdots \\
e(2): & 1 & 1 & \mathbf{1} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots \\
e(3): & 0 & 1 & 0 & \mathbf{1} & 1 & 0 & 1 & 1 & 1 & 1 & 0 & \cdots \\
e(4): & 1 & 1 & 0 & 1 & \mathbf{1} & 0 & 1 & 1 & 0 & 0 & 1 & \cdots \\
e(5): & 0 & 0 & 0 & 1 & 1 & \mathbf{0} & 1 & 1 & 0 & 0 & 1 & \cdots \\
e(6): & 0 & 1 & 0 & 1 & 1 & 1 & \mathbf{0} & 1 & 1 & 0 & 1 & \cdots \\
e(7): & 0 & 1 & 0 & 1 & 1 & 0 & 1 & \mathbf{0} & 1 & 1 & 0 & \cdots \\
e(8): & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & \cdots \\
e(9): & 1 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & \mathbf{0} & 0 & \cdots \\
\vdots & & & & & & & & & & & &
\end{array}$$

Now define a function $\chi : \mathbb{N} \to \{0, 1\}$ by taking the values on the diagonal and switching 0s to 1s and vice versa:

$$\chi(i) = \begin{cases} 1 & \text{if the } i\text{-th element of } e(i) \text{ is } 0 \\ 0 & \text{else.} \end{cases}$$

Then χ is different from all the $e(i)$ and hence corresponds to a subset not in the enumeration. Contradiction. □

The following is very similar and is left as an exercise:

Theorem XVIII.3.4. *The set $\mathbb{N}^{\mathbb{N}}$ is uncountable.*

Actually, the previous results can be generalized:

Theorem XVIII.3.5. *For any set A with $|A| > 1$, we have $|A| < |\mathcal{P}(A)|$ and $|A| < |A^A|$.*

The proof is again a diagonalization: if $|A| = |A^A|$ then there is a surjection $e : A \to A^A$. Thus for each $a \in A$, $e(a)$ is a function from A to A. Write $e_{a,b}$ for the element $e(a)(b)$. Also, fix two distinct elements a_0, a_1 of A. Now define a function $f : A \to A$ by

$$f(a) = \begin{cases} a_0 & \text{if } e_{a,a} \neq a_0 \\ a_1 & \text{else.} \end{cases}$$

XVIII.4 Summary

Two sets A, B are said to have the same *cardinality* when they are in bijective correspondence. Notation: $|A| = |B|$.

We found that $|\mathbb{N}| = |\mathbb{N} + \mathbb{N}| = |\mathbb{N} \times \mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}|$. Any set of the same cardinality or less as $\mathbb{N}$ is called *countable*. Sets which are not countable are called *uncountable*. A set A is countable if and only if there exists an *enumeration* of A, that is, a surjective function $\mathbb{N} \to A$.

Cantor's *diagonal argument* is used to show that

- $|\mathbb{N}| < |(0,1)^{\mathbb{N}}|$ (and hence $|\mathbb{N}| < |\mathbb{R}|$)
- $|A| < |\mathcal{P}(A)|$ for any set A
- $|A| < |A^A|$ for any set A with more than one element.

Finally, the Cantor-Bernstein Theorem states that $|A| \leq |B|$ and $|B| \leq |A|$ implies $|A| = |B|$.

XVIII.5 Exercises

Exercise 245. Prove the following statements about cardinalities:

(a) If $|A| = |A'|$ and $|B| = |B'|$ then $|A \times B| = |A' \times B'|$.

(b) If $|A| = |A'|$ and $|B| = |B'|$ then $|A + B| = |A' + B'|$.

(c) If $|A| = |A'|$ and $|B| = |B'|$ then $|A^B| = |A'^{B'}|$.

(d) If $|A| = |A'|$ then $|\mathcal{P}(A)| = |\mathcal{P}(A')|$.

Exercise 246. Prove that all of the following sets have the same cardinality:

- $\mathbb{N}$
- $\mathbb{N} + 1$
- $\mathbb{N} + k$, where k is a set with k elements.
- $\mathbb{N} + \mathbb{N} + \mathbb{N}$
- $\mathbb{Z} + \mathbb{Z}$
- $\mathbb{Z} \times \mathbb{Z}$
- $\mathbb{Z}^k$
- $\mathbb{Z} + (\mathbb{Q} \times \mathbb{N})^k$.
- The set of prime numbers (as a subset of $\mathbb{Z}$)
- The set of even numbers (as a subset of $\mathbb{Z}$)

Exercise 247. Suppose $a < b$ and $c < d$ are real numbers. Show that $(a, b) \cong (c, d)$ and that $[a, b] \cong [c, d]$. Finally, show that $(a, b) \cong [a, b]$. (Hint: pick $c < a, d > b$ and use the Cantor-Schröder-Bernstein Theorem.)

Exercise 248. Assume A is infinite. Use $|A| = |A \times A|$ to show that $|A| = |A + A|$.

Exercise 249. Suppose that A is uncountable, and that $B \subseteq A$ is countable. Show that $A - B$ is uncountable as well.

Exercise 250. Show that the set of irrational numbers is uncountable.

Exercise 251. Consider the set L of affine functions $\mathbb{R} \to \mathbb{R}$. (An affine function is one of the form $f(x) = ax + b$ for $a, b \in \mathbb{R}$.) Prove that $|L| = |\mathbb{R}|$.

Exercise 252. Let A be the set of all possible strings (of arbitrary but finite length) from the alphabet $\{a, b, \ldots, z\}$. Is this a countable set?

Exercise 253. Prove that the set of finite subsets of $\mathbb{N}$ is countable. (Hint: try to find a systematic way of enumerating such sets.)

Exercise 254. A subset $A \subseteq \mathbb{N}$ is called *cofinite* if $\mathbb{N} - A$ is finite. Prove that the set of cofinite subsets of $\mathbb{N}$ is countable.

Exercise 255. Consider the set of infinite subsets of $\mathbb{N}$. Is this a countable set?

Exercise 256. Show that if there exists a surjection $A \to \mathbb{N}$ then A is infinite. You may use AC.

Exercise 257. Let $\mathbb{Q}_n[x]$ be the set of n-th degree polynomials (in one variable) with coefficients in $\mathbb{Q}$. Thus, an element looks like $a_n x^n + \cdots + a_1 x + a_0$ where the $a_i \in \mathbb{Q}$. Show that $\mathbb{Q}_n[x]$ is countable. Generalize to polynomials in several variables.

Exercise 258. Write $\mathbb{R}_n[x]$ for the set of n-th degree polynomials in one variable over $\mathbb{R}$ (see previous exercise). Prove that $|\mathbb{R}_n[x]| = |\mathbb{R}|$. You may use that $|\mathbb{R} \times \mathbb{R}| = |\mathbb{R}|$.

Exercise 259. Show that $|\mathbb{C}| = |\mathbb{R}|$.

Exercise 260. Consider the equivalence relation on $\mathbb{R}$ given by

$$x \sim y \Leftrightarrow x - y \in \mathbb{Z}$$

and the quotient set $\mathbb{R}/\sim$. Show that each equivalence class is countable. Also show that the quotient set is uncountable by setting up a bijection with a set which is known to be uncountable.

Exercise 261. Let $(A_i)_{i \in I}$ be a family of sets where I is countable and where each A_i is countable. Show that $\coprod_{i \in I} A_i$ and $\bigcup_{i \in I} A_i$ are also countable.

Exercise 262. A *fixpoint* of a function $f : A \to A$ is an element $a \in A$ for which $f(a) = a$. Say f is *fixpoint-free* when it has no fixpoints, i.e., $f(a) \neq a$ for all $a \in A$. Verify that in each of the diagonal arguments above we used a fixpoint-free function.

LECTURE XIX

Ordered Sets

In earlier lectures we studied relations with special properties. Most notably, we considered functions (and in turn various special classes of functions, such as bijections), and equivalence relations. We now embark on the study of a third important class of relations called *orderings*.

XIX.1 Definition and Examples

We've already met several standard examples of orderings: the "less than or equal"-relation on $\mathbb{N}$ (or on $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$), and the subset relation (on $\mathcal{P}(A)$ for some set A). The following definition extracts what is common to these examples.

Definition XIX.1.1 (Ordering)**.** Let A be a set and R a relation on A. We say that R is a *pre-ordering relation* when R is reflexive and transitive. When R is also anti-symmetric, we call it an *ordering relation*

Typically, we use symbols such as $\leq$ or $\preceq$ to denote orderings. When A is a set and $\leq$ is an ordering on A, we say that $(A, \leq)$ is a *partially ordered set* (often abbreviated to *poset*).

Example XIX.1.2. Let $A = \{a, b, c, d\}$, and consider the following relations:

- $R_1 = \{(a,a), (b,b), (c,c), (d,d), (a,c)\}$
- $R_2 = \{(a,a), (b,b), (c,c), (d,d), (a,c), (a,d)\}$
- $R_3 = \{(a,a), (b,b), (c,c), (d,d), (a,c), (c,d)\}$
- $R_4 = \{(a,a), (b,b), (c,c), (d,d), (a,b), (a,c), (a,d)\}$

- $R_5 = \{(a,a),(b,b),(c,c),(d,d),(a,b),(b,c),(c,d),(b,c),(b,d),(a,d)\}$
- $R_6 = \{(a,a),(b,b),(c,c),(d,d),(a,b),(b,a),(a,d),(b,d)\}$

The relation R_1 is an ordering, as is R_2. However, R_3 is not, because it is not transitive. The relation R_4 is an ordering, and so is R_5. Finally, R_6 is a pre-ordering but not an ordering (it is not anti-symmetric).

One of the points to take away from this elementary example is that one given set can have many different orderings on it. (Even if there is one "obvious" one, such as on the natural numbers!) Here are some more examples, this time of a slightly more abstract nature.

Examples XIX.1.3.

1. Let A be any set. Then the diagonal relation Δ_A is an ordering. It is usually called the *discrete* ordering.
2. The maximal relation on a set A is not an ordering when A has more than one element: in that case the maximal relation is not anti-symmetric. However, it is a pre-order relation.
3. When $(A, \leq)$ is a partial order, we can consider the opposite relation $\leq^{op}$. By definition, we have $x \leq^{op} y$ if and only if $y \leq x$. Then $\leq^{op}$ is also an ordering relation.

Sometimes we're more interested in the "strictly less than"-relation than in the "less than or equal"-relation. (And even if we're not, listing all the pairs (x,x) in a relation gets cumbersome.) This is handled as follows:

Definition XIX.1.4 (Strict Ordering)**.** A relation R on a set A is a *strict ordering* when it is anti-reflexive, transitive and anti-symmetric.

For example, the relation $R = \{(a,b),(a,c),(d,c)\}$ is a strict ordering on $\{a,b,c,d,e\}$. The following easy lemma shows that the concepts of strict ordering and general ordering are intertranslatable:

Lemma XIX.1.5. *There is a one-one correspondence between ordering relations on A and strict ordering relations on A.*

Proof. Suppose first that $<$ is a strict ordering on A. Then define $\leq$ to be the following relation on A:

$$a \leq b \Leftrightarrow a < b \text{ or } a = b.$$

(Equivalently, $\leq$ is the union of $<$ and the diagonal.) Then $\leq$ is easily seen to be an ordering relation. Conversely, if $\leq$ is an ordering, define $<$ by

$$a < b \Leftrightarrow a \leq b \text{ and } a \neq b.$$

Then $<$ is a strict ordering. The constructions of $<$ from $\leq$ and vice versa are mutually inverse. □

To conclude this section, we give a few more examples from mathematical practice:

Examples XIX.1.6.

1. Define a relation on $\mathbb{Z}$ by
$$x \preceq y \Leftrightarrow_{\text{def}} x \text{ divides } y.$$
This is an ordering.

2. Consider the set $\mathbb{N}^{\mathbb{N}}$ of functions from $\mathbb{N}$ to $\mathbb{N}$. Define an ordering on this set by
$$f \leq g \Leftrightarrow_{\text{def}} f(x) \leq g(x) \text{ for all } x \in \mathbb{N}.$$
This is called the *pointwise ordering* on $\mathbb{N}^{\mathbb{N}}$. Note that the $\leq$ on the left hand side is the one we're defining, while the $\leq$ on the right hand side is the usual ordering on $\mathbb{N}$. This is overloading of notation and can be confusing, but is common practice.

3. Generalizing the previous example: let X be any set, and let $(A, \leq)$ be an ordered set. Define an ordering on A^X by
$$f \leq g \Leftrightarrow_{\text{def}} f(x) \leq g(x) \text{ for all } x \in X.$$
Again, we're using the same symbol $\leq$ for two different orderings.

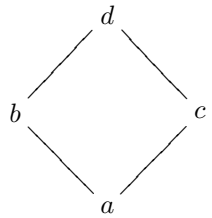
4. When A is any set, the powerset $\mathcal{P}(A)$ is ordered by inclusion. When $(A, \leq)$ is an ordered set, we can use the ordering on A to define another relation on $\mathcal{P}(A)$, namely
$$U \leq V \Leftrightarrow_{\text{def}} \forall x \in U \exists y \in V . x \leq y.$$
This defines a pre-ordering which is generally not anti-symmetric. (Try to find an example.)

XIX.2 Hasse Diagrams

Hasse diagrams are a tool for visualizing ordered sets. Of course, we already know how to draw pictures of relations, and in principle we could use that for ordered sets as well. However, the notion of Hasse diagram is easier and more intuitive when it comes to orderings.

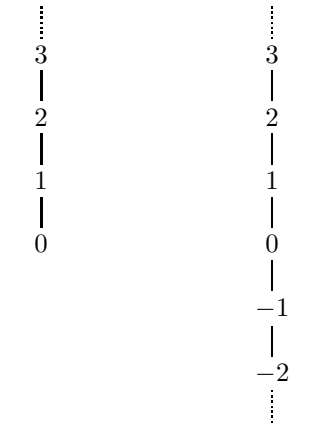
In a Hasse diagram, we draw the elements of the set as we normally do; however, when $a \leq b$, we draw a below b. The following example illustrates this:

Example XIX.2.1. Let $A = \{a, b, c, d\}$ and let R be the order relation specified by $\{(a, b), (a, c), (a, d), (b, d), (c, d)\}$. Then the corresponding Hasse diagram is



Note that we don't draw an edge from a to d, even though $a \leq d$. This isn't necessary, because we can see in the picture that d is above a. Also, we don't represent reflexivity of the relation. When two elements in a Hasse diagram are at the same height (such as b and c in the example) then that means that they are *incomparable*, in the sense that neither $b \leq c$ nor $c \leq b$.

For another example, here are (parts of) the Hasse diagrams of the usual orderings on $\mathbb{N}$ and on $\mathbb{Z}$.



$(\mathbb{N}, \leq)$ $(\mathbb{Z}, \leq)$

XIX.3 Constructions

In this section we discuss several methods for constructing new orderings from old ones.

First, we consider disjoint sums. Let $(A, \leq)$ and $(B, \leq)$ be orderings. We define an ordering on $A + B$ by

$$(x, i) \leq (y, j) \Leftrightarrow_{\text{def}} i = j \text{ and } x \leq y.$$

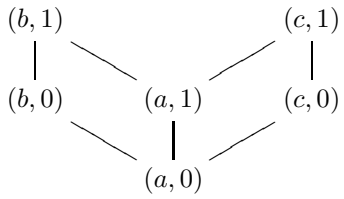
In terms of Hasse diagrams, this means that we simply place the diagrams for $(A, \leq)$ and for $(B, \leq)$ side by side: all elements of A are incomparable with all elements of B.

Next, we consider products of orderings. Suppose $(A, \leq)$ and $(B, \leq)$ are orderings. Define an ordering on $A \times B$ by

$$(a, b) \leq (a', b') \Leftrightarrow_{def} a \leq a' \text{ and } b \leq b'.$$

This is also called the *componentwise* ordering. For example, suppose we have $A = \{a, b, c\}$ with $a \leq b$ and $a \leq c$, and $B = \{0, 1\}$ with $0 \leq 1$. Then the following Hasse

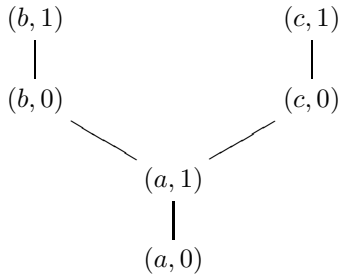
diagram depicts the product ordering on $A \times B$:



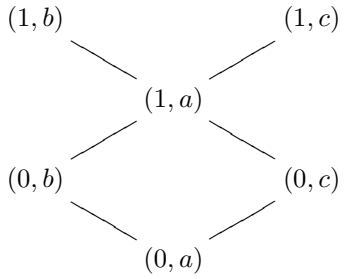
There is another possible way of ordering the product $A \times B$; this is called the *lexicographic ordering*. Define

$$(x, y) \leq_l (x', y') \Leftrightarrow_{\text{def}} \begin{cases} x < x' & \text{or} \\ x = x' \text{ and } y \leq y' \end{cases}$$

Taking the previous example, the lexicographic ordering on $A \times B$ is



We can also compute the lexicographic ordering on $B \times A$:



Note that these are quite different!

XIX.4 Summary

An *ordering* on a set is a reflexive, transitive and anti-symmetric relation. Orderings are often represented by *Hasse diagrams*, in which $x \leq y$ is expressed by drawing y higher than x.

It is important to keep in mind that a set may be ordered in different ways.

Orderings can be combined in various ways:

- Disjoint sum: given two orderings $(A, \leq$ and $(B, \leq)$ we can order $A + B$ via the relation $(x, i) \leq (y, j)$ iff $i = j$ and $x \leq y$.
- Product ordering: this is the ordering on $A \times B$ given by $(x, y) \leq (x', y')$ iff $x \leq x'$ and $y \leq y'$.
- Lexicographic ordering: this is another ordering on $A \times B$ given by $(x, x') \leq (y, y')$ iff $x < x'$ or $x = x'$ and $y \leq y'$.

XIX.5 Exercises

Exercise 263. List all possible orderings on the following sets and draw the corresponding Hasse diagrams:

(a) $\emptyset$

(b) 1 (a one-element set)

(c) $\{a, b\}$

(d) $\{a, b, c\}$

(e) $\{a, b, c, d\}$

Exercise 264. Draw the Hasse diagram of the subset ordering on $\mathcal{P}(A)$ where $A = \{1, 2, 3\}$.

Exercise 265. Give at least three different ordering relations on the set $\mathbb{Z}$.

Exercise 266. Draw the Hasse diagram of the divisibility relation on $\mathbb{N}$ and on $\mathbb{Z}$.

Exercise 267. Let PROP denote the set of all propositions. Is the relation $\vdash$ defined by

$$\phi \vdash \psi \Leftrightarrow (\phi \to \psi) \text{ is a tautology}$$

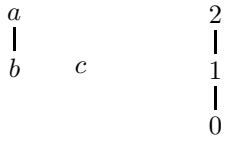
an ordering?

Exercise 268. Let $(A, \leq)$ be an ordering and let $U \subseteq A$ be a subset of A. Consider the relation $\leq \cap (U \times U)$ on U. Show that this is an ordering. (It is called the *restriction* of the ordering on A to U.)

Exercise 269. Let $A = \{a, b, c\}$ and $B = \{0, 1, 2\}$. Consider the following relations on A and on B:

Work out the product ordering on $A \times B$ and on $B \times A$. Also work out the lexicographic ordering on $A \times B$ and on $B \times A$.

Exercise 270. Same question, but now for



Exercise 271. Same question, but now for

Exercise 272. Let $A = \{1, 2, \ldots, 24\}$, and consider the divisibility relation on this set. Draw the Hasse diagram.

Exercise 273. Draw (part of) the Hasse diagram for the lexicographic ordering on $\mathbb{N} \times \mathbb{N}$. Same question for $\mathbb{N} \times \mathbb{Z}$.

Exercise 274. Let $f : A \to B$ be a function and suppose that $\leq$ is an ordering on B. Define a relation on A by

$$x \leq y \Leftrightarrow_{\text{def}} f(x) \leq f(y).$$

Show that this is a pre-ordering.

Exercise 275. Let A be a set, and consider $\mathcal{P}(A)/\cong$, the quotient of $\mathcal{P}(A)$ by the relation $\cong$. Show that cardinal inequality $|U| \leq |V|$ is an ordering on this set.

Exercise 276. Let A be a set and consider $\mathsf{Par}(A, B)$, the set of partial functions from A to B. (These are just single-valued relations.) Write $dom(f) = \{x \in A | \exists y. f(x) = y\}$, and define

$$f \leq g \Leftrightarrow \forall x \in dom(f). f(x) = g(x).$$

Show that this is an ordering on the set of partial functions.

Exercise 277. Let R be a pre-ordering on A; define a new relation $\sim$ on A by $x \sim y \Leftrightarrow xRy \wedge yRx$. Show that this is an equivalence relation and that the quotient $A/\sim$ carries a partial ordering induced by R.

LECTURE XX

FURTHER THEORY

In this lecture we look at some important special properties ordered sets may have. In particular, we study the classes of linearly ordered sets and of complete orderings. We also consider chains in orderings.

XX.1 LINEARITY

There is a distinct difference between the poset $\mathcal{P}(A)$ ordered by inclusion and the poset $\mathbb{Z}$, ordered as usual: in the first, we can find elements U, V such that U and V are *incomparable*, in the sense that neither is included in the other. In $\mathbb{Z}$ (or in $\mathbb{N}, \mathbb{Q}$ or $\mathbb{R}$ for that matter) this is not possible: any two elements are comparable. This is captured in the following definition:

Definition XX.1.1 (Linearity)**.** Let $(A, \leq)$ be a poset. Then $(A, \leq)$ is called *linear* when for all $x, y \in A$ we have $x \leq y$ or $y \leq x$.

Linearly ordered sets are also called *linear orders*; some texts use the terminology *total orders*, but note that this is in conflict with the use of the term "total" for relations.

I have already told you the standard examples of linear orderings: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$. Here's a more offbeat example of something which is almost linear, but not quite.

Example XX.1.2 (Line with two origins)**.** Consider the set $A = \mathbb{R} + \mathbb{R}$, two disjoint copies of the real line. Define an equivalence relation on A by

$$(x, 0) \sim (y, 1) \Leftrightarrow_{\text{def}} x = y \neq 0.$$

That is, we identify elements of the form $(a, 0)$ and $(a, 1)$, but we keep the two origins separate. (You may want to prove that this is indeed an equivalence relation.) Now

consider the quotient $A/\sim$. This set inherits an ordering from $\mathbb{R}$ which is not linear because the two origins are not comparable. (Work out the details!)

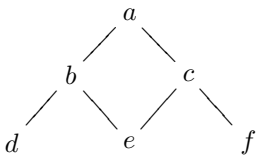
XX.2 Suprema and Infima

Definition XX.2.1 (Supremum)**.** Given a poset $(A, \leq)$ and a subset $U \subseteq A$, a *supremum* of U is an element $a \in A$ such that $u \leq a$ for all $u \in U$, and whenever y is an element of A with $u \leq y$ for all $u \in U$, then $a \leq y$.

A supremum is also called a *join*, or a *least upper bound*: it is an upper bound in the sense that it is above all elements of U, and it is the least such. Because a set U can have at most one supremum in $(A, \leq)$ (check this) we write $\bigvee U$ for the supremum of U if it exists. If the set U has exactly two elements, say $U = \{p, q\}$, then we write $p \vee q$ for $\bigvee\{p, q\}$. Such a supremum is called a *binary supremum*.

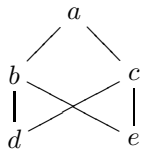
Examples XX.2.2.

1. Consider the Hasse diagram



   We have $a = b \vee c$, $b = d \vee e$, $c = e \vee f$, $a = d \vee f = \bigvee\{d, e, f\} = \bigvee\{d, c, f\}$.

2. In the Hasse diagram



   the set $\{d, e\}$ has no supremum; it has three upper bounds (namely a, b and c) but none of them is the smallest.

3. In the closed unit interval, ordered as usual, we have $\bigvee\{1 - \frac{1}{n+1} | n \in \mathbb{N}\} = 1$. Indeed, 1 is an upper bound for this set since $1 - \frac{1}{n+1} < 1$ for all $n \in \mathbb{N}$. Moreover, it is the least such bound: if $y < 1$ would be a bound, pick n large enough so that $1 - \frac{1}{n+1} > y$. Contradiction.

4. In $\mathcal{P}(A)$ the supremum of a set of subsets is simply their union.

5. Some sets don't have a supremum, simply because they are not bounded above: for example the set of prime numbers (regarded as a subset of $\mathbb{N}$) has no upper bound, let alone a least upper bound.

6. Even if a set has an upper bound, it may not have a least upper bound. Consider the rational numbers, and the subset $U = \{x \in \mathbb{Q} | x < \pi\}$. This set is bounded above, but there is no least upper bound, because π is not an element of $\mathbb{Q}$.

If we replace $\leq$ by $\geq$ we get:

Definition XX.2.3 (Infimum)**.** Given a poset $(A, \leq)$ and a subset $U \subseteq A$, an *infimum* of U is an element $a \in A$ such that $u \geq a$ for all $u \in U$, and whenever y is an element of A with $u \geq y$ for all $u \in U$, then $a \geq y$.

Infima are also called *meets*, or *greatest lower bounds*, and are unique when they exist. We denote the meet of U by $\bigwedge U$. Meets in $(A, \leq)$ are simply joins in the opposite ordering $(A, \leq^{op})$.

A special case of joins and meets is worth mentioning. When, in a poset $(A, \leq)$, the join $\bigvee A$ exists, it is called the *top element* (also: *maximum element*, or *greatest element*) of A. Dually, $\bigwedge A$ is called a *bottom element* (or *least element*, or *minimum element*).

For example, 0 is the least element of $\mathbb{N}$; there is no greatest element. The integers don't have a greatest or least element. In $\mathcal{P}(X)$ under the subset ordering, the empty set is the least element, while X itself is the top element.

The following lemma is an exercise in vacuity, but nevertheless useful.

Lemma XX.2.4. *In any poset $(A, \leq)$ with a least element $\bot$ we have $\bigvee \emptyset = \bot = \bigwedge A$. Dually, if $\top$ is a greatest element we have $\bigwedge \emptyset = \top = \bigvee A$.*

Proof. Suppose $\bot$ is a least element. We must show that it is a supremum of the empty set and an infimum of A. The latter is clear. So consider any element $a \in A$. Trivially, a is an upper bound for the empty set. Hence all elements of A are upper bounds for $\emptyset$. A supremum for $\emptyset$ is the least of them all, that is, it is the least element of A. The dual statement is proved analogously. □

Definition XX.2.5. When in a partial ordering $(A, \leq)$ each subset has a supremum and an infimum, we say that the ordering is *complete*.

The two canonical examples of complete orderings are:

Example XX.2.6. The powerset $\mathcal{P}(X)$ with the subset ordering is complete: joins are unions, and meets are intersections.

Example XX.2.7. The closed unit interval $[0, 1]$ is complete.

The following are not complete:

Examples XX.2.8.

1. The poset $\mathbb{Q} \cap [0, 1]$ with the usual ordering is not complete.
2. The poset $\mathcal{P}_f(X)$ of finite subsets of X is not complete when X is an infinite set.
3. Since a complete poset has both a least and a greatest element (see Lemma XX.2.4), posets such as $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ are not complete.

The importance of the completeness of $[0,1]$ (or any closed, bounded subset of $\mathbb{R}$) is hard to overstate. It is in fact a *defining* property of the real numbers, which sets them apart from the rationals.[1]

While suprema and infima are technically different, they are not unrelated, as the following lemma shows.

Lemma XX.2.9. *Suppose a poset has all infima. Then it also has all suprema. (The converse is true as well.)*

Proof. Suppose $(A, \leq)$ has infima. Then it has a top element by Lemma XX.2.4. Thus any subset $U \subseteq A$ has an upper bound. Now for a subset U of A, consider the set

$$V = \{x \in A | x \text{ is an upper bound of } U\}$$

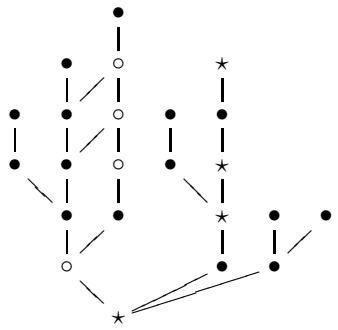
and form $b = \bigwedge V$. Claim: b is the supremum of U. First, note that if $u \in U$, we have $u \leq b$ since u is a lower bound for V and since b is the greatest lower bound of V. Thus b is an upper bound for U. Now suppose we had any other upper bound x of U. Then $x \in V$. Hence $b \leq x$ by definition of b, and we're done. □

XX.3 Chains

Chains are special subsets of orders which are of particular importance in practice.

Definition XX.3.1 (Chain)**.** A *chain* in a poset $(A, \leq)$ is a subset U of A which is linearly ordered (as a subposet of A).

The following picture illustrates: the elements denoted $\circ$ form a chain, as do the elements denoted $\star$.



Note that a chain may "skip" elements.

[1] Technically, the real numbers are defined in terms of rational numbers; the real numbers are what you get when you "add the suprema which are missing in $\mathbb{Q}$". In your analysis course, this will be made precise.

Example XX.3.2. If $A = \mathcal{P}(\mathbb{N})$, then the subset

$$U = \{\{0, \ldots, n\} | n \in \mathbb{N}\}$$

is a chain. Another chain in this poset is the set $\{\mathbb{N} - \{0, \ldots, n\} | n \in \mathbb{N}\}$.

A slightly more refined concept is that of an increasing or decreasing sequence of elements. When A is any set, a *sequence* of elements of A is simply a function $f : \mathbb{N} \to A$. We sometimes write $a_0, a_1, a_2, \ldots$ for the sequence $f(0), f(1), f(2), \ldots$. For example, the function $f(n) = \frac{1}{n+1}$ defines a sequence in $\mathbb{R}$, which we can also write $1, \frac{1}{2}, \frac{1}{3}, \ldots$.

When A carries an ordering we can ask for a sequence to take this into account, in the following sense:

Definition XX.3.3 (Increasing Sequence)**.** Let $(A, \leq)$ be a poset, and $f : \mathbb{N} \to A$ a sequence in A. Then f is an *increasing sequence* when $i \leq j$ implies $f(i) \leq f(j)$. If $i < j$ implies $f(i) < f(j)$ then f is said to be *strictly increasing*.

There is an obvious dual concept of decreasing sequence, the precise definition of which we leave to you. Increasing sequences are also called *monotone*.

Examples XX.3.4.

1. The sequence $0, \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots$ is strictly increasing in $\mathbb{Q}$.
2. The identity function on $\mathbb{N}$ is a strictly increasing sequence.
3. The function $f(x) = \lfloor \frac{x}{2} \rfloor$ is an increasing sequence in $\mathbb{N}$. (Note that it "repeats" elements; this is allowed in the definition of an increasing sequence, but it makes it non-strict.)
4. The sequence $a_n = -2^n$ is strictly decreasing in $\mathbb{Z}$.
5. Any constant function is an increasing as well as a decreasing sequence (but not strict).
6. The sequence $\{0\}, \{0, 2\}, \{0, 2, 4\}, \{0, 2, 4, 6\}, \ldots$ is strictly increasing in $\mathcal{P}(\mathbb{N})$.
7. The function $f : \mathbb{N} \to \mathbb{R}$ given by $f(n) = 2^{-n}$ is a strictly decreasing sequence in $\mathbb{R}$.

Earlier we stated that the closed unit interval $[0, 1]$ is complete, in the sense of having all suprema and infima. We can now state a variation of that:

Proposition XX.3.5. *Let $a_0, a_1, \ldots$ be a sequence in $\mathbb{R}$ which is* bounded above*, in the sense that there exists q with $a_i \leq q$ for all $i \in \mathbb{N}$. Then the chain $\{a_i | i \in \mathbb{N}\}$ has a supremum.*

Proof. The values a_i need not lie in a closed bounded interval $[p, q]$ because the sequence may actually contain arbitrarily small elements. Let $p = a_0$, and consider the subset U of $\{a_i | i \in \mathbb{N}\}$ of those a_i with $p \leq a_i$. This is a non-empty subset, and it lies within the closed and bounded interval $[p, q]$, which is complete. Then $\bigvee U$ is the supremum of $\{a_i | i \in \mathbb{N}\}$. □

The completeness of closed intervals in $\mathbb{R}$ has another interesting consequence:

Archimedean Property:
There do not exist positive real numbers x, y such that for all $n \in \mathbb{N}$ we have $nx < y$.

This is proved as follows. First, call an element x *infinitesimal* w.r.t. an element y when for all $n \in \mathbb{N}$ we have $nx < y$. Let I be the set of all such positive infinitesimal x. We make a few observations about I. First, note that when $x \in I$, then in fact $nx \in I$ for every positive natural number n. Also note that if $x \in I$, then also $\frac{x}{n} \in I$ for every positive natural number n. Finally, if $0 < x' < x$ and $x \in I$ then also $x' \in I$. Suppose now that I is nonempty. Then we have $I \subseteq (0, 1)$, since clearly a number $x \geq 1$ cannot be infinitesimal. Let $a = \bigvee I$. Then a is positive. Question: is a itself infinitesimal? If so, then so is $2a$. That is a contradiction, because $2a > a$ and a is the supremum of I. Therefore a is not infinitesimal. But then $\frac{a}{2}$ also cannot be infinitesimal. That is also a contradiction, since that would mean that it is a smaller upper bound of I. The only way out is to admit that $I = \emptyset$.

XX.4 Summary

We considered two common types of orderings: *linear* orders (in which any two elements are comparable) and *complete* orders (in which any subset has an infimum and a supremum).

- A *supremum* (least upper bound) $\bigvee U$ of a subset U of a poset A is an element satisfying the statement

$$\forall x \in U.x \leq \bigvee U \quad \wedge \quad \forall y \in A[\forall x \in U.x \leq y \rightarrow \bigvee U \leq y].$$

- An *infimum* (greatest lower bound) $\bigwedge U$ of a subset U of a poset A is an element satisfying the statement

$$\forall x \in U.x \geq \bigwedge U \quad \wedge \quad \forall y \in A[\forall x \in U.x \geq y \rightarrow \bigwedge U \geq y].$$

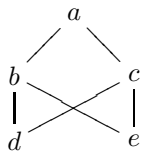
Finally, we considered *increasing sequences* (and their duals, decreasing sequences): these are functions $f : \mathbb{N} \rightarrow A$ satisfying $x \leq y \Rightarrow f(x) \leq f(y)$.

The *Archimedean Property* (possessed by the real numbers) states that there are no positive x, y such that for all natural numbers $n \in \mathbb{N}$ we have $nx < y$.

XX.5 Exercises

Exercise 278. Consider the set $A = \{a, b, c, d\}$ Find all linear orderings on A.

Exercise 279. Consider the Hasse diagram



Is $\{d, c, a\}$ a chain? What about $\{d, a\}$? And $\{d, b, c\}$?

Exercise 280. Prove that a finite linear ordering has a least and a greatest element.

If $(A, \leq)$ is linear and finite, then given an element $a \in A$, we either have that a is a least element (in which case we're done) or there is an element a' strictly below a (by linearity). Repeat this for a'; if it is not the least element we get a strictly smaller a'' and so on. If A had no least element we would obtain a strictly descending sequence of elements, which is impossible since A is finite.

Exercise 281. Show that in a linear ordering, any finite subset has a supremum and an infimum.

Exercise 282. Consider the function $a : \mathbb{N} \to \mathbb{R}$ defined by $a(n) = \sin(2\pi n(1 + \frac{1}{10^{100}})$. Is this an increasing sequence?

No, the first $10^{100}/4$ terms are increasing, but after that the sequence decreases.

Exercise 283. Show that if $(A, \leq)$ is a linear ordering, then for any $U \subseteq A$, the ordering on A restricted to U is also linear.

Exercise 284. Show that there exist increasing sequences in $\mathbb{Q}$ which have a supremum in $\mathbb{R}$ but not in $\mathbb{Q}$.

Consider $\pi = 3.1415\ldots$. Define a sequence of rational numbers by

$$a_0 = 3, a + 1 = 3.1, a_2 = 3.14, a_3 = 3.141, a_4 = 3.1415, \ldots$$

Clearly this sequence increases and converges to π.

Exercise 285. Explain the difference in terms of the existence of suprema and infima between the open unit interval $(0, 1)$ and the closed unit interval $[0, 1]$.

Exercise 286. Suppose $(A, \leq)$ and $(B, \leq)$ are partially ordered sets. For each of the following, either give a proof or a counterexample.

(a) If A and B each have a least element, then so does the product ordering on $A \times B$.

(b) If A and B each have a least element, then so does the lexicographic ordering on $A \times B$.

(c) If A and B are linear, then so is the product ordering on $A \times B$.

(d) If A and B are linear, then so is the lexicographic ordering on $A \times B$.

If A, B have least elements $\bot_A, \bot_B$ then $(\bot_A, \bot_B)$ is the least element of $A \times B$ both in the product ordering and in the lexicographic ordering.

Product orderings are generally not linear even if A, B are (take $\mathbb{N} \times \mathbb{N}$ for example), but the lexicographic ordering is: given any two elements (x, y) and (u, v), we either have $x < u$ (in which case $(x, y) \leq_l (u, v)$) or we have $u < x$ (in which case $(u, v) \leq_l (x, y)$) or $x = y$. In the last case we have $(x, y) \leq_l (u, v)$ or $(u, v) \leq_l (x, y)$ depending on whether $y \leq v$ or $v \leq y$.

Exercise 287. Prove in detail the claims made about the set I of infinitesimal numbers.

Exercise 288. Show that the Archimedean property is equivalent to the statement that for each positive real number x there exists $n \in \mathbb{N}$ with $\frac{1}{n} < x$.

Exercise 289. In a poset with suprema, show that $U \subseteq V$ implies $\bigvee U \leq \bigvee V$. Does the converse hold?

Assume $U \subseteq V$. The element $\bigvee V$ satisfies $x \leq \bigvee V$ for all $x \in V$, hence in particular $x \leq \bigvee V$ for all $x \in U$. But then by definition $\bigvee U \leq \bigvee V$.

Exercise 290. Consider $\mathbb{N} + \{\infty\}$, ordered by setting $x \leq \infty$ for all x. Show that this is a complete poset.

Consider a subset U of $\mathbb{N} \cup \{\infty\}$. If U is a finite set not containing ∞ then the supremum of U is simply the maximal element in U. (And it is 0 if U is empty.). In all other cases, we have $\bigvee U = \infty$.

Exercise 291. Show that if a poset satisfies $\{a \wedge b, a \vee b\} = \{a, b\}$ then it is linear. Does the converse hold?

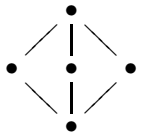
Given a, b, the assumption gives that $a \wedge b = a$ or that $a \wedge b = b$. But $a \wedge b = a$ iff $a \leq b$ and $a \wedge b = b$ iff $b \leq a$. Thus $a \leq b$ or $b \leq a$, hence A is linear. The converse holds as well; in a linear ordering we may take $a \wedge b = min(a, b)$ and $a \vee b = max(a, b)$.

Exercise 292. Give an example of a poset which is complete but which doesn't satisfy the distributive law

$$a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c).$$

Hint: you only need 5 elements.

Take



Exercise 293. The infinite distributive law is

$$a \wedge \bigvee U = \bigvee \{a \wedge u | u \in U\}.$$

Prove that $\mathcal{P}(X)$ satisfies this law. Which predicate-logical principle is the analogue of this law?

LECTURE XXI

Well-Orderings

When we discussed the Axiom of Choice, we found that in some sets the task of choosing an element from a non-empty subset is easy, because the set carries an ordering: given a subset, simply choose the least element in that subset.

Of course, that doesn't always work: $\mathbb{R}$ carries an ordering, but not every non-empty subset has a least element. The problem is that the ordering is not nice enough (even though it's linear). In this lecture we study well-orderings, which are those orderings for which we can solve our choice problem. We also consider applications which are often used in practice, most notably Zorn's Lemma.

XXI.1 Well-ordered sets

We begin with the expected definition:

Definition XXI.1.1 (Well-ordering)**.** A poset $(A, \leq)$ is a *well-order* when every non-empty subset of A has a least element.

We have already mentioned some examples, but here is a more precise account:

Examples XXI.1.2.

1. The natural number $\mathbb{N}$ with the usual ordering is the archetypical example of a well-ordering.
2. The integers $\mathbb{Z}$ are *not* well-ordered. For example, the subset $\mathbb{Z}$ does not have a least element. Same problem for $\mathbb{Q}$ and $\mathbb{R}$.
3. The closed rational unit interval $\{x \in \mathbb{Q} | 0 \leq x \leq 1\}$ is not a well-order: the subset $\{\frac{1}{n+1} | n \in \mathbb{N}\}$ is non-empty but doesn't have a least element. Same argument for the real closed unit interval.

4. Any finite linear ordering is a well-order.

5. Consider $A = \{0, 1\} \times \mathbb{N}$ in the lexicographic ordering (which in this case looks like two copies of $\mathbb{N}$ stacked on top of each other). This is a well-order: given a non-empty subset U of A, first ask whether U contains any elements of the form $(0, n)$. If so, there is a least such n. If not, all elements are of the form $(1, m)$, and there must be a least such m.

The following lemma, whose proof we leave as a useful exercise, lays out some of the basic properties of well-orders.

Lemma XXI.1.3. *If $(A, \leq)$ is a well-order, then*

- *it has a least element;*
- *it is linear;*
- *any $Y \subseteq A$ is again a well-order in the restricted ordering on Y.*

Well-orderings can also be characterized in terms of decreasing sequences. We use the following terminology: a decreasing sequence $a_0, a_1, a_2, \ldots$ is said to *become stationary* when there is an $N \in \mathbb{N}$ such that $a_i = a_N$ for all $i \geq N$.

Proposition XXI.1.4. *A linear ordering $(A, \leq)$ is a well-ordering if and only if every decreasing sequence becomes stationary.*

Proof. Suppose A is well-ordered and $a_0, a_1, \ldots$ is a decreasing sequence. Then the set $\{a_i | i \in \mathbb{N}\}$ has a least element, a_N. Because the sequence decreases we have $a_i \leq a_N$ for all $i \geq N$. But since $a_N \leq a_i$ for all $i \in \mathbb{N}$, we see that the sequence becomes stationary after N steps.

Conversely, suppose each decreasing sequence in A becomes stationary and that $U \subseteq A$ is non-empty. Let a_0 be any element in U. If there exists an element strictly below a_0 in U, let a_1 be such element; if not, let $a_1 = a_0$. Continue this way to obtain a decreasing sequence $a_0, a_1, a_2, \ldots$. Since this sequence becomes stationary after N steps for some N, this means that there is no element in U strictly below a_N, that is, that a_N is the least element of U. [Note: this last step uses linearity of $(A, \leq)$.] □

XXI.2 Choice, Order and Zorn's Lemma

As stated earlier, the main point about well-ordered sets is that they give us a way to define a choice function.[1] This leads to the question whether we can always equip a given set with a well-order. For example, can we well-order the real numbers?

Definition XXI.2.1 (Well-ordering Axiom)**.** Every set can be well-ordered. That is, for every set X there exists an ordering $\leq$ of X which is a well-order.

[1] Another important aspect is that well-ordered sets admit an induction principle, generalizing the usual induction principle for natural numbers.

Proposition XXI.2.2. *The well-ordering axiom implies the Axiom of Choice.*

Proof. We show that if X can be well-ordered, then a choice function for X exists. So suppose that $(X, \leq)$ is a well-ordering of X. We need to define a choice function $s : \mathcal{P}_+(X) \to X$. Put, for a non-empty $U \subseteq X$,

$$s(U) = \text{the least element of } U.$$

This makes sense, because by definition of well-order U has a least element (and such a least element is necessarily unique). Clearly we have $s(U) \in U$, so s is a choice function. □

In fact, the converse implication also holds, as we shall see in a moment. However, at this stage it is convenient to introduce a third principle called Zorn's Lemma. Before we state it, we have a bit of terminology:

Definition XXI.2.3 (Maximal element)**.** An element a in a poset $(A, \leq)$ is *maximal* when there is no $b \in A$ with $a < b$.

Each maximum element is maximal, but not the other way around. A counterexample is given by the discrete ordering on a set, where every element is maximal but no element is the maximum. Note also that a poset may have more than one maximal element.

Zorn's Lemma now reads:

Zorn's Lemma:
If $(A, \leq)$ is a poset in which each chain has an upper bound, then $(A, \leq)$ has a maximal element.

Before we prove Zorn's Lemma, we give a typical application. Suppose we want to show that every vector space V has a basis. Intuitively, we start by picking any non-zero vector a_0 in V; if this vector already spans V we're done. Otherwise, we can choose a vector a_1 which is linearly independent of a_0. If $\{a_0, a_1\}$ spans V, we're done, otherwise we continue. If V is finite-dimensional this stops after finitely many steps. But what if V is really big? For example, V could be the vector space of all functions from $\mathbb{R}$ to $\mathbb{R}$. Then a basis for V is uncountable, and adding elements one by one is going to take a while.

Zorn's Lemma "solves" this problem. Consider the poset P whose elements are pairs (U, A) where U is a subspace of V and A is a basis for U. The ordering is given by $(U, A) \leq (V, B)$ iff $U \subseteq V$ and $A \subseteq B$. If we have a chain $K = \{(U_i, A_i) | i \in I\}$ in P, there exists an upper bound for it: take the least subspace of V spanned by the union $U = \bigcup_{i \in I} U_i$. Then the set $A = \bigcup_{i \in I} A_i$ is a basis for U, so that (U, A) is an upper bound for the chain. Zorn's Lemma now asserts that P has a maximal element (W, C). Claim: we have $U = W$, so that C is the sought-after basis. Proof: suppose $W \neq V$.

Then there exists a vector $v \in V$ which is not in W (and hence not a linear combination of vectors in W). Consider the subspace of V defined as $W' = span(W \cup \{v\})$. Then $C' = C \cup \{v\}$ is a basis for W'. But that means that $(W, C) < (W', C')$, contradicting the maximality of (W, C).

Many mathematical applications follow this pattern: they invoke Zorn's Lemma to show that some mathematical object exists, rather than using the Axiom of Choice directly.

Next, we state the main result:

Theorem XXI.2.4. *The Axiom of Choice, the Well-ordering Axiom and Zorn's Lemma are equivalent.*

Proof. We have already explained how the Axiom of Choice follows from the Well-ordering Axiom. It therefore suffices to show that AC implies Zorn's Lemma, and that Zorn's Lemma implies the Well-ordering Axiom.

A rigorous proof that AC implies Zorn uses transfinite induction, but the general idea is relatively simple. Suppose we have a poset $(A, \leq)$ in which every chain has an upper bound. Suppose that no maximal element exists. Then given any chain C in A, we can find an upper bound a_C for C, and then a strictly bigger element $f(C) > a_C$. Choose, for every chain C, such element $f(C)$. Now inductively define a chain by letting a_0 be arbitrary, and by letting $a_k = f\{a_i | i < k\}$. This gives a contradiction, because at some point the chain becomes "too big": its cardinality can never exceed that of A.

Finally, we prove that Zorn's Lemma implies the Well-ordering Axiom. Consider a set A which we want to well-order. Define a poset P whose elements are pairs $(U, \leq_U)$, where U is a subset of A and $\leq_U$ a well-ordering of U. The ordering is given as follows: $(U, \leq_U) \leq (V, \leq_V)$ iff $U \subseteq V$ and $\leq_U$ is $\leq_V$ restricted to U. Consider a chain $C = \{(U_i, \leq_{U_i}) | i \in I\}$ in P. Then $(U, \leq_U)$, with $U = \bigcup_{i \in I} U_i$ with the obvious ordering is an upper bound of C. Thus the condition of Zorn's Lemma is met, and there exists a maximal element $(W, \leq_W)$ in P. Claim: $W = A$. Suppose not. Then there exists $a \in A$ with $a \notin W$. Extend $(W, \leq_W)$ by placing a below every element of W. Then we have a well-ordering of $W \cup \{a\}$, contradicting maximality of W. □

XXI.3 Summary

A *well-order* is a poset in which each non-empty subset has a least element. Well-orders are linear and have a least element, but the converse is false. A well-order on a set gives a choice function for that set.

The *Well-ordering Axiom* states that every set can be well-ordered. (Note: this is not to say that every ordering is a well-ordering.)

Zorn's Lemma states that if every chain in a poset has an upper bound, the poset has at least one maximal element. Zorn's Lemma, the Axiom of Choice and the Well-ordering Axiom are equivalent.

Zorn's Lemma is most frequently used in mathematical practice, for example to show that every vector space has a basis, or that every field has an algebraic closure.

XXI.4 Exercises

Exercise 294. Show that any finite linear order is a well-order.

Exercise 295. Give an example to show that a poset without a greatest element can have the property that every chain has an upper bound. Can the poset be infinite?

Exercise 296. Consider the poset $\mathsf{Par}(A, B)$ of partial functions from A to B (ordered by $f \leq g$ iff $f(x) = g(x)$ for all $x \in dom(f)$. Show that every chain in this poset has an upper bound. What do the maximal elements look like? Is there a maximum element?

Exercise 297. Consider the poset $\mathsf{Rel}(A, B)$ of relations from A to B, ordered by inclusion. Show that every chain has an upper bound. What are the maximal elements? Is there a maximum element?

Exercise 298. In a partially ordered set $(X, \leq)$, we say that an element y is a *successor* of an element x if $x < y$, and moreover $x \leq z \leq y$ implies $x = z$ or $x = y$. Intuitively, y is above x but there are no elements in between.

(a) In the partial order $(\mathbb{N}, \leq)$, what is the successor of an element n?

(b) In the partial order $(\mathbb{Q}, \leq)$, does the element 3 have a successor?

(c) Show that if a partial order has finitely many elements, then every element is either maximal (has no elements above it) or has a successor.

(d) Give an example of a partially ordered set in which there is an element with more than one successor.

(e) Show that if x is a non-maximal element of a well-ordering, then it has a unique successor.

(f) Give an example of a well-order containing an element which is not the least element but which also is not a successor of anything.

APPENDIX A

Resources and suggested reading

There are many introductory textbooks to logic and set theory. Some possible texts you may wish to look into after this course include:

1. One of the standard introductory set theory textbooks is P.R. Halmos' *Naive Set Theory* (Undergraduate Texts in Mathematics, Springer, 1998). It explains many aspects of set theory in much more detail and finesse than we have covered in this course.

2. If you're interested in the development of formal set theory (i.e., the formal system ZFC and related systems), there is the old but useful text by P. Suppes *Axiomatic Set Theory* (Dover Books, 1962).

3. A concise discussion of many interesting aspects of logic and set theory is P.T. Johnstone's *Notes on Logic and Sets* (Cambridge University Press, 1987). It covers classical first order logic and culminates in the treatment of the famous Gödel Incompleteness Theorems.

4. For serious mathematics students who wish to get a deeper and more conceptual insight into some of the material introduced in this course (and many others which pervade mathematics) we recommend the book *Sets for Mathematics*, by F.W. Lawvere and R. Rosebrugh (Cambridge University Press, 2003), see

   `http://www.mta.ca/~rrosebru/setsformath/`

5. A webpage by Eric Schechter discussing many typical mistakes seen on the undergraduate level is

   `http://www.math.vanderbilt.edu/~schectex/commerrs/`

6. And just for fun, here is a short expository paper concerning the Axiom of Choice, Trichotomy and the Generalized Continuum Hypothesis:

   `http://www.maa.org/news/monthly544-553.pdf`

appendix B

Guide to writing proofs

In everyday mathematical usage, by a proof we mean a justification or demonstration of the truth of a mathematical statement. It is important to realize that what counts as a convincing justification or demonstration is not absolute, but highly dependent on the social context[1]. For example, an expert in a certain area of mathematics may be convinced of the truth of a proposition after reading about the key idea in the proof, which a novice may need to see many of the intermediate steps spelled out in full detail before he or she accepts the proof as correct. Also, in one class a professor may give full marks to a student whose work demonstrates that she understands enough of the material, even though the way it's written down is not flawless, while in another class the same professor will subtract marks for imprecisions in the writing of the proof. Thus, the level of detail and precision in a proof should be tailored to those who will read it.

For everyday mathematics (and in almost all of your courses), the goal is therefore to write proofs in such a way that your audience can easily follow them, gets confidence that you know what you're talking about, doesn't get the impression that you are skipping over things which need more attention, and that you are not unnecessarily elaborating on things which everyone knows are trivial or irrelevant to the problem at hand. Thus, proofs should be precise, readable, and to the point.

There is, however, a branch of mathematics called *proof theory*, which has among its aims to make precise what exactly counts as a proof, what are the limits to what we can prove, and how we can justify the various methods and techniques we use in daily life when proving theorems. On the one hand, this is a foundational enterprise, because it seeks to provide a solid technical foundation and justification for mathematics and mathematical reasoning. On the other hand, it can also be regarded as *meta-mathematics*, because it studies features of mathematics using mathematical techniques. If you think you're interested in this sort of stuff, sign up for a logic course (such as MAT3361).

[1] The highest level of rigour is obtained through computer verification.

Learning to write good proofs is not easy. The best way to learn is to carefully study examples of good proofs, to try lots of proofs yourself and make sure you get feedback on those. Keep in mind at all times that understanding how the proof works is necessary, but by no means a guarantee that you write it correctly. Also, after writing a proof, try to analyse and criticise it. How and where did you use the assumptions? What is the key step in the proof? What is the overall logical structure? In what follows below, I'll discuss several aspects of proofs and writing proofs which should make it easier for you to focus on the important things.

B.1 What are we trying to prove?

The things we wish to prove can have various forms and shapes. They all have one thing in common: they must be unambiguous, and they must have a truth value (be either true or false). You can't prove the number 12, or the function $\sin(x)$. There's nothing true or false about them; they are mathematical *objects*, not propositions. Here are some examples of the various forms mathematical statements can take:

- $a = b$ (here, the statement is that one thing equals another thing)
- aRb (one thing stands in relation R to another thing)
- If p then q (something implies something else)
- p if and only if q (something is equivalent to something else)
- There exists an x such that $P(x)$ (here, P is a property)
- There does not exist an x such that $P(x)$ and $Q(x)$
- For all x, $P(x)$
- If $x \in A$ satisfies $P(x)$ then it also satisfies $Q(x)$.

In the next section we explain how the logical form of a statement can guide you in understanding how to set up a proof. For now, we point out that what seems, from a logical point of view, the simplest type of statement, namely $a = b$, can actually be very hard to prove. This is because a and b can be very complicated objects; it can also be because their descriptions are difficult to reason about. For example, when a and b are sets, we need to prove that they have the same elements: we must show that $\forall x.(x \in a \leftrightarrow x \in b)$. But in a concrete case, say

$$a = \{n \in \mathbb{N} | n > 2, x^n + y^n = z^n\}, \quad b = \emptyset$$

(where x, y, z are some fixed natural numbers) you would have difficulty figuring out the elements of the set a. Regardless, the first thing you should do when you're facing the task of proving something is to reflect on the form of the statement.

B.2 Logical Aspects of Proofs

Let's look at some basic guidelines for turning the logical form of a statement into a strategy for proving it. Equally important is the question of how the logical form of the information given to us (in the form of axioms, hypotheses or given data) can be exploited. For each possible connective ($\wedge, \vee, \neg, \rightarrow, \leftrightarrow, \forall, \exists$) we will discuss these two aspects. We also explain how to disprove statements of the given form. In each of the cases, we give the schematics of the proof strategy; the ellipsis ($\dot{:}$) indicates that the remaining task is to fill in the dots by other valid reasoning steps.

B.2.1 Implication

Proving an implication

Schematically, this is what a proof of $p \rightarrow q$ looks like:

Proof of $p \rightarrow q$

Assume that p holds.

$\vdots$

Thus it follows that q also holds.

Conclusion: $p \rightarrow q$ holds.

Thus, the strategy is to show that *if* p holds, *then* q holds as well. Since an implication is always true when the antecedent is false, this is sufficient: if p is false then the implication is *vacuously true*. (Meaning: there is nothing to prove.) It may also happen that you don't need p to prove q. For example, if I ask you to prove that if p is odd then so is $2p+1$, the assumption is not needed. We can simply show that $2p+1$ is odd.

Frequently, an implication $p \rightarrow q$ is established by setting up a chain of intermediate implications. You might not be able to prove $p \rightarrow q$ directly, but maybe you can prove $p \rightarrow r_1$. Then you can prove $r_1 \rightarrow r_2$, $r_2 \rightarrow r_3$, and so on. And finally you get $r_n \rightarrow q$. This is an important idea: try to break down a difficult proof into smaller, manageable steps.

Refuting an implication

An implication $p \rightarrow q$ is false precisely when p is true and q is false. Thus to refute $p \rightarrow q$ you need to establish two things:

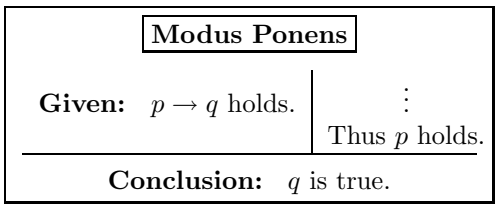
Refutation of $p \rightarrow q$

$\vdots$	$\vdots$
Thus p is true.	Thus q is false.

Conclusion: $p \rightarrow q$ is false.

Using an implication

The most common way to use an implication (which might be given to us as hypothesis, or which might have been established already) is *Modus Ponens*:

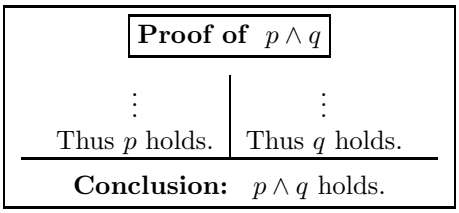


Hence Modus Ponens allows us to infer the consequent q from the implication $p \to q$ provided we can prove the antecedent p.

B.2.2 Conjunction

Proving a conjunction

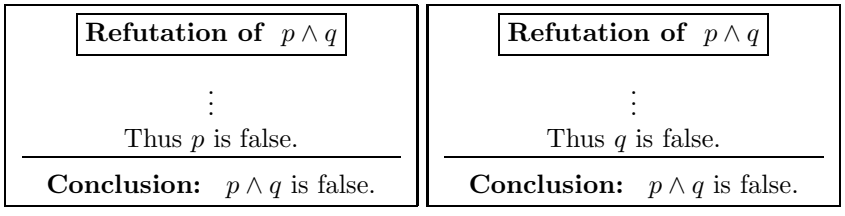
Schematically, this is what a proof of $p \wedge q$ looks like:



Thus, to show a statement of the form $p \wedge q$, simply both show p and show q.

Refuting a conjunction

A conjunction $p \wedge q$ is false precisely when at least on of p, q is false. Thus there are two strategies:



Using a conjunction

When you are given (or have previously proved) a conjunction $p \wedge q$, then you can infer p and you can infer q.

Specialization	**Specialization**
Given: $p \wedge q$.	**Given:** $p \wedge q$.
Conclusion: p.	**Conclusion:** q.

B.2.3 Disjunction

Proving a disjunction

There are two ways to prove $p \vee q$:

Proof of $p \vee q$	**Proof of** $p \vee q$
$\vdots$	$\vdots$
Thus, p holds.	Thus, q holds.
Conclusion: $p \vee q$ holds.	**Conclusion:** $p \vee q$ holds.

Thus, to show a statement of the form $p \vee q$, simply one or the other. Sometimes, things are not that straightforward, however, because it may actually be impossible to show one of the two disjunct. For example, suppose I ask you to prove that a certain positive integer n either is prime or is is divisible by a prime number strictly smaller than itself. The number may be too large to search for prime factors, or the way the number is presented to you this may be too difficult (e.g. I could have told you that n is a root of some very complicated equation). But surely the statement is true! And the obvious way to prove it is by considering cases. Case 1: the number is prime and we're done. Case 2: the number isn't prime; but then by definition it has a prime divisor m with $2 \leq m < n$. Thus in either case the statement holds.

While this example seems silly, it actually illustrates an important strategy: if you can't prove your claim directly, try to consider separate cases. If you know that the cases are exhaustive (meaning that one of them must hold, even if you don't know which one) and that the statement is true in all of those cases, then you may conclude that the statement holds regardless.

Refuting a disjunction

A disjunction $p \vee q$ is false precisely when p and q are both false. Thus:

Refutation of $p \vee q$	
$\vdots$	$\vdots$
Thus p is false.	Thus q is false.
Conclusion: $p \vee q$ is false.	

Using a disjunction

When you are given (or have previously proved) a disjunction $p \vee q$, then you can use reasoning by cases: to draw a conclusion r from $p \vee q$, show that p implies r and that q implies r.

Reasoning by cases

Assume p.	Assume q.
$\vdots$	$\vdots$
Thus r holds.	Thus r holds.

Conclusion: $p \vee q$ implies r.

Bi-implication

Proving a bi-implication

Schematically, this is what a proof of $p \leftrightarrow q$ looks like:

Proof of $p \leftrightarrow q$

Assume p.	Assume q.
$\vdots$	$\vdots$
Thus q holds.	Thus p holds.

Conclusion: $p \leftrightarrow q$ holds.

Thus, to show a statement of the form $p \leftrightarrow q$, you prove $p \to q$ and you prove $q \to p$.

Refuting a bi-implication

A bi-implication $p \leftrightarrow q$ is false precisely when $p \to q$ is false or when $q \to p$ is false. Thus there are two strategies:

Refutation of $p \leftrightarrow q$

$\vdots$	$\vdots$
Thus p is true.	Thus q is false.

Conclusion: $p \leftrightarrow q$ is false.

Refutation of $p \leftrightarrow q$

$\vdots$	$\vdots$
Thus p is fakse.	Thus q is true.

Conclusion: $p \leftrightarrow q$ is false.

Using a bi-implication

When you are given (or have previously proved) a $p \leftrightarrow q$, then you use it as you would use the implications $p \to q$ and $q \to p$.

Modus Ponens

Given: $p \leftrightarrow q$.

$\vdots$

Thus p holds.

Conclusion: q is true.

Modus Ponens

Given: $p \leftrightarrow q$.

$\vdots$

Thus q holds.

Conclusion: p is true.

B.2.4 Negation

Proving a Negation

Schematically, this is what a proof of $\neg p$ looks like:

Proof of $\neg p$

Assume p.

$\vdots$

Contradiction.

Conclusion: $\neg p$ holds.

Thus, the strategy is to show that *if* p holds, *then* we obtain a contradiction. Effectively this shows $p \to \bot$, which is equivalent to $\neg p$. Alternatively, you can think of it this way: either p is true or $\neg p$ is true. If you can eliminate the first possibility, then the only remaining possibility is the one you want to prove.

Refuting a negation

How can we prove that $\neg p$ is false? Of course, one way is to show that p is true.

Refutation of $\neg p$

$\vdots$

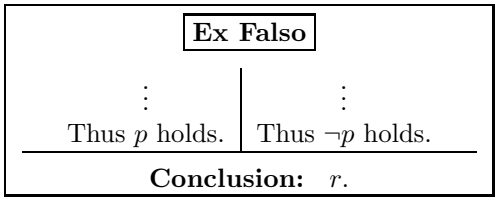
Thus p is true.

Conclusion: $\neg p$ is false.

Of course, you can also assume $\neg p$ and derive a contradiction, see next item.

Using a negation

If you are given a statement $\neg p$ then often the way to draw conclusions from this is to derive a contradiction.



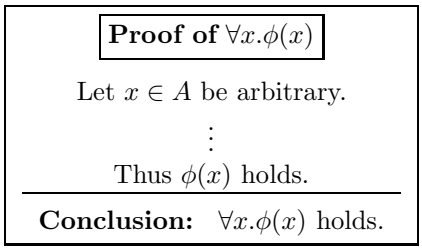
Here, r may be any statement you like. Literally, this rule says that if you can derive a contradiction you may conclude whatever you like. The reason this is valid is because of the fact that an implication is true whenever the antecedent is false; clearly, the statements p and $\neg p$ together are false.

Of course, this type of reasoning will typically occur as part of a larger proof in which you have made some assumptions. Without assumptions, you will never be able to prove a contradiction!

B.2.5 Universal Quantification

Proving a universally quantified statement

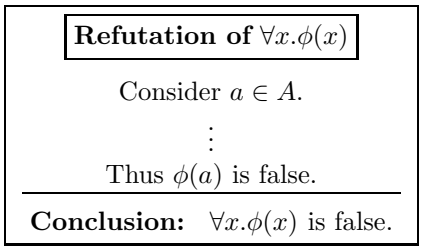
Schematically, this is what a proof of $\forall x \in A.\phi(x)$ looks like:



It is important to note that in this prove, you may not assume anything about the element x (except of course that it is an element of A); otherwise you wouldn't be proving that ϕ holds for *all* elements $x \in A$.

Refuting a universally quantified statement

How can we prove that $\forall x \in A.\phi(x)$ is false? By finding a counterexample. (Thus you must specify a concrete element a in A for which $\phi(a)$ is false.)



Using a universally quantified formula

If you are given a statement $\forall x \in A.\phi(x)$ you may instantiate it to show that any element $a \in A$ satisfies ϕ.

Instantiation

Given: $\forall x \in A.\phi(x)$ | $a \in A$

Conclusion: $\phi(a)$ holds.

B.2.6 Existential Quantification

Proving an existentially quantified statement

Schematically, this is what a proof of $\exists x.\phi(x)$ looks like:

Proof of $\exists x \in A.\phi(x)$

Consider the following element $a \in A$:

$\vdots$

Thus $\phi(a)$ holds.

Conclusion: $\exists x \in A.\phi(x)$ holds.

Thus to prove $\exists x \in A.\phi(x)$ it suffices to give a concrete example of an element $a \in A$ with $\phi(a)$. (Often, it is difficult to find an example. In that case indirect reasoning is sometimes required. For example, you can assume the statement is false and try to derive a contradiction.)

Refuting an existentially quantified statement

How can we prove that $\exists x \in A.\phi(x)$ is false? By showing that all elements of A fail to satisfy ϕ.

Refutation of $\exists x.\phi(x)$

Let $x \in A$ be arbitrary.

$\vdots$

Thus $\phi(x)$ is false.

Conclusion: $\exists x \in A.\phi(x)$ is false.

Essentially, this shows $\forall x \in A.\neg\phi(x)$, which is logically equivalent to $\neg\exists x \in A.\phi(x)$.

Using an existentially quantified formula

If you are given a statement $\exists x \in A.\phi(x)$ and you want to draw a conclusion r you face the problem that you may not know for which particular $a \in A$ the statement ϕ holds. The reasoning pattern is then as follows: we assume that we are given an element $x \in A$ for which $\phi(x)$ holds (even though we don't know which one); we then try to draw the desired conclusion r. If r doesn't involve the unknown x, then we have shown that $\exists x.\phi(x)$ implies r. Compare this with the rule for using a disjunction (reasoning by cases). In effect, to know $\exists x.\phi(x)$ is also to say that one of the cases $\phi(a), \phi(b), \phi(c), \ldots$ must hold.

Existential Elimination

Let $x \in A$ satisfy $\phi(x)$.

$\vdots$

Thus r holds.

Conclusion: $\exists x \in A.\phi(x)$ implies r.

There is an important side condition here, namely that the conclusion r does not mention the unknown element x as a free variable.

Example B.2.1. Given: $\forall x.(A(x) \to B(x))$ and $\exists x.(A(x) \wedge C(x))$. To prove: $\exists x.(B(x) \wedge C(x))$.

Solution. Take an element a for which $A(a) \wedge C(a)$ is true. Now use instantiation: from $\forall x.(A(x) \to B(x))$ we conclude $A(a) \to B(a)$.

Next, $A(a) \wedge C(a)$ implies $A(a)$, and together with $A(a) \to B(a)$ this gives $B(a)$. Also, $A(a) \wedge C(a)$ gives $C(a)$, and hence $B(a) \wedge C(a)$. This shows that $\exists.(B(a) \wedge C(a))$ holds.

B.3 What does a good proof look like?

Now that we have discussed the logical facets of proofs, it is time to turn to more practical aspects of writing proofs.

First, a proof should have a *clear structure*. At each stage of the proof, the reader should be able to figure out what is happening and why. On the coarsest level, this is what a proof looks like:

Beginning:	Lists the assumptions made or the data given, as well as pertinent definitions.
Middle:	Consists of a series of statements, each of which follows logically from the earlier ones. Justification for each step should be provided.
End:	Concludes with that which was to be proved.

Thus, a proof should be a bit like an essay, in that the real work goes into the middle part, but there is an introduction which sets the stage, and a conclusion which wraps up. This seems rather obvious, but getting this part right is often half of the work (and will earn you part marks), while getting it wrong is a recipe for disaster. One reason why people go wrong here is because the way you originally thought of the proof may be very different from the way you should present it. Often, you worked backwards from the conclusion, trying to replace the ultimate goal by simpler subgoals, until you reached a point where it is obvious that each of the subgoals follow from the assumptions. However, this is not the order in which the proof should be presented. Your written proof shouldn't be a literal recording of your train of thought, but a reconstruction of it.

Second, writing mathematical proofs requires using *precise and unambiguous language*. Except for formal proofs which are intended to be checked by a computer, mathematical proofs are a mixture of mathematical symbolism (formulas, equations, diagrams) and plain (English) language. It is obvious that the formulas should be clear and correct, and that changing an x into a y can have rather devastating consequences. However, using incorrect language to tie the formulas together can be equally bad. The main function of English in a proof is to explain how the various mathematical formulas, equations, et cetera, relate to each other. Does equation 1 follow from equation 2 by means of an algebraic manipulation? Or is equation 2 an instance of an assumption which has nothing to do with equation 1? Or is equation 2 assumed to be false in order to derive a contradiction when combining it with equation 1? There is a sense in which writing a proof is analogous to cooking: you may have the right ingredients (equations) but you can't just put them on the table like that: you need to combine them in the right order using the right techniques.

Many students begin to shy away from providing text in their proofs, because they figure out that the requirements on the text in their proofs are quite stringent. If I don't write anything except for true equations there's nothing the professor can hold against me, they seem to think. Unfortunately, that's not how it works. Your job is to convince me that you understand how all the pieces of the puzzle fit together. And in most cases, the only way to do that is to explain in words what is going on, and why you're writing the formulas you're writing. Here's an example of a "proof" which only involves symbolism and no supporting text.

Example Show that the square of an odd number is odd.

"Proof". $(2p+1)^2 = 4p^2 + 4p + 1$, $4p^2 + 4p = 4(p^2 + p)$. QED.

Critique: The key idea is there, but the reader is not given any guidance, and there are no words to indicate the structure.[2] Here's how to take the idea of the proof and dress it up so that it becomes a readable story:

Proof. Suppose n is an odd number. By definition, this means it can be written as $n = 2p + 1$ for some p. We now get

$$n^2 = (2p+1)^2 = 4p^2 + 4p + 1 = 2(2p^2 + 2p) + 1.$$

[2] In this particular example the math is rather simple so you might be forgiven for being succinct, but in more complicated cases this is no longer so.

The last step simply factors out a factor of 2. These equations show that n^2 can be written in the form $2m + 1$, and hence that it is odd. □

(Incidentally, some people err in the direction of too much text; usually, that is a way of compensating for a lack of understanding of how the proof really works. A clear proof has a good balance of symbolism and text.)

B.4 Examples

Here are some exemplary proofs. I have stated explicitly which part is the beginning, which part is the middle and which part is the end, but this is just for clarity; in practice, you don't write this (although it should be clear from the text which part is which). Keep in mind also that the proof given here is not the only one: there are many correct solutions.

Example B.4.1. Show that for all sets A, B, if $A \subseteq \mathcal{P}(B))$ then $\bigcup A \subseteq B$.

Proof.

[Beginning] We are given two arbitrary sets A, B, and the assumption is that $A \subseteq \mathcal{P}(B)$. The relevant definitions are those of the powerset: $\mathcal{P}(X) = \{U | U \subseteq X\}$ and the union $\bigcup A = \{u \in x | x \in A\}$.

[Middle] We first spell out the assumption. By definition of subset, every $x \in A$ satisfies $x \in \mathcal{P}(B)$. Using the definition of powerset, this becomes:

$$\text{If } x \in A \text{ then } x \subseteq B. \tag{B.1}$$

We now show the inclusion $\bigcup A \subseteq B$. That is, we take an arbitrary $u \in \bigcup A$ and show that $u \in B$. Since $\bigcup A = \{u \in x | x \in A\}$, we may pick $x \in A$ such that $u \in x$. By equation (B.1) this gives us $x \subseteq B$. Hence $u \in x$ implies $u \in B$.

[End] Therefore $u \in \bigcup A$ implies $u \in B$, which means $\bigcup A \subseteq B$, and the proof is complete. □

This was a relatively straightforward proof, in that when we unpack the assumptions and spell out what the statement we're trying to prove means, there is little work in connecting the two. Note however, that a proof like this one becomes a mess if you're not very precise about the difference between elements of a set and subsets of a set.

Example B.4.2. Show that there exist infinitely many prime numbers. You may use the fact that every positive integer factors uniquely (up to reordering) as a product of primes.

Proof.

[Beginning] Relevant definition: a prime number is a number which has exactly two divisors: 1 and itself. Given: every positive integer admits a unique prime factorization.

[Middle] Suppose that we have a finite set of prime numbers $\{p_1, \ldots, p_k\}$. We will show that we can find a prime number q which is not in this set. To this end, consider

$a = p_1 \cdots p_k + 1$. If this number is prime, then it is certainly different from all the p_i, so we're done. If it is not prime, then we may use the fact that every integer has a prime factorization to write $a = q_1 \cdots q_l$, where the q_i are prime numbers. Now none of the primes $p_1, \ldots, p_k$ can divide a, because the remainder of a divided by p is 1. So none of the p_i is a prime factor of a, and hence the prime factors q_j of a must be different from the p_i.

[End] Since we have shown that any finite set of prime numbers can always be extended to a larger set, there must be infinitely many primes.

In case you didn't know, this result is called Euclid's Theorem. The first thing to note about the proof is that it first reformulates the problem "There exist infinitely many primes" to "Any finite set of primes can be enlarged". This is a useful reformulation, because it suggests a proof strategy: take a finite set of primes and then find a prime not in this set. Also note that the second part (finding a prime outside the given set) uses a case distinction.

Example B.4.3. Show using the axioms of a linear map that there does not exist a linear map $\mathbb{R} \to \mathbb{R}$ whose range consists exactly of the rational numbers.

Proof.

[Beginning] The axioms for a linear map $f : \mathbb{R} \to \mathbb{R}$ are:

$$f(x + y) = f(x) + f(y) \qquad f(ax) = af(x) \qquad \text{for all } a, x, y \in \mathbb{R}$$

The range of a function f is the set $\{y | \exists x. f(x) = y\}$.

[Middle] We assume, towards a contradiction, that a linear map f whose range is equal to $\mathbb{Q}$ exists. By definition of range, that implies that for every rational number $q \in \mathbb{Q}$ there exists $x \in \mathbb{R}$ with $f(x) = q$. In particular, there exists $x \in \mathbb{R}$ for which $f(x) = 1$, since $1 \in \mathbb{Q}$.

Using the second axiom of linear maps, we now find that for any $a \in \mathbb{R}$, we have $f(ax) = af(x) = a1 = a$. Hence every $a \in \mathbb{R}$ is of the form $f(y)$. This shows that the range of f equals $\mathbb{R}$. This is a contradiction, because the range of f was assumed to contain only rational numbers.

[End] Therefore our assumption leads to a contradiction, and hence no linear map with range $\mathbb{Q}$ exists. □.

Here we were asked to prove that something does not exist. Thus we assume that we have an f with the listed properties and derive a contradiction. It turns out that we didn't need the first axiom of linearity, only the part about scalar multiplication. When it turns out that your proof doesn't use one of the assumptions, be extra alert: it means you've either done a very efficient job, or there's something you've missed! Note also that we used the assumption about the range of f in two places, first to find an x with $f(x) = 1$, and then to get a contradiction with $f(y) = a$.

Example B.4.4. Prove that $\sqrt{2}$ is irrational.

Proof.

[Beginning] Definition of rational number: a number of the form $\frac{p}{q}$, where $p, q \in \mathbb{Z}$. Definition of irrational number: a number which is not rational.

[Middle] Towards a contradiction, suppose that $\sqrt{2}$ is rational. Then by definition of rational number, we may write $\sqrt{2} = \frac{p}{q}$, for integers p, q. By dividing out common factors in p and q, we may assume that p and q are not both even. (For otherwise they would have a factor 2 in common, and we would divide out by that factor.) Now we have

$$2 = (\sqrt{2})^2 = \left(\frac{p}{q}\right)^2 = \frac{p^2}{q^2}.$$

This shows that $p^2 = 2q^2$, so that p^2 is even. But the square of an odd number is always odd, so p itself must have been even, i.e. we have $p = 2n$ for some integer n. This gives

$$2 = \frac{p^2}{q^2} = \frac{4n^2}{q^2}$$

so that $q^2 = 2n^2$. This shows that q^2, and hence also q is even as well. But then both p and q are even. Contradiction, because we assumed that not both are even.

[End] Thus $\sqrt{2}$ is not rational, and hence is irrational. □.

This is a textbook proof which you need to know by heart. We must show that something is not rational. Thus we assume it is and derive a contradiction. In the beginning, a clever trick is used, namely making sure that we divide out common even factors so that p and q are not both even. (Make sure you understand why this is valid - there is a big difference by simply assuming something which was not given at all (not acceptable!) and showing that we can, without loss of generality, assume that we are in a special situation.)

Example B.4.5. Prove from the definition of convergence that the sequence $(1/n)_{n>0}$ converges.

Proof.

[Beginning] Definition of convergence: a sequence (a_n) converges to a when for every $\epsilon > 0$ there exists N such that $|a_n - a| < \epsilon$ for all $n \geq N$. We are given the sequence $(1/n)_{n>0}$.

[Middle] By inspection of the first few terms $1/1, 1/2, 1/3, 1/4, 1/5, \ldots$ of the sequence, we see that it approaches 0. Thus we will prove that the sequence converges to 0. To this end, suppose we are given $\epsilon > 0$. We need to find a number N such that for each $n \geq N$, the distance between $1/n$ and 0 is less than ϵ. That is, we need to find N such that $1/n < \epsilon$ for all $n \geq N$. The condition $1/n < \epsilon$ can be rewritten as $1/\epsilon < n$, since both n and ϵ are strictly positive. Therefore, if we take N to be $1 + 1/\epsilon$, we get, for each $n \geq N$:

$$1/n \leq 1/N = \frac{1}{1 + 1/\epsilon} = \frac{1}{(1+\epsilon)/\epsilon} = \epsilon/(1+\epsilon) < \epsilon.$$

[End] This proves that the sequence $(1/n)$ converges to 0. □.

This type of statement is complex because it contains three quantifiers.[3] However, being systematic helps. We assume we are given an arbitrary positive ϵ. Now we need

[3] Humans are notoriously bad at reasoning with more than three quantifier alternations, and you must force yourself to carefully follow the rules for reasoning with quantifiers.

to find N with a certain property, namely that for all $n \geq N$, $1/n < \epsilon$. We need to reflect on the latter inequality a bit before we realize that $N = 1 + 1\epsilon$ does the trick, but once we see that the rest of the proof is easy. (Note that you could take different N, e.g. $N = 1 + 1/2\epsilon$.)

B.5 Common mistakes in proofs

A lot of things can go wrong in writing proofs, ranging from small mistakes which don't affect the overall soundness of the proof to fundamental errors which invalidate it entirely. I'm listing the most common problems.

1. **Having the proof backwards.** This usually results from writing down the steps in the order in which you thought of them. Here's an example of a backward "proof" of the statement that $A \subseteq \mathcal{P}(B)$ implies $\bigcup A \subseteq B$.

   Incorrect Proof. We need to prove $\bigcup A \subseteq B$. Using the definition of union and subset, this means that $\{u \in x | x \in A\}$ is a subset of B, i.e. for each $u \in x$ with $x \in A$ we have $u \in B$. So for any $x \in A$ we have $x \subseteq B$, because $u \in x$ implies $u \in B$ (as shown in the previous line). But that means that $x \in A$ implies $x \in \mathcal{P}(B)$, which just says that $A \subseteq \mathcal{P}(B)$. This is is indeed the case by assumption so we're done. QED.

   What went wrong here? All definitions were expanded correctly, but the order of reasoning is backwards. Instead of beginning with the assumptions and ending with the conclusion we started by expanding the conclusion and showing that the assumptions were true. But there's no point in showing something which you are given to be true already... The remedy is to separate the way you first thought of the proof from the way you are going to present it, and to pay careful attention to whether the result starts with the assumptions and ends with the desired conclusion.

   (By the way, if you look carefully, the above incorrect proof is almost a correct proof of another statement. Which?)

   Here's another example, taken from an exam paper of an anonymous 2362 student[4]. To prove: the relation $\sim$ on $\mathbb{Q}$ defined by $a \sim b \Leftrightarrow a - b \in \mathbb{N}$ is an equivalence relation.

   "Proof" Reflective: let $a \sim a$ then $a - a$ stands. Symmetric: let $a \sim b$ then $a - b$, if $b \sim a$ then $b - a$ still holds. Transitive: let $c \in \mathbb{Z}$. If $a - b$ and $b - c$ are true then $a - c$ holds therefore $a \sim c$.

   Critique: It appears the student thought about what (s)he had to do to prove $a \sim a$, then realized this had something to do with $a - a$, looked at $a - a$ and thought it was a nice looking expression (perhaps, but we can't say for sure, even

[4] Copied verbatim from an anonymous 2362 student's final exam paper; I should stress that the purpose here is not to make fun of a student making a mistake but to expose the mistake so that we can learn from it.

realized that it was equal to 0) and then concluded (s)he had proved reflexivity. However, as written down things don't make any sense. First of all, the proof starts with "Let $a \sim a$". (This means that $a \sim a$ is introduced as an assumption.) However, this is exactly what we're supposed to prove! The proof should start with "Let $a \in \mathbb{Q}$", then observe that $a - a = 0$ and hence that $a \sim a$. The same problem occurs in the attempted proof of symmetry. One final point of criticism concerning this proof: it consistently confuses *numbers* with *propositions*. The expression $a - a$ denotes a number. It can't be true or false, and it cannot hold (nor stand for that matter). By contrast, $a - a = 0$ is a proposition which has a truth value.

2. **Type-checking errors; incorrect use of notation and terminology.** Let me give an example first, again taken verbatim from the 2362 exam of an unfortunate student. To prove: if $(X, \leq_X)$ and $(Y, \leq_Y)$ are linear orders, then so is the product ordering on $X \times Y$.

   "Proof" Given $(X, \leq_X)$ and $(Y, \leq_Y)$ are linear, then $\exists x \in X$ and $\exists a \in X$ such that $x \leq a$, and $\exists y \in Y$ and $\exists b \in Y$ such that $y \leq b$. If $x \leq a$ and $y \leq b$ then $x \times y \leq a \times b$. So $x \times y$ is linear. Since $x \in X$ and $y \in Y$, $x \times y \in X \times Y$. Thus $X \times Y$ is linear.

   Critique: On the positive side, this proof starts by mentioning the assumption and ends with the conclusion. Aside from that, it's mostly nonsense. First of all, the first claims (about the existence of elements x, a, y, b) are simply false, for the trivial reason that X and/or Y could be empty. But the real problem is with the $x \times y$. What is this supposed to mean? It is supposed to be an element in $X \times Y$. But the elements of a product are *pairs* (x, y), not products. But even if we generously assume this is what the student had in mind, and we try to replace $x \times y$ by (x, y), it still doesn't make sense, because now we see the statement "So (x, y) is linear." What on earth does it mean to say that an element of $X \times Y$ is linear?? Linearity is a property of an ordered set as a whole, not of its individual elements. [This is what programmers call a *type-checking error*, meaning that we try to apply functions or properties to entities for which this doesn't make sense.]

3. **Logical mistakes.** The most common errors are:

   (a) *Incorrect negation of a statement.* You need to find the negation of a statement when you do a proof by contradiction (or when you try to refute a statement). Mistakes are most often made with statements involving quantifiers and implications. Here's an example. What is the negation of the statement that for each $x \in A$ there exists $y \in B$ such that $x = y$?

   1. There is no $x \in A$ such that if $y \in B$ we have $x = y$.
   2. There is an $x \in A$ such that for all $y \in B$ we have $x = y$.
   3. For each $x \in A$ there exists $y \in B$ such that $x \neq y$.
   4. There exists an $x \in A$ such that for all $y \in B$ we have $x \neq y$.

   (The correct answer is 4.)

   (b) *Proof by Example.* This means that you prove only one special case, while you're supposed to prove a general statement. For example, here's a "proof" that if n is divisible by 3, then so is n^2.

Incorrect Proof. Let n be an arbitrary number which is divisible by 3, such as 12. Then $n^2 = 12^2 = 144 = 3 \cdot 48$, and we're done.

The problem of course is that 12 is far from arbitrary. The proof only shows that the statement holds for $n = 12$.

(c) *Confusing the converse, negation and contrapositive of an implication.* Suppose we are asked to prove an implication "If p then q". A common mistake is to prove the converse instead (the converse of this implication would be $q \to p$). What is correct is to prove the contrapositive statement $\neg q \to \neg p$, because the latter is logically equivalent to the original statement. However, don't confuse the contrapositive (or the converse) with the negation of the statement, which reads $\neg(p \to q)$.

Consider the following example, again taken from a 2362 exam. Question: is the set $\mathbb{Q}^+ = \{x \in \mathbb{Q} | x \geq 0\}$ with the restricted ordering a well-order?

Wrong answer: Yes: any well-order has to have a least element. $(Q, \leq)$ is not a well-order because it doesn't have a least element, but $\mathbb{Q}^+$ does.

Critique: The student confuses the true implication

$$\text{If } (X, \leq) \text{ is a well-order then it has a least element}$$

with the manifestly false converse

$$\text{If } (X, \leq) \text{ has a least element then it is a well-order.}$$

To put the same thing in slightly different terms: having a least element is a necessary, but not a sufficient condition for being a well-order.

4. **Incorrect manipulations and abuse of notation.** These include manipulations of formulas and equations which are not allowed, even though they might accidentally lead to the correct outcome. Most of the time these are grounded in a lack of understanding of what the formulas and the symbols in them mean, and what rules we are allowed to use to manipulate them.

   For example, let's prove that the set of odd numbers is empty:

   Incorrect Proof. We have $\mathbb{Z} = E + O$ where E is the set of even numbers and O is the set of odd numbers. Then also $O = \mathbb{Z} - E$, so

$$\begin{aligned} |O| &= |\mathbb{Z} - E| \\ &= |\mathbb{Z}| - |E| \\ &= |\mathbb{N}| - |\mathbb{N}| \\ &= \emptyset \end{aligned}$$

   The mistakes here (besides not recognizing the absurdity of the conclusion) is that the equation $|A - B| = |A| - |B|$ doesn't make sense. Some similar looking equations (which are part of what is called *cardinal arithmetic*) do hold, e.g. $|A + B| = |A| + |B|$. However, these things break down when we use the difference of sets.

5. **Using the wrong definitions.** This one speaks for itself. If you don't use the correct definitions to start with, then the rest of the proof is irrelevant.

6. **Insufficient justification of intermediate steps.** This includes leaving out intermediate steps which are not obvious, jumping to conclusions, or using results along the way which are correct but which haven't been established yet. Also bad is using existing results without mentioning it.

   By way of example, let's see what's wrong with the following proof that there are infinitely many prime numbers.

   Handwaving Proof. To prove that there are infinitely many primes, consider an arbitrary prime p. Now we can always pick a prime q which is strictly larger than p. This shows that we can build a sequence of larger and larger primes, and hence there are infinitely many of them.

   Where is the leap of faith here? It's in the claim that we can always find a bigger prime. That's true, but it's at the same time the key part and the only non-obvious part of the proof. (Compare with the correct proof in the previous section.) Note also that this proof attempt is suspicious in that it consists of prose only. This can be a symptom of the author not quite identifying the crux of the matter.

B.6 Practical tips

We have discussed how to write and how not to write a proof. But we haven't said anything about how to come up with a proof. You need to have some strategies for attacking proofs which are not immediate. Some of these may seem obvious from what we said before, but it's important to try easy and obvious things before spending lots of time on sophisticated techniques.

1. Before you begin, make sure you understand the logical format of the thing you're trying to prove. This format often dictates the proof strategy.

2. The easy parts of the proof are the beginning and the end. So start by writing out very carefully what is given, what the assumptions are, and unpack what these mean. Also write out carefully what exactly you're supposed to prove, and unpack that statement as well. Now your job is to fill in the gap in the middle. Sometimes that will stare you in the face. Other times you may need to do some work on either end in order to make things fit. (Of course, you have to make sure at the end that you write it down in the correct order.)

3. Another way to bridge the beginning and the end is by trying to find an intermediate statement, and then later figuring out how to get there from the beginning and how to get the end from the intermediate statement. (Of course, this can applied iteratively, finding a sequence of intermediate statements.)

4. If a direct approach turns out to be (too) difficult, it may help to find a logically equivalent formulation of the statement. For example, if you're trying to prove an implication $p \to q$, you may instead prove the equivalent $\neg q \to \neg p$.

5. Sometimes you will need to do a proof by contradiction. The reason why this sometimes works is that it gives you an extra assumption to work with.

6. If you don't have a strong intuition yet about why the statement you're trying to prove is true (or whether it is true at all), try looking at some special cases first to get a better insight into the problem.

7. Sometimes techniques and ideas from one area can be used in unexpected ways in other areas. Be open-minded about trying various techniques, especially if you have the feeling that there is a similarity with problems you've seen before.

8. After writing down your proof, check it carefully for the following:

   - Is the order of the proof correct?
   - Are you using what is given, and nothing else?
   - Does each line parse, and is each of the individual statements correct?
   - Do you provide justification for each of the steps?
   - Is there a good balance between mathematical symbolism and plain English?

9. After you think you've completed your proof, don't rush to the next one, but make sure you learn something from it. Do you understand the structure of your proof? How does it work in special cases? Where do you use the assumptions, and how?

10. Try to explain your proofs to others to see if you understand them well enough to communicate them clearly. Make sure you get lots of feedback from your teachers and fellow students.

APPENDIX C

Solutions to selected exercises

Lecture I

Exercise 1:

(a) Yes (this is a simple proposition).

(b) Yes (this is a compound proposition).

(c) No (this is a question, not a statement).

(d) No (this is an imperative).

(e) Yes.

(f) No; the first part is an imperative, the second part a proposition.

(g) Yes.

(h) No (again an imperative).

(i) It is a proposition, but a false one. (And one you should be wary of, given the self-referential nature.)

Exercise 2: *Definitely* not well-formed are (c), (d), (g), (i), and (j). By convention we accept (b). Note that (d) doesn't fall under the convention because $(p \leftrightarrow q) \leftrightarrow r$ is not equivalent to $p \leftrightarrow (q \leftrightarrow r)$. (Most of the time, when people write $p \leftrightarrow q \leftrightarrow r$ they actually mean $(p \leftrightarrow q) \wedge (q \leftrightarrow r)$.)

Exercise 5:

(a) $\neg p \rightarrow \neg r$

(b) $r \rightarrow p$

(c) $p \wedge q$ (the "even though" suggests contrast, but doesn't have a logical meaning)

(d) $\neg q \wedge (\neg p \to \neg q)$ (also possible: $(p \to \neg q) \wedge (\neg p \to \neg q)$)

(e) $r \to (\neg p \wedge \neg q)$

(f) $(\neg q \vee \neg p) \to \neg r$

Lecture II

Exercise 9:

(a) $v(p \wedge (q \vee r)) = \mathsf{f}$ because $v(q \vee r) = \mathsf{f}$.

(b) $v(p \to (\neg q \to \neg r)) = \mathsf{t}$ because $v(\neg q) = v(\neg r) = \mathsf{t}$, whence $v(\neg q \to \neg r) = \mathsf{t}$.

(c) $v(\neg p \to \bot) = \mathsf{t}$ because $v(\neg p) = \mathsf{f}$.

(d) $v(p \vee (q \wedge (\neg r \to q))) = \mathsf{t}$ because $v(p) = \mathsf{t}$.

Exercise 10: Consider first the possibility that $v(p) = \mathsf{f}$. In that case, the third statement gives $v(q \vee r) = \mathsf{f}$, whence $v(q) = \mathsf{f} = v(r)$. This assignment makes all three statements true.

Consider now the possibility that $v(p) = \mathsf{t}$. Then $v(\neg q \to r) = \mathsf{f}$, so that $v(q) = v(r) = \mathsf{f}$. This also makes the statements true. Hence we can conclude that either way $v(q) = v(r) = \mathsf{f}$, and $v(p)$ can be anything.

Exercise 11:

(a) The most straightforward solution is $(\neg p \wedge q) \vee (p \wedge \neg q)$. Also possible is $p \leftrightarrow \neg q$.

(b) Use the laws of boolean algebra to transform the formula as follows:

$$(\neg p \wedge q) \vee (p \wedge \neg q) \equiv (\neg p \wedge \neg\neg q) \vee (\neg\neg p \wedge \neg q) \equiv \neg(p \vee \neg q) \vee \neg(\neg p \vee q).$$

(c) Further manipulate to get $\neg(\neg(\neg p \to \neg q)) \to \neg(\neg\neg p \to q)$. This may be simplified to $\neg(\neg(q \to p)) \to \neg(p \to q)$.

Exercise 14: Use the legenda

p – We support the candidate.

q – She will be elected.

r – Taxes will increase.

s – We can afford to buy a new car.

Then the translation becomes:

$$\frac{\begin{array}{c} p \to q \\ q \to r \\ \neg r \to s \end{array}}{\neg p \to s}$$

This is not valid. When $v(p) = v(s) = v(q) = \mathtt{f}$ and $v(r) = \mathtt{t}$ all assumptions are true but the conclusion is false.

Exercise 18: We have (find out for yourself which rules were used!)

$$\begin{aligned}(\neg p \to q) \to r &\equiv (\neg\neg p \vee q) \to r \\ &\equiv (p \vee q) \to r \\ &\equiv \neg(p \vee q) \vee r \\ &\equiv (\neg p \wedge \neg q) \vee r \\ &\equiv r \vee (\neg p \wedge \neg q) \\ &\equiv (r \vee \neg p) \wedge (r \vee \neg q) \\ &\equiv (\neg p \vee r) \wedge (\neg q \vee r) \\ &\equiv (p \to r) \wedge (q \to r)\end{aligned}$$

Exercise 21: No, this is not possible. A knight cannot say this because it would be a lie, and a knave cannot say this because it would be true.

Exercise 22: First, the speaker cannot be a knight, for then he would be lying. Thus the speaker is a knave and the statement is false. The only way it can be false is when the other person is a knight.

Exercise 24: If the first person is a knight, then his statement is true, hence they are both knights. But then the second statement would be a lie. Contradiction. Thus the first person is a knave and hence his statement is false, i.e., the two are of different type. Thus the second person is a knight. (And the second person correctly states that the first statement is false.)

Exercise 27: Recall that an implication is always true when the consequent is true. Thus the statements by A and C are true statements, and hence must have been spoken by knights.

As for B, note that he can't be a knight. For then the implication would have a true antecedent and false conclusion, meaning that the whole implication would be a lie. Can B be a knave? Then the implication would be false, which is only possible if the antecedent were true, which it is not. Conclusion: B can never say this.

Finally, D. This can be said by a knight (for then both antecedent and consequent are false, hence the entire implication is true). However, it can also be said by a knave: then the antecedent is true and the conclusion false, hence the entire implication false. We can't deduce anything here.

Exercise 28: Suppose A is a knight. Then his statement is true, hence B is a knave. Then B's statement is false, meaning that A and C have opposite type. Therefore C is a knave. Suppose next that A is a knave. Then his statement is false, hence B is a knight. Then B's statement is true, meaning that A and C have the same type. Thus C is a knave.

In either case, C is a knave.

Lecture III

Exercise 32: Throughout, we use $P(x)$ for x is prime, $E(x)$ for x is even, and $O(x)$ for x is odd.

(a) $\exists x \in \mathbb{N}.E(x) \wedge \exists x.P(x)$.

(b) $\neg\forall x.(O(x) \to P(x))$.

(c) $\exists x.(E(x) \wedge P(x))$.

(d) $\neg\forall x.(P(x) \to O(x))$.

(e) $\forall x.[P(x) \wedge x \neq 2 \to O(x)]$

(f) $\exists x.[\neg P(x) \wedge \neg E(x)]$

(g) $\forall x.[\neg P(x) \to (\neg O(x) \vee x \neq 2)]$.

Exercise 33:

(a) This is false. A counterexample is: $x = 2, y = 0$.

(b) Still false: take $x = 1, y = 1$.

(c) True.

(d) True: given $x \in \mathbb{N}$ with $x > 0$, you can take $y = \frac{1}{n}$.

(e) False: for example, when $x = \pi$, the only number y with $xy = 1$ is $y = \frac{1}{\pi}$. But that number is not in $\mathbb{N}$.

(f) True: for example, take $x = \frac{1}{2}$. Then if y is any real number there are two cases: either $y \leq 0$, in which case we're done, or $y > 0$, in which case we have $\frac{1}{2}y < y$.

(g) True. (This statement expresses that there exists a unique element x for which $xy = y$ for all y. Of course, $x = 1$ is this element.)

(h) True: take $x = 2$. Then given y, z, there are two cases. Case 1: $y < z$. Then also $y < 2z$, hence $zx > y$. Case 2: $z \leq y$. Then $z > 2y$ as well, and hence $xy > z$.

Exercise 35: The statement says that in every row of the matrix there is an entry which is either ≤ -5 or is the square of a positive natural number. In matrix A, this fails for the second (and the third) row. In matrix B, each row has a 1 in it, which is a square of a positive natural number. Thus the statement is true for B. It is true for C as well: rows 1 and 3 have entries smaller than -5 and row 2 has a square, namely $(3+1)^2 = 16$. In matrix D, every row has a 1, so the statement is true. For E it is true as well, and it is false for F because row 3 doesn't have an entry which is less than -5 or and entry of the form $(n+1)^2$.

Exercise 36: It's Donny who is out of his element here. While it is true that for $x = 1$ the statement is incorrect, all one needs is one example of an x for which the statement holds. As Walter correctly points out, the statement holds for $x = -1$,

simply because then for any y the antecedent of the implication is false, and hence the whole implication is true. Another witnessing example would be $x = \frac{1}{2}$, because then for $0 < y \leq \frac{1}{2}$ we have $y^2 < y$.

Exercise 37: We have

$$\begin{aligned}
\neg\exists x \in \mathbb{R}\forall y \in \mathbb{R}.((0 < y \wedge y \leq x) \rightarrow y^2 < y) &\equiv \forall x \in \mathbb{R}\neg\forall y \in \mathbb{R}.((0 < y \wedge y \leq x) \rightarrow y^2 < y) \\
&\equiv \forall x \in \mathbb{R}\exists y \in \mathbb{R}.\neg((0 < y \wedge y \leq x) \rightarrow y^2 < y) \\
&\equiv \forall x \in \mathbb{R}\exists y. \in \mathbb{R}.((0 < y \wedge y \leq x) \wedge \neg(y^2 < y) \\
&\equiv \forall x \in \mathbb{R}\exists y \in \mathbb{R}.((0 < y \wedge y \leq x) \wedge y \leq y^2)
\end{aligned}$$

Exercise 39: We use the following dictionary:

$D(x)$ – x is a dog

$C(x)$ – x is a cat

$B(x)$ – x is big

$S(x)$ – x is small

$R(x, y)$ – x chases y

(a) $\forall x\forall y.(D(x) \wedge C(y) \rightarrow R(x, y))$

(b) $\exists x.(D(x) \wedge \neg\exists y.(C(y) \wedge R(x, y)))$

(c) "Some dogs chase some cats" would be translated by $\exists x\exists y.(D(x) \wedge C(y) \wedge R(x, y))$. However, the "only" suggests that they don't chase *all* cats. If that's your interpretation, you should translate as $\exists x\exists y.(D(x) \wedge C(y) \wedge R(x, y) \wedge \neg\forall z.(C(z) \rightarrow R(x, z)))$.

(d) $\forall x.(D(x) \wedge B(x) \rightarrow \forall y.(C(y) \wedge S(y) \rightarrow \neg R(x, y)))$.

(e) $\forall x\forall y((C(y) \wedge B(y) \wedge R(x, y)) \rightarrow (D(x) \wedge B(x)))$.

(f) $\exists x.(C(x) \wedge B(x) \wedge \neg\exists y.(D(y) \wedge R(x, y)))$.

(g) $\neg\forall x.((D(x) \wedge S(x)) \rightarrow \forall y.((C(y) \wedge B(y)) \rightarrow R(x, y)))$.

(h) $\forall x\forall y.((C(x) \wedge D(y) \wedge R(x, y)) \rightarrow (B(x) \wedge S(y)))$.

Exercise 41: This is not a tautology. Consider the domain {John,Mary}, and let $L(x, y)$ stand for x loves y. Then the statement says that for each person, either that person loves someone or that person is loved by everyone. When John doesn't love anyone the statement is false.

Exercise 43:

(a) $\forall x\forall y.\neg P(x, y)$

(b) $\forall x\forall y.(\neg P(x, y) \wedge \neg Q(x, y))$

(c) $\forall x \forall y.(\neg P(x,y) \vee \neg Q(x,y))$

(d) $\exists x \exists y.\neg P(x,y)$

(e) $\exists x \exists y.(P(x,y) \wedge \neg Q(x,y)$

(f) $\forall x \exists y.\neg P(x,y)$

(g) $\forall x \exists y.(P(x,y) \wedge \neg Q(x,y)$

(h) $\exists x \forall y.\neg P(x,y)$

(i) $\exists x.(\forall y.\neg P(x,y) \wedge \forall z.\neg Q(x,z))$

(j) $\exists x.(P(x) \wedge \exists y.\neg P(y))$

(k) $\exists x.(\exists y.P(x,y) \wedge \exists z.\neg Q(x,y))$

Exercise 46: With the usual dictionary, we get

$$\frac{\begin{array}{c}\forall x.[L(J,x) \leftrightarrow x \neq J] \\ \forall x.[L(J,x) \rightarrow L(M,x)]\end{array}}{L(J,M) \wedge L(M,J)}$$

This argument is invalid: $L(J,M)$ follows from the first premise, but $L(M,J)$ can be false while the premises are true.

Exercise 49: First, note that B can't be a knight; since the first part of his conjunctive statement is true the second part must be false. Hence B is sober. This, however, is a counterexample to A's first statement. Therefore A is a knave as well, and his professed sobriety is a lie as well.

Lecture V

Exercise 50: Here's a hint: ask whether this barber shaves himself.

Exercise 51: If we have $X \in X$ then we get an *infinite regress*

$$\cdots X \in X \in X \in X \in X.$$

Put more precisely, this would mean that the membership relation would not be well-founded, in the sense that there would be infinite descending chains. In standard set theory, this can't happen, but there are versions of set theory (called *non well-founded set theory*) where it is allowed.

Exercise 52: We would get $U \in U$. See exercise 51.

Lecture VI

Exercise 57: $\mathcal{P}(X) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}, \{a,b,c\}\}$.

Exercise 59: When X has n elements (where $n \in \mathbb{N}$), the set $\mathcal{P}(X)$ has 2^n elements. A rigorous proof requires induction on n.

Exercise 65: True. Note first that $A \subseteq C$ as before. To show that $A \neq C$ take $x \in B$ with $x \notin A$. Then also $x \in C$. This shows that $A \neq C$.

Exercise 68: No. $A \subset B$ implies $A \subseteq B$. Similarly, $B \subset A$ implies $B \subseteq A$. Together they give $A = B$. But $A \subset B$ implies $A \neq B$. Contradiction.

Exercise 70: Take $A = \emptyset, B = \{\emptyset\}, C = \{\emptyset, \{\emptyset\}\}$. Generally, $A \in B, B \in C$ does not imply $A \in C$, as the example

$$A = \emptyset, B = \{\emptyset\}, C = \{\{\emptyset\}\}$$

shows.

Lecture VII

Exercise 75:

(a) $\emptyset \in \mathcal{P}(A)$ is true for any set A, since $\emptyset \subseteq A$.

(b) $\emptyset \subseteq A$ for any set A, so in particular for $A = \mathcal{P}(\emptyset)$.

(c) $A \in \mathcal{P}(A)$ for any set A, so in particular for $A = \mathcal{P}(\emptyset)$.

(d) We have $\emptyset \in \mathcal{P}\mathcal{P}(\emptyset)$, hence $\{\emptyset\} \in \mathcal{P}\mathcal{P}\mathcal{P}(\emptyset)$. Also $\emptyset \in \mathcal{P}\mathcal{P}\mathcal{P}(\emptyset)$. Thus both elements of $\mathcal{P}(\emptyset)$ are also elements of $\mathcal{P}\mathcal{P}\mathcal{P}(\emptyset)$.

(e) False. Take for example $X = \{\emptyset\}$. Then $\{X\} = \{\{\emptyset\}\}$, which is not an element of $\mathcal{P}(\{\emptyset\})$.

(f) True, since $X \in \mathcal{P}(X)$.

(g) True, since $A \in \mathcal{P}(A)$ for any A, in particular for $A = \{X\}$.

(h) False. Take for example $X = \{\emptyset\}$. Then $\mathcal{P}(\{\{\emptyset\}\}) = \{\emptyset, \{\{\emptyset\}\}\}$. Then $X \notin \mathcal{P}(\{X\})$.

Exercise 76: Suppose $A = B$. Then $A \subseteq B$ and $B \subseteq A$. By Proposition VII.3.1, this gives $\mathcal{P}(A) \subseteq \mathcal{P}(B)$ and $\mathcal{P}(B) \subseteq \mathcal{P}(A)$, hence $\mathcal{P}(A) = \mathcal{P}(B)$.

The converse is true as well. If $\mathcal{P}(A) = \mathcal{P}(B)$ then on the one hand $A \in \mathcal{P}(A)$, so also $A \in \mathcal{P}(B)$, meaning $A \subseteq B$. By symmetry, $B \subseteq A$ also follows. Hence $A = B$.

Exercise 78: Since $X \in \mathcal{P}(X)$, this would imply that $X \in X$.

Exercise 79: Consider an arbitrary set X. Then let

$$\emptyset = \{x \in X | x \neq x\}.$$

This defines the empty set. Note that this only works when we already have shown some set X to exist.

Lecture VIII

Exercise 82:

(a) $A \cap B = \{2\}$

(b) $A \cup B = \{0, 1, 2, 3, 4, 5\}$

(c) $A \cap B \cap C = \emptyset$

(d) $A - B = \{0, 1\}$

(e) $A^c - B^c = \{3, 4, 5\}$

(f) $(A - B)^c = \{x \in \mathbb{N} | x > 1\}$

(g) $(A^c \cap C^c)^c = \{0, 1, 2, 3, 5\}$

Exercise 83:

(a) $\{x \in \mathbb{R} | x \notin \mathbb{Q} \text{ and } x \leq 0\}$ (i.e., the set of all negative irrational numbers).

(b) $\{x \in \mathbb{R} | x \geq 0 \text{ and } x \in \mathbb{Q} \rightarrow x \geq 1\}$. (All real numbers ≥ 1 and all irrational numbers between 0 and 1.)

(c) $\{1\}$.

Exercise 85: First assume $A \subseteq B$. We always have $A \cap B \subseteq A$. Thus we need to show that $A \subseteq A \cap B$. To this end, it suffices to show that each element of A is also an element of $A \cap B$, that is, that it is both in A and in B. The first is trivial, and the second follows from $A \subseteq B$.

Conversely, assume that $A \cap B = A$. Then in particular $A \subseteq A \cap B$, meaning that each element in A is also an element of B. Hence $A \subseteq B$.

Exercise 90: Let $x \in (A - B) \cap (B - A)$. Then $x \in A - B$ and $x \in B - A$. The first means that $x \in A$ but $x \notin B$. The second means that $x \in B$ and $x \notin A$. Contradiction. Hence there are no elements in $(A - B) \cap (B - A)$.

Exercise 92: We have

$$\begin{aligned} u \in \mathcal{P}(A \cap B) &\Leftrightarrow u \subseteq A \cap B \\ &\Leftrightarrow u \subseteq A \text{ and } u \subseteq B \\ &\Leftrightarrow u \in \mathcal{P}(A) \text{ and } u \in \mathcal{P}(B) \\ &\Leftrightarrow u \in \mathcal{P}(A) \cap \mathcal{P}(B). \end{aligned}$$

Lecture IX

Exercise 95: $\{\emptyset\} \times \{\emptyset\} = \{(\emptyset, \emptyset)\}$. (Which, using the definition of ordered pair, is $\{\emptyset, \{\emptyset\}\}$.)

Exercise 96: Yes, it is the product of $\{\emptyset, \{\emptyset\}\}$ with itself.

Exercise 98: We show that $A \times B \neq B \times A$. Take $A = \{\emptyset\}$ and take $B = \{\{\emptyset\}\}$. Then $A \times B = \{(\emptyset, \{\emptyset\})\}$ while $B \times A = \{(\{\emptyset\}, \emptyset)\}$. Since $\emptyset \neq \{\emptyset\}$, these two are distinct. However, in case where $A = \emptyset$ or $B = \emptyset$, then equality does hold.

Exercise 102: Assume $A \subseteq A', B \subseteq B'$. Take $(x, y) \in A \times B$. Then $x \in A, y \in B$. Hence also $x \in A', y \in B'$. This means that $(x, y) \in A' \times B'$, showing that $A \times B \subseteq A' \times B'$.

Exercise 104(a): Consider first an element $(x, y) \in A \times (B \cup C)$. Then $x \in A$ and $y \in B \cup C$.

Case 1: $y \in B$. Then $(x, y) \in A \times B$, and hence also $(x, y) \in (A \times B) \cup (A \times C)$.

Case 2: $y \in C$. Then $(x, y) \in A \times C$, and hence also $(x, y) \in (A \times B) \cup (A \times C)$.

Either way, we have $(x, y) \in (A \times B) \cup (A \times C)$, so that $A \times (B \cup C) \subseteq (A \times B) \cup (A \times C)$.

For the converse, assume $(x, y) \in (A \times B) \cup (A \times C)$.

Case 1: $(x, y) \in A \times B$. Then $x \in A, y \in B$, so also $y \in B \cup C$. Hence $(x, y) \in A \times (B \cup C)$.

Case 2: $(x, y) \in A \times C$. Then $x \in A, y \in C$, so also $y \in B \cup C$. Hence $(x, y) \in A \times (B \cup C)$.

In both cases the result holds, and we conclude $(A \times B) \cup (A \times C) \subseteq A \times (B \cup C)$.

Lecture X

Exercise 109:

(a) $R^\circ = \{(b, a), (d, c), (d, a)\}$

(b) $R \circ R = \emptyset$

(c) $R \circ R^\circ = \{(b, b), (b, d), (d, b), (d, d)\}$

(d) $R^\circ \circ R = \{(a, a), (a, c), (c, c), (c, a)\}$

(e) $S \circ R = \{(a, p), (a, r), (c, p), (c, r)\}$

(f) $S \circ R^\circ = \{(b, p), (b, q), (d, p), (d, q)\}$

Exercise 110: The following relations are those satisfying $R = R^\circ$:

$$R_1 = \emptyset \qquad R_2 = \{(0, 0)\}$$
$$R_3 = \{(1, 1)\} \qquad R_4 = \{(0, 0), (1, 1)\}$$
$$R_5 = \{(0, 1), (1, 0)\} \qquad R_6 = \{(0, 0), (0, 1), (1, 0)\}$$
$$R_7 = \{(1, 1), (0, 1), (1, 0)\} \qquad R_8 = \{(0, 0), (1, 1), (0, 1), (1, 0)\}$$

Exercise 113: Yes, let $A = \{a, b\}$ and $R = S = \{(a, b)\}$. Then $R \circ S = \emptyset$.

Exercise 116: We have $(x, y) \in (< \circ <)$ iff there exists z with $x < y$ and $y < z$. That in turn is the same as saying that $x < y+1$. Thus $< \circ <$ is the relation $\{(x, y) | x < y+1\}$.

Next, $(x, y) \in (< \circ\circ <)$ iff there exists z with $x < z$ and $y < z$. But for any x, y we can always find z with $x < z$ and $y < z$ (take $z = x + y + 1$ for example). Thus the relation $<^\circ \circ <)$ is the maximal relation.

Finally, $(x, y) \in (< \circ <^\circ)$ iff there exists z with $z < x$ and $z < y$. Such z exists precisely when x and y are both non-zero. Thus we see that $< \circ <^\circ$ is the relation $\{(x, y) | x \neq 0, y \neq 0\}$.

Lecture XI

Exercise 119: A counterexample to all three is given by $A = \{0, 1\}$, $R = \{(0, 1), (1, 1)\}$. This is a functional relation from A to itself, but the converse is neither total nor single-valued.

Exercise 121: This relation is the graph of $f(x) = x^3$, and hence functional. Normally the converse of a graph is not functional (see exercise 119) but in this case it is: for every $y \in \mathbb{R}$ there exists x with $y = x^3$, so that the relation is total. What's more, such x is necessarily unique, so the relation is single-valued.

Exercise 123: Every relation $R \subseteq A \times \emptyset$ is necessarily empty, hence a fortiori so is every function. But functions are total, hence A must be empty. (For otherwise given $a \in A$ we should have $b \in \emptyset$ with $f(a) = b$.)

Exercise 126: This is a bijective function: it has inverse $f^{-1}(x) = \sqrt[3]{x}$.

Exercise 128: An affine function is bijective iff it is injective iff it is surjective iff its slope a is non-zero; the value b is irrelevant.

Exercise 130: Suppose $(g \circ f)(x) = (g \circ f)(y)$, that is, $g(f(x)) = g(f(y))$. Then since g is injective, we get $f(x) = f(y)$. Next, since f is injective, we get $x = y$. Thus $g \circ f$ is injective.

Exercise 133: It suffices to show that f^{-1} has an inverse. But its inverse is simply f, as we have $f \circ f^{-1} = 1$ and $f^{-1} \circ f = 1$. (Thus the statement that f^{-1} is the inverse of f is essentially the same statement that f is the inverse of f^{-1}.)

Exercise 136:

$$\begin{array}{ll} a \mapsto a, b \mapsto b, c \mapsto c & a \mapsto a, b \mapsto c, c \mapsto b \\ a \mapsto b, b \mapsto a, c \mapsto c & a \mapsto c, b \mapsto b, c \mapsto a \\ a \mapsto b, b \mapsto c, c \mapsto a & a \mapsto c, b \mapsto a, c \mapsto b \end{array}$$

Exercise 141: Yes, the function $f(x) = x+1$ is injective when regarded as a function from $\mathbb{N}$ to $\mathbb{N}$. The absolute value function on $\mathbb{Z}$ (or on $\mathbb{Q}$ or $\mathbb{R}$) is surjective but not injective.

Lecture XII

Exercise 149: (cf. exercise 136)

$$\begin{array}{ll} a \mapsto 0, b \mapsto 1, c \mapsto 2 & a \mapsto 0, b \mapsto 2, c \mapsto 1 \\ a \mapsto 1, b \mapsto 0, c \mapsto 2 & a \mapsto 2, b \mapsto 1, c \mapsto 0 \\ a \mapsto 1, b \mapsto 2, c \mapsto 0 & a \mapsto 2, b \mapsto 0, c \mapsto 1 \end{array}$$

Exercise 153: $\chi_P : \mathbb{N} \to \{0,1\}$ is given by

$$\chi_P(x) = \begin{cases} 1 & \text{if } x \text{ is prime} \\ 0 & \text{otherwise.} \end{cases}$$

Exercise 154: We need to find the elements of $\{0,1\} \times \{0,1\}$ which are sent by the max function to 1. We have $\max(\mathrm{x}, \mathrm{y}) = 1$ iff $x = 1$ or $y = 1$ (or both). Thus the subset with characteristic function max is $\{(0,1),(1,0),(1,1)\}$.

Exercise 155: Calculate

$$\begin{aligned} \chi_{U \cap V}(x) = 1 &\Leftrightarrow x \in U \cap V \\ &\Leftrightarrow x \in U \text{ and } x \in V \\ &\Leftrightarrow \chi_U(x) = 1 \text{ and } \chi_V(x) = 1 \\ &\Leftrightarrow \chi_U(x)\chi_V(x) = 1 \end{aligned}$$

Exercise 158: Given $a \in A$, we have $r(a) \neq \emptyset$, because there exists $b \in B$ with $R(a,b)$, i.e., $b \in r(a)$. Thus the empty set is not of the form $r(a)$ for any $a \in A$.

Exercise 161: We need to define $\phi : A \times (B+C) \to (A \times B) + (A \times C)$. An element of $A \times (B+C)$ is a pair (x,y) where $x \in A$ and $y \in B+C$. Thus y is itself a pair of the form $y = (z,0)$ with $z \in B$ or $y = (z,1)$ with $z \in C$. On the other hand, an element of the form $(A \times B) + (A \times C)$ is of the form $((x,z),0)$ where $(x,z) \in A \times B$ or $((x,z),1)$ where $(x,z) \in A \times C$. We thus define

$$\phi(x,(z,i)) =_{\text{def}} ((x,z),i) \qquad \text{for } i = 0,1.$$

An inverse for ϕ may be given by

$$\psi((x,z),i) =_{\text{def}} (x,(z,i)) \qquad \text{for } i = 0,1.$$

The remaining verification that ϕ and ψ are inverse to each other is now straightforward.

Exercise 162 (d): Define $\phi : (A \times B)^C \to A^C \times B^C$ by $\phi(f) = (\pi_A \circ f, \pi_B \circ f)$, where $\pi_A : A \times B \to A$ is the first projection and $\pi_B : A \times B \to B$ the second. Define an inverse by $\psi(h,k)(c) = (h(c),k(c))$. Then

$$\psi\phi(f)(c) = \psi(\pi_A f, \pi_B f)(c) = (\pi_A f(c), \pi_B f(c)) = f(c)$$

and

$$\phi\psi(h,k) = (\pi_A \psi(h,k), \pi_B \psi(h,k)),$$

and $\pi_A \psi(h,k)(c) = \pi_A(h(c),k(c)) = h(c)$, and similarly $\pi_B \psi(h,k)(c) = k(c)$.

Exercise 165: We have

$$\mathsf{Rel}(A \times B, C) =_{\text{def}} \mathcal{P}((A \times B) \times C) \cong \mathcal{P}(A \times (B \times C)) =_{\text{def}} \mathsf{Rel}(A, B \times C).$$

The middle isomorphism is a consequence of the fact that $(A \times B) \times C \cong A \times (B \times C)$, together with the fact that $X \cong Y$ implies $\mathcal{P}(X) \cong \mathcal{P}(Y)$.

Lecture XIII

Exercise 170:

(a) $\Delta_A \subseteq R$ means that for all $x \in A$ we have $(x, x) \in R$. This says precisely that R is reflexive.

(b) $\Delta_A \cap R = \emptyset$ means that $(x, x) \notin R$ for all $x \in A$. This says that R is irreflexive.

(c) $R^\circ \subseteq R$ means that for all $x, y \in A$, if yRx then xRy. This means R is symmetric. In general, we have $R \subseteq S$ implies $R^\circ \subseteq S^\circ$. Together with $R^{\circ\circ} = R$ this gives the equivalence with the other conditions.

(d) $R \cap R^\circ \subseteq \Delta_A$ means that $(x, y) \in R$ and $(y, x) \in R$ implies that $(x, y) \in \Delta_A$, i.e., that $x = y$. This says that R is anti-symmetric.

(e) $R \circ R \subseteq R$ means that for all $x, z \in A$, if there is some y with xRy, yRz then xRz. This means that R is symmetric.

Exercise 175: Define $f : \mathbb{Z} \times \mathbb{Z} \to \mathbb{Z}$ by $f(x, y) = x + y$. Then the relation is of the form $(x, y) \sim (u, v) \Leftrightarrow f(x, y) = f(u, v)$, and we proved that this is an equivalence relation. The equivalence class of a point (a, b) is the (integer points on the) line $y = -x + (a + b)$ through (a, b) with slope -1.

Exercise 177: The equivalence class of 0 is $\{0\}$, the equivalence class of 1 is the set of all positive numbers, and the equivalence class of $\{-1\}$ is the set of all negative numbers. Thus there are three equivalence classes, and the quotient $\mathbb{Z}/\sim$ is in bijective correspondence with $\{-1, 0, 1\}$.

Exercise 178 (a): $R \cap S$ is reflexive: given $x \in A$, we have xRx and xSx because R and S are reflexive. Hence $(x, x) \in (R \cap S)$. $R \cap S$ is symmetric: given $(x, y) \in (R \cap S)$, we have xRy and xSy. Because R and S are symmetric, we get yRx and ySx. Hence $(y, x) \in (R \cap S)$. Finally, $R \cap S$ is transitive: given $(x, y) \in (R \cap S)$ and $(y, z) \in (R \cap S)$, we have xRy, yRz, xSy, ySz. By transitivity of R and S, we get xRz and xSz. Thus $(x, z) \in (R \cap S)$ as required.

An alternative proof uses the calculus of relations and exercise 170. Reflexivity is established as follows: $\Delta_A \subseteq R, \Delta_A \subseteq S$, hence $\Delta_A \subseteq R \cap S$. Symmetry is proved by $(R \cap S)^\circ = R^\circ \cap S^\circ = R \cap S$. Finally, transitivity follows from

$$(R \cap S) \circ (R \cap S) \subseteq (R \circ R) \cap (R \circ S) \cap (S \circ R) \cap (S \circ S) = R \cap (R \circ S) \cap (S \circ R) \cap S \subseteq R \cap S.$$

(The first containment follows from a statement about the interaction between composition and intersection, which you may wany to formulate and verify separately.)

Exercise 179: Define the function ϕ by $\phi[x]_R = [x]_S$. We must show that this is well-defined. Consider y with yRx. Then $\phi[y]_R = [y]_S$, so we must show that $[y]_S = [x]_S$. But $R \subseteq S$, so yRx implies ySx, whence $[y]_S = [x]_S$.

Lecture XIV

Exercise 183: First, each set of the form $[n, n+1)$ is non-empty (it contains the point $n + 1/2$ for example). Next, given two sets of the form $[n, n+1)$ and $[m, m+1)$ we either have $n = m$, in which case the two sets are identical, or $n \neq m$, in which case the two sets are disjoint. Finally, given any $x \in \mathbb{R}$, let n be the largest integer with $n \leq x$. Then $x \in [n, n+1)$.

Exercise 184: No, because these sets are not pairwise disjoint, e.g., $(-1, 1) \cap (-2, 2) = (-1, 1) \neq \emptyset$.

Exercise 186: Each $[x]$ corresponds to the set of points on the graph of f which also lie on the horizontal line $y = f(x)$.

Exercise 188: There is exactly one partitioning of $\emptyset$, namely the empty family of subsets of $\emptyset$.

Exercise 189:

(a) $0000, 0001, 0010, 0011, 0100, 0101, 0110, 0111, 1000, 1001, 1010, 1011, 1100, 1101, 1110, 1111$.

(b) s is a permutation of t in this case simply means that s and t have the same number of 0s. Clearly "having the same number of 0s" is an equivalence relation.

(c) There are five equivalence classes, namely

$$\{0000\},$$
$$\{0001, 0010, 0100, 1000\},$$
$$\{0011, 0101, 0110, 1001, 1010, 1100\},$$
$$\{0111, 1011, 1101, 1110\},$$
$$\{1111\}$$

Lecture XV

Exercise 192: The union is $\mathbb{R}$, since for each $x \in \mathbb{R}$ there is an $n \in \mathbb{N}$ with $-n < x < n$, so that $x \in (-n, n)$. The intersection is empty since when $a = 0$, the set $(-a, a)$ is empty.

Exercise 193: The union is $\mathbb{R} - \mathbb{Z}$: each $x \in \mathbb{R}$ is in one of the intervals $(a, a+1)$, except for the integers. The intersection is empty.

Exercise 194: The union is $(-1, 1)$. (When $n = 0$ we get this interval, and the subsequent intervals are smaller.) The intersection is $\{0\}$, since that is the only element common to all intervals of the given form.

Exercise 195: Note first that if $n = 1$ we have $\log_2 n = 0$, so that $[0, \log_2 0) = \emptyset$. Thus the intersection is empty. As $n \to \infty$ we have $\log_2 n \to \infty$, which shows that the union is $[0, \infty)$.

Exercise 197: It is the relation $x \sim y \Leftrightarrow x - y$ divides 2012. This is an equivalence relation which clearly contains R. Now suppose that S is any equivalence relation containing R; then we want to show that S contains $\sim$. So suppose $x \sim y$. Then $x - y = k(2012)$ for some integer k. Assume $k \geq 0$. Then we can prove by induction on k that xSy: if $k = 0$ this follows from reflexivity of S. If we have proved the statement for k, then IH gives $(x, k(2012) \cdot x) \in S$. Since $R \subseteq S$, we also have $(k(2012) \cdot x, (k + 1)(2012) \cdot x) \in S$. Transitivity now gives the desired result. Finally, if $k < 0$, use symmetry of S.

Exercise 199: It is the relation $xRy \Leftrightarrow |x| = |y|$. The relation $\sim$ is symmetric, but not reflexive and also not transitive. However, making it reflexive also makes it transitive. Since R is obtained by adding the diagonal to $\sim$ this shows that R is the smallest equivalence relation containing $\sim$.

Lecture XVI

Exercise 203:

$$\begin{array}{ll} f[U_0] = \emptyset & f^{-1}[V_0] = \emptyset \\ f[U_1] = \{7\} & f^{-1}[V_1] = \emptyset \\ f[U_2] = \{6, 7\} & f^{-1}[V_2] = \{1, 4\} \\ f[U_3] = \{6, 7, 8\} & f^{-1}[V_3] = \{0, 1, 4\} \\ f[U_4] = \{6\} & f^{-1}[V_4] = \{0, 2, 3\} \\ f[U_5] = \{6, 7, 8\} & f^{-1}[V_5] = \{0, 1, 2, 3, 4\} \end{array}$$

Exercise 205: Assume that $U \subseteq U'$. To show $f[U] \subseteq f[U']$, consider an element $y \in f[U]$. By definition of direct image, this means that $y = f(x)$ for some $x \in U$. Since $U \subseteq U'$, this implies that $x \in U'$. Consequently, $f(x) \in f[U']$, again by definition of direct image. We have shown that $x \in f[U]$ implies $x \in f[U']$, which proves the claim.

Exercise 207: A nice solution is $f(x) = x \sin x$. (Sketch the graph to see why for any y-value there are infinitely many x-values with $f(x) = y$.)

Exercise 210: The statement is false: the smallest counterexample is $f : \emptyset \to \{*\}$, the empty function (although the same idea works for any non-surjective function). We have, for $U = \emptyset$, $f[U]^c = \emptyset^c = \{*\}$, while $f[U^c] = f[\emptyset] = \emptyset$.

Exercise 212: For variation, we give a proof using the laws of predicate logic.

$$\begin{array}{rcl} y \in f[U \cup U'] & \equiv & \exists x.[y = f(x) \wedge x \in U \cup U'] \\ & \equiv & \exists x.[y = f(x) \wedge (x \in U \vee x \in U')] \\ & \equiv & \exists x.[(y = f(x) \wedge x \in U) \vee (y = f(x) \wedge x \in U')] \\ & \equiv & \exists x.[y = f(x) \wedge x \in U] \vee \exists x.[y = f(x) \wedge x \in U'] \\ & \equiv & y \in f[U] \vee y \in f[U'] \\ & \equiv & y \in f[U] \cup f[U'] \end{array}$$

Exercise 216: There are two fibres, namely $f^{-1}(p) = \{a, b\}$ and $f^{-1}(q) = \{c\}$. The partitioning is $\{\{a, b\}, \{c\}\}$.

Exercise 217: The fibres are precisely the singleton subsets of A. In fact, this is true not just for the identity but for any bijective function out of A.

Exercise 218: Note first that for $a \in A$ we can describe the fibre over a as

$$\pi_A^{-1}(a) = \{(x,y)|\pi_A(x,y) = a\} = \{(a,b)|b \in B\}.$$

This set is in bijective correspondence with B via the bijection which sends (a,b) to b.

Exercise 219: We get $A = \{(a,1),(b,1),(b,2),(c,2),(d,2),(c,3)\}$. The function f is defined by

$$f(a,1) = f(b,1) = 1, f(b,2) = f(c,2) = f(d,2) = 2, f(c,3) = 3.$$

There are three fibres (because I has three elements): $\{(a,1),(b,1)\}$, $\{(b,2),(c,2),(d,2)\}$ and $\{(c,3)\}$.

Exercise 220: The domain is the set

$$A = \{(n-1,n)|n \in \mathbb{Z}\} \cup \{(n,n)|n \in \mathbb{Z}\} \cup \{(n+1,n)|n \in \mathbb{Z}\}.$$

The function $A \to \mathbb{Z}$ sends (x,y) to y. The fibre over $n \in \mathbb{Z}$ is the set

$$\{(n-1,n),(n,n),(n+1,n)\}.$$

Exercise 222: There are three fibres, because the codomain $\mathbb{Z}/3$ has three elements. The fibre over $[0]$ is the set $\{3k|k \in \mathbb{Z}\}$, the fibre over $[1]$ is $\{3k+1|k \in \mathbb{Z}\}$ and the fibre over $[2]$ is $\{3k+2|k \in \mathbb{Z}\}$.

Lecture XVII

Exercise 225:

- $s(1) = a, s(2) = b, s(3) = c$
- $s(1) = a, s(2) = c, s(3) = c$
- $s(1) = a, s(2) = d, s(3) = c$
- $s(1) = b, s(2) = b, s(3) = c$
- $s(1) = b, s(2) = c, s(3) = c$
- $s(1) = b, s(2) = d, s(3) = c$

Exercise 228: We have $\mathcal{P}_+\{0,1\} = \{\{0\},\{1\},\{0,1\}\}$. Choice functions are uniquely determined on singletons, so we get $s\{0\} = 0, s\{1\} = 1$. We only need to specify the possible values of s at $\{0,1\}$, and thus there are two possibilities:

- $s\{0\} = 0, s\{1\} = 1, s\{0,1\} = 0$
- $s\{0\} = 0, s\{1\} = 1, s\{0,1\} = 1$

Exercise 229:

(a) The empty set has no non-empty subsets, so $\mathcal{P}_+(\emptyset) = \emptyset$. Therefore there is only one choice function, namely the unique function $\mathcal{P}_+(\emptyset) \to \emptyset$.

(b) A singleton set has a unique choice function $\mathcal{P}\{\emptyset\} \to \{\emptyset\}$.

(c) This is the same for any singleton set.

(d) A possible choice function is to set $s(U) =$ the least element of U.

(e) From a subset U of primes we may choose the one closest to 0. (If this doesn't determine a unique element, take the positive one.)

(f) Given a subset U, write all elements of U as $\frac{m}{n}$ where $n > 0$ and n, m relatively prime. Then first take the elements for which n is minimal; next, from these elements, take the element with the largest m.

Exercise 231:

- $s(p) = e, s(q) = a, s(r) = b$
- $s(p) = e, s(q) = a, s(r) = d$
- $s(p) = e, s(q) = c, s(r) = b$
- $s(p) = e, s(q) = c, s(r) = d$

Exercise 232: A bijection has exactly one section, namely its inverse. Clearly this is a section because of $ff^{-1} = 1_B$. However, for any section s of f we have $s = 1_A s = f^{-1}fs = f^{-1}1_B = f^{-1}$.

Exercise 234: We must have $s(n) \in \{n-1, n, n+1\}$ for all $n \in \mathbb{Z}$. There are three obvious choices:

- $s(n) = n - 1$
- $s(n) = n$
- $s(n) = n + 1$.

(Of course there are many others.)

Exercise 235: For each $i \in I$ we must choose an element of $A_i = \{i\}$. But we don't have a choice: we can only take i. Thus $c(i) = i$ is the unique choice function for this family.

Exercise 237: Under the correspondence between relations from A to B and functions $A \to \mathcal{P}(B)$, total relations R correspond to functions $r : A \to \mathcal{P}_+(B)$. Hence if we had a choice function $c : \mathcal{P}_+(B) \to B$, we could form $cr : A \to B$, which is contained in R.

Exercise 239: The obvious option is $s(x) = \arcsin x$, but for any integer k you may take $s(x) = \arcsin x + 2k\pi$, since then $\sin(\arcsin x + 2k\pi) = \sin \arcsin x = x$ for any $x \in [-1.1]$.

Exercise 241: The domain is the disjoint union of the sets A_i. This gives

$$X = \{(a,1),(b,1),(c,1),(d,1),(p,2),(q,2),(c,3),(p,4),(a,4)\}.$$

The function $X \to \bigcup_{i\in I} A_i$ sends (x,l) to x.

One possible choice function is $s(1) = a, s(2) = q, s(3) = c, s(r) = a$. The corresponding section is $m(1) = (a,1), m(2) = (q,2), m(3) = (c,3), m(4) = (a,4)$.

Exercise 242: The obvious choice is to choose, given $[x] \in \mathbb{R}/\sim$, the least real number $y \geq 0$ for which $x \sim y$. (For example, you would choose to represent $[14.3322222]$ by 0.3322222, and so on.) But you can add any integer to this function and still have a choice function.

Lecture XVIII

Exercise 245: Throughout, fix bijections $\phi_0 : A \to A'$ and $\phi_1 : B \to B'$.

(a) $\phi_0 \times \phi_1 : A \times A' \to B \times B'$ is a bijection.

(b) Let $\gamma : A + B \to A' + B'$ be defined by $\gamma(x,i) = (\phi_i(x), i)$. (You may want to spell this out to see why it works.) Then γ is a bijection.

(c) Let $\delta : A^B \to A'^{B'}$ be defined by $\delta(f) = \phi_0 \circ f \circ \phi_1^{-1}$. Then δ is a bijection.

(d) Follows from (c) by taking $B = B' = \{0,1\}$ and $\phi_1 = 1$.

Exercise 249: We have $A = B \cup (A - B)$. Suppose that $A - B$ were countable. Then, since the union of two countable sets is countable, this would imply that A is countable. Contradiction.

Exercise 251: Clearly $|\mathbb{R}| \leq |L|$, because for each $r \in \mathbb{R}$ we have a constant function f_r with value r, and the function $\mathbb{R} \to L$ which sends r to f_r is injective. On the other hand, the function $L \to \mathbb{R} \times \mathbb{R}$ which sends $f(x) = ax + b$ to the pair (a,b) is also injective, so $|L| \leq |\mathbb{R} \times \mathbb{R}| = |\mathbb{R}|$.

Exercise 253: You could use the pairing $\mathbb{N}\times\mathbb{N} \to \mathbb{N}$ here but there is a nicer method. Consider a finite set $A = \{a_1, \ldots, a_k\}$, where we may assume without loss of generality that $a_1 < a_2 < \cdots < a_k$. Consider the number

$$\phi(A) = p_1^{a_1} p_2^{a_2} \cdots p_k^{a_k}$$

where p_i is the i-th prime number. (So for example if $A = \{0,3,6\}$ we get

$$\phi(A) = 2^0 \cdot 3^3 \cdot 5^6$$

and so on.) If $A = \emptyset$ set $\phi(A) = 0$. Then ϕ is an injection from $\mathcal{P}_{fin}(\mathbb{N}) \to \mathbb{N}$ by uniqueness of prime factorization.

Exercise 255: This set is not countable; you can use the same diagonal argument which shows that $\mathcal{P}(\mathbb{N})$ is not countable, replacing arbitrary sequences of 0s and 1s by sequences which have infinitely many 1s.

Exercise 256: If there is a surjection $A \to \mathbb{N}$, we may take a section $s : \mathbb{N} \to A$, showing $|\mathbb{N}| \leq |A|$.

Exercise 258: This is similar to exercise 251: identify a polynomial $f(x) = a_n x^n + \cdots + a_1 x + a_0$ with the element $(a_n, \ldots, a_0) \in \mathbb{R}^{n+1}$ and use $|\mathbb{R}^{n+1}| = |\mathbb{R}|$.

Exercise 260: The equivalence class of a real number a is the set

$$[a] = \{a + k | k \in \mathbb{Z}\}.$$

The function $f : [a] \to \mathbb{Z}$ defined by $f(a + k) = k$ is a bijection, so that $|[a]| = |\mathbb{Z}|$.

The quotient set $\mathbb{R}/\sim$ is in bijection with the half open interval $[0, 1)$ (which we know to be uncountable): the function

$$\phi : [0, 1) \to \mathbb{R}\sim; \qquad \phi(a) = [a]$$

is a bijection. Indeed, if $[a] = [b]$ for $a, b \in [0, 1)$ then since $|a - b| < 1$ we must have $a = b$, so that ϕ is injective. And for any $[x] \in \mathbb{R}/! \sim$, we can write $x = a + k$ with $a \in [0, 1)$ and $k \in \mathbb{Z}$, which shows that ϕ is surjective.

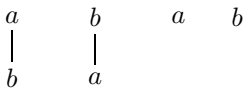
Lecture XIX

Exercise 263:

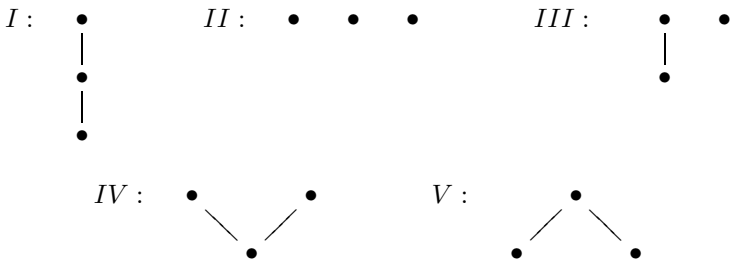
(a) The empty relation is an ordering of the empty set.

(b) There is also a unique ordering relation on a singleton set, namely the diagonal.

(c) For the set $\{a, b\}$ there are three possibilities:



(d) On $\{a, b, c\}$, we have orderings of the following shapes:

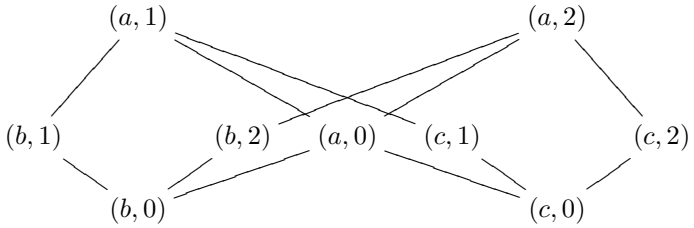


There are then 6 orderings of type I, 1 of type II, 6 of type III, 3 of type IV and 3 of type V. This gives 19 possible orderings in total.

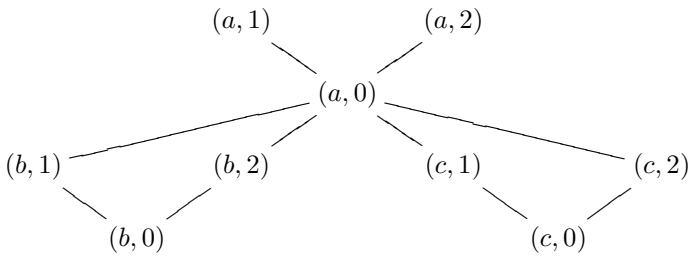
Exercise 265: There is the usual less than or equal-ordering; the discrete ordering $x \leq y$ iff $x = y$; and there is $x \leq y$ iff $x = y$ or $-x = y \in \mathbb{N}$. (There are tons of other possibilities.)

Exercise 267: It is a pre-ordering but not a partial order, since it is possible to have $\phi \vdash \psi$ and $\psi \vdash \phi$ without having $\phi = \psi$.

Exercise 269: The product ordering is



The lexicographic ordering on $A \times B$ is



Exercise 274: Since $f(x) \leq f(x)$ for all x, we get $x \leq x$, so the relation is reflexive. And $x \leq y \leq z$ means $f(x) \leq f(y) \leq f(z)$, which implies $f(x) \leq f(z)$, whence $x \leq z$, so that the relation is transitive.

Note that we cannot expect anti-symmetry: $f(x) = f(y)$ gives $x \leq y$ and $y \leq x$ but unless f is injective this doesn't force $x = y$.

Lecture XX

Exercise 280: If $(A, \leq)$ is linear and finite, then given an element $a \in A$, we either have that a is a least element (in which case we're done) or there is an element a' strictly below a (by linearity). Repeat this for a'; if it is not the least element we get a strictly smaller a'' and so on. If A had no least element we would obtain a strictly descending sequence of elements, which is impossible since A is finite.

Exercise 282: No, the first $10^{100}/4$ terms are increasing, but after that the sequence decreases.

Exercise 284: Consider $\pi = 3.1415\ldots$. Define a sequence of rational numbers by

$$a_0 = 3, a + 1 = 3.1, a_2 = 3.14, a_3 = 3.141, a_4 = 3.1415, \ldots$$

Clearly this sequence increases and converges to π.

Exercise 286: If A, B have least elements $\bot_A, \bot_B$ then $(\bot_A, \bot_B)$ is the least element of $A \times B$ both in the product ordering and in the lexicographic ordering.

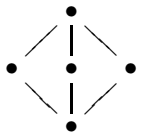
Product orderings are generally not linear even if A, B are (take $\mathbb{N} \times \mathbb{N}$ for example), but the lexicographic ordering is: given any two elements (x, y) and (u, v), we either have $x < u$ (in which case $(x, y) \leq_l (u, v)$) or we have $u < x$ (in which case $(u, v) \leq_l (x, y)$) or $x = y$. In the last case we have $(x, y) \leq_l (u, v)$ or $(u, v) \leq_l (x, y)$ depending on whether $y \leq v$ or $v \leq y$.

Exercise 289: Assume $U \subseteq V$. The element $\bigvee V$ satisfies $x \leq \bigvee V$ for all $x \in V$, hence in particular $x \leq \bigvee V$ for all $x \in U$. But then by definition $\bigvee U \leq \bigvee V$.

Exercise 290: Consider a subset U of $\mathbb{N} \cup \{\infty\}$. If U is a finite set not containing ∞ then the supremum of U is simply the maximal element in U. (And it is 0 if U is empty.). In all other cases, we have $\bigvee U = \infty$.

Exercise 291: Given a, b, the assumption gives that $a \wedge b = a$ or that $a \wedge b = b$. But $a \wedge b = a$ iff $a \leq b$ and $a \wedge b = b$ iff $b \leq a$. Thus $a \leq b$ or $b \leq a$, hence A is linear. The converse holds as well; in a linear ordering we may take $a \wedge b = min(a, b)$ and $a \vee b = max(a, b)$.

Exercise 292: Take



Lecture XXI

Exercise 295: Take any set with more than one element, discretely ordered. This also works for infinite sets.

Exercise 291:

(a) The successor of n is $n + 1$.

(b) In $\mathbb{Q}$ *no* element has a successor, because if $x < y$ we can always consider $\frac{y-x}{2}$, which lies strictly in between.

(c) If x is not maximal, then there is an element y with $x < y$. Such y need not be a successor of x, because there may be a y' with $x < y' < y$. Similarly, such z need not be a successor, because we can have y'' with $x < y'' < y'$. But since A is finite the sequence $y, y', y'', \ldots$ must come to a halt, and therefore x has a successor.

(d) $A = \{a, b, c\}$ with $a \leq b, a \leq c$.

(e) Let x be non-maximal, and let U be the set of all elements strictly above x. Then U is non-empty and therefore has a least element y. Then y must be the successor of x: since $y \in U$ we have $x < y$, and if $x < z$ then $z \in U$, whence $z \leq y$.

(f) In $\mathbb{N} \cup \{\infty\}$ the element ∞ is not a successor of any element.

Index