

# **Discrete Mathematics**

*Solutions with Commentary*

---

Marcelo Fiore

Ohad Kammar

Dima Szamozvancev

2022

# Contents

|       |   |    |
|-------|---|----|
| 1.    | On proofs . . . . .                     | 1  |
| 1.1.  | Basic exercises . . . . .               | 1  |
| 1.2.  | Core exercises . . . . .                | 4  |
| 1.3.  | Optional exercises . . . . .            | 12 |
| 2.    | On numbers . . . . .                    | 15 |
| 2.1.  | Basic exercises . . . . .               | 15 |
| 2.2.  | Core exercises . . . . .                | 20 |
| 2.3.  | Optional exercises . . . . .            | 25 |
| 3.    | More on numbers . . . . .               | 28 |
| 3.1.  | Basic exercises . . . . .               | 28 |
| 3.2.  | Core exercises . . . . .                | 29 |
| 3.3.  | Optional exercises . . . . .            | 40 |
| 4.    | On induction . . . . .                  | 42 |
| 4.1.  | Basic exercises . . . . .               | 42 |
| 4.2.  | Core exercises . . . . .                | 45 |
| 4.3.  | Optional exercises . . . . .            | 57 |
| 5.    | On sets . . . . .                       | 64 |
| 5.1.  | Basic exercises . . . . .               | 64 |
| 5.2.  | Core exercises . . . . .                | 69 |
| 5.3.  | Optional advanced exercises . . . . .   | 76 |
| 6.    | On relations . . . . .                  | 78 |
| 6.1.  | Basic exercises . . . . .               | 78 |
| 6.2.  | Core exercises . . . . .                | 80 |
| 7.    | On partial functions . . . . .          | 85 |
| 7.1.  | Basic exercises . . . . .               | 85 |
| 7.2.  | Core exercises . . . . .                | 87 |
| 8.    | On functions . . . . .                  | 88 |
| 8.1.  | Basic exercises . . . . .               | 88 |
| 8.2.  | Core exercises . . . . .                | 89 |
| 8.3.  | Optional advanced exercise . . . . .    | 90 |
| 9.    | On bijections . . . . .                 | 91 |
| 9.1.  | Basic exercises . . . . .               | 91 |
| 9.2.  | Core exercises . . . . .                | 92 |
| 10.   | On equivalence relations . . . . .      | 95 |
| 10.1. | Basic exercises . . . . .               | 95 |
| 10.2. | Core exercises . . . . .                | 97 |
| 11.   | On surjections and injections . . . . . | 99 |
| 11.1. | Basic exercises . . . . .               | 99 |


---

|       |                                      |     |
|-------|--------------------------------------|-----|
| 11.2. | Core exercises . . . . .             | 99  |
| 12.   | On images . . . . .                  | 101 |
| 12.1. | Basic exercises . . . . .            | 101 |
| 12.2. | Core exercises . . . . .             | 103 |
| 13.   | On countability . . . . .            | 105 |
| 13.1. | Basic exercises . . . . .            | 105 |
| 13.2. | Core exercises . . . . .             | 106 |
| 13.3. | Optional advanced exercise . . . . . | 109 |
| 14.   | On inductive definitions . . . . .   | 109 |
| 15.   | On regular expressions . . . . .     | 113 |
| 16.   | On finite automata . . . . .         | 114 |
| 17.   | On regular languages . . . . .       | 116 |
| 18.   | On the Pumping Lemma . . . . .       | 119 |

# 1. On proofs

## 1.1. Basic exercises

*The main aim is to practice the analysis and understanding of mathematical statements (e.g. by isolating the different components of composite statements), and exercise the art of presenting a logical argument in the form of a clear proof (e.g. by following proof strategies and patterns).*


The solutions will consist of the proper formal proof, showing the required level of detail and precision – you should try to present your answers in a similar form. Of course, a formal written proof is just a polished facade of the (usually more difficult) process of finding the correct sequence of reasoning steps and proof techniques, which often constitutes the “scratchwork”. Therefore, most of the formal proofs will be accompanied with notes (marked ) on how the problem was approached, what guided the reasoning process and what mistakes should be avoided. Mastering the art of formal proof may take some practice, but it is a very important skill to acquire both for this course and your whole scientific education.


Prove or disprove the following statements.

Some fairly simple statements, but they showcase a wide range of proposition types and proof techniques. Accordingly, the proof notes can apply to most of the statements you will encounter, no matter how complicated.

1. Suppose  $n$  is a natural number larger than 2, and  $n$  is not a prime number. Then  $2 \cdot n + 13$  is not a prime number.

The statement is false. Choose  $n = 9$ . Then  $n = 3 \cdot 3$  isn't prime, yet  $2 \cdot n + 13 = 31$  is prime, and we disproved the statement by a counterexample.

 “Prove or disprove” questions should usually start with a sanity check: try a few numbers, and if things seem to work, try a formal proof. Unfortunately, this is not a sure-fire technique, as you may need to backtrack after realising that the statement is false after all. If you realise this in the middle of writing the formal proof (rather than just the scratchwork), you need to cross everything out and start again: there is no space for “plot twists” in a proof attempt, and you should state if the statement is true or false right away.

 To disprove a statement, all we need to present is a counterexample which falsifies it. There is no need to explain the general situation where the statement doesn't work, or try to prove the negation of the statement. The counterexample doesn't have to be very elaborate, often edge cases like 0 or the empty set do the job perfectly. However, we have to make sure that our counterexample falls under the consideration of the statement: 0 will not falsify a proposition that starts with “for every positive integer”.

2. If  $x^2 + y = 13$  and  $y \neq 4$  then  $x \neq 3$ .

We equivalently prove that if  $x^2 + y = 13$  then  $y \neq 4$  implies  $x \neq 3$ . Assume that  $x^2 + y = 13$ . We establish the contrapositive of the goal, i.e. if  $x = 3$  then  $y = 4$ . Indeed,

assume  $x = 3$ . Then,  $y = 13 - x^2 = 13 - 9 = 4$ , as required.

♪ The statement is of the form  $(P \wedge Q) \Rightarrow R$ , so our proof algorithm dictates that we should start by assuming  $P$  and  $Q$ , and prove  $R$ . However, in this case,  $Q$  and  $R$  are *negative* assertions: knowing that  $y$  could be anything other than 4 is less useful than knowing that it is equal to something. Since the consequent  $R$  is also negated, we may instead consider proving a contrapositive. However, the contrapositive of  $(P \wedge Q) \Rightarrow R$  is  $\neg R \Rightarrow \neg(P \wedge Q)$ , which will turn our useful assumption  $P$  into a negated goal. Instead, we perform a *partial assumption*: we assume  $P$  only, and prove  $Q \Rightarrow R$ ; this can now be done easily, since the contrapositive will give us the additional assumption  $x = 3$  and the goal will be  $y = 4$ . Logically, this technique follows from the equivalence  $(P \wedge Q) \Rightarrow R \simeq P \Rightarrow (Q \Rightarrow R)$ , which can be seen as “currying” the proposition  $Q$ .

3. For an integer  $n$ ,  $n^2$  is even if and only if  $n$  is even.

( $\Rightarrow$ ) We prove the following more general result: The product of an even integer with any integer is an even integer. The required proposition follows as a corollary.

Consider any two integers  $m, n$  and assume that  $m$  is even. By definition of even integers,  $m = 2k$  for some integer  $k$ . Therefore,  $m \cdot n = 2k \cdot n = 2(k \cdot n)$  and thus, by definition,  $m \cdot n$  is an even integer.

( $\Leftarrow$ ) We prove the contrapositive; i.e.,  $n$  odd implies  $n^2$  odd. Assume  $n$  is odd, then by [Proposition 8](#) (product of odd numbers is odd) of the notes,  $n \cdot n = n^2$  is odd.

♪ As soon as we see the phrase “if and only if”, we need to remember not to move on to the next question halfway through the proof. Most of the time such proofs will consist of two parts, so stating which direction is being proved is helpful for the reader (and a good reminder to you to complete both directions).

♪ Both directions follow as *corollaries* (logical consequences which are important in their own right) of more general statements about multiplying two different even/odd numbers. The advantage of making theorem statements as general as possible is that they can be applied in many contexts and give rise to useful corollaries; we could have proved that the square of an odd number is odd directly, but the underlying proof approach would have been exactly the same as Proposition 8 so we might as well use it!

4. For all real numbers  $x$  and  $y$  there is a real number  $z$  such that  $x + z = y - z$ .

Consider arbitrary real numbers  $x, y$ , and choose  $z = \frac{y-x}{2}$ . Then,  $z$  is a real number satisfying  $x + z = x + \frac{y-x}{2} = \frac{y+x}{2} = y + \frac{y-x}{2} = y - z$ . Therefore, there exists a real number  $z$  satisfying  $x + z = y - z$ .

♪ This is an example of an existence proof, which tend to look a bit backwards when written formally: rather than deriving the witness as part of the proof, we “give away” the answer right away, then show that it satisfies the required property. Of course, we don’t just pluck the witness out of thin air, or resort to a lucky guess. We find what it should be from

the required properties via some sort of calculation, and once we found an answer, we present it as a witness at the very beginning of the formal proof. Showing that it satisfies the properties is of course straightforward (but can't be omitted), since that's how we found the witness in the first place. In this specific example, the answer  $z = \frac{y-x}{2}$  can be found by simple rearrangement of the condition  $x + z = y - z$ , but, to present this as a formal existence proof, we need to state the witness and rigorously demonstrate that it satisfies the requirement.

5. For all integers  $x$  and  $y$  there is an integer  $z$  such that  $x + z = y - z$ .

The statement is false. Indeed, for the integers  $x = 0$  and  $y = 1$  we will prove that there does not exist an integer  $z$  satisfying  $x + z = y - z$ ; i.e., equivalently, such that  $z = 1 - z$ . Assume to the contrary that such an integer, say  $z_0$ , existed. Then, we would have  $2 \cdot z_0 = 1$  and hence  $z_0 = \frac{1}{2}$ ; which is absurd as  $\frac{1}{2}$  is not an integer. Therefore, there are integers  $x$  and  $y$  for which there is no integer  $z$  such that  $x + z = y - z$ .

♪ This proof may seem more verbose than it needs to be: surely we can just say “let  $x = 0$  and  $y = 1$ , then  $z = \frac{y-x}{2} = \frac{1}{2}$  which is not an integer”. The problem with this reasoning is that it is not a nonexistence proof: we showed that the specific  $z$  that can be computed with the method above is not an integer, but that doesn't mean there cannot be any other  $z$  that works. To be completely rigorous, we need to show that the existence of any  $z$  that satisfies the property is a logical absurdity – from this follows the lengthy but airtight proof of the answer.

♪ Note how in two lines we got to a statement that is *obviously* false: that there exists an integer  $z$  such that  $z = 1 - z$ . We need to resist the temptation to take logical shortcuts and appeal to the intuition of the reader to fill in the holes of our argument: anyone with familiarity of basic arithmetic will recognise this as false (just rearrange the equation to get  $z = \frac{1}{2}$ , which is not an integer), but this will not convince someone who reasons purely by logic (and, unfortunately, supervisors and examiners are such people). As explained in the previous point, the easiest logically rigorous way to show that such an integer  $z$  does not exist is by contradiction.

6. The addition of two rational numbers is a rational number.

Consider any two rational numbers  $r, s$ . By definition, there exist some integers  $a, c$  and some nonzero integers  $b, d$  such that  $r = \frac{a}{b}$  and  $s = \frac{c}{d}$ . Then,  $r + s = \frac{a \cdot d + b \cdot c}{b \cdot d}$  is a quotient of an integer (namely  $a \cdot d + b \cdot c$ ) by a nonzero integer (namely  $b \cdot d$ ), and hence a rational number.

♪ A large part of writing formal proofs is just expanding definitions: rather than trying to reason about rational numbers, we use their formal definition to transition into a proof about integers. The more abstract the statement (quite common in set theory), the more layers of definitions we may need to unwrap. However, this can allow us to prove some rather difficult-looking propositions with a very simple, low-level reasoning step.

7. For every real number  $x$ , if  $x \neq 2$  then there is a unique real number  $y$  such that  $2 \cdot y / (y + 1) = x$ .

We need to show that for every real number  $x$ , if  $x \neq 2$  then there exists a real number  $y$  satisfying: ①  $\frac{2y}{y+1} = x$  and ② for all real numbers  $z$ , if  $\frac{2z}{z+1} = x$  then  $y = z$ .

Consider an arbitrary real number  $x$ , and assume  $x \neq 2$ . Then,  $y = \frac{x}{2-x}$  is a real number satisfying ①, and if  $z$  is any real number satisfying  $\frac{2z}{z+1} = x$  then  $2 \cdot z = z \cdot x + x$ . Hence,  $(2 - x) \cdot z = x$ . As  $x \neq 2$ ,  $z = \frac{x}{2-x} = y$ .

♪ This is a unique existence proof, so requires two separate arguments: existence and uniqueness. The standard way of proving uniqueness is to assume another value with the same property, and show that it must be equal to the existing witness. Uniqueness may seem like a relatively unimportant result, but in fact, it forms the basis of powerful proof techniques which we'll see later on.

8. For all integers  $m$  and  $n$ , if  $m \cdot n$  is even, then either  $m$  is even or  $n$  is even.

One may prove the contrapositive of the statement; i.e. that if  $m$  and  $n$  are odd then  $m \cdot n$  is odd. But this is nothing but [Proposition 8](#) of the notes.

♪ Negation-based proof techniques (contradiction or contraposition) are often used to avoid awkward proof patterns, usually involving existence or disjunction. Rather than have a disjunctive goal (which requires some sort of case-splitting), we negate it to turn (via the de Morgan laws) into a conjunctive assumption.

## 1.2. Core exercises

*Having practised how to analyse and understand basic mathematical statements and clearly present their proofs, the aim is to get familiar with the basics of divisibility.*

1. Characterise those integers  $d$  and  $n$  such that:

a)  $0 \mid n$

We prove that an integer  $n$  satisfies  $0 \mid n$  iff  $n = 0$ .

( $\Rightarrow$ ) Assume  $0 \mid n$ . By definition, for some integer  $l$ ,  $n = l \cdot 0 = 0$ .

( $\Leftarrow$ ) Assume  $n = 0$ . Then,  $n = 0 \cdot 0$  and, by definition,  $0 \mid n$ .

♪ A good example of the need to be precise when applying definitions. We may intuitively interpret  $d \mid n$  (" $d$  divides  $n$ ") as " $\frac{n}{d}$  is an integer", and conclude that  $0 \mid n$  is impossible because  $\frac{n}{0}$  is undefined. However, the formal definition of  $d \mid n$  makes no mention of the division operator: it is an algebraically more fundamental concept which only requires multiplication to express. Strictly speaking, we haven't yet formally defined division in the course – sure, you *know* what division is from school, but giving a precise and rigorous definition is more difficult than it may seem! If we use the proper definition of divisibility for this exercise, we do find an appropriate value for  $n$ , namely 0: zero divides zero because there exists an integer  $l$  (any integer

will work) such that  $0 = l \cdot 0$ .

b)  $d \mid 0$

We prove that  $d \mid 0$  for all integers  $d$ . Indeed, let  $d$  be an arbitrary integer. Then,  $0 = 0 \cdot d$  and hence  $d \mid 0$ .

2. Let  $k, m, n$  be integers with  $k$  positive. Show that:

$$(k \cdot m) \mid (k \cdot n) \iff m \mid n$$

Consider any positive integer  $k$  and any two integers  $m, n$ .

( $\Rightarrow$ ) Assume  $(k \cdot m) \mid (k \cdot n)$ . Then,  $k \cdot n = l \cdot (k \cdot m)$ . As  $k > 0$ , we can cancel  $k$  and deduce  $n = l \cdot m$ . Hence,  $m \mid n$ .

( $\Leftarrow$ ) Assume  $m \mid n$ . Then,  $n = a \cdot m$  for some integer  $a$ ; and multiplying by  $k$ , we have  $k \cdot n = a \cdot (k \cdot m)$ . Hence,  $(k \cdot m) \mid (k \cdot n)$ .

🎵 “Cancelling things” on both sides of an equation is a very standard process in elementary (“high-school”) algebra. While in many cases it is still allowed in this course, you should pay extra attention to any side-conditions required for the cancellation, or if cancellation is even possible for the algebraic structure you’re working with! It does hold for addition and multiplication (with a side-condition), but, for example, an equation between function composites  $f \circ g = f \circ e$  cannot be simplified to  $g = e$  in general (only if  $f$  is an *injection* – see later). Cancellability may be a property of the structure, or particular elements in a structure, rather than something you can just do arbitrarily.

🎵 The ( $\Rightarrow$ ) direction of this proof relied on the fact that  $k$  is positive, and in particular, nonzero – otherwise we wouldn’t be able to cancel the  $k$ s (and the property wouldn’t actually hold). We did not require any assumptions on  $k$  in the ( $\Leftarrow$ ) direction, so we could extract a weaker form of the theorem, stating that for every integer  $k, m$  and  $n$ ,

$$m \mid n \implies (k \cdot m) \mid (k \cdot n)$$

In some cases you may not have the assumption that  $k$  is positive but may still be able to apply this weaker form to make progress. However, this is technically not a corollary of the stronger statement, because that requires an unneeded assumption on  $k$ .

3. Prove or disprove that: For all natural numbers  $n$ ,  $2 \mid 2^n$ .

This is false, as  $2 \nmid 2^0$ .


🎵 This is just a gentle reminder that 0 is a natural number!

4. Show that for all integers  $l, m, n$ ,

$$l \mid m \wedge m \mid n \implies l \mid n$$




Consider any integers  $l, m, n$ , and assume  $l \mid m \wedge m \mid k$ . As  $l \mid m$ ,  $m = a \cdot l$  for some integer  $a$ . As  $m \mid n$ ,  $n = b \cdot m$  for some integer  $b$ . But then:  $n = b \cdot m = b \cdot (a \cdot l) = (b \cdot a) \cdot l$  and, as  $b \cdot a$  is an integer, we have  $l \mid n$ .

 An example of a proof which is not particularly difficult or illuminating, but it's still presented in a clear, structured, formal manner. It should take about a line of scratchwork to convince yourself that the statement is true, but that is only the first step: next, you need to convince the reader of the proof, who may not find your sketch clear or rigorous enough. Learning how to present even the simplest arguments in a formal, systematic manner will massively aid you in tackling more difficult propositions which may seem very daunting at first, but are actually much easier to digest connective-by-connective, definition-by-definition.

5. Find a counterexample to the statement: For all positive integers  $k, m, n$ ,

$$(m \mid k \wedge n \mid k) \implies (m \cdot n) \mid k$$

Choose  $k = m = n = 2$ . Then,  $k, m, n$  are positive integers. As  $2 \mid 2$ , we have  $m \mid k \wedge n \mid k$  yet  $(2 \cdot 2) \nmid 2$ .

 While questions like this don't explicitly ask for it, you need to find a counterexample and also show that it is a counterexample, i.e. that it contradicts the statement. Only writing  $k = m = n = 2$  is not enough; you need to justify your answer.

6. Prove that for all integers  $d, k, l, m, n$ ,

a)  $d \mid m \wedge d \mid n \implies d \mid (m + n)$


Assume  $d \mid m \wedge d \mid n$ . As  $d \mid m$ ,  $m = a \cdot d$  for some integer  $a$ . As  $d \mid n$ ,  $n = b \cdot d$  for some integer  $b$ . Therefore,  $m + n = a \cdot d + b \cdot d = (a + b) \cdot d$ . As  $a + b$  is an integer, we have  $d \mid (m + n)$  as required.

b)  $d \mid m \implies d \mid k \cdot m$

Assume  $d \mid m$ ; i.e.  $m = a \cdot d$  for some integer  $a$ . Then,  $k \cdot m = k \cdot (a \cdot d) = (k \cdot a) \cdot d$ . As  $k \cdot a$  is an integer,  $d \mid (k \cdot m)$ .

c)  $d \mid m \wedge d \mid n \implies d \mid (k \cdot m + l \cdot n)$

Assume  $d \mid m \wedge d \mid n$ . As  $d \mid m$ , by 6(b) above,  $d \mid (k \cdot m)$ . Analogously, from  $d \mid n$  we have  $d \mid (l \cdot n)$ . Thus,  $d \mid (k \cdot m) \wedge d \mid (l \cdot n)$  so that applying 6(a) we conclude  $d \mid (k \cdot m + l \cdot n)$  as required.

 Science is about building on the shoulders of giants – even if that giant is us, ten minutes ago. After proving two useful properties of divisibility in parts 6(a) and 6(b), they are now part of our “knowledge base” and we can refer back to them freely, without having to reprove them again.

Mathematics and computer science are all about decomposition and composition (also known as divide-and-conquer). Faced with a complicated proposition/problem, we break it up into smaller components which are much easier to reason about. Then, we find ways to solve the subproblems: prove lemmas and sub-theorems or write functions, classes and methods to perform well-defined tasks. Finally, we combine the sub-solutions and reap the rewards. In practice, however, the challenge is not always in solving the subproblems from scratch, but figuring out which existing elements of the knowledge base/programming library can be glued together to give the desired results: after all, if we or someone else has solved some difficult problem already, we shouldn't need to do it again! There may be a striking one-liner proof/program that does the job, but finding it may take significantly more effort than just solving the problem manually. But, seeing how programmers can spend hours finding the shortest, simplest, fastest, most space-efficient algorithms, there is a lot of enjoyment to be had in crafting concise and elegant proofs that combine clever reasoning techniques with existing propositions in satisfying ways. We will hopefully see examples of this in the course so you can appreciate proof-writing not as a chore, but something intellectually stimulating and often quite addictive!

7. Prove that for all integers  $n$ ,


$$30 \mid n \iff (2 \mid n \wedge 3 \mid n \wedge 5 \mid n)$$

( $\Rightarrow$ ) Assume  $30 \mid n$ . Then,  $n = 30 \cdot a$  for some integer  $a$ . Thus,  $n = 2 \cdot (15 \cdot a)$  and so  $2 \mid n$ . Similarly,  $n = 3 \cdot (10 \cdot a)$  and therefore  $3 \mid n$ . And, as  $n = 5 \cdot (6 \cdot a)$ , we also deduce  $5 \mid n$ . Therefore  $2 \mid n \wedge 3 \mid n \wedge 5 \mid n$ .

( $\Leftarrow$ ) Assume  $2 \mid n \wedge 3 \mid n \wedge 5 \mid n$ . As  $2 \mid n$  and  $3 \mid n$  and  $5 \mid n$ , we have  $n = 2 \cdot a$  and  $n = 3 \cdot b$  and  $n = 5 \cdot c$  for some integers  $a, b, c$ . Moreover, we have:

$$30 \cdot (-a + b + c) = (-15) \cdot 2 \cdot a + 10 \cdot 3 \cdot b + 6 \cdot 5 \cdot c = (-15) \cdot n + 10 \cdot n + 6 \cdot n = n$$

Thus,  $n = 30 \cdot k$  for the integer  $k = -a + b + c$ , as required.

 The ( $\Leftarrow$ ) direction of this proof is more subtle than it may look – we can't just multiply 2, 3 and 5 together (see §1.2.5 above). Instead, we know that  $30 \mid 30a$ , so  $30 \mid 15n$ ; similarly,  $30 \mid 10n$  and  $30 \mid 6n$ . We need to put these together to get  $30 \mid n$ , for which we make use of §1.2.6(a) above to find a linear combination of  $15n$ ,  $10n$  and  $6n$  that adds up to  $n$ . After some thinking, we find that  $(-1) \times 15 + 10 + 6$  works, giving us the desired coefficients of  $a$ ,  $b$  and  $c$ .

8. Show that for all integers  $m$  and  $n$ ,

$$(m \mid n \wedge n \mid m) \implies (m = n \vee m = -n)$$

Consider any pair of integers  $m, n$ , and assume that  $m \mid n$  and that  $n \mid m$ . If  $m = 0$  then, by §1.2.1(a) above,  $n = 0$  and we have  $m = n$ .

Consider henceforth the case  $m \neq 0$ . As  $m \mid n$  and  $n \mid m$ , there are integers  $a, b$  such that  $n = a \cdot m$  and  $m = b \cdot n$ . Thus,  $m = b \cdot a \cdot m$  and, as  $m \neq 0$ , we have  $b \cdot a = 1$ . Then, since  $a$  and  $b$  are integers, either  $a = b = 1$  or  $a = b = -1$  (otherwise, one would have  $a \cdot b \geq 2$  or  $a \cdot b \leq -2$ ). Finally, if  $a = b = 1$  then  $m = n$ , and if  $a = b = -1$  then  $m = -n$ . Either way,  $m = n$  or  $m = -n$  as required.

♪ You may have started the proof from the second paragraph, without assuming that  $m \neq 0$ . Then, at the step  $m = b \cdot a \cdot m$ , you would be stuck (if you're being careful): you can't divide by  $m$  because it may be 0. In such cases a common solution is to handle the problematic case ( $m = 0$ ) separately, then have the desired extra assumption  $m \neq 0$  in the main proof and continue from there.

9. Prove or disprove that: For all positive integers  $k, m, n$ ,

$$k \mid (m \cdot n) \implies k \mid m \vee k \mid n$$

We disprove it by means of a counterexample. Choose  $m = n = 2$  and  $k = 4$ . Then  $k \mid m \cdot n$ , yet neither  $k \mid m$  nor  $k \mid n$ .

♪ It may sometimes be easier to disprove the contrapositive statement, since an implication holds if and only if its contrapositive holds.

10. Let  $P(m)$  be a statement for  $m$  ranging over the natural numbers, and consider the following derived statement (with  $n$  also ranging over the natural numbers):

$$P^\#(n) \triangleq \forall k \in \mathbb{N}. 0 \leq k \leq n \implies P(k)$$

- a) Show that, for all natural numbers  $\ell$ ,  $P^\#(\ell) \implies P(\ell)$ .

Let  $\ell$  be a natural number, and assume that

$$P^\#(\ell) = (\forall \text{ natural number } k. 0 \leq k \leq \ell \implies P(k))$$

holds.

Since  $\ell$  is a natural number, it follows by instantiation that

$$0 \leq \ell \leq \ell \implies P(\ell)$$

and, since  $0 \leq \ell \leq \ell$  is true by reflexivity of  $\leq$ , it follows by Modus Ponens that  $P(\ell)$  holds as required.

♪ This last exercise starts to trip some students up, understandably: so far we've been proving properties about numbers and divisibility, while now we're proving things about seemingly nothing in particular. Such abstract proofs are very common

in mathematics, for the very simple reason that they can be applied in a huge number of ways – in this case,  $P(m)$  can be *any* logical statement about natural numbers, and the propositions will hold no matter how simple or complicated the definition of  $P$  is! You may think of this as a “polymorphic” theorem, since we are proving something about an arbitrary predicate  $P$  (any first-class function `nat -> bool`, if you will), but as a consequence, we cannot assume anything about how it’s defined.

Thinking abstractly takes some getting used to, as you may feel like there isn’t anything to go on or any familiar notion to grasp in order to build intuition. However, abstractness has the major benefit of avoiding any distracting details and low-level “fluff” that could lead the proof attempt astray. If the above proposition was specialised to  $P(m)$  meaning “ $m$  is even”, you might start by unwrapping the definition of evenness and incorporate it into the proof somehow, despite the property holding no matter what  $P(m)$  actually is. Abstract proofs like this often involve purely logical reasoning, without invoking any number theory or algebra – and logical reasoning is often easier, since we essentially have an algorithm for proving logical statements. Thus, when you are faced with an incomprehensible jumble of logical symbols, the task may well be easier than proving a simple statement about natural numbers!

- b) Exhibit a concrete statement  $P(m)$  and a specific natural number  $n$  for which the following statement *does not* hold:

$$P(n) \implies P^\#(n)$$

Let  $P(m) \triangleq (m = 1)$  and  $n = 1$ . Then  $P(1)$  is the true proposition  $(1 = 1)$ , but  $P^\#(1) \iff P(0) \wedge P(1)$  is equivalent to  $(0 = 1) \wedge (1 = 1)$  which is false.

♪ Here we actually needed to “decode” the definition of  $P^\#$  in order to find a way to falsify the above statement. Fortunately this is not too difficult in this case:  $P^\#(n)$  holds if  $P(k)$  holds for all naturals less than or equal to  $n$ , essentially turning a predicate  $P$  about naturals into a predicate  $P^\#$  about a finite collection of naturals (similarly to how `map` turns a function on values into a function on lists of values). Then, we need to find a predicate  $P$  and  $n \in \mathbb{N}$  that does not satisfy  $P(n) \implies P^\#(n)$ . This is trickier than just finding a number, since there are lots of ways we could define  $P$ . But, once again, we try something very simple ( $P(m)$  holds for  $m = 1$  only) and find that it can easily be turned into a counterexample. There are lots of other options for  $P$  of course, but there’s no need to try something convoluted or interesting to get a contradiction (and equally, there’s no need to spend time finding the *simplest* counterexample if you’ve already found a more complicated one).

- c) Prove the following:

- $P^\#(0) \iff P(0)$

( $\Rightarrow$ ) Assume  $P^\#(0)$ ; that is, for all  $0 \leq k \leq 0$ ,  $P(k)$ . As  $0 \leq 0 \leq 0$ ,  $P(0)$  holds.  
 ( $\Leftarrow$ ) Assume  $P(0)$ . Consider any  $k$ , and assume  $0 \leq k \leq 0$ . Then,  $k = 0$  and  $P(k)$  holds by assumption.

$$\bullet \forall n \in \mathbb{N}. (P^\#(n) \Rightarrow P^\#(n+1)) \Leftrightarrow (P^\#(n) \Rightarrow P(n+1))$$

( $\Rightarrow$ ) Assume that  $(P^\#(n) \Rightarrow P^\#(n+1))$ , and further assume that  $P^\#(n)$  holds. Then, it follows that also  $P^\#(n+1)$  holds; i.e. that

$$\forall \text{ natural number } k. 0 \leq k \leq n+1 \Rightarrow P(k).$$

In particular, by instantiation, we have that

$$0 \leq n+1 \leq n+1 \Rightarrow P(n+1)$$

and since the antecedent of this implication is true, we deduce that  $P(n+1)$  holds, as required.


( $\Leftarrow$ ) Assume that ①  $(P^\#(n) \Rightarrow P(n+1))$ , and further assume that ②  $P^\#(n)$  holds. We need show that  $P^\#(n+1)$  also holds; i.e. that

$$\forall \text{ natural number } k. 0 \leq k \leq n+1 \Rightarrow P(k).$$

or, equivalently, that

$$P^\#(n) \wedge P(n+1)$$

hold, which is indeed the case because  $P^\#(n)$  holds by assumption ② and  $P(n+1)$  follows by Modus Ponens from assumptions ① and ②.

 This is the most complicated statement in this exercise sheet, so do not worry if had difficulties with it. We have universal quantification, bi-implication, nested implication, and unwrapping the definition of  $P^\#$  gives another layer of quantification and implication. You may take a minute to get a feel for what the statement is saying, but the nice thing about purely logical proofs is that you can often dive in head-first without really thinking about what you're proving!

Look at the top-level construct (universal quantification, bi-implication, etc.), apply the proof pattern for that construct (often giving you some assumptions), and continue until your goal becomes some atomic statement like  $P(n+1)$  (which you can't unwrap further, since you don't know what  $P$  is). After "digesting" the proof goal, you should have a bunch of assumptions that you can work with: unwrapping some definitions, instantiating universals, applying Modus Ponens. Eventually you should end up with an assumption that matches the atomic proof goal, and that's enough to conclude the proof.

At first, doing the "assume | prove"-style scratchwork is very helpful for practicing proof patterns and keeping track of goals and assumptions. It should mostly

feel like an algorithmic process: with a few rule applications you can turn a very scary-looking formula into a primitive goal and a lot of assumptions to work with, and the task usually boils down to finding a way of combining assumptions on the LHS to get something that matches the RHS. Occasionally there is a small bit of actual “thinking” required to make progress, such as finding an appropriate value to instantiate a  $\forall$  with, or transforming an assumption in some useful way: in the ( $\Leftarrow$ ) direction above, really the only clever bit was figuring out that

$$\forall \text{ natural number } k. 0 \leq k \leq n + 1 \implies P(k).$$

is equivalent to

$$P^\#(n) \wedge P(n + 1)$$

but even this step was not made in a vacuum, since we already had the assumptions  $P^\#(n)$  and  $P(n + 1)$ . With some practice these methodical proofs should become second nature, and you will be able to keep track of things in your head, directly writing down the formal proof without any prior scratchwork.

•  $(\forall m \in \mathbb{N}. P^\#(m)) \iff (\forall m \in \mathbb{N}. P(m))$

( $\Rightarrow$ ) Assume that  $\forall$  natural number  $m. P^\#(m)$ , and let  $n$  be an arbitrary natural number. Then, by assumption,  $P^\#(n)$  holds; that is

$$\forall \text{ natural number } k. 0 \leq k \leq n \implies P(k)$$

and, by instantiation,  $0 \leq n \leq n \implies P(n)$  so that  $P(n)$  holds. Thus, we have shown

$$\forall \text{ natural number } m. P(m)$$

( $\Leftarrow$ ) Assume that  $\textcircled{1} \forall$  natural number  $m. P(m)$ . We need show that for all natural numbers  $m$  and  $k$ ,

$$0 \leq k \leq m \implies P(k)$$

To this end, let  $m$  and  $k$  be arbitrary natural numbers, and assume  $0 \leq k \leq m$ . Since  $k$  is a natural number, we may instantiate assumption  $\textcircled{1}$  with it yielding  $P(k)$  as required.

$\square$  This theorem may seem both surprising and unsurprising. Even though  $P^\#(m)$  is definitely more general than  $P(m)$  (since  $P^\#(m)$  implies  $P(m)$  but not vice versa), in the “limit” of quantifying over *all* natural numbers, they become equivalent. Then again, if  $P(m)$  holds for all natural numbers  $m$ , of course it would hold for all natural numbers smaller than any  $n$ ! This theorem (and the properties proved as part of this exercise) form the basis of an important proof technique which will be discussed later in the course.

### 1.3. Optional exercises

1. A series of questions about the properties and relationship of triangular and square numbers (adapted from David Burton).

- a) A natural number is said to be *triangular* if it is of the form  $\sum_{i=0}^k i = 0 + 1 + \dots + k$ , for some natural  $k$ . For example, the first three triangular numbers are  $t_0 = 0$ ,  $t_1 = 1$  and  $t_2 = 3$ .

Find the next three triangular numbers  $t_3$ ,  $t_4$  and  $t_5$ .

$$t_3 = 6, t_4 = 10, t_5 = 15.$$

- b) Find a formula for the  $k^{\text{th}}$  triangular number  $t_k$ .

**Geometric approach.**

$$2 \cdot t_k = \begin{array}{c} \circ \\ \circ \quad \circ \\ \vdots \\ \circ \quad \dots \quad \circ \end{array} + \begin{array}{c} \bullet \quad \dots \quad \bullet \\ \vdots \\ \bullet \quad \bullet \\ \bullet \end{array} = \begin{array}{c} \circ \quad \bullet \quad \bullet \quad \dots \quad \bullet \\ \circ \quad \circ \quad \bullet \quad \dots \quad \bullet \\ \vdots \\ \circ \quad \dots \quad \circ \quad \bullet \quad \bullet \\ \circ \quad \dots \quad \circ \quad \circ \quad \bullet \end{array} = k \cdot (k+1)$$

**Algebraic approach.**

Note that, on the one hand,

$$\begin{aligned} \sum_{i=0}^k (i+1)^2 - \sum_{i=0}^k i^2 &= (k+1)^2 + \left( \sum_{i=0}^{k-1} (i+1)^2 \right) - \left( \sum_{i=1}^k i^2 \right) - 0^2 \\ &= (k+1)^2 \end{aligned}$$

and that, on the other,

$$\begin{aligned} \sum_{i=0}^k (i+1)^2 - \sum_{i=0}^k i^2 &= \sum_{i=0}^k ((i+1)^2 - i^2) \\ &= \sum_{i=0}^k (2 \cdot i + 1) \\ &= \left( 2 \cdot \sum_{i=0}^k i \right) + \sum_{i=0}^k 1 \\ &= 2 \cdot t_k + k + 1 \end{aligned}$$

$$\text{so that } t_k = \frac{k^2+k}{2}.$$

- c) A natural number is said to be *square* if it is of the form  $k^2$  for some natural number  $k$ .

Show that  $n$  is triangular iff  $8 \cdot n + 1$  is a square. (Plutarch, circ. 100BC)

( $\Rightarrow$ ) Assume  $n$  is triangular; i.e.  $n = t_k$  for some natural number  $k$ . By the previous item,  $n = \frac{k \cdot (k+1)}{2}$  and one has that  $8 \cdot n + 1 = (2 \cdot k + 1)^2$  is a square number.

( $\Leftarrow$ ) Assume that  $8 \cdot n + 1$  is a square number; i.e.  $8 \cdot n + 1 = a^2$  for some natural number  $a$ . Then  $a^2$  is odd and, by [Proposition 12](#) of the notes, thus so is  $a$ . Therefore,  $a = 2 \cdot k + 1$  for some natural number  $k$ . Finally, since  $8 \cdot n + 1 = a^2 = (2 \cdot k + 1)^2 = 4 \cdot k^2 + 4 \cdot k + 1$

one has  $n = \frac{k^2+k}{2} = t_k$  as required.

- d) Show that the sum of every two consecutive triangular numbers is square. (Nicomachus, circ. 100BC)

Consider any two consecutive triangular numbers  $t_k$  and  $t_{k+1}$ . Then, a simple calculation shows that the sum  $t_k + t_{k+1}$  equals  $(k+1)^2$  and hence is square:

$$\frac{k^2+k}{2} + \frac{(k+1)^2+k+1}{2} = \frac{2k^2+4k+2}{2} = k^2+2k+1 = (k+1)^2$$

- e) Show that, for all natural numbers  $n$ , if  $n$  is triangular, then so are  $9 \cdot n + 1$ ,  $25 \cdot n + 3$ ,  $49 \cdot n + 6$  and  $81 \cdot n + 10$ . (Euler, 1775)

Consider any natural number  $n$ , and assume that  $n$  is triangular; i.e.  $n = \frac{k \cdot (k+1)}{2}$  for some natural number  $k$ . Then, calculate that  $9 \cdot n + 1 = t_{3k+1}$ :

$$9 \frac{k^2+k}{2} + 1 = \frac{9k^2+9k+2}{2} = \frac{9k^2+6k+1+3k+1}{2} = \frac{(3k+1)^2+3k+1}{2} = t_{3k+1}$$

Similarly, by completing the square, we can show that  $25 \cdot n + 3 = t_{5k+2}$ ,  $49 \cdot n + 6 = t_{7k+3}$ , and  $81n + 10 = t_{9k+4}$ .

- f) Prove the generalisation: For all  $n$  and  $k$  natural numbers, there exists a natural number  $q$  such that  $(2n+1)^2 \cdot t_k + t_n = t_q$ . (Jordan, 1991, attributed to Euler)

Here's a proof by a 2014/15 student (who wished to remain anonymous). Let  $n$  and  $k$  be arbitrary natural numbers. We know that:

$$t_k = \frac{k(k+1)}{2} \quad \text{and} \quad t_n = \frac{n(n+1)}{2}$$

Choose  $q = 2nk + n + k$ , and calculate:

$$\begin{aligned} t_q &= \frac{q(q+1)}{2} = \frac{(2nk+n+k) \cdot (2nk+n+k+1)}{2} \\ &= \frac{4n^2k^2 + 4n^2k + 4nk^2 + 4nk + k^2 + k + n^2 + n}{2} \\ &= \frac{(4n^2+4n+1)(k^2+k) + n^2+n}{2} \\ &= (2n+1)^2 \cdot \frac{k(k+1)}{2} + \frac{n(n+1)}{2} \\ &= (2n+1)^2 t_k + t_n \end{aligned}$$

Therefore we are done.

2. Let  $P(x)$  be a predicate on a variable  $x$  and let  $Q$  be a statement not mentioning  $x$ . Show that the following equivalence holds:

$$\left( (\exists x. P(x)) \implies Q \right) \iff \left( \forall x. (P(x) \implies Q) \right)$$



$(\Rightarrow)$  Assume  $(\exists x. P(x)) \Rightarrow Q$ . We need show  $\forall x. (P(x) \Rightarrow Q)$ . We do this by considering an arbitrary  $a$  and showing that  $P(a) \Rightarrow Q$ , for which in turn we further assume  $P(a)$  and finally show  $Q$ .

To recap, then, we are in the following situation:

| Assumptions  | Goal |
|--|------|
| $(\exists x. P(x)) \Rightarrow Q$<br>for arbitrary $a$<br>$P(a)$ | $Q$  |


Then, by the last assumption,  $\exists x. P(x)$  and from this and the first assumption, by Modus Ponens, we deduce  $Q$  as required.

$(\Leftarrow)$  Assume  $\forall x. (P(x) \Rightarrow Q)$ . We need show  $(\exists x. P(x)) \Rightarrow Q$ . For which we further assume  $\exists x. P(x)$  and show  $Q$

To recap, then, we are in the following situation:

| Assumptions  | Goal |
|--|------|
| $\forall x. (P(x) \Rightarrow Q)$<br>$\exists x. P(x)$ | $Q$  |

From the second assumption, there is an  $a$  for which ①  $P(a)$  holds and, by instantiation from the first assumption, ②  $P(a) \Rightarrow Q$ . By Modus Ponens from ② and ①,  $Q$  follows as required.

 This is a very important duality that crops up in many different forms in mathematics and computer science (and you will certainly encounter variants of it in future courses). Despite this, it may seem quite unintuitive: it almost seems to say that we can convert existential quantification into universal! Of course, we can't ignore the shifting of the parentheses: it's certainly *not* the case that

$$(\exists x. P(x) \Rightarrow Q) \iff (\forall x. P(x) \Rightarrow Q)$$

A good way to get an intuition for this property is as a generalisation of case analysis. If a property  $Q$  depends on the existence of a witness  $x$  satisfying  $P(x)$ , but not  $x$  itself, we need to prove  $Q$  no matter what  $x$  is. That is, our proof must hold for any actual value of the witness, so we can instead look at what possible values can  $x$  take, and show  $Q$  by assuming  $P(x)$  for *all* values  $x$ . We are case analysing the potential values of the witness, and proving  $Q$  no matter what it is.

Alternatively, we can look at the contrapositives of both sides:

$$(\neg Q \Rightarrow \neg(\exists x. P(x))) \iff (\forall x. (\neg Q \Rightarrow \neg P(x)))$$

The LHS becomes  $\neg Q \Rightarrow (\forall x. \neg P(x))$  using the de Morgan rule for quantifiers; but now

the universal can be extended over the whole implication, since assuming  $\neg Q$  first and then taking an arbitrary  $x$  is the same as taking an arbitrary  $x$  and then assuming  $\neg Q$  (which doesn't say anything about  $x$ ).

## 2. On numbers

### 2.1. Basic exercises

1. Let  $i, j$  be integers and let  $m, n$  be positive integers. Show that:

a)  $i \equiv i \pmod{m}$

By §1.2.1(b), every number divides  $i - i = 0$ , so  $m \mid i - i$ .

b)  $i \equiv j \pmod{m} \implies j \equiv i \pmod{m}$

Assume  $i \equiv j \pmod{m}$ . Then  $m \mid i - j$ ; i.e.  $i - j = k \cdot m$  for some integer  $k$ . Thus,  $j - i = (-k) \cdot m$ , and as  $-k$  is an integer  $m \mid j - i$ ; i.e.  $j \equiv i \pmod{m}$ .

c)  $i \equiv j \pmod{m} \wedge j \equiv k \pmod{m} \implies i \equiv k \pmod{m}$

Assume  $i \equiv j \pmod{m} \wedge j \equiv k \pmod{m}$ . Then,  $m \mid i - j$  and  $m \mid j - k$ . Hence, by §1.2.6(a),  $m \mid (i - j) + j - k = i - k$  and thus  $i \equiv k \pmod{m}$ .

When working with congruence, we have three layers of definitions:  $i \equiv j \pmod{m}$ , defined as  $m \mid i - j$ , defined as  $\exists k \in \mathbb{Z}. i - j = k \cdot m$ . To prove fundamental properties about congruence (symmetry or transitivity), we usually need to go “down a level” and reason about divisibility. At this level, we may be able to use known properties of divisibility, such as in part (c); other times it may be easier to go further down, and talk about the primitive definition of divisibility, such as in part (b). In the second case we are essentially proving a lemma about divisibility “inline”: that  $d \mid m$  implies  $d \mid -m$ . Alternatively, we may notice that this property follows as a direct corollary of §1.2.6(b), with the multiplicative constant  $k = -1$ . The statement we prove is valid either way, but in some cases writing a quick inline proof may be easier or harder than finding if it is an instance of some existing property.

2. Prove that for all integers  $i, j, k, l, m, n$  with  $m$  positive and  $n$  nonnegative,

a)  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m} \implies i + k \equiv j + l \pmod{m}$

Assume  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m}$ . Then,  $m \mid i - j$  and  $m \mid k - l$ . Hence, by §1.2.6(a),  $m \mid (i - j) + (k - l) = (i + k) - (j + l)$  and  $i + k \equiv j + l \pmod{m}$ .

b)  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m} \implies i \cdot k \equiv j \cdot l \pmod{m}$

Assume  $i \equiv j \pmod{m} \wedge k \equiv l \pmod{m}$ . Then,  $m \mid (i - j)$  and  $m \mid (k - l)$ . By §1.2.6(b),  $m \mid i \cdot (k - l)$  and  $m \mid l \cdot (i - j)$ ; and, by §1.2.6(a),  $m \mid i \cdot (k - l) + l \cdot (i - j) = i \cdot k - j \cdot l$ . Hence,  $i \cdot k \equiv j \cdot l \pmod{m}$ .

c)  $i \equiv j \pmod{m} \implies i^n \equiv j^n \pmod{m}$

For  $n = 0$ ,  $i^n \equiv j^n \pmod{m}$  always. Assume now ①  $i \equiv j \pmod{m}$ . Then, for  $n = 1$ , we are done by assumption. For  $n = 2$ , by the previous item, we have ②  $i^2 \equiv j^2 \pmod{m}$ . From ① and ②, again by the previous item, we have  $i^3 \equiv j^3 \pmod{m}$ . Iterating this process we get  $i^n \equiv j^n \pmod{m}$  for every value of  $n$ .

🎵 If you're familiar with it, you may be screaming "induction!" – indeed, a formal proof requires the mathematical Principle of Induction, which will be studied later in the course.

🎵 These properties of congruence are fairly simple to state and prove, but combined with the previous exercise they form the basis of equational proofs about congruence. They allow us to extend a congruence between two integers into a congruence between two algebraic (polynomial) expressions of arbitrary nesting which differ in those two integers. For example, if we know that  $i \equiv j \pmod{m}$ , we also know  $(3i^2 + 5i - 7)^4 \equiv (3j^2 + 5j - 7)^4 \pmod{m}$  by repeatedly applying the properties proved in this exercise:  $i \equiv j \pmod{m}$  implies  $i^2 \equiv j^2 \pmod{m}$  implies  $3i^2 \equiv 3j^2 \pmod{m}$  and so on. This is really helpful in equational proofs in modular arithmetic, because we can rewrite parts of an expression not only if they are equal, but also when they are merely congruent. We will see examples of this shortly.

3. Prove that for all natural numbers  $k, l$  and positive integers  $m$ ,

a)  $\text{rem}(k \cdot m + l, m) = \text{rem}(l, m)$

By the Division Theorem,

$$l = \text{quo}(l, m) \cdot m + \text{rem}(l, m)$$

and hence

$$k \cdot m + l = (k + \text{quo}(l, m)) \cdot m + \text{rem}(l, m)$$

from which it follows by the Division Theorem that

$$\text{quo}(k \cdot m + l, m) = k + \text{quo}(l, m) \quad \text{and} \quad \text{rem}(k \cdot m + l, m) = \text{rem}(l, m).$$

🎵 The [Division Theorem](#) may seem like a dramatic name for a fairly obvious and unremarkable statement: that numbers can be divided with a remainder. But, in fact, the theorem is quite powerful and allows one to prove properties surprisingly easily. Let's remind ourselves of the full statement:

For every natural number  $m$  and positive natural number  $n$ , there exists a unique pair of integers  $q$  and  $r$  such that  $0 \leq q, 0 \leq r < n$ , and  $m = q \cdot n + r$ .

This is a *unique existence* statement, a form very common in mathematics. The associated proof technique relies both on the existence and uniqueness components. To highlight the former, consider the following alternative statement of the Division

**Theorem:**

Given any natural number  $m$  and for any choice of positive integer  $n$ , we can write  $m$  as  $m = q \cdot n + r$  where  $q$  and  $r$  are unique integers satisfying  $0 \leq q$  and  $0 \leq r < n$ .

This form emphasises the fact that if we have a natural number  $m$ , we can choose *any* natural number  $n$ , and the theorem guarantees that it's possible to write  $m$  in terms of  $n$  in the specific form  $m = q \cdot n + r$  for two unique naturals satisfying  $0 \leq q$  and  $0 \leq r < n$ . In essence, we get immediate “access” to two naturals  $q$  and  $r$  and two new assumptions about these naturals, as well as their uniqueness proofs.

Since  $q$  and  $r$  are uniquely determined by  $m$  and  $n$ , we can write them as  $\text{quo}(m, n)$  and  $\text{rem}(m, n)$  as if  $\text{quo}$  and  $\text{rem}$  were functions. In reality, they are just shorthands for “the natural  $q$  (resp.  $r$ ) determined from  $m$  and  $n$  by the Division Theorem”. With these, you can succinctly state the Division Theorem as

Any natural number  $m$  can be expressed as  $m = \text{quo}(m, n) \cdot n + \text{rem}(m, n)$  for any choice of positive integer  $n$ , with  $\text{quo}(m, n), \text{rem}(m, n) \in \mathbb{N}$  and  $\text{rem}(m, n) < n$ .

You may well ask “why go through all this when we have the integer division and remainder operators”? Well, we haven't formally defined them yet (and one way to define them formally is precisely via  $\text{quo}$  and  $\text{rem}$ !), but even ignoring that, proofs using uniqueness wouldn't really work if we just treated  $\text{rem}$  and  $\text{quo}$  as operators. To see how this works, let's expand the solution to the question above.

We are required to show that for all natural numbers  $k, l$  and positive integers  $m$ ,  $\text{rem}(k \cdot m + l, m) = \text{rem}(l, m)$  – any multiple of  $m$  can be cancelled out in a remainder by  $m$ . If we think of  $\text{rem}$  as the remainder operator (e.g. `%` in Java), this seems obvious – but other than spelling out the details of division as repeated subtraction (the Division Algorithm), it's quite tricky to prove! Instead, as we said above,  $\text{rem}(l, m)$  should be treated as “the unique  $r$  determined by  $l$  and  $m$  by the Division Theorem”. This is where uniqueness comes in: we know that for any other expansion  $l = \text{quo}(l, m) \cdot m + r'$  with  $r' < m$ ,  $r'$  must be equal to  $\text{rem}(l, m)$ . Thus, equality of remainders can be derived from showing that they satisfy the same property: that they can appear in the same expansion of  $l$  (via  $m$ ) and are both strictly less than  $m$ .

The question is exactly a proof of equality of two remainders:  $\text{rem}(k \cdot m + l, m)$  and  $\text{rem}(l, m)$ . If we show that they appear in two “different” expansions of the same natural number, they must be equal. What expansion would  $\text{rem}(l, m)$  appear in? Easy: the Division Theorem tells us that  $l$  can always be rewritten in terms of  $m$  as

$$l = \text{quo}(l, m) \cdot m + \text{rem}(l, m)$$

Similarly,  $\text{rem}(k \cdot m + l, m)$  appears in the expansion

$$k \cdot m + l = \text{quo}(k \cdot m + l, m) \cdot m + \text{rem}(k \cdot m + l, m)$$

All we did is apply the streamlined form of the Division Theorem, expanding both  $l$  and  $k \cdot m + l$  in terms of  $m$ . They can't directly be compared yet, because they are expansions of different naturals. To resolve that, we just add  $k \cdot m$  to the first equation and factorise to get:

$$k \cdot m + l = (k + \text{quo}(l, m)) \cdot m + \text{rem}(l, m)$$

And with that, we are done! How? Well, we have two different expansions of the number  $k \cdot m + l$ : it's equal both to

$$\text{quo}(k \cdot m + l, m) \cdot m + \text{rem}(k \cdot m + l, m) \quad \text{and} \quad (k + \text{quo}(l, m)) \cdot m + \text{rem}(l, m)$$

where both  $\text{rem}(k \cdot m + l, m)$  and  $\text{rem}(l, m)$  are less than  $m$ . But the Division Theorem tells us that there is exactly one such expansion of  $k \cdot m + l$  possible, so these two remainders *cannot* be different! That is to say,

$$\text{rem}(k \cdot m + l, m) = \text{rem}(l, m)$$

which was precisely our proof goal.


Such surprising and abrupt conclusions are very much characteristic of proofs by *universal properties*: rather than proving equality directly, we show that both remainders satisfy the universal property (specified by the Division Theorem) of the same number  $k \cdot m + l$  and therefore must be equal. We will see a lot of examples of this in the course and the exercises: while many statements can be proved by alternative means, proofs by universal properties are often remarkably compact and elegant, achieving the same goal with only a few clever reasoning steps.

b)  $\text{rem}(k + l, m) = \text{rem}(\text{rem}(k, m) + l, m)$

Because

$$\begin{aligned} \text{rem}(k + l, m) &= \text{rem}(\text{quo}(k, m) \cdot m + \text{rem}(k, m) + l, m) && \text{(by DT on } k \text{ with } m) \\ &= \text{rem}(\text{rem}(k, m) + l, m) && \text{(by §2.1.3(a))} \end{aligned}$$

Note that, as a corollary,  $\text{rem}(k + l, m) = \text{rem}(\text{rem}(k, m) + \text{rem}(l, m), m)$ .

 The previous property of remainders is quite useful, especially in combination with the Division Theorem: since we have a choice of expanding  $k$  in terms of any positive integer, we can choose  $m$  to then ensure that the term  $\text{quo}(k, m) \cdot m$  – being a multiple of  $m$  – can be cancelled out.

c)  $\text{rem}(k \cdot l, m) = \text{rem}(k \cdot \text{rem}(l, m), m)$

Because

$$\begin{aligned} \text{rem}(k \cdot l, m) &= \text{rem}(k \cdot \text{quo}(l, m) \cdot m + k \cdot \text{rem}(l, m), m) && \text{(by DT on } l \text{ with } m) \\ &= \text{rem}(k \cdot \text{rem}(l, m), m) && \text{(by §2.1.3(a))} \end{aligned}$$

Note that, as a corollary,  $\text{rem}(k \cdot l, m) = \text{rem}(\text{rem}(k, m) \cdot \text{rem}(l, m), m)$ .

♪ Once again, we start by expanding a natural in terms of  $m$ , then use part §2.1.3(a) to cancel the whole term. In this case, we choose  $l$ : this was guided by the need to end up with a  $\text{rem}(l, m)$ , which we wouldn't get by expanding  $k$ .

4. Let  $m$  be a positive integer.

a) Prove the associativity of the addition and multiplication operations in  $\mathbb{Z}_m$ ; that is:

$$\forall i, j, k \in \mathbb{Z}_m. (i +_m j) +_m k = i +_m (j +_m k) \quad \text{and} \quad (i \cdot_m j) \cdot_m k = i \cdot_m (j \cdot_m k)$$

Consider arbitrary  $i, j, k$  in  $\mathbb{Z}_m$ , and calculate as follows:

$$\begin{aligned} (i +_m j) +_m k &= [[i + j]_m + k]_m && \text{(by definition of } +_m) \\ &= \text{rem}(\text{rem}(i + j, m) + k, m) && \text{(by definition of } [\cdot]_m) \\ &= \text{rem}((i + j) + k, m) && \text{(by §2.1.3(b))} \\ &= \text{rem}(i + (j + k), m) && \text{(by associativity of addition)} \\ &= \text{rem}(i + \text{rem}(j + k, m), m) && \text{(by §2.1.3(b))} \\ &= [i + [j + k]_m]_m && \text{(by definition of } [\cdot]_m) \\ &= i +_m (j +_m k) && \text{(by definition of } +_m) \end{aligned}$$

Similarly, consider arbitrary  $i, j, k$  in  $\mathbb{Z}_m$ , and calculate as follows:

$$\begin{aligned} (i \cdot_m j) \cdot_m k &= [[i \cdot j]_m \cdot k]_m && \text{(by definition of } \cdot_m) \\ &= \text{rem}(\text{rem}(i \cdot j, m) \cdot k, m) && \text{(by definition of } [\cdot]_m) \\ &= \text{rem}((i \cdot j) \cdot k, m) && \text{(by §2.1.3(c))} \\ &= \text{rem}(i \cdot (j \cdot k), m) && \text{(by associativity of multiplication)} \\ &= \text{rem}(i \cdot \text{rem}(j \cdot k, m), m) && \text{(by §2.1.3(c))} \\ &= [i \cdot [j \cdot k]_m]_m && \text{(by definition of } [\cdot]_m) \\ &= i \cdot_m (j \cdot_m k) && \text{(by definition of } \cdot_m) \end{aligned}$$

♪ When defining something in terms of an existing construction, its properties will often directly follow from the known properties of the underlying definition. In this case, associativity of  $+_m$  relies on the associativity of  $+$  in terms of which  $+_m$  is defined. However, we needed a lemma about addition and remainders to simplify the expressions until we can directly appeal to the associativity of  $+$ .

♪ These are examples of *equational proofs*, a very common and useful technique for mathematical reasoning, generalising the algebraic calculations you are familiar with from school. Whenever we need to prove equality or equivalence of two mathematical objects (numbers, sets, functions, etc.), we can build it up as a chain of equalities, each rewriting some part of the expression via some known property, definition, or

lemma. There's often a symmetry in the proofs, nicely showcased in this exercise: the first half unwraps several layers of definitions and simplifies the resulting expressions; the second half does the same in reverse. Indeed, it's often helpful to write equational proofs starting from both ends, until they meet in the middle.


b) Prove that the additive inverse of  $k$  in  $\mathbb{Z}_m$  is  $[-k]_m$ .

We need show that  $k +_m [-k]_m \equiv 0 \pmod{m}$ ; and indeed, since

$$l \equiv [l]_m \pmod{m} \text{ for all } l \in \mathbb{Z}$$

one has that

$$k +_m [-k]_m = [k + [-k]_m]_m \equiv k + [-k]_m \equiv k + (-k) = 0 \pmod{m}$$

 This is an example of a *congruence proof*: a weaker form of an equational proof where some of the steps are not strict equalities, but congruences modulo  $m \in \mathbb{Z}^+$ . Since congruence is a so-called *equivalence relation* (it's reflexive, symmetric, and transitive, all proved in §2.1.1), a chain of congruences establishes a congruence between the endpoints. Reflexivity allows us to strengthen some of the congruences into equalities: in the example above,  $k +_m [-k]_m = [k + [-k]_m]_m$  is a strict equality, since it is the definition of  $+_m$ . Importantly, all congruences must be modulo the same  $m \in \mathbb{Z}^+$ , which is denoted at the end of the proof, ranging over the entire chain of congruences.

## 2.2. Core exercises

1. Find an integer  $i$ , natural numbers  $k, l$  and a positive integer  $m$  for which  $k \equiv l \pmod{m}$  holds while  $i^k \equiv i^l \pmod{m}$  does not.

Take  $i = 2, k = 0, l = 3$ , and  $m = 3$ . Then,  $k = 0 \equiv 3 = l \pmod{3}$ , yet  $2^0 = 1 \not\equiv 8 = 2^3 \pmod{3}$ .

2. Formalise and prove the following statement: A natural number is a multiple of 3 iff so is the number obtained by summing its digits. Do the same for the analogous criterion for multiples of 9 and a similar condition for multiples of 11.

For all natural numbers  $n$  and digits  $a_0, \dots, a_n$ ,

- $\left(\sum_{i=0}^n a_i \cdot 10^i\right) \equiv 0 \pmod{3} \iff \left(\sum_{i=0}^n a_i\right) \equiv 0 \pmod{3}$
- $\left(\sum_{i=0}^n a_i \cdot 10^i\right) \equiv 0 \pmod{9} \iff \left(\sum_{i=0}^n a_i\right) \equiv 0 \pmod{9}$
- $\left(\sum_{i=0}^n a_i \cdot 10^i\right) \equiv 0 \pmod{11} \iff \left(\sum_{i=0}^n (-1)^i \cdot a_i\right) \equiv 0 \pmod{11}$

The above follow from the following stronger statements

- $\left(\sum_{i=0}^n a_i \cdot 10^i\right) \equiv \left(\sum_{i=0}^n a_i\right) \pmod{3}$
- $\left(\sum_{i=0}^n a_i \cdot 10^i\right) \equiv \left(\sum_{i=0}^n a_i\right) \pmod{9}$

$$\cdot \left( \sum_{i=0}^n a_i \cdot 10^i \right) \equiv \left( \sum_{i=0}^n (-1)^i \cdot a_i \right) \pmod{11}$$

The rule for 3 uses the fact that  $10 \equiv 1 \pmod{3}$ , which, by the exponentiation property shown in §2.1.2(c), implies  $10^l \equiv 1 \pmod{3}$  for all  $l \in \mathbb{Z}^+$ . This can be applied in every term of the sum (since congruences can be applied within sums and products as shown in §2.1.2, reducing the  $10^l$  coefficients to 1. The technique works the same for divisibility by 9, since  $10 \equiv 1 \pmod{9}$ ; for 11, we notice that  $10^{2n} \equiv 1 \pmod{11}$ , but  $10^{2n+1} \equiv 10 \equiv -1 \pmod{11}$  for all  $n \in \mathbb{N}$ .

There are also other proofs. Below is one based on the Binomial Theorem, rather than on the theory of divisibility and/or congruences for the case of divisibility by 11. Please study it and re-adapt it to the cases of divisibility by 3 and by 9.

First we calculate that

$$\begin{aligned} \sum_{i=0}^n a_i \cdot 10^i &= \sum_{i=0}^n a_i \cdot (11-1)^i \\ &= \sum_{i=0}^n a_i \cdot \sum_{j=0}^i \binom{i}{j} \cdot 11^j \cdot (-1)^{i-j} \\ &= \sum_{i=0}^n a_i \cdot \left[ (-1)^i + 11 \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \\ &= \left( \sum_{i=0}^n (-1)^i \cdot a_i \right) + 11 \cdot \left[ \sum_{i=1}^n a_i \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \end{aligned}$$

and then argue as follows:

( $\Rightarrow$ ) Assume  $11 \mid \left( \sum_{i=0}^n a_i \cdot 10^i \right)$ ; so that  $\sum_{i=0}^n a_i \cdot 10^i = 11 \cdot k$  for some integer  $k$ . Then,

$$\sum_{i=0}^n (-1)^i \cdot a_i = 11 \cdot \left( k - \left[ \sum_{i=1}^n a_i \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \right)$$

showing that  $11 \mid \left( \sum_{i=0}^n (-1)^i \cdot a_i \right)$ .


( $\Leftarrow$ ) Assume  $11 \mid \left( \sum_{i=0}^n (-1)^i \cdot a_i \right)$ ; so that  $\sum_{i=0}^n (-1)^i \cdot a_i = 11 \cdot l$  for some integer  $l$ . Then,

$$\sum_{i=0}^n a_i \cdot 10^i = 11 \cdot \left( l + \left[ \sum_{i=1}^n a_i \cdot \sum_{j=1}^i \binom{i}{j} \cdot 11^{j-1} \cdot (-1)^{i-j} \right] \right)$$

showing that  $11 \mid \sum_{i=0}^n a_i \cdot 10^i$ .

3. Show that for every integer  $n$ , the remainder when  $n^2$  is divided by 4 is either 0 or 1.

This is [Lemma 26](#) of the notes.

 The question here refers to the “intuitive” notions of division and remainder, but by recognising their connection to congruence we can refer to the known number-theoretic properties of modular arithmetic.

4. What are  $\text{rem}(55^2, 79)$ ,  $\text{rem}(23^2, 79)$ ,  $\text{rem}(23 \cdot 55, 79)$  and  $\text{rem}(55^{78}, 79)$ ?





•  $\mathbb{Z}_6$

| + | 0 | 1 | 2 | 3 | 4 | 5 | · | 0 | 1 | 2 | 3 | 4 | 5 | −(·) | (·) <sup>−1</sup> |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------|-------------------|
| 0 | 0 | 1 | 2 | 3 | 4 | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0    | 0                 |
| 1 | 1 | 2 | 3 | 4 | 5 | 0 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 1    | 1                 |
| 2 | 2 | 3 | 4 | 5 | 0 | 1 | 2 | 0 | 2 | 4 | 0 | 2 | 4 | 2    |                   |
| 3 | 3 | 4 | 5 | 0 | 1 | 2 | 3 | 0 | 3 | 0 | 3 | 0 | 3 | 3    |                   |
| 4 | 4 | 5 | 0 | 1 | 2 | 3 | 4 | 0 | 4 | 2 | 0 | 4 | 2 | 4    |                   |
| 5 | 5 | 0 | 1 | 2 | 3 | 4 | 5 | 0 | 5 | 4 | 3 | 2 | 1 | 5    | 5                 |

•  $\mathbb{Z}_7$

| + | 0 | 1 | 2 | 3 | 4 | 5 | 6 | · | 0 | 1 | 2 | 3 | 4 | 5 | 6 | −(·) | (·) <sup>−1</sup> |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|------|-------------------|
| 0 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0    | 0                 |
| 1 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 1    | 1                 |
| 2 | 2 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 0 | 2 | 4 | 6 | 1 | 3 | 5 | 2    | 4                 |
| 3 | 3 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 0 | 3 | 6 | 2 | 5 | 1 | 4 | 3    | 5                 |
| 4 | 4 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 0 | 4 | 1 | 5 | 2 | 6 | 3 | 4    | 2                 |
| 5 | 5 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 0 | 5 | 3 | 1 | 6 | 4 | 2 | 5    | 3                 |
| 6 | 6 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 0 | 6 | 5 | 4 | 3 | 2 | 1 | 6    | 6                 |

Great demonstration of the property that every element of  $\mathbb{Z}_p$  has a multiplicative inverse if  $p$  is a prime. Algebraically, this makes  $\mathbb{Z}_p$  a *field*: a place where you can do division.

7. Let  $i$  and  $n$  be positive integers and let  $p$  be a prime. Show that if  $n \equiv 1 \pmod{p-1}$  then  $i^n \equiv i \pmod{p}$  for all  $i$  not multiple of  $p$ .

Assume that  $i$  and  $n$  are positive integers and that  $p$  is a prime. Assume further that  $n \equiv 1 \pmod{p-1}$ ; so that  $n-1 = k \cdot (p-1)$  for some *natural number*  $k$ . Then, for  $i$  not a multiple of  $p$ , we have that

$$\begin{aligned} i^n &= i \cdot (i^{p-1})^k \\ &\equiv i \cdot 1^k \pmod{p} && \text{(by Fermat's Little Theorem)} \\ &= i \end{aligned}$$

When the question involves prime numbers, you should expect to require properties and theorems specific to primes. In this course – which is only an introduction to number theory – this will quite often be Fermat's Little Theorem.

8. Prove that  $n^3 \equiv n \pmod{6}$  for all integers  $n$ .

We can proceed by case analysis: since either  $n \equiv 0 \pmod{6}$ , or  $n \equiv 1 \pmod{6}$ , or  $n \equiv 2 \pmod{6}$ , or  $n \equiv 3 \pmod{6}$ , or  $n \equiv 4 \pmod{6}$ , or  $n \equiv 5 \pmod{6}$ , we check that  $n^3 \equiv n \pmod{6}$  in each case.

- Case  $n \equiv 0 \pmod{6}$ :  $n^3 \equiv 0^3 = 0 \equiv n \pmod{6}$ .
- Case  $n \equiv 1 \pmod{6}$ :  $n^3 \equiv 1^3 = 1 \equiv n \pmod{6}$ .
- Case  $n \equiv 2 \pmod{6}$ :  $n^3 \equiv 2^3 = 8 \equiv 2 \equiv n \pmod{6}$ .
- Case  $n \equiv 3 \pmod{6}$ :  $n^3 \equiv 3^3 = 27 \equiv 3 \equiv n \pmod{6}$ .
- Case  $n \equiv 4 \pmod{6}$ :  $n^3 \equiv 4^3 = 64 \equiv 4 \equiv n \pmod{6}$ .
- Case  $n \equiv 5 \pmod{6}$ :  $n^3 \equiv 5^3 = 125 \equiv 5 \equiv n \pmod{6}$ .

Of course, this wouldn't really work for larger moduli – see next question. A more elegant solution is proving  $6 \mid n^3 - n$ , which, by the well-known divisibility rule for 6, follows from showing  $3 \mid n^3 - n$  and  $2 \mid n^3 - n$ . Now,

$$n^3 - n = n \cdot (n^2 - 1) = (n - 1) \cdot n \cdot (n + 1);$$


but this is a product of three consecutive integers, so at least one of them must be even and one must be divisible by 3. That is,  $n^3 - n = 2 \cdot 3 \cdot k$  for some  $k \in \mathbb{Z}$ , so  $n^3 \equiv n \pmod{6}$ .

Yet another approach is formally establishing the lemma (which can be seen as the generalisation of the divisibility rule of 6):

$$(a \equiv b \pmod{2} \wedge a \equiv b \pmod{3}) \iff a \equiv b \pmod{6}$$

In one direction, we have that  $a = 2k + b = 3l + b$ , so  $2k = 3l$  for integers  $k$  and  $l$ ; since  $3l$  must be even and 3 is odd,  $l$  must itself be even:  $l = 2m$  for some  $m \in \mathbb{Z}$ . Substituting back, we have  $a = 3 \cdot 2m + b$ , so  $a - b = 6m$ . In the opposite direction,  $a - b = 6k = 2 \cdot 3 \cdot k$ , which immediately implies  $2 \mid a - b$  and  $3 \mid a - b$ .

Now, it is sufficient to prove that  $n^3 \equiv n \pmod{2}$  and  $n^3 \equiv n \pmod{3}$ . The latter is a direct instance of Fermat's Little Theorem for the prime 3; the former holds by the congruence chain  $n^3 \equiv n^2 \equiv n \pmod{2}$ , with both steps using Fermat's Little Theorem  $n^2 \equiv n \pmod{2}$ , multiplied by  $n$  in the first step using the product property of §2.1.2.

 There are usually many ways of approaching a proof, ranging from “brute force” methods to elegant and concise number-theoretic arguments. It doesn't technically matter what you do as long as the proof is correct – but just like how “working” code doesn't always mean “neat and readable” code, you should strive to make your proofs as streamlined as possible. It's also very useful to practice recognising patterns and realising where some known lemma or property can be applied, since they often end up doing the bulk of the work: you shouldn't need to reprove a specific case of a known, more general statement.

### 9. Prove that $n^7 \equiv n \pmod{42}$ for all integers $n$ .

An exhaustive case analysis would be impractical in this case. Instead, we adapt our more conceptual solutions above.

First, we use a very similar proof as above for the lemma

$$(a \equiv b \pmod{6} \wedge a \equiv b \pmod{7}) \iff a \equiv b \pmod{42}$$

(notice how the crucial step is  $6k = 7l$  implying that  $6 \mid l$ , because  $6 \nmid 7$  – the lemma wouldn't hold for non-coprime numbers (see §1.2.5). Another trick in this case is recognising that  $a - b = 7(a - b) - 6(a - b)$ , and, since by assumption  $a - b = 6k = 7l$ , we have  $a - b = 7 \cdot 6k - 6 \cdot 7l = 42 \cdot (k - l)$ ).

Now,  $n^7 \equiv n \pmod{7}$  holds by Fermat's Little Theorem. To show  $n^7 \equiv n \pmod{6}$ , we can equivalently show  $n^7 \equiv n \pmod{2}$  and  $n^7 \equiv n \pmod{3}$ ; both follow by repeated applications of Fermat's Little Theorem.

### 2.3. Optional exercises

1. Prove that for all integers  $n$ , there exist natural numbers  $i$  and  $j$  such that  $n = i^2 - j^2$  iff either  $n \equiv 0 \pmod{4}$  or  $n \equiv 1 \pmod{4}$  or  $n \equiv 3 \pmod{4}$ .

Consider an arbitrary integer  $n$ .

( $\Rightarrow$ ) Assume there exist natural numbers  $i$  and  $j$  such that  $n = i^2 - j^2$ . By [Proposition 25](#) of the notes, we have that

$$\text{either } i^2 \equiv 0 \pmod{4} \text{ or } i^2 \equiv 1 \pmod{4}$$

and

$$\text{either } j^2 \equiv 0 \pmod{4} \text{ or } j^2 \equiv 1 \pmod{4}$$

We therefore have four cases:

- $i^2 \equiv 0 \pmod{4}$  and  $j^2 \equiv 0 \pmod{4}$ , in which case  $n \equiv 0 \pmod{4}$ ;
- $i^2 \equiv 0 \pmod{4}$  and  $j^2 \equiv 1 \pmod{4}$ , in which case  $n \equiv -1 \equiv 3 \pmod{4}$ ;
- $i^2 \equiv 1 \pmod{4}$  and  $j^2 \equiv 0 \pmod{4}$ , in which case  $n \equiv 1 \pmod{4}$ ;
- $i^2 \equiv 1 \pmod{4}$  and  $j^2 \equiv 1 \pmod{4}$ , in which case  $n \equiv 0 \pmod{4}$ ;

Hence, either  $n \equiv 0 \pmod{4}$ , or  $n \equiv 1 \pmod{4}$ , or  $n \equiv 3 \pmod{4}$  as required.

( $\Leftarrow$ ) Assume that either  $n \equiv 0 \pmod{4}$ , or  $n \equiv 1 \pmod{4}$ , or  $n \equiv 3 \pmod{4}$ . We need to find natural numbers  $i$  and  $j$  such that  $n = i^2 - j^2$ .

Graphically, we want to show that one can distribute any number of balls (as long as it's congruent to 0, 1 or 3 modulo 4) in a square grid leaving an empty square sub-grid, for instance as follows (for  $i = 7$ ,  $j = 3$ , and  $n = 40$ ):

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
| • | • | • | • | • | • | • |
| • | • | • | • | • | • | • |
| • | • | • | • | • | • | • |
| • | • | • |   |   |   | • |
| • | • | • |   |   |   | • |
| • | • | • |   |   |   | • |
| • | • | • | • | • | • | • |

We split our analysis in three cases.

- Case  $n$  is zero.

There are natural numbers  $i = j = 0$  such that  $n = i^2 - j^2$ , and we are done.

- Case  $n$  is a non-zero even integer.

As  $\text{rem}(n, 4) = n - \text{quo}(n, 4) \cdot 4$  (by the Division Theorem), it follows that  $\text{rem}(n, 4)$  is even and since hence it is necessarily 0. Thus,  $n$  is in fact a non-zero multiple of 4; say of the form  $4 \cdot k$  for some non-zero integer  $k$ . Then,

$$n = (k+1)^2 - (k-1)^2 = (-k-1)^2 - (1-k)^2$$

and since either

$$k+1 \text{ and } k-1 \text{ are natural numbers}$$

or

$$-k-1 \text{ and } 1-k \text{ are natural numbers}$$

there are natural numbers  $i, j$  such that  $n = i^2 - j^2$ . (Note that this argument slightly generalises that of Proposition 22 of the notes.)

Graphically, we are in the following kind of situation:

|                |                |                |                |                |                |                |                |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>3</sub> |
| • <sub>1</sub> |                |                |                |                |                |                | • <sub>3</sub> |
| • <sub>1</sub> |                |                |                |                |                |                | • <sub>3</sub> |
| • <sub>1</sub> |                |                |                |                |                |                | • <sub>3</sub> |
| • <sub>1</sub> |                |                |                |                |                |                | • <sub>3</sub> |
| • <sub>1</sub> |                |                |                |                |                |                | • <sub>3</sub> |
| • <sub>1</sub> |                |                |                |                |                |                | • <sub>3</sub> |
| • <sub>1</sub> | • <sub>4</sub> | • <sub>4</sub> | • <sub>4</sub> | • <sub>4</sub> | • <sub>4</sub> | • <sub>4</sub> | • <sub>4</sub> |

- Case  $n$  is odd.

Then  $n = 2 \cdot k + 1$  for some integer  $k$ , and

$$n = (k+1)^2 - k^2 = (-k-1)^2 - (-k)^2 .$$

Since either

$$k+1 \text{ and } k \text{ are natural numbers}$$

or

$$-k-1 \text{ and } -k \text{ are natural numbers}$$

there are natural numbers  $i, j$  such that  $n = i^2 - j^2$ .

Graphically, we are in the following kind of situation:

|                |                |                |                |                |                |                |
|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| •              | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> | • <sub>2</sub> |
| • <sub>1</sub> |                |                |                |                |                |                |
| • <sub>1</sub> |                |                |                |                |                |                |
| • <sub>1</sub> |                |                |                |                |                |                |
| • <sub>1</sub> |                |                |                |                |                |                |
| • <sub>1</sub> |                |                |                |                |                |                |
| • <sub>1</sub> |                |                |                |                |                |                |

🎵 Graphical proofs are great for intuition: so called “proofs without words” are often as illuminating as they are beautiful. However, they are not (usually) a substitute for a formal proof by logical reasoning, especially if the proposition to be shown is more general than what could be encoded graphically. In this case, the statement is about all *integers*  $n$ , while the graphical proof can only work for a *natural number*  $n$ .

2. A *decimal* (respectively *binary*) *repunit* is a natural number whose decimal (respectively binary) representation consists solely of 1's.

a) What are the first three decimal repunits? And the first three binary ones?

The first three decimal repunits are 1, 11, and 111; while the first three binary repunits are 1, 3, and 7.

b) Show that no decimal repunit strictly greater than 1 is a square, and that the same holds for binary repunits. Is this the case for every base? *Hint*: Use [Lemma 26](#) of the notes.

Let  $n$  be a decimal repunit greater than 1; that is,  $n = \sum_{i=0}^l 10^i$  for some  $l \geq 1$ . Then,

$$n \equiv \sum_{i=0}^l 2^i \equiv 1 + 2 = 3 \pmod{4}$$

and, by [Proposition 25](#) of the notes, we deduce that  $n$  is not square.

Incidentally, the calculation above already contains the proof of the property for binary repunits, since they are of the form  $n = \sum_{i=0}^l 2^i$

The statement:

For every base  $r$ , there are no  $r$ -ary repunits greater than 1 that are square.


is false. As a counterexample, take the base  $r = 3$  and the 3-ary repunit 4 consisting of two 1's.

### 3. More on numbers

#### 3.1. Basic exercises

1. Calculate the set  $\text{CD}(666, 330)$  of common divisors of 666 and 330.

We have that  $666 = 2 \cdot 3^2 \cdot 37$  and  $330 = 2 \cdot 3 \cdot 5 \cdot 11$ . Hence,  $\text{CD}(666, 330) = \{1, 2, 3, 2 \cdot 3\} = \{1, 2, 3, 6\}$ .

 You may be familiar with this method of computing the common divisors of two numbers using their prime factorisation – this of course relies on the Fundamental Theorem of Arithmetic, introduced later in the course.

2. Find the gcd of 21212121 and 12121212.


We run [Euclid's Algorithm](#):

$$\begin{aligned} \text{gcd}(21212121, 12121212) &= \text{gcd}(12121212, 9090909) \\ &= \text{gcd}(9090909, 3030303) \\ &= 3030303 \end{aligned}$$

3. Prove that for all positive integers  $m$  and  $n$ , and integers  $k$  and  $l$ ,

$$\text{gcd}(m, n) \mid (k \cdot m + l \cdot n)$$

Let  $m, n$  be positive integers and  $k, l$  be integers. As  $\text{gcd}(m, n) \mid m$  and  $\text{gcd}(m, n) \mid n$  it follows from §1.2.6(a) that  $\text{gcd}(m, n) \mid k \cdot m$  and  $\text{gcd}(m, n) \mid l \cdot n$ ; from which it further follows by §1.2.6(b) that  $\text{gcd}(m, n) \mid (k \cdot m + l \cdot n)$ .

 Like  $\text{rem}$ , we can treat  $\text{gcd}(m, n)$  as a function of two positive integers  $m$  and  $n$ , or as a symbol for the greatest common divisor of  $m$  and  $n$  defined using the universal property of gcds. For example, we make use of the fact that  $\text{gcd}(m, n)$  is a common divisor of  $m$  and  $n$ , so we “automatically” get  $\text{gcd}(m, n) \mid m$  and  $\text{gcd}(m, n) \mid n$ . We will see more examples of this in the upcoming exercises.

4. Find integers  $x$  and  $y$  such that  $x \cdot 30 + y \cdot 22 = \text{gcd}(30, 22)$ . Now find integers  $x'$  and  $y'$  with  $0 \leq y' < 30$  such that  $x' \cdot 30 + y' \cdot 22 = \text{gcd}(30, 22)$ .

Run the [Extended Euclid's Algorithm](#) to find that  $\text{gcd}(30, 22) = 2$  and  $x \cdot 30 + y \cdot 22 = 2$  for  $x = 3$  and  $y = -4$ . To get a  $y'$  between the range  $0 \leq y' < 30$ , we notice that

$$(x + 11 \cdot l) \cdot 30 + (y - 15 \cdot l) \cdot 22 = 2$$

for all integers  $l$  ([Slide 219](#)), and find a value  $l_0$  such that  $0 \leq y - 15 \cdot l_0 < 30$  setting  $x' = x + 11 \cdot l_0$  and  $y' = y - 15 \cdot l_0$ . The two options are  $l_0 = -1$  for  $(-8) \cdot 30 + 11 \cdot 22 = 2$ , and  $l_0 = -2$  for  $(-19) \cdot 30 + 26 \cdot 22 = 2$ .

5. Prove that for all positive integers  $m$  and  $n$ , there exists integers  $k$  and  $l$  such that  $k \cdot m + l \cdot n = 1$  iff  $\text{gcd}(m, n) = 1$ .

( $\Rightarrow$ ) By [Corollary 62](#) of the notes: if 1 can be expressed as a linear combination of  $m$  and  $n$ , and  $\gcd(m, n)$  must divide any linear combination of  $m$  and  $n$ , we must have  $\gcd(m, n) = 1$ .

( $\Leftarrow$ ) By [Theorem 70](#) of the notes:  $\gcd(m, n)$  is a linear combination of  $m$  and  $n$ .

6. Prove that for all integers  $n$  and primes  $p$ , if  $n^2 \equiv 1 \pmod{p}$  then either  $n \equiv 1 \pmod{p}$  or  $n \equiv -1 \pmod{p}$ .

Assume  $n^2 \equiv 1 \pmod{p}$ . Then  $p$  divides  $n^2 - 1 = (n - 1) \cdot (n + 1)$ . By [Euclid's Theorem](#),  $p \mid (n - 1)$  or  $p \mid (n + 1)$ ; that is, either  $n \equiv 1 \pmod{p}$  or  $n \equiv -1 \pmod{p}$ .

### 3.2. Core exercises

1. Prove that for all positive integers  $m$  and  $n$ ,  $\gcd(m, n) = m$  iff  $m \mid n$ .

Let  $m$  and  $n$  be arbitrary positive integers.

( $\Rightarrow$ ) Assume that  $\gcd(m, n) = m$ . Then  $m$  is the greatest common divisor of both  $m$  and  $n$ , and in particular a divisor of  $n$ .

( $\Leftarrow$ ) Assume  $m \mid n$ .

Here are two arguments.

- a) We have that  $n = k \cdot m$  for some positive integer  $k$ , and hence that


$$\gcd(m, n) = \gcd(m, k \cdot m) = m \cdot \gcd(1, k) = m$$

where the second equality is a consequence of the linearity property ([Lemma 63\(3\)](#) of the notes) of  $\gcd$ .

- b) By [Theorem 61](#) of the notes, it suffices to prove that

- $m \mid m$  and  $m \mid n$ , and
- for all positive integers  $d$  such that  $d \mid m$  and  $d \mid n$  it necessarily follows that  $d \mid m$ ;

all of which hold trivially.

 It's worth analysing the second approach, as it's quite characteristic of proofs by universal properties: the proof just "pops out" without us having to do a whole lot of work, similar to our use of the Division Theorem in [§2.1.3\(a\)](#).

As mentioned in [§3.1.3](#), there are several equivalent ways of thinking about gcds. One is as a function of two positive integers  $m$  and  $n$ , computed via [Euclid's Algorithm](#); another is as a label for a unique number characterised by the universal property of being the greatest common divisor of  $m$  and  $n$ . The difference may seem insignificant, but that is precisely because of [Theorem 61](#), which states that the value computed by Euclid's Algorithm coincides with the greatest common divisor. The universal property of gcds (which we'll get to shortly) is the *specification* of what it is to be a greatest common divisor;



Theorem 61 states that Euclid's Algorithm satisfies the specification. We don't *define* the greatest common divisor of  $m$  and  $n$  as "the number returned by Euclid's Algorithm"; just as how we don't define a sorted list as "the list returned by the quicksort algorithm" or a lasagna as "the dish you get by following this specific recipe in this specific cookbook". We already know what a gcd/sorted list/lasagna is supposed to be, and we can then ask whether some algorithm computes the gcd or some recipe makes a lasagna, or it doesn't. Of course, what makes a lasagna and what is the *best* lasagna is entirely subjective, while mathematical concepts can be unambiguously characterised using universal properties.

Universal properties have two parts: the *property* and the *universality*. The former characterises the set of candidates for the concept we are considering; the latter selects a specific candidate which is "better" than all the other ones. In the case of the greatest common divisor of  $m$  and  $n$ , the property is that of being a common divisor of  $m$  and  $n$ : the set of candidates that satisfy this property is  $\text{CD}(m, n)$ . The "best" such candidate that we are looking for is the one which is greater than all the other ones, and since  $\text{CD}(m, n)$  is a finite non-empty set of natural numbers, it must have a unique greatest element  $\max(\text{CD}(m, n))$ . We can denote this element (which depends entirely on  $m$  and  $n$ ) as  $\text{gcd}(m, n)$  and call it the greatest common divisor of  $m$  and  $n$ .

From this description (or, really, definition) of  $\text{gcd}(m, n)$  as the greatest element of the set of common divisors, we can directly extract two "axioms":  $\text{gcd}(m, n) \in \text{CD}(m, n)$  (since it is a common divisor), and for all  $d \in \text{CD}(m, n)$ ,  $d \leq \text{gcd}(m, n)$  (since it is the greatest common divisor). In fact, we can state something stronger: not only are all other common divisors numerically smaller than  $\text{gcd}(m, n)$ , they also all divide it:  $\forall d \in \text{CD}(m, n). d \mid \text{gcd}(m, n)$ . Expanding these, we universally characterise  $\text{gcd}(m, n)$  as the unique natural number  $g$  satisfying the properties of being a common divisor and a multiple of all common divisors:

$$\textcircled{1} g \mid m \wedge g \mid n \quad \textcircled{2} \forall d \in \mathbb{Z}^+. (d \mid m \wedge d \mid n) \implies d \mid g$$

Using the transitivity of divisibility (§1.2.4), we can combine these into the concise specification of the universal property of greatest common divisors:

$$\textcircled{3} \forall d \in \mathbb{Z}^+. (d \mid m \wedge d \mid n) \iff d \mid \text{gcd}(m, n)$$

It's easy to show that gcds are unique: if we had two gcds, both would have to satisfy  $\textcircled{2}$  and, in particular, they must divide each other; but divisibility (on positive integers) is antisymmetric (§1.2.8), so the two gcds must be equal. Uniqueness in turn gives rise to the following important proof principle:

To prove that a number  $g \in \mathbb{Z}^+$  is equal to  $\text{gcd}(m, n)$ ,  
it is sufficient to show that  $g$  satisfies  $\textcircled{1}$  and  $\textcircled{2}$ .

This is similar to the approach we used with the Division Theorem: to prove that a number  $r$  is equal to  $\text{rem}(m, n)$ , it was sufficient to show that it is less than  $n$  and it can appear in an expansion  $m = q \cdot k + r$  with  $q \in \mathbb{N}$ . Adapting this technique to the combined form  $\textcircled{3}$ ,

we get a useful and particularly simple variation:

To prove that a number  $d$  divides  $\gcd(m, n)$ , it's sufficient to show that  $d \mid m$  and  $d \mid n$ .

This, combined with the antisymmetry of divisibility (on positive integers), allows us to prove equality of gcds, as shown in the example proofs of [Lemma 63](#) in the notes. In essence, the first step in proving something about  $\gcd(m, n)$  or  $\text{rem}(n, m)$  is “forgetting” about the gcd or rem and approach the proof via the universal property; it may seem like a very roundabout technique (as opposed to, for example, a direct chain of equalities ending in  $\gcd(m, n)$ ), but it often leads to short and straightforward proofs. However, it's definitely not the case that *all* proofs about gcds have to be done this way, and we'll see more examples later!

To conclude the discussion, let us expand on proof (b) of this exercise, which uses the UP of gcds. To recap, in the ( $\Leftarrow$ ) direction we need to show:

$$\forall m, n \in \mathbb{Z}^+. m \mid n \implies \gcd(m, n) = m$$

As always, assume  $m, n \in \mathbb{Z}^+$  and  $m \mid n$ . The proof goal  $\gcd(m, n) = m$  asks us to show that  $m$  is equal to  $\gcd(m, n)$ ; but, by the proof principle above, it is sufficient to show that  $m$  satisfies ① and ②. That is,

$$\textcircled{1} m \mid m \wedge m \mid n \quad \textcircled{2} \forall d \in \mathbb{Z}^+. (d \mid m \wedge d \mid n) \implies d \mid m$$

① holds by reflexivity of  $\mid$  and our assumption  $m \mid n$ ; ② is a direct implication. And that's it! The proof (a) wasn't exactly complicated either, but (b) was rightly labelled as “trivial”.

The beautiful thing about this characterisation of gcds is that it is an instance of a much more general mathematical notion called a *greatest lower bound* (with the dual *least upper bound* being the least common multiple). These concepts appear all over mathematics and computer science, and you will encounter many examples in this course as well; accordingly, the proof technique described above can be (and will be, and has already been!) applied in several seemingly different contexts. As a teaser, see if you can spot the similarity between statement ③ above, and the pattern for proving a conjunction of two statements  $P$  and  $Q$  given any set  $A$  of assumptions:

$$\forall A. (A \Rightarrow P) \wedge (A \Rightarrow Q) \iff A \Rightarrow (P \wedge Q)$$

2. Let  $m$  and  $n$  be positive integers with  $\gcd(m, n) = 1$ . Prove that for every natural number  $k$ ,

$$m \mid k \wedge n \mid k \iff m \cdot n \mid k$$

Let  $m$  and  $n$  be arbitrary positive integers, and assume that ①  $\gcd(m, n) = 1$ . Further, let  $k$  be a natural number.

( $\Rightarrow$ ) Assume that ②  $m \mid k$  and ③  $n \mid k$ .

It follows from ① that

$$m \cdot i + n \cdot j = 1 \quad \text{④}$$

for some integers  $i, j$ ; and it follows from ② and ③ that

$$k = a \cdot m = b \cdot n \quad \text{⑤}$$


for some natural numbers  $a, b$ .

Multiplying ④ by  $k$  on both sides and using ⑤, we therefore have

$$k = b \cdot n \cdot m \cdot i + a \cdot m \cdot n \cdot j = (b \cdot i + a \cdot j) \cdot (m \cdot n)$$

showing that  $(m \cdot n) \mid k$ .

( $\Leftarrow$ ) Assume that  $(m \cdot n) \mid k$ . Then, since both  $m \mid (m \cdot n)$  and  $n \mid (m \cdot n)$ , by the transitivity of divisibility, we are done.

 The ( $\Rightarrow$ ) direction of this proof used another characterisation of  $\gcd(m, n)$  as the *least positive linear combination of  $m$  and  $n$* . (NB: “Least” here means “lowest”, not the superlative of “less positive”.) Now that we are more familiar with universal properties, we can decode this description as ①  $\gcd(m, n)$  is a linear combination of  $m$  and  $n$ , and ②  $\gcd(m, n)$  divides all linear combinations of  $m$  and  $n$ :

$$\text{① } \exists k_0, l_0 \in \mathbb{Z}. k_0 \cdot m + l_0 \cdot n = \gcd(m, n) \quad \text{② } \forall k, l \in \mathbb{Z}. \gcd(m, n) \mid k \cdot m + l \cdot n$$

This characterisation is especially useful if we are able to express 1 as a linear combination of  $m$  and  $n$ , since ② means they must be *coprime*, i.e.  $\gcd(m, n) = 1$ . Another common use of an assumption of coprimality  $\gcd(m, n) = 1$  is that multiplication by  $\gcd(m, n)$  is a no-op, so we can freely introduce  $\gcd(m, n)$  or  $k_0 \cdot m + l_0 \cdot n$  for some  $k_0, l_0 \in \mathbb{Z}$  into any expression. This is what we make use of in the question when multiplying ④ and ⑤.

3. Prove that for all positive integers  $a, b, c$ , if  $\gcd(a, c) = 1$  then  $\gcd(a \cdot b, c) = \gcd(b, c)$ .

Below are three different proofs of the property.

#### Proof by equational reasoning

For  $a, b, c$  positive integers such that  $\gcd(a, c) = 1$ , we have

$$\begin{aligned} \gcd(b, c) &= \gcd(\gcd(a, c) \cdot b, c) && \text{(since } \gcd(a, c) = 1\text{)} \\ &= \gcd(\gcd(a \cdot b, c \cdot b), c) && \text{(by linearity)} \\ &= \gcd(a \cdot b, \gcd(c \cdot b, c)) && \text{(by associativity)} \\ &= \gcd(a \cdot b, c) && \text{(by §3.2.1)} \end{aligned}$$

#### Proof by universality

Let  $a, b, c$  positive integers such that  $\gcd(a, c) = 1$ . We need to prove that  $\gcd(a \cdot b, c) = \gcd(b, c)$ , or equivalently, that  $\gcd(a \cdot b, c) \mid \gcd(b, c)$  and  $\gcd(b, c) \mid \gcd(a \cdot b, c)$ . By the

universal property of gcds, it is sufficient to show the following two properties:

- $\gcd(a \cdot b, c) \mid b$  and  $\gcd(a \cdot b, c) \mid c$ . The latter holds since  $\gcd(a \cdot b, c)$  is a divisor of  $c$ . To establish the former, we note that  $b = \gcd(a, c) \cdot b$  (since  $a$  and  $c$  are coprime), and by distributivity,  $\gcd(a \cdot b, c \cdot b)$ . Thus, we can show that  $\gcd(a \cdot b, c) \mid \gcd(a \cdot b, c \cdot b)$ , or equivalently,  $\gcd(a \cdot b, c) \mid a \cdot b$  and  $\gcd(a \cdot b, c) \mid c \cdot b$ , both of which follow from  $\gcd(a \cdot b, c)$  being a common divisor of  $a \cdot b$  and  $c$ .
- $\gcd(b, c) \mid a \cdot b$  and  $\gcd(b, c) \mid c$ . Both follow from  $\gcd(b, c)$  being a divisor of  $b$  and  $c$ .

### Proof using the Fundamental Theorem of Arithmetic

The [Fundamental Theorem of Arithmetic](#) states that every positive integer is expressible as the product of a unique finite sequence of ordered primes. If two integers are coprime, their unique prime factorisations must be disjoint: that is, there is no prime  $p$  that appears in the factorisation of both  $a$  and  $c$ . For any  $b \in \mathbb{Z}^+$ , the prime factorisation of  $a \cdot b$  will be the product of those of  $a$  and  $b$ . Therefore the common prime factors of  $a \cdot b$  and  $c$  must be the common factors of  $b$  and  $c$ , since there are no common factors of  $a$  and  $c$  by assumption. Since the greatest common divisor is the product of the common prime factors, we must have  $\gcd(a \cdot b, c) = \gcd(b, c)$ .

♪ These are three fairly different proofs of the same (relatively simple) theorem: one uses equational reasoning and some properties of gcds, the second makes use of universality, while the third relies on a powerful and general theorem rather than gcd properties. The first is probably the most concise form, but of course it relies on us having established all the required properties of gcds already.

4. Prove that for all positive integers  $m$  and  $n$ , and integers  $i$  and  $j$ :

$$n \cdot i \equiv n \cdot j \pmod{m} \iff i \equiv j \pmod{\frac{m}{\gcd(m, n)}}$$

We have:

$$\begin{aligned} n \cdot i \equiv n \cdot j \pmod{m} &\iff k \cdot m = n(i - j) \\ &\iff k \cdot \frac{m}{\gcd(m, n)} = \frac{n}{\gcd(m, n)} \cdot (i - j) \\ &\iff \frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j) \end{aligned}$$

Now we show that

$$\frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j) \iff i \equiv j \pmod{\frac{m}{\gcd(m, n)}}$$

( $\Leftarrow$ ) We have  $\frac{m}{\gcd(m, n)} \mid i - j$  by assumption, and from the multiplication property of divisibility (§1.2.6(b)), we have  $\frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j)$ .

( $\Rightarrow$ ) We first establish that  $\frac{m}{\gcd(m,n)}$  and  $\frac{n}{\gcd(m,n)}$  are coprime using linearity:


$$\gcd(m, n) = \gcd\left(\frac{m \cdot \gcd(m, n)}{\gcd(m, n)}, \frac{n \cdot \gcd(m, n)}{\gcd(m, n)}\right) = \gcd(m, n) \cdot \gcd\left(\frac{m}{\gcd(m, n)}, \frac{n}{\gcd(m, n)}\right)$$


Since  $\gcd(m, n)$  is a positive integer, this equality can only hold if  $\gcd\left(\frac{m}{\gcd(m, n)}, \frac{n}{\gcd(m, n)}\right) = 1$ .

This assumption of coprimality can then be used in [Euclid's Theorem](#) to conclude

$$\frac{m}{\gcd(m, n)} \mid \frac{n}{\gcd(m, n)} \cdot (i - j) \implies \frac{m}{\gcd(m, n)} \mid (i - j)$$

as required.

 The inspiration for the first “creative” step (dividing both sides by  $\gcd(m, n)$ ) comes from seeing the term  $\frac{m}{\gcd(m, n)}$  in the proof goal.

 A very useful corollary of this theorem is that we can always divide both sides of a congruence by a positive integer that is coprime with the modulus. Similarly, we can divide both sides of the congruence *and* the modulus with any positive integer that divides all three. The general theorem handles the case “in between”, when a positive integer divides both sides of the congruence, but not the modulus.

5. Prove that for all positive integers  $m, n, p, q$  such that  $\gcd(m, n) = \gcd(p, q) = 1$ , if  $q \cdot m = p \cdot n$  then  $m = p$  and  $n = q$ .

Let  $m, n, p, q$  be positive integers. Assume that  $\gcd(m, n) = \gcd(p, q) = 1$  and further that

$$\textcircled{1} \quad q \cdot m = p \cdot n.$$

Multiplying both sides of the identity  $1 = \gcd(m, n)$  by  $p$  and using the linearity property of  $\gcd$  we have that

$$p = p \cdot \gcd(m, n) = \gcd(p \cdot m, p \cdot n) \quad \textcircled{2}$$

Now, from  $\textcircled{1}$  and the linearity property of  $\gcd$ , we also have that

$$\gcd(p \cdot m, p \cdot n) = \gcd(p \cdot m, q \cdot m) = \gcd(p, q) \cdot m \quad \textcircled{3}$$

Finally, since  $\gcd(p, q) = 1$ , one has  $p = m$  from  $\textcircled{2}$  and  $\textcircled{3}$ .

We can show with an analogous argument that  $n = q$  as well.

6. Prove that for all positive integers  $a$  and  $b$ ,  $\gcd(13 \cdot a + 8 \cdot b, 5 \cdot a + 3 \cdot b) = \gcd(a, b)$ .

### Calculational proof

For all positive integers  $a$  and  $b$ , one has

$$\begin{aligned} \gcd(13 \cdot a + 8 \cdot b, 5 \cdot a + 3 \cdot b) &= \gcd((13 \cdot a + 8 \cdot b) - (5 \cdot a + 3 \cdot b), 5 \cdot a + 3 \cdot b) \\ &= \gcd(8 \cdot a + 5 \cdot b, 5 \cdot a + 3 \cdot b) \\ &= \gcd((8 \cdot a + 5 \cdot b) - (5 \cdot a + 3 \cdot b), 5 \cdot a + 3 \cdot b) \end{aligned}$$

$$\begin{aligned}
&= \gcd(3 \cdot a + 2 \cdot b, 5 \cdot a + 3 \cdot b) \\
&= \gcd(3 \cdot a + 2 \cdot b, (5 \cdot a + 3 \cdot b) - (3 \cdot a + 2 \cdot b)) \\
&= \gcd(3 \cdot a + 2 \cdot b, 2 \cdot a + b) \\
&= \gcd((3 \cdot a + 2 \cdot b) - (2 \cdot a + b), 2 \cdot a + b) \\
&= \gcd(a + b, 2 \cdot a + b) \\
&= \gcd(a + b, (2 \cdot a + b) - (a + b)) \\
&= \gcd(a + b, a) \\
&= \gcd((a + b) - a, a) \\
&= \gcd(b, a) \\
&= \gcd(a, b)
\end{aligned}$$

### Conceptual proof (advanced)

We prove following general statement (see [2018/P8/Q9 exam question](#)):

$$\forall n \in \mathbb{N}. \gcd(a \cdot F_{n+3} + b \cdot F_{n+2}, a \cdot F_{n+1} + b \cdot F_n) = \gcd(a, b)$$

where  $F_n$  is the  $n^{\text{th}}$  Fibonacci number, defined recursively as

$$F_0 = 0 \quad F_1 = 1 \quad F_{n+2} = F_{n+1} + F_n$$

For  $n \in \mathbb{N}$ , we prove the following two properties, which, by the universal property of gcds, will imply the required equality.

- Both  $\gcd(a, b) \mid (aF_{n+3} + bF_{n+2})$  and  $\gcd(a, b) \mid (aF_{n+1} + bF_n)$ .

$\gcd(a, b)$  divides both  $a$  and  $b$ , so it divides every integer linear combination of them (§1.2.6(c)).

- For all positive integers  $d$ ,

$$\text{if } d \mid (aF_{n+3} + bF_{n+2}) \text{ and } d \mid (aF_{n+1} + bF_n) \text{ then } d \mid \gcd(a, b).$$

Let  $d$  be a positive integer such that  $d \mid (aF_{n+3} + bF_{n+2})$  and  $d \mid (aF_{n+1} + bF_n)$ ; so that  $di = aF_{n+3} + bF_{n+2}$  and  $dj = aF_{n+1} + bF_n$  for (positive) integers  $i$  and  $j$ .

It follows that

$$\begin{aligned}
d \cdot (iF_n - jF_{n+2}) &= (F_n \cdot F_{n+3} - F_{n+2} \cdot F_{n+1}) \cdot a \\
&= (F_n \cdot F_{n+2} + F_n \cdot F_{n+1} - F_n \cdot F_{n+1} - F_{n+1} \cdot F_{n+1}) \cdot a \\
&= (F_n \cdot F_{n+2} - F_{n+1}^2) \cdot a \\
&= (-1)^{n+1} a && \text{(Cassini's Identity)}
\end{aligned}$$

so that  $d \mid a$ ; and, analogously,

$$\begin{aligned} d \cdot (iF_{n+1} - jF_{n+3}) &= (F_{n+1} \cdot F_{n+2} - F_{n+3} \cdot F_n) \cdot a \\ &= (F_{n+1} \cdot F_{n+1} + F_n \cdot F_{n+1} - F_n \cdot F_{n+1} - F_n \cdot F_{n+2}) \cdot b \\ &= (F_{n+1}^2 - F_n \cdot F_{n+2}) \cdot b \\ &= (-1)^n b \end{aligned} \quad \text{(Cassini's Identity)}$$

so that  $d \mid b$ . Thus,  $d \mid \gcd(a, b)$  as required.

 You will learn more about Fibonacci numbers in the next set of exercises.

7. Let  $n$  be an integer.

a) Prove that if  $n$  is not divisible by 3, then  $n^2 \equiv 1 \pmod{3}$ .

This is an instance of [Fermat's Little Theorem](#).

b) Show that if  $n$  is odd, then  $n^2 \equiv 1 \pmod{8}$ .

Let  $n$  be an odd integer, and thereby let  $k$  be an integer such that  $n = 2 \cdot k + 1$ .

We consider two cases.

- Case  $k$  is even.

Then,  $k = 2 \cdot l$  for some integer  $l$ , and  $n^2 = 8 \cdot l \cdot (2 \cdot l + 1) \equiv 1 \pmod{8}$ .

- Case  $k$  is odd.

Then,  $k = 2 \cdot l + 1$  for some integer  $l$ , and  $n^2 = 8 \cdot (2 \cdot l + 1) \cdot (l + 2) + 1 \equiv 1 \pmod{8}$ .

Either way  $n^2 \equiv 1 \pmod{8}$ , as required.

c) Conclude that if  $p$  is a prime number greater than 3, then  $p^2 - 1$  is divisible by 24.

Let  $p$  be a prime greater than 3. Then,  $p$  is an odd integer not divisible by 3 and it follows from part (a) that: ①  $3 \mid (p^2 - 1)$ . Moreover, as  $p$  is odd, we have from part (b) that: ②  $8 \mid (p^2 - 1)$ .

Finally, since  $\gcd(3, 8) = 1$ , by [§3.2.2](#) one has that ① and ② imply  $24 \mid (p^2 - 1)$  as required.

8. Prove that  $n^{13} \equiv n \pmod{10}$  for all integers  $n$ .

To show  $n^{13} \equiv n \pmod{10}$ , by the direct corollary of [§3.2.2](#) it is sufficient to show  $n^{13} \equiv n \pmod{2}$  and  $n^{13} \equiv n \pmod{5}$ . Both hold by successive applications of Fermat's Little Theorem, repeatedly reducing  $n^2$  or  $n^5$  to  $n$  until we reach  $n$ . For example:

$$n^{13} = n^5 \cdot n^5 \cdot n^3 \equiv n \cdot n \cdot n^3 = n^5 \equiv n \pmod{5}$$

9. Prove that for all positive integers  $l$ ,  $m$  and  $n$ , if  $\gcd(l, m \cdot n) = 1$  then  $\gcd(l, m) = 1$  and  $\gcd(l, n) = 1$ .

Let  $l$ ,  $m$ , and  $n$  be arbitrary positive integers, and assume that  $\gcd(l, m \cdot n) = 1$ .

By §3.1.5( $\Leftarrow$ ), there exist integers  $i$  and  $j$  such that  $i \cdot l + j \cdot m \cdot n = 1$ . Thus, we have that

$$\text{there exist integers } i \text{ and } a \text{ such that } i \cdot l + a \cdot m = 1$$

and

$$\text{there exist integers } i \text{ and } b \text{ such that } i \cdot l + b \cdot n = 1.$$

Therefore, by §3.1.5( $\Rightarrow$ ) one has that  $\gcd(l, m) = 1$  and  $\gcd(l, n) = 1$ .

10. Solve the following congruences:

a)  $77 \cdot x \equiv 11 \pmod{40}$

By §3.2.4, a solution will satisfy the congruence iff it satisfies  $\textcircled{\dagger} 7 \cdot x \equiv 1 \pmod{40}$  ( $\gcd(40, 11) = 1$  so the modulus does not change). As 7 and 40 are coprime, this amounts to finding the multiplicative inverse of 7 in  $\mathbb{Z}_{40}$  (Corollary 75), which is the second coefficient in the expression of 1 as a linear combination of 40 and 7. We run the Extended Euclid's Algorithm to find that  $40 \cdot 3 + 7 \cdot (-17) = 1$ . Thus,  $x_0 = -17$  is a solution to  $\textcircled{\dagger}$ , and therefore to  $77 \cdot x_0 \equiv 11 \pmod{40}$ . To find the general form of solutions, we note that the linear combination of 40 and 7 is not unique (Slide 219), so  $x$  can have the general form  $x = -17 + 40n \equiv 23 + 40n$  for any integer  $n$ .

b)  $12 \cdot y \equiv 30 \pmod{54}$

By §3.2.4, a solution will satisfy the congruence iff it satisfies  $\textcircled{\dagger} 2 \cdot y \equiv 5 \pmod{9}$ , that is,  $2 \cdot y + 9 \cdot k = 5$  for some  $k \in \mathbb{Z}$ . Now, since 2 and 9 are coprime, we can express 1 as their linear combination, computing the coefficients using the Extended Euclid's Algorithm:  $2 \cdot (-4) + 9 \cdot 1 = 1$ . Multiplying both sides by 5 gives us  $2 \cdot (-20) + 9 \cdot 5 = 5$ , which is a solution to  $\textcircled{\dagger}$  with  $y_0 = -20$ . To generate all the solutions, we note that  $\textcircled{\dagger}$  is satisfied by  $y_0 + 9n$  for any  $n$ , so  $y$  can have the general form  $y = -20 + 9n \equiv 7 + 9n$  for any integer  $n$ .

c) 
$$\begin{cases} 13 \equiv z \pmod{21} \\ 3 \cdot z \equiv 2 \pmod{17} \end{cases}$$

To solve a system of congruences, we find the general form of solutions for the congruences individually, then find the ones that satisfy both.

All solutions to the first congruence are of the form  $z_1 = 13 + 21k$  for  $k \in \mathbb{Z}$ .

Solutions of the congruence  $3 \cdot z \equiv 2 \pmod{17}$  satisfy  $\textcircled{\dagger} 3 \cdot z + 17 \cdot n = 2$ . Since 3 and 17 are coprime, we can express 1 as their linear combination using EEA:  $3 \cdot 6 + 17 \cdot (-1) = 1$ . Multiplying by 2 on both sides gives a solution to  $\textcircled{\dagger}$ , and from there, we get the general form of solutions as  $z_2 = 12 + 17l$  for  $l \in \mathbb{Z}$ .

The solutions for the congruence system will be those which are both of the form  $z_1$



and  $z_2$  simultaneously:

$$13 + 21 \cdot k = 12 + 17 \cdot l$$

Albeit this looks like one equation with two unknowns, we can rearrange it to the form

$$21 \cdot (-k) + 17 \cdot l = 1 \quad \textcircled{\ddagger}$$


which we can solve using EEA, since 21 and 17 are coprime:

$$21 \cdot (-4) + 17 \cdot 5 = 1$$

Thus,  $\textcircled{\ddagger}$  has general solutions  $k = 4 + 17i$  and  $l = 5 + 21j$  for  $i, j \in \mathbb{Z}$ ; at these specific values of  $k$ , the general solution  $z_1 = 13 + 21 \cdot k$  for the first congruence also satisfies the second congruence (and similarly for  $z_2$ ). Substituting  $k$  into  $z_1$  or  $l$  into  $z_2$  gives

$$z = 97 + 357i \quad \forall i \in \mathbb{Z}.$$

which is the general form of solutions that satisfy the system of congruences.

 This question shows the usefulness of the characterisation of gcds via linear combinations: it allows us to solve one equation with two unknowns, as long as the RHS is a multiple of the gcd of the coefficients (so if the coefficients are coprime, the RHS can be any positive integer). Solving a congruence  $ax \equiv b \pmod{m}$  amounts to characterising the integer solutions of the equation  $ax - my = b$  (known as a linear Diophantine equation), which exist only if  $\gcd(a, m) \mid b$ .

If a congruence  $ax \equiv b \pmod{m}$  has one solution  $x_0$  (i.e. if  $\gcd(a, m) \mid b$ ), it has an infinite number of solutions of the form  $x = x_0 + pk$  for  $k \in \mathbb{Z}$ , all separated by a “period”  $p$ . In some cases (such as part (a)), the period coincides with the modulus, so all possible solutions can be derived from a single integer  $x_0 \in \mathbb{Z}_m$ . In other cases (such as part (b)) the solutions may be more “frequent” due to the period being a fraction of the modulus:  $m = dp$ . Then, the solutions  $x_0 + pk$  can be split into  $d$  classes, all with the period  $m$ , but different initial values  $x_0, x_1, \dots, x_{d-1} \in \mathbb{Z}_m$ . One such class  $\{\dots, x - 2m, x - m, x, x + m, x + 2m, \dots\}$  is often called the *congruence class of  $x$  modulo  $m$*  (denoted  $\overline{x}_m$  or sometimes  $[x]_m$ , although this course uses the latter notation to refer to the least positive element of  $\overline{x}_m$  in  $\mathbb{Z}_m$ ), so in essence, an infinite number of integer solutions to a congruence can be characterised by a finite number of congruence classes. With this interpretation, part (a) had only one solution  $\overline{23}_{40}$ , while part (b) had six:

$$\overline{7}_{54} \quad \overline{16}_{54} \quad \overline{25}_{54} \quad \overline{34}_{54} \quad \overline{43}_{54} \quad \overline{52}_{54}$$

By considering a solution to be a congruence class modulo  $m$ , we can show that a congruence  $ax \equiv b \pmod{m}$  has exactly  $\gcd(a, m)$  solutions if  $\gcd(a, m) \mid b$ , and 0 otherwise. Of course, the  $d = \gcd(a, m)$  congruence classes modulo  $m$  can be combined into one congruence class modulo  $m/d$  – the two representations are equivalent, but one may be

more useful in some contexts than the other. As an example, compare the phrases “every 8 hours starting at 1am” and “every day at 1am, 9am, and 5pm”, and how we must use the latter form to refer to events repeating regularly several times a week because 7 prime.

Since integer solutions of a congruence are not unique, we can ask which solutions of one congruence also satisfy another – that is, solve a *system of congruences*. These are quite different from the systems of equations you are familiar with, which involve  $n$  unknowns and  $n$  independent equations, and the solution is found by expressing one variable in terms of the others and performing substitutions. Congruence systems involve only one unknown, and the individual congruences are independent constraints on this one unknown. Rather than trying to combine the congruences via substitution, we solve each of them independently, getting sets of congruence classes for each individual congruence. Then, the task is finding the common elements of the congruence classes (their intersection), which therefore must satisfy the whole system of congruences simultaneously. If the individual solutions have the form  $x + pk$  and  $y + ql$ , the congruence classes  $\bar{x}_p$  and  $\bar{y}_q$  will intersect when  $x + pk = y + ql$ ; this now becomes another linear Diophantine equation of the form  $pk - ql = y - x$  that can be solved if  $\gcd(p, q) \mid y - x$ . The resulting integer values for  $k$  and  $l$  tell us the number of periods one needs to offset  $x$  and  $y$  by until they coincide, and since all solutions are uniformly periodic,  $k$  and  $l$  will themselves be periodic congruence classes. The general expressions can then be substituted back into either  $x + pk$  or  $y + ql$  to find an initial value and a larger period for the solutions that satisfy both parts of the congruence system.

As a simple example, consider the congruence classes  $\bar{1}_2$ ,  $\bar{2}_3$  and  $\bar{2}_4$ . The classes  $\bar{1}_2$  and  $\bar{2}_3$  will intersect whenever  $1 + 2n = 2 + 3k$ , and the linear Diophantine equation  $2n - 3k = 1$  has solutions  $n = 3m + 2$  and  $k = 2m + 1$ . What this means is that every 3<sup>rd</sup> ● starting from the second one (using 0-indexed counting) will coincide with every 2<sup>nd</sup> ■ starting from the first one, as can be seen below at step 5 (when  $m = 0$ ) and 11 (when  $m = 1$ ). To figure out what “every 3<sup>rd</sup> ● starting from the second one” means on the resolution of the integers, we substitute the solution for  $n$  back into  $1 + 2n$ , which combines the periods of “there is a solution at every 3<sup>rd</sup> circle” and “there is a circle every 2 steps” into “there is a solution every 6 steps” and similarly for the offset. Thus, the intersection of  $\bar{1}_2$  and  $\bar{2}_3$  will be  $\bar{5}_6$ . We can do a similar procedure to find the intersection of  $\bar{2}_3$  and  $\bar{2}_4$  to be  $\bar{2}_{12}$ . However,  $\bar{1}_2$  and  $\bar{2}_4$  will never intersect, since the Diophantine equation  $2n - 4l = 1$  has no solutions –  $\gcd(2, 4) = 2 \nmid 1$ . Congruence systems often arise from the interaction of periodic events: examples are scheduling, polyrhythms, predator-prey life cycles, etc.

|          | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|----------|---|---|---|---|---|---|---|---|---|----|----|----|----|----|
| $1 + 2n$ | ● |   | ● |   | ● |   | ● |   | ● |    | ●  |    | ●  |    |
| $2 + 3k$ |   | ■ |   |   | ■ |   |   | ■ |   |    | ■  |    |    | ■  |
| $2 + 4l$ |   | ▲ |   |   |   | ▲ |   |   |   | ▲  |    |    |    | ▲  |

11. What is the multiplicative inverse of: (a) 2 in  $\mathbb{Z}_7$ , (b) 7 in  $\mathbb{Z}_{40}$ , and (c) 13 in  $\mathbb{Z}_{23}$ ?

We apply [Corollary 75](#) of the notes, which states that if  $\gcd(m, n) = 1$ , the multiplicative inverse of  $[n]_m$  is  $[lc_2(m, n)]_m$ , where  $lc_2(m, n)$  is the second coefficient of the expression of 1 as a linear combination of  $m$  and  $n$  using EEC. With this, we get that:

- a)  $1 \cdot 7 + (-3) \cdot 2 = 1$ , so  $2^{-1} \equiv 4 \pmod{7}$   
 b)  $3 \cdot 40 + (-17) \cdot 7 = 1$ , so  $7^{-1} \equiv 23 \pmod{40}$   
 c)  $4 \cdot 23 + (-7) \cdot 13 = 1$ , so  $13^{-1} \equiv 16 \pmod{23}$

12. Prove that  $[22^{12001}]_{175}$  has a multiplicative inverse in  $\mathbb{Z}_{175}$ .

We first establish the following lemma:

For every pair of positive integers  $m$  and  $n$ , we have that  $[n]_m$  has a multiplicative inverse in  $\mathbb{Z}_m$  iff  $\gcd(m, n) = 1$ .

( $\Rightarrow$ ) Let  $m$  and  $n$  be arbitrary positive integers, and assume that  $[n]_m$  has a multiplicative inverse in  $\mathbb{Z}_m$ , say  $l$ . Then,

$$n \cdot l \equiv [n \cdot l]_m = [n]_m \cdot_m l = 1 \pmod{m}$$

and thus there exists an integer  $k$  such that  $n \cdot l + m \cdot k = 1$ . Thus, from [§3.1.5](#)( $\Rightarrow$ ),  $\gcd(m, n) = 1$ .

( $\Leftarrow$ ) By [Corollary 75\(2\)](#) of the notes.

Now,  $\gcd(22^{12001}, 175) = \gcd(2^{12001} \cdot 11^{12001}, 5^2 \cdot 7)$ , and since the two numbers have no prime factors in common, they must be coprime. By the above lemma,  $\gcd(22^{12001}, 175) = 1$  implies that  $[22^{12001}]_{175}$  has a multiplicative inverse, as required.

### 3.3. Optional exercises

1. Let  $a$  and  $b$  be natural numbers such that  $a^2 \mid b \cdot (b + a)$ . Prove that  $a \mid b$ .

*Hint:* For positive  $a$  and  $b$ , consider  $a_0 = \frac{a}{\gcd(a, b)}$  and  $b_0 = \frac{b}{\gcd(a, b)}$  so that  $\gcd(a_0, b_0) = 1$ , and show that  $a^2 \mid b(b + a)$  implies  $a_0 = 1$ .

If either  $a$  or  $b$  are 0 the result is straightforward. Consider thus the case in which both  $a$  and  $b$  are positive integers, and assume that  $a^2 \mid b(b + a)$ .

Then, for  $a_0 = \frac{a}{\gcd(a, b)}$  and  $b_0 = \frac{b}{\gcd(a, b)}$ , we have that  $a_0 \mid b_0(b_0 + a_0)$  and, since  $\gcd(a_0, b_0) = 1$ , that  $a_0 \mid (b_0 + a_0)$  so that  $a_0 \mid b_0$  and thus  $a_0 = \gcd(a_0, b_0) = 1$ . Therefore,  $\gcd(a, b) = a$  and we are done.

2. Prove the converse of [§1.3.1\(f\)](#): For all natural numbers  $n$  and  $s$ , if there exists a natural number  $q$  such that  $(2n + 1)^2 \cdot s + t_n = t_q$ , then  $s$  is a triangular number. (49<sup>th</sup> Putnam, 1988)

*Hint:* Recall that if  $\oplus q = 2nk + n + k$  then  $(2n + 1)^2 t_k + t_n = t_q$ . Solving for  $k$  in  $\oplus$ , we get that  $k = \frac{q-n}{2n+1}$ ; so it would be enough to show that the fraction  $\frac{q-n}{2n+1}$  is a natural number.

Suggested by a 2014/15 student (who wished to remain anonymous).

Assume  $(2n + 1)^2 s + t_n = t_q$ . Then,  $t_n \equiv t_q \pmod{(2n + 1)^2}$ ; so that  $n(n + 1) \equiv q(q + 1) \pmod{(2n + 1)^2}$  and hence  $(q - n)(q - n + 2n + 1) \equiv 0 \pmod{(2n + 1)^2}$ .

Therefore  $(2n + 1)^2 \mid (q - n)(q - n + 2n + 1)$ , and it follows from the previous item that  $(2n + 1) \mid (q - n)$ .

As  $t_q \geq t_n$ , we have that  $q \geq n$ , and therefore that  $k = \frac{q-n}{2n+1}$  is a natural number. By assumption and the solution to §1.3.1(f), we then have:

$$(2n + 1)^2 s + t_n = t_q = (2n + 1)^2 t_k + t_n$$

and so that  $s = t_k$ , as required.

3. Informally justify the correctness of the following alternative algorithm for computing the gcd of two positive integers:

```
let rec gcd0(m, n) = if m = n then m
                    else let p = min m n
                        and q = max m n
                        in gcd0(p, q - p)
```


The partial correctness of the algorithm follows from [Corollary 58\(2\)](#) of the notes. To establish the termination of `gcd0` on a pair of positive integers  $(m, n)$  we consider and analyse the computations arising from the call `gcd0(m, n)`. We consider two cases:

- Case  $m = n$ .

The computation of `gcd0(m, n)` reduces in one step to  $m$ , and therefore terminates.

- Case  $m \neq n$ .

The computation of `gcd0(m, n)` reduces in one step to that of `gcd0(p, q - p)`, where  $p = \min(m, n)$  and  $q = \max(m, n)$ . Thus, the passage of computing `gcd0(m, n)` by means of computing `gcd0(p, q - p)` maintains the invariant of having both components of the pair being positive integers; but, crucially, strictly decreases the sum of the pairs in each recursive call (as  $m + n > \max(m, n) = p + (q - p)$  because both  $m$  and  $n$  are positive). As this process cannot go on forever (the sum is of two strictly positive integers but decreases at every step, so the lowest it can go is  $1 + 1 = 2$ , at which point  $m = n$ ), the recursive calls must eventually stop and the overall computation terminate (in fact, in a number of steps necessarily less than or equal to the sum of the input pair).

 We can use induction to make this argument formal; see §4.3.1.

## 4. On induction

### 4.1. Basic exercises

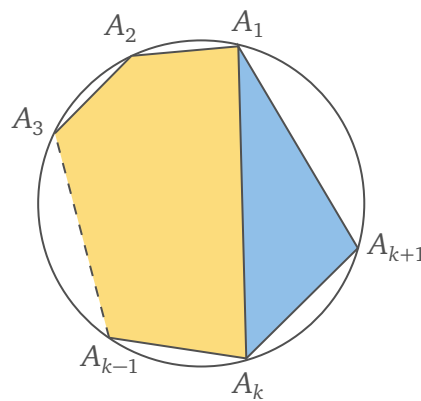
1. Prove that for all natural numbers  $n \geq 3$ , if  $n$  distinct points on a circle are joined in consecutive order by straight lines, then the interior angles of the resulting polygon add up to  $180 \cdot (n - 2)$  degrees.

We prove this property  $P(n)$  of all  $n \geq 3$  by mathematical induction from basis 3.

**Base case:**  $n = 3$ . Three connected points on a circle must form a triangle: since they are distinct, they cannot be collinear. The sum of internal angles of a triangle is  $180^\circ$ , which is  $180 \cdot (3 - 2)$  degrees.

**Inductive case:**  $n = k + 1$ . Assume that  $\textcircled{\text{H}} P(k)$  holds and take an arbitrary polygon constructed from  $k + 1$  points  $A_1, \dots, A_{k+1}$  on a circle. The  $(k + 1)$ -gon can be separated into a  $k$ -gon and a triangle with a line segment connecting  $A_1$  and  $A_k$ . By the induction hypothesis  $\textcircled{\text{H}}$ , the interior angles of the  $k$ -gon add up to  $S_k = 180 \cdot (k - 2)$  degrees. The sum of angles of the whole polygon is  $S_{k+1} = S_k + \angle A_k A_1 A_{k+1} + \angle A_1 A_{k+1} A_k + \angle A_1 A_k A_{k+1}$ , where the angle terms belong to the triangle  $\triangle A_1 A_k A_{k+1}$ . Since its interior angles must add up to  $180^\circ$ , we have the expression for the sum of internal angles of the  $(k + 1)$ -gon:

$$S_{k+1} = S_k + 180^\circ = 180 \cdot (k - 2) + 180^\circ = 180 \cdot ((k + 1) - 2)$$



$\textcircled{\text{J}}$  While the formula holds for any polygon, working with points on a circle makes the inductive proof easier, since we never need to worry about three points being on the same line and only making up one side.

$\textcircled{\text{J}}$  It may be tempting to approach the inductive step by starting with a  $k$ -gon, then *adding* a new point to turn it into a  $(k + 1)$ -gon and increasing the sum of internal angles by  $180^\circ$ . The problem with this is that we are *given* a  $(k + 1)$ -gon to start with, and its vertices are predetermined: we need to split it up into a triangle and a  $k$ -gon, no matter what the points are. This distinction is fairly minor in this case and doesn't cause any difficulties (any line segment connecting two vertices one point apart will split do the job), but remembering what parameters we have control over vs. what we are given (that is, what we need to assume as being arbitrary) is very important in proofs, especially inductive ones. We will

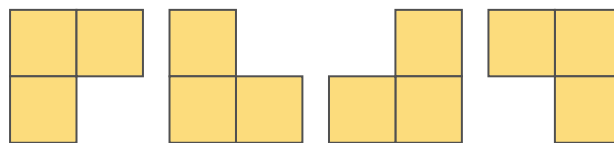
see examples of this throughout this sheet.

2. Prove that, for any positive integer  $n$ , a  $2^n \times 2^n$  square grid with any one square removed can be tiled with L-shaped pieces consisting of 3 squares.

We prove the property  $P(n)$  of all  $n \geq 1$  by mathematical induction from basis 1:

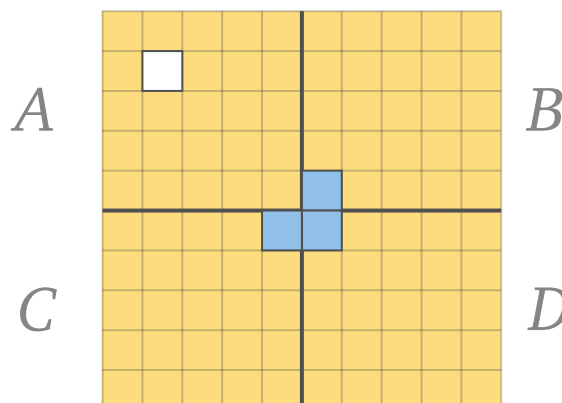
$$P(n) = \forall 0 \leq i, j \leq n. \text{ a } 2^n \times 2^n \text{ grid } A \text{ with square } A_{i,j} \text{ missing can be tiled}$$


**Base case:**  $n = 1$ . Take a  $2^1 \times 2^1 = 2 \times 2$  grid and assume one of the squares is missing. This must be one of the following four situations, depending on which one of the 4 squares was removed:



All resulting shapes can be tiled with one L-shaped piece consisting of three squares.

**Inductive step:**  $n = k + 1$ . Assume  $\textcircled{H} P(k)$ : a  $2^k \times 2^k$  grid with any square missing can be tiled with L-shaped pieces. Take a  $2^{k+1} \times 2^{k+1}$  grid with any one square missing. The grid can be split into four  $2^k \times 2^k$  quarters which we label by  $A, B, C$  and  $D$ ; assume, without loss of generality, that the missing square is in quarter  $A$  at position  $A_{i,j}$ . By the  $\textcircled{H}$  applied to  $i$  and  $j$ , the quarter  $A$  can be tiled with  $A_{i,j}$  missing. Next, we use the  $\textcircled{H}$  applied to  $i = k$  and  $j = 1$  to tile quarter  $B$  with the bottom left square missing. Similarly, we tile  $C$  and  $D$  with two applications of the induction hypothesis ( $\textcircled{H}(1, k)$  and  $\textcircled{H}(1, 1)$ , respectively) with the top right and left corners missing. The three missing corners form an L-shaped hole of 3 squares in the middle of the  $2^{k+1} \times 2^{k+1}$  grid, which can be filled in with one additional tile. This leaves only one missing square  $A_{i,j}$  with the rest of the grid tiled with L-shaped pieces, so we are done.



 This is an example of an inductive proof where the proposition  $P(n)$  is itself a universally quantified statement: we state property for all grid size parameters  $n$ , and within a particular grid of size  $2^n \times 2^n$ , for all possible grid cells that could be missing. Thus, after case-splitting on  $n$ , we still have a universally quantified proof obligation; however, in the inductive case,

we also have a universally quantified inductive assumption.

While the general pattern for proofs like this is just an instance of the standard induction principle, it is worth analysing nevertheless:

To prove a property of the form

$$\forall n \in \mathbb{N}. \forall x \in A. P(n, x)$$

it is sufficient to prove

$$\forall x \in A. P(0, x) \quad \text{and} \quad \forall k \in \mathbb{N}. (\forall y \in A. P(k, y)) \implies (\forall x \in A. P(k + 1, x))$$

The base case – which is usually seen as the “trivial” step – is now itself a universally quantified statement which may not necessarily be easy to establish. Indeed, if the inner quantification is over natural numbers as well, we may end up having to do *another* inductive proof of  $\forall m \in \mathbb{N}. P(0, m)$  if a direct proof (“Let  $m$  be an arbitrary natural number and prove  $P(0, m)$ ...”) is not possible.

The inductive step highlights the interplay between the two quantifications. Unwrapping the formula, we get three assumptions: an arbitrary natural number  $k$ , an arbitrary element  $x \in A$ , and a *proof* that  $P(k, y)$  holds for any choice of  $y \in A$ . In the process of the proof, this induction hypothesis can be applied to any element  $y \in A$ , be it  $x \in A$ , a value computed from  $x$ , or any other value arbitrarily chosen by us. There is a significant difference between the inductive step above, and a formula such as

$$\forall k \in \mathbb{N}. \forall x \in A. P(k, x) \implies P(k + 1, x)$$

which leaves us no flexibility in “tailoring” the IH to our needs by choosing an appropriate value for  $x$ .

The question above had an inner universal quantification over the position of the missing cell, so the proof cannot depend on any particular choice of position in the  $2^{k+1} \times 2^{k+1}$  grid. However, we do have control over the position of the missing cell when applying the induction hypothesis to the  $2^k \times 2^k$  quarter grids: we can essentially think of the  $\textcircled{\text{IH}}$  as a “function” from coordinates  $(i, j)$  to the proof of “tileability”. To complete the inductive step, we first apply the IH to the coordinates of the actual hole in the  $2^{k+1} \times 2^{k+1}$  within the  $A$  quarter, then select the appropriate locations for the holes in the quarters  $B$ ,  $C$  and  $D$  to leave an L-shaped hole in the middle. We apply the  $\textcircled{\text{IH}}$  both to the unknown values  $(i, j)$  given to us by the universal quantifier on the LHS of the implication, as well as values that we select deliberately to create space for an extra L-shaped tile.

$\textcircled{\text{J}}$  The phrase “without loss of generality” is often used to reduce repetition or make simplifying assumptions that do not change the strength of the result. It is usually understood that if the assumption is violated, it can be altered in an obvious way to make the rest of the proof go through. It is important to ensure that the assumption really doesn’t affect

the generality of the statement: saying things like “w.l.o.g., assume  $n$  is even/nonzero/a power of two” is sometimes tempting, but it’s rarely clear how the proof could be extended to numbers which are odd/zero/not a power of two, and proving these cases may require entirely different approaches to the one considered. Above, we assumed w.l.o.g. that the hole is in quarter  $A$  so we don’t need to repeat the proof for all four quarters. The proofs would not be exactly the same (e.g. if the hole was in quarter  $B$ , the IH would need to be applied to the  $(i - k, j)$  coordinates of the  $2^k \times 2^k$  grid), but it’s clear that the general idea would work in each case.

🎵 The proof above doesn’t just show that a tiling is possible, it gives a concrete algorithm for constructing it. Proofs like this are – unsurprisingly – called *constructive* proofs (also known as *effective* proofs to avoid confusion with constructive mathematics), as opposed to *nonconstructive* or *pure existence* proofs which show that a mathematical object exists, but doesn’t give a concrete example or way of computing one. Constructive proofs by induction naturally give rise to recursive algorithms, where the application of the  $\textcircled{\text{H}}$  corresponds to a recursive call. Of course, when implementing the recursive algorithm, we don’t have the luxury of saying that “without loss of generality, assume the user will never call the function with the hole outside of quarter  $A$ ” – we have to explicitly handle all four possibilities and slightly different recursive calls to cover any possible input.

## 4.2. Core exercises

1. Establish the following:

(a) For all positive integers  $m$  and  $n$ ,

$$(2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} = 2^{m \cdot n} - 1$$

The first thing to note is that an inductive proof is not really necessary. Indeed, for arbitrary positive integers  $m$  and  $n$ , one can calculate that

$$\begin{aligned} (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} &= \sum_{i=0}^{m-1} 2^{(i+1) \cdot n} - \sum_{i=0}^{m-1} 2^{i \cdot n} \\ &= \sum_{i=1}^{m-1} 2^{i \cdot n} + 2^{((m-1)+1) \cdot n} - 2^{0 \cdot n} - \sum_{i=1}^{m-1} 2^{i \cdot n} \\ &= 2^{m \cdot n} - 1 \end{aligned}$$

However, as it is very instructive, two inductive proofs follow. Note the different, though subtle, ways in which the inductive hypothesis is used in each proof.

For the *first proof*, we show

$$\forall m \in \mathbb{Z}^+. P(m)$$



for  $P(m)$  the statement

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} = 2^{m \cdot n} - 1$$

by the Principle of Induction.

**Base case:**  $m = 1$ . The statement  $P(1)$  amounts to

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot 2^{1 \cdot n} = 2^{1 \cdot n} - 1$$

which is vacuously true.

**Inductive step:**  $m = k + 1$ . Let  $k$  be an arbitrary positive integer, and assume that the Inductive Hypothesis  $P(k)$  holds for it; i.e. that

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot n} = 2^{k \cdot n} - 1 \quad \textcircled{\text{IH}}_1$$

We need show that  $P(k + 1)$  follows; i.e. that

$$\forall n \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{(k+1)-1} 2^{i \cdot n} = 2^{(k+1) \cdot n} - 1$$

To this end, we let  $l$  be an arbitrary positive integer and proceed to show that

$$(2^l - 1) \cdot \sum_{i=0}^k 2^{i \cdot l} = 2^{(k+1) \cdot l} - 1 \quad \textcircled{1}$$

Indeed, instantiating the  $\textcircled{\text{IH}}_1$ , we have that

$$(2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} = 2^{k \cdot l} - 1 \quad \textcircled{2}$$

and so that

$$\begin{aligned} (2^l - 1) \cdot \sum_{i=0}^k 2^{i \cdot l} &= \left( (2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} \right) + (2^l - 1) \cdot 2^{k \cdot l} \\ &= 2^{k \cdot l} - 1 + (2^l - 1) \cdot 2^{k \cdot l} && \text{(by } \textcircled{2}) \\ &= 2^{(k+1) \cdot l} - 1 \end{aligned}$$

establishing  $\textcircled{1}$  as required. □

For the *second proof*, to show

$$\forall n \in \mathbb{Z}^+. \forall m \in \mathbb{Z}^+. (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n} = 2^{m \cdot n} - 1$$

we let  $l$  be an arbitrary positive integer and prove

$$\forall m \in \mathbb{Z}^+. Q(l, m)$$

for  $Q(l, m)$  the statement

$$(2^l - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot l} = 2^{m \cdot l} - 1$$

by the Principle of Induction.

**Base case:**  $m = 1$ . The statement  $Q(l, 1)$  amounts to

$$(2^l - 1) \cdot 2^{0 \cdot l} = 2^{1 \cdot l} - 1$$

which is vacuously true.

**Inductive step:**  $m = k + 1$ . Let  $k$  be an arbitrary positive integer, and assume that the Inductive Hypothesis  $Q(l, k)$  holds for it; i.e. that

$$(2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} = 2^{k \cdot l} - 1 \quad \textcircled{\text{IH}}_2$$


We need show that  $Q(l, k + 1)$  follows; i.e. that

$$(2^l - 1) \cdot \sum_{i=0}^{(k+1)-1} 2^{i \cdot l} = 2^{(k+1) \cdot l} - 1 \quad \textcircled{1}$$

Indeed,

$$\begin{aligned} (2^l - 1) \cdot \sum_{i=0}^k 2^{i \cdot l} &= \left( (2^l - 1) \cdot \sum_{i=0}^{k-1} 2^{i \cdot l} \right) + (2^l - 1) \cdot 2^{k \cdot l} \\ &= 2^{k \cdot l} - 1 + (2^l - 1) \cdot 2^{k \cdot l} && \text{(by } \textcircled{\text{IH}}_2 \text{)} \\ &= 2^{(k+1) \cdot l} - 1 \end{aligned}$$

establishing  $\textcircled{1}$  as required.

 The core of the proof is the same in both cases; the difference is how they set up the induction hypothesis. The first proof includes the quantification over  $n$  in the  $\textcircled{\text{IH}}_1$  and applies it to the arbitrary  $l$  in the proof to get a specific instance  $\textcircled{2}$ . The second proof fixes this  $l$  right from the start, introducing it as a new arbitrary variable in the standard manner of proving universal quantification. Then, the predicate to be established by inductively is “parameterised” by this  $l$ , so the statement  $Q(l, m)$  doesn’t actually need a nested quantification. Despite  $\textcircled{\text{IH}}_2$  not containing a universal quantification, the proof only requires it at the specific  $l$  we already introduced. This makes the second proof slightly simpler, but it would not work if we ever needed the induction hypothesis at any other value of  $n$ .

(b) Suppose  $k$  is a positive integer that is not prime. Then  $2^k - 1$  is not prime.

Let  $k$  be an arbitrary positive integer. We consider two cases:

- $k = 1$ . The statement holds because  $2^1 - 1 = 1$  is not prime.
- $k \geq 2$ . Assume that  $k \geq 2$  is not prime. Hence, it is of the form  $m \cdot n$  for natural numbers  $m, n$  greater than or equal to 2. It follows from the previous item that  $2^k - 1 = 2^{m \cdot n} - 1 = (2^n - 1) \cdot \sum_{i=0}^{m-1} 2^{i \cdot n}$ ; and, since  $2^n - 1 \geq 2^2 - 1 = 3$  and  $\sum_{i=0}^{m-1} 2^{i \cdot n} \geq 1 + 4 = 5$ , we have that  $2^k - 1$  has a non-trivial decomposition. Hence it is not prime.

2. Prove that

$$\forall n \in \mathbb{N}. \forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^n \geq 1+n \cdot x$$

We prove  $\forall n \in \mathbb{N}. P(n)$  for  $P(n)$  the statement

$$\forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^n \geq 1+n \cdot x$$

by the Principle of Induction.

**Base case:**  $n = 0$ . The statement  $P(0)$  reduces to

$$\forall x \in \mathbb{R}. x \geq -1 \implies 1 \geq 1$$

and holds vacuously.

**Inductive step:**  $n = k+1$ . Let  $k$  be an arbitrary natural number, and assume  $P(k)$ ; i.e. assume the Inductive Hypothesis

$$\forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^k \geq 1+k \cdot x \quad \textcircled{\text{IH}}$$

We need show that  $P(k+1)$  also holds; i.e. that

$$\forall x \in \mathbb{R}. x \geq -1 \implies (1+x)^{k+1} \geq 1+(k+1) \cdot x$$

To this end, we let  $y$  be an arbitrary real number, assume further that

$$y \geq -1 \quad \textcircled{1}$$

and proceed to show that

$$(1+y)^{k+1} \geq 1+(k+1) \cdot y \quad \textcircled{2}$$

From  $\textcircled{\text{IH}}$ , by instantiation and Modus Ponens using  $\textcircled{1}$ , one concludes that

$$(1+y)^k \geq 1+k \cdot y$$

and from this, since by  $\textcircled{1}$  we have  $1+y \geq 0$ , it follows that

$$(1+y)^{k+1} = (1+y)^k \cdot (1+y) \geq (1+k \cdot y) \cdot (1+y) = 1+(k+1) \cdot y + k \cdot y^2$$

Thus, from the fact that  $k \cdot y^2 \geq 0$ , ② holds.

3. Recall that the Fibonacci numbers  $F_n$  for  $n \in \mathbb{N}$  are defined recursively by  $F_0 = 0$ ,  $F_1 = 1$ , and  $F_{n+2} = F_n + F_{n+1}$  for  $n \in \mathbb{N}$ .

- a) Prove Cassini's Identity: For all  $n \in \mathbb{N}$ ,

$$F_n \cdot F_{n+2} = F_{n+1}^2 + (-1)^{n+1}$$

We prove

$$\forall n \in \mathbb{N}. F_n \cdot F_{n+2} = F_{n+1}^2 + (-1)^{n+1}$$

by the Principle of Induction.

**Base case:**  $n = 0$ . We have that

$$F_0 \cdot F_2 = F_1^2 + (-1)^1$$

because  $F_0 = 0$  and  $F_1 = 1$ .

**Inductive step:**  $n = k + 1$ . For any natural number  $k$ , assume the Induction Hypothesis

$$F_n \cdot F_{k+2} = F_{k+1}^2 + (-1)^{k+1}$$

which can be rearranged to the following form by subtracting  $(-1)^{k+1}$ :

$$F_{k+1}^2 = (-1)^k + F_n \cdot F_{k+2} \quad \text{Ⓜ}$$

We need show that

$$F_{k+1} \cdot F_{(k+1)+2} = F_{(k+1)+1}^2 + (-1)^{(k+1)+1}$$

i.e. that

$$F_{k+1} \cdot F_{k+3} = F_{k+2}^2 + (-1)^k$$

for which one calculates as follows:

$$\begin{aligned} F_{k+1} \cdot F_{k+3} &= F_{k+1}^2 + F_{k+1} \cdot F_{k+2} && (F_{k+3} = F_{k+1} + F_{k+2}) \\ &= (-1)^k + F_n \cdot F_{k+2} + F_{k+1} \cdot F_{k+2} && \text{(by Ⓜ)} \\ &= (-1)^k + F_{k+2}^2 && (F_{k+2} = F_k + F_{k+1}) \end{aligned}$$

- b) Prove that for all natural numbers  $k$  and  $n$ ,

$$F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k$$

A standard one-step induction proof is possible, but the task becomes quite a bit simpler if we have two induction hypotheses.

**One-step induction proof**

We prove that

$$\forall k \in \mathbb{N}. P(k)$$

for  $P(k)$  the statement

$$\forall n \in \mathbb{N}. F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k$$

by the Principle of Induction.

**Base case:** We need show that

$$\forall n \in \mathbb{N}. F_{n+1} = F_{n+1} \cdot F_1 + F_n \cdot F_0$$

which holds because  $F_1 = 1$  and  $F_0 = 0$ .

**Inductive step:** For an arbitrary natural number  $k$ , assume the Induction Hypothesis

$$\forall n \in \mathbb{N}. F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k \quad \textcircled{\text{IH}}$$

We need show that

$$\forall n \in \mathbb{N}. F_{n+(k+1)+1} = F_{n+1} \cdot F_{(k+1)+1} + F_n \cdot F_{k+1}$$

i.e. that

$$\forall n \in \mathbb{N}. F_{n+k+2} = F_{n+1} \cdot F_{k+2} + F_n \cdot F_{k+1} \quad \textcircled{1}$$

To this end, we let  $m$  be an arbitrary natural number and proceed to show the equivalent identity:

$$F_{(m+1)+k+1} = F_{m+1} \cdot F_{k+2} + F_m \cdot F_{k+1} \quad \textcircled{2}$$


Indeed, instantiating the universally-quantified Induction Hypothesis  $\textcircled{\text{IH}}$  for the natural number  $m + 1$ , one has that

$$F_{(m+1)+k+1} = F_{(m+1)+1} \cdot F_{k+1} + F_{m+1} \cdot F_k$$

from which one further calculates as follows:

$$\begin{aligned} & F_{(m+1)+1} \cdot F_{k+1} + F_{m+1} \cdot F_k \\ &= F_m \cdot F_{k+1} + F_{m+1} \cdot F_{k+1} + F_{m+1} \cdot F_k && (F_{(m+1)+1} = F_m + F_{m+1}) \\ &= F_m \cdot F_{k+1} + F_{m+1} \cdot F_{k+2} && (F_{k+2} = F_k + F_{k+1}) \end{aligned}$$

to conclude  $\textcircled{2}$ .

 This is an example of a proposition that could also be established by nested induction: rather than show  $\textcircled{1}$  directly for an arbitrary  $n \in \mathbb{N}$ , we could do another base case for  $n = 0$  and inductive case for  $n = m + 1$ . It's not always obvious when this is required, but quite often results in a lengthier, but simpler proof.

**Two-step induction proof**

We prove that

$$\forall k \in \mathbb{N}. P(k)$$

for  $P(k)$  the statement

$$\forall n \in \mathbb{N}. F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k$$

by the Principle of Induction with two induction hypotheses.

**Base case 1:**  $k = 0$ . We need show that

$$\forall n \in \mathbb{N}. F_{n+1} = F_{n+1} \cdot F_1 + F_n \cdot F_0$$

which holds because  $F_1 = 1$  and  $F_0 = 0$ .

**Base case 2:**  $k = 1$ . We need show that

$$\forall n \in \mathbb{N}. F_{n+2} = F_{n+1} \cdot F_2 + F_n \cdot F_1$$

which holds because  $F_2 = 1$ ,  $F_1 = 1$  and  $F_{n+1} + F_n = F_{n+2}$ .

**Inductive step:** Assume the following two Induction Hypotheses:

$$F_{n+k+1} = F_{n+1} \cdot F_{k+1} + F_n \cdot F_k \quad (\text{IH}_1)$$


$$F_{n+k+2} = F_{n+1} \cdot F_{k+2} + F_n \cdot F_{k+1} \quad (\text{IH}_2)$$


We need to prove that

$$F_{n+(k+2)+1} = F_{n+1} \cdot F_{k+3} + F_n \cdot F_{k+2}$$

One calculates as follows:

$$\begin{aligned} & F_{n+k+3} \\ &= F_{n+k+1} + F_{n+k+2} \\ &= (F_{n+1} \cdot F_{k+1} + F_n \cdot F_k) + (F_{n+1} \cdot F_{k+2} + F_n \cdot F_{k+1}) \quad ((\text{IH}_1 \text{ and } \text{IH}_2)) \\ &= F_{n+1} \cdot (F_{k+1} + F_{k+2}) + F_n \cdot (F_k + F_{k+1}) \\ &= F_{n+1} \cdot F_{k+3} + F_n \cdot F_{k+2} \quad (F_{k+3} = F_{k+1} + F_{k+2} \text{ and } F_{k+2} = F_k + F_{k+1}) \end{aligned}$$

 Recognising the value of two induction hypotheses leads to a significantly simpler and more elegant proof. It is important to remember that if we go back  $k$  induction steps, we also need to prove the first  $k$  base cases.

 If either of  $k$  or  $n$  is positive, this identity gives a way of expanding  $F_{n+k}$  as a sum of products of Fibonacci numbers – a useful property whenever the index is a sum.

c) Deduce that  $F_n \mid F_{l \cdot n}$  for all natural numbers  $n$  and  $l$ .

We prove that

$$\forall l \in \mathbb{N}. P(l)$$

for  $P(l)$  the statement

$$\forall n \in \mathbb{N}. F_n \mid F_{l \cdot n}$$

by the Principle of Induction.

**Base case:** We need to show that

$$\forall n \in \mathbb{N}. F_n \mid F_{0 \cdot n}$$

i.e. that

$$\forall n \in \mathbb{N}. F_n \mid 0$$

which holds because we know that every integer divides 0 from §1.2.1(b).

**Inductive step:** For an arbitrary natural number  $l$ , assume the Induction Hypothesis

$$\forall n \in \mathbb{N}. F_n \mid F_{l \cdot n} \quad (\text{IH})$$

We need to show that

$$\forall n \in \mathbb{N}. F_n \mid F_{(l+1) \cdot n}$$


i.e. that

$$\forall n \in \mathbb{N}. F_n \mid F_{l \cdot n + n}$$

To this end, let  $n \in \mathbb{N}$  be an arbitrary natural number. We first consider the case when  $n = 0$ : we have  $F_0 \mid F_{l \cdot 0 + 0}$  from the fact that  $0 \mid 0$  (see §1.2.1(a)). Otherwise, we can express  $F_{l \cdot n + n}$  as  $F_{l \cdot n + (n-1)+1}$  and expand using §4.2.3(b) as follows:

$$\begin{aligned} & F_{l \cdot n + (n-1)+1} \\ &= F_{l \cdot n + 1} \cdot F_{(n-1)+1} + F_{l \cdot n} \cdot F_{n-1} && \text{(by §4.2.3(b))} \\ &= F_{l \cdot n + 1} \cdot F_n + k \cdot F_n \cdot F_{n-1} && \text{(by (IH), } \exists k \in \mathbb{Z}. F_{l \cdot n} = k \cdot F_n) \\ &= F_n \cdot (F_{l \cdot n + 1} + k \cdot F_{n-1}) \end{aligned}$$

Thus,  $F_{(l+1) \cdot n} = k' \cdot F_n$  for  $k' = F_{l \cdot n + 1} + k \cdot F_{n-1}$ , showing that  $F_n \mid F_{(l+1) \cdot n}$ , as required.

 Words like “deduce” and “conclude” are a dead giveaway that you should be using properties you showed in a previous part of the question, so you should always try to transform the proposition or play around with your assumptions until a previous lemma could be applied – this step often takes care of the “hard part” of the proof. In this exercise the inductive step gave us  $F_{l \cdot n + n}$ ; since the index is a sum of two natural numbers with  $n$  positive, we notice that the previous identity can be applied to expand the term into two more “manageable” subterms.

- d) Prove that  $\text{gcd}(F_{n+2}, F_{n+1})$  terminates with output 1 in  $n$  steps for all positive integers  $n$ .

We prove that

$$\forall n \in \mathbb{N}. \text{gcd}(F_{n+2}, F_{n+1}) \text{ terminates with output 1 in } n \text{ steps}$$

by the Principle of Induction.

**Base case:** We need to show that

$$\text{gcd}(F_3, F_2) \text{ terminates with output 1 in 1 step}$$

Since  $F_3 = 2$  and  $F_2 = 1$  and  $1 \mid 2$ , the algorithm terminates with the base case of  $F_2 = 1$  after one step.

**Inductive step:** For an arbitrary natural number  $k$ , assume the Induction Hypothesis

$$\text{gcd}(F_{k+2}, F_{k+1}) \text{ terminates with output 1 in } k \text{ steps} \quad \textcircled{\text{IH}}$$

We need to prove that

$$\text{gcd}(F_{k+3}, F_{k+2}) \text{ terminates with output 1 in } k + 1 \text{ steps}$$

By the definition of Fibonacci numbers,  $F_{k+3} = F_{k+2} + F_{k+1}$ . Since  $F_{k+2} \geq F_{k+1}$ , this is a valid quotient-remainder decomposition of  $F_{k+3}$  so by the Division Theorem we have that  $\text{quo}(F_{k+3}, F_{k+2}) = 1$  and  $\text{rem}(F_{k+3}, F_{k+2}) = F_{k+1}$ . As  $F_{k+1}$  is positive,  $F_{k+2} \nmid F_{k+3}$  and  $\text{gcd}(F_{k+3}, F_{k+2})$  steps to  $\text{gcd}(F_{k+2}, \text{rem}(F_{k+3}, F_{k+2})) = \text{gcd}(F_{k+2}, F_{k+1})$ . By the  $\textcircled{\text{IH}}$ , this terminates with output 1 in  $k$  steps; thus, starting with the additional computation step,  $\text{gcd}(F_{k+3}, F_{k+2})$  terminates with output 1 in  $k + 1$  steps.

e) Deduce also that:

$$(i) \text{ for all positive integers } n < m, \text{gcd}(F_m, F_n) = \text{gcd}(F_{m-n}, F_n),$$

and hence that:

$$(ii) \text{ for all positive integers } m \text{ and } n, \text{gcd}(F_m, F_n) = F_{\text{gcd}(m,n)}.$$

Firstly, we prove the following statement equivalent to (i):

For all positive integers  $n$  and natural numbers  $k$ ,

$$\text{gcd}(F_{n+k+1}, F_n) = \text{gcd}(F_{k+1}, F_n)$$

We make use of the following corollary/restatement of [Theorem 61](#), which allows us to use properties of Euclid's Algorithm in reasoning about gcds:

$$\text{For all positive integers } m \text{ and } n, \text{gcd}(m, n) = \text{gcd}(m, n).$$

In particular, we can adapt the recursive case of the definition of  $\text{gcd}$  into:

$$\forall m, n \in \mathbb{Z}^+. \text{gcd}(m, n) = \text{gcd}(\text{rem}(m, n), n) \quad \textcircled{1}$$

and the previous part [§4.2.3\(d\)](#) (shifted to positive integers) into:

$$\forall m \in \mathbb{Z}^+. \text{gcd}(F_{m+1}, F_m) = 1 \quad \textcircled{2}$$



Now, let  $n$  be a positive integer and  $k$  a natural number. Then,

$$\begin{aligned}
 \gcd(F_{n+k+1}, F_n) &= \gcd(F_{n+1} \cdot F_{k+1} + F_n \cdot F_k, F_n) && \text{(by §4.2.3(b))} \\
 &= \gcd(\text{rem}(F_{n+1} \cdot F_{k+1} + F_n \cdot F_k, F_n), F_n) && \text{(by ①)} \\
 &= \gcd(F_{n+1} \cdot F_{k+1}, F_n) && \text{(by §2.1.3(a))} \\
 &= \gcd(F_{k+1}, F_n) && \text{(by §3.2.3 and ②)}
 \end{aligned}$$

Secondly, we prove the following statement from which (ii) follows:

for all positive integers  $l, P(l)$

where  $P(l)$  is the statement:

for all positive integers  $m, n$ ,  
if  $\text{gcd}\theta(n, m)$  terminates in  $l$  steps then  $\text{gcd}(F_m, F_n) = F_{\text{gcd}(m, n)}$

for  $\text{gcd}\theta$  the function from §3.3.3. The proof is by the Principle of Induction.

**Base case:** Let  $m, n$  be arbitrary positive integers. Assume that  $\text{gcd}\theta(m, n)$  terminates in 1 step. Then  $m = n$  and  $\text{gcd}(F_m, F_n) = F_m = F_{\text{gcd}(m, n)}$ .

**Inductive step:** Let  $l$  be an arbitrary positive integer, and assume the Induction Hypothesis  $P(l)$ . Further, let  $m, n$  be arbitrary positive integers, and assume that  $\text{gcd}\theta(m, n)$  terminates in  $l + 1$  steps. Then, for  $p = \min(m, n)$  and  $q = \max(m, n)$ ,  $\text{gcd}\theta(m, n) = \text{gcd}\theta(p, q - p)$  and  $\text{gcd}\theta(p, q - p)$  terminates in  $l$  steps. Thus, by the Induction Hypothesis, we have that  $\text{gcd}(F_{q-p}, F_p) = F_{\text{gcd}(q-p, p)}$ . Finally, since by the previous item,  $\text{gcd}(F_m, F_n) = \text{gcd}(F_q, F_p) = \text{gcd}(F_{q-p}, F_p)$  and  $F_{\text{gcd}(q-p, p)} = F_{\text{gcd}(q, p)} = F_{\text{gcd}(m, n)}$  we are done.

One can intuitively deduce that property (ii) holds because we are performing the simplified Euclid's Algorithm (with repeated subtraction rather than remainder) on the indices of the Fibonacci number via a repeated application of property (i). This is indeed the case, but formulating this into a proof is far from obvious. Given that this is an exercise sheet on inductive proofs, we could try doing induction on  $m$  or  $n$ , only to notice that we can't make use of the inductive hypothesis in any meaningful way. Indeed, the "repetition" that we're trying to capture has nothing to do with the numerical value of  $m$  or  $n$  directly, but rather the number of times we have to apply property (i) to compute their gcd. Given  $m, n \in \mathbb{Z}^+$ , we either cannot apply (i) because  $m$  and  $n$  are equal, or we can apply it once to get  $\text{gcd}(F_{m-n}, F_n)$ , recursively apply it  $l$  more times to get  $F_{\text{gcd}(m-n, n)}$ , and then "unapply" one step of  $\text{gcd}\theta$  to get  $F_{\text{gcd}(m, n)}$ .

Extracting a strong enough induction hypothesis from this intuition is still nontrivial and requires us to explicitly refer to the termination of  $\text{gcd}\theta$ . Moreover,  $m$  and  $n$  are universally quantified in the induction statement and the required property  $\text{gcd}(F_m, F_n) = F_{\text{gcd}(m, n)}$  is made dependent on a termination hypothesis that refers to the induction variable  $l$ , rather than relating the two with a conjunction. This means

that when proving the inductive case, we can *assume* that  $\text{gcd}(n, m)$  terminates in more than one step, and execute one step of the algorithm manually by applying property (i). It may take several attempts to construct sufficiently strong induction hypotheses, and as this exercise shows, they are not always as direct as case-analysing on a positive/nonnegative integer that is quantified over in the proposition.

f) Show that for all positive integers  $m$  and  $n$ ,  $(F_m \cdot F_n) \mid F_{m \cdot n}$  if  $\text{gcd}(m, n) = 1$ .

Since  $m$  and  $n$  are coprime, §4.2.3(e) gives:

$$\text{gcd}(F_m, F_n) = F_{\text{gcd}(m, n)} = F_1 = 1$$

implying that  $F_m$  and  $F_n$  are themselves coprime. From §4.2.3(c) we know that  $F_m \mid F_{m \cdot n}$  and  $F_n \mid F_{m \cdot n}$ . This, together with coprimality of  $F_m$  and  $F_n$  and §3.2.2 implies that  $F_m \cdot F_n \mid F_{m \cdot n}$ , as required.

g) Conjecture and prove theorems concerning the following sums for any natural number  $n$ :

(i)  $\sum_{i=0}^n F_{2 \cdot i}$

After some test cases we conjecture the following identity:

$$\sum_{i=0}^n F_{2 \cdot i} = F_{2n+1} - 1$$

and prove it by the Principle of Induction.

**Base case:**  $n = 0$ . The sum consists of a single term  $F_{2 \cdot 0} = F_0 = 0$ , which equals  $F_{2 \cdot 0 + 1} - 1 = F_1 - 1 = 0$ .

**Inductive step:**  $n = k + 1$ . We assume the Induction Hypothesis

$$\sum_{i=0}^k F_{2 \cdot i} = F_{2k+1} - 1 \quad \text{(IH)}$$

and prove that

$$\sum_{i=0}^{k+1} F_{2 \cdot i} = F_{2(k+1)+1} - 1$$

We can calculate as follows:

$$\begin{aligned} \sum_{i=0}^{k+1} F_{2 \cdot i} &= F_{2 \cdot (k+1)} + \sum_{i=0}^k F_{2 \cdot i} \\ &= F_{2k+2} + F_{2k+1} - 1 \\ &= F_{2k+3} - 1 = F_{2(k+1)+1} - 1 \end{aligned} \quad \text{(by (IH))}$$

(ii)  $\sum_{i=0}^n F_{2 \cdot i+1}$

We conjecture the following identity:

$$\sum_{i=0}^n F_{2 \cdot i+1} = F_{2n+2}$$

and prove it by the Principle of Induction.

**Base case:**  $n = 0$ . The sum consists of a single term  $F_{2 \cdot 0+1} = F_1 = 1$ , which equals  $F_{2 \cdot 0+2} = F_2 = 1$ .

**Inductive step:**  $n = k + 1$ . We assume the Induction Hypothesis

$$\sum_{i=0}^k F_{2 \cdot i+1} = F_{2k+2} \quad \text{(IH)}$$

and prove that

$$\sum_{i=0}^{k+1} F_{2 \cdot i+1} = F_{2(k+1)+2}$$

We can calculate as follows:

$$\begin{aligned} \sum_{i=0}^{k+1} F_{2 \cdot i+1} &= F_{2 \cdot (k+1)+1} + \sum_{i=0}^k F_{2 \cdot i+1} \\ &= F_{2k+3} + F_{2k+2} \\ &= F_{2k+4} = F_{2(k+1)+2} \end{aligned} \quad \text{(by (IH))}$$

(iii)  $\sum_{i=0}^n F_i$

We conjecture the following identity:

$$\sum_{i=0}^n F_i = F_{n+2} - 1$$

We can prove this by induction as before. Instead, we derive it from the previous two results by case-analysis on  $n$ :

**Case**  $n = 2k$ . If  $k$  is 0, the sum is  $0 = F_{0+2} - 1$ . Otherwise, the sum consists of the first  $k$  even Fibonacci numbers plus the first  $(k - 1)$  odd Fibonacci numbers:

$$\sum_{i=0}^{2k} F_i = \left( \sum_{i=0}^k F_{2 \cdot i} \right) + \left( \sum_{i=0}^{k-1} F_{2 \cdot i+1} \right) = F_{2k+1} + F_{2k} - 1 = F_{2k+2} - 1$$

**Case**  $n = 2k + 1$ . The sum consists of the sum of the first  $k$  even Fibonacci numbers plus the first  $k$  odd Fibonacci numbers:

$$\sum_{i=0}^{2k+1} F_i = \left( \sum_{i=0}^k F_{2 \cdot i} \right) + \left( \sum_{i=0}^k F_{2 \cdot i+1} \right) = F_{2k+1} - 1 + F_{2k+2} = F_{(2k+1)+2} - 1$$

### 4.3. Optional exercises

1. Recall the  $\text{gcd}$  function from §3.3.3. Use the Principle of Mathematical Induction from basis 2 to formally establish the following correctness property of the algorithm:

For all natural numbers  $l \geq 2$ , we have that for all positive integers  $m, n$ , if  $m + n \leq l$  then  $\text{gcd}(m, n)$  terminates.

As suggested, we proceed by Mathematical Induction from basis 2.

**Base case:** We need show that for all positive integers  $m, n$ , if  $m + n \leq 2$  then  $\text{gcd}(m, n)$  terminates. To this end, we let  $m$  and  $n$  be arbitrary positive integers, and assume that  $m + n \leq 2$ . Then,  $m = n = 1$  and  $\text{gcd}(m, n)$  terminates.

**Inductive step:** Let  $l$  be an arbitrary natural number greater than or equal 2, and assume the Induction Hypothesis

For all positive integers  $m, n$ , if  $m + n \leq l$  then  $\text{gcd}(m, n)$  terminates. Ⓜ

We need show that for all positive integers  $m, n$ , if  $m + n \leq l + 1$  then  $\text{gcd}(m, n)$  terminates. To this end, we let  $a, b$  be arbitrary positive integers, assume that  $a + b \leq l + 1$ , and proceed to prove that  $\text{gcd}(a, b)$  terminates.

We consider three cases.

- If  $a = b$ , then  $\text{gcd}(a, b)$  terminates.
- If  $a < b$ , then  $\text{gcd}(a, b) = \text{gcd}(a, b - a)$ . Moreover, by the Inductive Hypothesis Ⓜ, we have that

if  $a + (b - a) \leq l$  then  $\text{gcd}(a, b - a)$  terminates,

and since

$$a + (b - a) = b \leq l + 1 - a \leq l$$

it follows that  $\text{gcd}(a, b - a)$  terminates and therefore that so does  $\text{gcd}(a, b)$ .

- If  $b < a$ , then  $\text{gcd}(a, b) = \text{gcd}(a, a - b)$ . Moreover, by the Inductive Hypothesis Ⓜ, we have that

if  $b + (a - b) \leq l$  then  $\text{gcd}(a, a - b)$  terminates,

and since

$$b + (a - b) = a \leq l + 1 - b \leq l$$

it follows that  $\text{gcd}(a, a - b)$  terminates and therefore that so does  $\text{gcd}(a, b)$ .

2. The set of *univariate polynomials* (over the rationals) on a variable  $x$  is defined as that of arithmetic expressions equal to those of the form  $\sum_{i=0}^n a_i \cdot x^i$ , for some  $n \in \mathbb{N}$  and some coefficients  $a_0, a_1, \dots, a_n \in \mathbb{Q}$ .

(a) Show that if  $p(x)$  and  $q(x)$  are polynomials then so are  $p(x) + q(x)$  and  $p(x) \cdot q(x)$ .

Let  $p(x) = \sum_{i=0}^m a_i \cdot x^i$  and  $q(x) = \sum_{j=0}^n b_j \cdot x^j$  be polynomials, and assume without loss of generality that  $m > n$ . For simplicity, we extend the coefficients  $a_i$  and  $b_j$  to all natural indices, with  $a_i = 0$  for  $m < i$  and  $b_j = 0$  for  $n < j$ . Then, the sum  $p(x) + q(x)$  is a polynomial (of degree  $m$ ) because it is of the form:

$$p(x) + q(x) = \sum_{i=0}^m (a_i + b_i) \cdot x^i$$

where the coefficients  $a_i + b_i$  are rational numbers since  $\mathbb{Q}$  is closed under addition.

For the product  $p(x) \cdot q(x)$ , we calculate using the distributivity of multiplication over addition:

$$\begin{aligned} p(x) \cdot q(x) &= \left( \sum_{i=0}^m a_i \cdot x^i \right) \cdot \left( \sum_{j=0}^n b_j \cdot x^j \right) \\ &= \sum_{i=0}^m \left( a_i \cdot x^i \cdot \sum_{j=0}^n b_j \cdot x^j \right) \\ &= \sum_{i=0}^m \sum_{j=0}^n a_i \cdot x^i \cdot b_j \cdot x^j \\ &= \sum_{i=0}^m \sum_{j=0}^n a_i \cdot b_j \cdot x^{i+j} \end{aligned}$$

The number of terms in the sum of a fixed degree  $d$  will be equal to the number of ways one can construct  $d$  as a sum of an  $i \leq m$  and a  $j \leq n$ ; for example there will be at most one term of degree 0 or  $m+n$ , two terms of degree  $1 = 1 + 0 = 0 + 1$  and  $m+n-1 = m + (n-1) = (m-1) + n$ , three of degree 2 and  $m+n-2$  and so on. Terms of the same degree can be combined, with their coefficients getting added together. Using our extended coefficient indexing, the coefficient of the term of degree  $k$  can be concisely expressed as:

$$c_k = \sum_{j=0}^k a_j \cdot b_{k-j}$$

As expected,  $c_0 = a_0 \cdot b_0$  (the constant terms),  $c_{m+n} = a_0 \cdot b_{m+n} + \dots + a_m \cdot b_n + \dots + a_{m+n} b_0 = 0 + \dots + a_m \cdot b_n + \dots + 0$  (most of the coefficients are “out of range” and are 0) and  $c_n = a_0 \cdot b_n + a_1 \cdot b_{n-1} + \dots + a_n \cdot b_0$  ( $n$  nonzero coefficients). Since these are all rational numbers, the product of two polynomials is indeed a polynomial (of degree  $m+n$ ) because it is of the form:

$$p(x) \cdot q(x) = \sum_{k=0}^{m+n} c_k \cdot x^k$$

- (b) Deduce as a corollary that, for all  $a, b \in \mathbb{Q}$ , the linear combination  $a \cdot p(x) + b \cdot q(x)$  of two polynomials  $p(x)$  and  $q(x)$  is a polynomial.

Every rational number  $a$  can be seen as a polynomial of degree 0, with its only coefficient being  $a$ . Thus,  $a \cdot p(x)$  is a product of polynomials and hence is a polynomial. The sum of two such expressions is still a polynomial, so we can conclude that the linear combination  $a \cdot p(x) + b \cdot q(x)$  of two polynomials for  $a, b \in \mathbb{Q}$  is a polynomial.

- (c) Show that there exists a polynomial  $p_2(x)$  such that  $p_2(n) = \sum_{i=0}^n i^2 = 0^2 + 1^2 + \dots + n^2$  for every  $n \in \mathbb{N}$ .<sup>1</sup>

*Hint:* Note that for every  $n \in \mathbb{N}$ ,

$$(n+1)^3 = \sum_{i=0}^n (i+1)^3 - \sum_{i=0}^n i^3$$

The required polynomial is

$$p_2(n) = \frac{1}{3}n^3 + \frac{1}{2}n^2 + \frac{1}{6}n$$

We show that this is a sum of squares for any  $n \in \mathbb{N}$  by induction.

**Base case:**  $n = 0$ . The polynomial reduces to 0, which is the sum of the square number  $0 = 0^2$ .

**Inductive step:**  $n = k + 1$ . Assume the Induction Hypothesis:

$$p_2(k) = \frac{1}{3}k^3 + \frac{1}{2}k^2 + \frac{1}{6}k = \sum_{i=0}^k i^2 \quad \textcircled{\text{IH}}$$

We need to prove that

$$p_2(k+1) = \sum_{i=0}^{k+1} i^2$$

The polynomial expands as follows:

$$\begin{aligned} p_2(k+1) &= \frac{1}{3}(k+1)^3 + \frac{1}{2}(k+1)^2 + \frac{1}{6}(k+1) \\ &= \frac{1}{3}k^3 + k^2 + k + \frac{1}{3} + \frac{1}{2}k^2 + k + \frac{1}{2} + \frac{1}{6}k + \frac{1}{6} \\ &= \left( \frac{1}{3}k^3 + \frac{1}{2}k^2 + \frac{1}{6}k \right) + k^2 + 2k + \frac{1}{3} + \frac{1}{2} + \frac{1}{6} \\ &= \sum_{i=0}^k i^2 + (k^2 + 2k + 1) \quad \text{(by } \textcircled{\text{IH}}) \\ &= \sum_{i=0}^k i^2 + (k+1)^2 = \sum_{i=0}^{k+1} i^2 \end{aligned}$$

Thus  $p_2(k+1)$  is the sum of consecutive squares, as required.

<sup>1</sup>Chapter 2.5 of *Concrete Mathematics* by R.L. Graham, D.E. Knuth and O. Patashnik looks at this in great detail.

♪ As is usual with existence proofs, the hard work is done behind the scenes and we start off the formal proof by magically producing a witness that just so happens to satisfy the required property. The required witness for the existence was calculated from the supplied hint:

$$\begin{aligned}
 (n+1)^3 &= \sum_{i=0}^n (i+1)^3 - \sum_{i=0}^n i^3 \\
 &= \sum_{i=0}^n (i^3 + 3i^2 + 3i + 1) - \sum_{i=0}^n i^3 \\
 &= \left( \sum_{i=0}^n 3i^2 + 3i + 1 \right) + \sum_{i=0}^n i^3 - \sum_{i=0}^n i^3 \\
 &= \sum_{i=0}^n 3i^2 + 3i + 1 = 3 \cdot \sum_{i=0}^n i^2 + \sum_{i=0}^n 3i + 1
 \end{aligned}$$

Rearranging, we get that

$$\begin{aligned}
 \sum_{i=0}^n i^2 &= \frac{1}{3} \left( (n+1)^3 - \sum_{i=0}^n 3i + 1 \right) \\
 &= \frac{1}{3} \left( n^3 + 3n^2 + 3n + 1 - \left( n + 1 + \frac{3}{2}(n^2 + n) \right) \right) \\
 &= \frac{1}{3} n^3 + n^2 + n + \frac{1}{3} - \frac{1}{3} n - \frac{1}{3} - \frac{1}{2} n^2 + \frac{1}{2} n \\
 &= \frac{1}{3} n^3 + \frac{1}{2} n^2 + \frac{1}{6} n
 \end{aligned}$$

Now, we suspect that this is the right answer, but the formal proof should start with the statement of the answer followed by a proof that it satisfies the required property. This is especially important in this case, when the proposed witness was calculated using the (unverified) hint; separately proving that the polynomial is a sum of squares makes our answer independent of the hint. The formal proof may well be done using a different technique (in this case, induction), but it should not present any unpleasant surprises since our proposed witness is almost certainly correct.

Of course, the statement for this question is a rather obfuscated way of saying “find a formula for the sum of the first  $n$  square numbers”. You may already have it memorised as

$$\sum_{i=0}^n i^2 = \frac{n(n+1)(2n+1)}{6}$$

Multiplying things out indeed leads to the formula for the polynomial  $p_2(n)$  we had above. Even if we recognise this shortcut (instead of deriving it from the hint), we still need to prove that the formula works – this is still best accomplished using induction.

- (d) Show that, for every  $k \in \mathbb{N}$ , there exists a polynomial  $p_k(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_k(n) = \sum_{i=0}^n i^k = 0^k + 1^k + \dots + n^k$ .

*Hint:* Generalise the hint above, and the similar identity

$$(n+1)^2 = \sum_{i=0}^n (i+1)^2 - \sum_{i=0}^n i^2$$

For  $k \in \mathbb{N}$ ,  $P(k)$  be the statement

There exists a polynomial  $p_k(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_k(n) = \sum_{i=0}^n i^k$ .

We prove this by the Principle of Strong Induction.

**Base case:** The polynomial needs to satisfy  $p_0(n) = \sum_{i=0}^n i^0$ ; since  $i^0 = 1$ , this is simply equal to  $p_0(n) = n + 1$ , which is a polynomial.

**Inductive step:** Assume the Strong Induction Hypothesis: for all  $0 \leq l \leq k$ ,

there exists a polynomial  $p_l(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_l(n) = \sum_{i=0}^n i^l$ .  $\textcircled{\text{IH}}_S$

We need to show that  $P(k+1)$  holds, that is

there exists a polynomial  $p_{k+1}(x)$  such that, for all  $n \in \mathbb{N}$ ,  $p_{k+1}(n) = \sum_{i=0}^n i^{k+1}$

The required witness of existence is

$$p_{k+1}(n) = \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n) \right) \quad \textcircled{\text{E}}$$

This is indeed a polynomial since:

- $p_j(n)$  is a polynomial for all  $0 \leq j \leq k$  by the Strong Induction Hypotheses, and  $\sum_{j=0}^k \binom{k+2}{j} p_j(n)$  is a linear combination of polynomials which is a polynomial;
- $(n+1)^{k+2}$  can be expanded using the Binomial Theorem into a sum of powers of  $n$  with binomial coefficients, so it too is a polynomial;
- the sum of two polynomials is a polynomial, and  $\frac{1}{k+2}$  is a rational coefficient.

We prove that  $p_{k+1}(n) = \sum_{i=0}^n i^{k+1}$  for all  $n \in \mathbb{N}$  by induction on  $n$ .

**Base case:** As before,  $p_{k+1}(0) = 0$ .

**Inductive step:** Assume the Induction Hypothesis

$$p_{k+1}(n) = \sum_{i=0}^n i^{k+1} \quad \textcircled{\text{IH}}$$

and prove that

$$p_{k+1}(n+1) = \sum_{i=0}^{n+1} i^{k+1}$$



First, we note the following two calculations:

$$(n+2)^{k+2} = ((n+1)+1)^{k+2} = \sum_{i=0}^{k+2} \binom{k+2}{i} (n+1)^i \quad (\text{Binomial Theorem})$$

$$= (n+1)^{k+2} + (k+2) \cdot (n+1)^{k+1} + \sum_{i=0}^k \binom{k+2}{i} (n+1)^i \quad (\text{extract two summands})$$

$$\sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n+1)$$

$$= \sum_{j=0}^k \binom{k+2}{j} \cdot \sum_{a=0}^{n+1} a^j = \sum_{a=0}^{n+1} \sum_{j=0}^k \binom{k+2}{j} \cdot a^j \quad (\text{by } \textcircled{\text{IH}}_S \text{ and distributivity})$$

$$= \sum_{j=0}^k \binom{k+2}{j} \cdot (n+1)^j + \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j \quad (\text{extract last summand})$$

Combining the two, we have that

$$\begin{aligned} (n+2)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n+1) \\ = (n+1)^{k+2} + (k+2) \cdot (n+1)^{k+1} - \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j \end{aligned} \quad \textcircled{1}$$

Now we are ready to expand the polynomial of the inductive step:

$$\begin{aligned} p_{k+1}(n+1) &= \frac{1}{k+2} \left( (n+2)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n+1) \right) \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} + (k+2) \cdot (n+1)^{k+1} - \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j \right) \quad (\text{by } \textcircled{1}) \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{a=0}^n \sum_{j=0}^k \binom{k+2}{j} \cdot a^j \right) + (n+1)^{k+1} \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot \sum_{a=0}^n a^j \right) + (n+1)^{k+1} \\ &= \frac{1}{k+2} \left( (n+1)^{k+2} - \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n) \right) + (n+1)^{k+1} \quad (\text{by } \textcircled{\text{IH}}_S) \\ &= p_{k+1}(n) + (n+1)^{k+1} = \sum_{i=0}^n i^{k+1} + (n+1)^{k+1} = \sum_{i=0}^{n+1} i^{k+1} \quad (\text{by } \textcircled{\text{E}} \text{ and } \textcircled{\text{H}}) \end{aligned}$$

Thus, we have shown (by the nested Mathematical Induction) that our definition of

$p_{k+1}(n)$  by  $\textcircled{E}$  indeed satisfies  $p_{k+1}(n) = \sum_{i=0}^n i^{k+1}$  for all  $n \in \mathbb{N}$ . Then, by the outer Strong Induction, we can conclude that there exists a polynomial  $p_k(n)$  for all  $k \in \mathbb{N}$  that satisfies  $p_k(n) = \sum_{i=0}^n i^k$  for all  $n \in \mathbb{N}$ .

$\textcircled{J}$  Once again, we found the witness  $\textcircled{E}$  by calculating backwards from the (conjectured) generalisation of the hint

$$(n+1)^k = \sum_{i=0}^n (i+1)^k - \sum_{i=0}^n i^k$$

We *could* prove that this holds, but we can also use it without proof to derive the witness, as long as we then formally show that the witness is correct. Given that the property is only used behind the scenes as an “educated guess”, it will not invalidate the proof even if the conjecture is actually incorrect. The calculation of the witness is as follows:

$$\begin{aligned} (n+1)^{k+2} &= \sum_{m=0}^n (m+1)^{k+2} - \sum_{m=0}^n m^{k+2} \\ &= \left( \sum_{m=0}^n \sum_{j=0}^{k+2} \binom{k+2}{j} \cdot m^j \right) - \sum_{m=0}^n m^{k+2} && \text{(Binomial Theorem)} \\ &= \left( \sum_{j=0}^{k+2} \sum_{m=0}^n \binom{k+2}{j} \cdot m^j \right) - \sum_{m=0}^n m^{k+2} && \text{(commute summation)} \\ &= \sum_{j=0}^{k+1} \sum_{m=0}^n \binom{k+2}{j} \cdot m^j && \text{(subtract last summand)} \\ &= \sum_{m=0}^n \binom{k+2}{k+1} \cdot m^{k+1} + \sum_{j=0}^k \sum_{m=0}^n \binom{k+2}{j} \cdot m^j && \text{(extract last summand)} \\ &= (k+2) \sum_{m=0}^n m^{k+1} + \sum_{j=0}^k \binom{k+2}{j} \cdot \sum_{m=0}^n m^j && \text{(binom. coefficient)} \\ &= (k+2) \sum_{m=0}^n m^{k+1} + \sum_{j=0}^k \binom{k+2}{j} \cdot p_j(n) && \text{(\textcircled{H}_S)} \end{aligned}$$

We rearrange this to get  $\sum_{m=0}^n m^{k+1}$  and set that as the witness formula for  $p_{k+1}(n)$ .

$\textcircled{J}$  This proof is a rather involved example of a nested, mixed induction proof: we do strong induction over  $k \in \mathbb{N}$  and mathematical induction over  $n \in \mathbb{N}$  when proving that our proposed witness  $\textcircled{E}$  for  $p_{k+1}(n)$  (the inductive case of the outer induction) is correct. The strong induction hypothesis  $\textcircled{H}_S$  is used throughout the proof, both in the derivation of the witness and the proof of its correctness.

$\textcircled{J}$  Note that we haven’t actually constructed a closed-form expression for  $p_k(n)$ , but a recursive algorithm for computing it from formulae for lower degrees. Importantly, we established that the recursive expression is indeed a polynomial using the clos-

ure properties proved in earlier parts. This is sufficient to prove that there exists a polynomial expression for  $\sum_{i=0}^n i^k$ , but of course one has to do quite some additional work to extract the degree and the coefficients of the polynomial from the recursive construction. The general, closed-form expression is known as Faulhaber's Formula and features the Bernoulli numbers, a rather irregular-looking sequence of rational numbers used throughout mathematics; for instance,  $B_{14} = \frac{7}{6}$ ,  $B_{15} = 0$ ,  $B_{16} = -\frac{3617}{510}$ .

## 5. On sets

### 5.1. Basic exercises

1. Prove that  $\subseteq$  is a partial order, that is, it is:

a) reflexive:  $\forall$  sets  $A$ .  $A \subseteq A$


Let  $A$  be a set; we need to show that for all  $x \in A$ ,  $x$  is in  $A$ , which follows immediately.

b) transitive:  $\forall$  sets  $A, B, C$ .  $(A \subseteq B \wedge B \subseteq C) \implies A \subseteq C$

Let  $A, B, C$  be sets and  $x \in A$  an element. We need to show that  $x \in C$ . Since  $A \subseteq B$ ,  $x \in B$ ; and since  $B \subseteq C$ ,  $x \in C$ , as required.

c) antisymmetric:  $\forall$  sets  $A, B$ .  $(A \subseteq B \wedge B \subseteq A) \iff A = B$

Let  $A, B$  be sets and suppose  $A \subseteq B$  and  $B \subseteq A$ . Then, if  $x \in A$  then  $x \in B$ , and conversely, if  $x \in B$  then  $x \in A$ . That is,  $x \in A$  if and only if  $x \in B$ , which implies that  $A$  and  $B$  are equal sets.

 Straightforward properties of the subset relation that follow from the fact that implication (in terms of which  $\subseteq$  is defined) is itself a partial order. The first two properties enable partial order reasoning to establish  $A \subseteq B$  as a chain of subset relations starting at  $A$  and ending at  $B$ ; antisymmetry gives rise to a proof technique for showing that two sets are equal iff they are both subsets of each other.

2. Prove the following statements:

a)  $\forall$  sets  $A$ .  $\emptyset \subseteq A$

Let  $A$  be a set. We need to show that every element of  $\emptyset$  is in  $A$ , but since there are no elements in  $\emptyset$ , this vacuously holds.

b)  $\forall$  sets  $A$ .  $(\forall x. x \notin A) \iff A = \emptyset$

Let  $A$  be a set.

( $\implies$ ) Assume  $\forall x. x \notin A$ . We need to show that  $A = \emptyset$ , or equivalently,  $\emptyset \subseteq A$  and  $A \subseteq \emptyset$ . The former holds by the previous property, and to show the latter, we need to prove that for all  $x$ , if  $x$  is in  $A$  then  $x$  is in  $\emptyset$ . But, by assumption,  $x \notin A$ , so the rest follows vacuously.

( $\Leftarrow$ ) Assume  $A = \emptyset$  and  $x$  an element. We need to show that  $x$  is not in  $A$ . But  $A$  is the empty set, and by definition, it has no elements; therefore  $x$  is not in  $A$ .

♪ Properties like this are sometimes harder to prove than more complicated set-theoretic statements – they just seem *too obvious* to warrant or require a proof, and attempting one feels like circular reasoning. However, if something is obvious, it should have an accompanying formal proof built from first principles (otherwise we really can't call the property obvious)! The first part of this exercise used proof by "vacuous truth", which is based on the logical principle that falsity implies anything. If we have an assumption which is false (such as that an element  $x$  is in the empty set), any conclusion could follow vacuously. A related principle is that every element in the empty set satisfies any property  $P$ :  $\forall x \in \emptyset. P(x)$ . The second part of the question could be established by simply saying that it follows from the defining property of the empty set.

3. Find the union, and intersection of:

a)  $\{1, 2, 3, 4, 5\}$  and  $\{-1, 1, 3, 5, 7\}$

Let  $A = \{1, 2, 3, 4, 5\}$  and  $B = \{-1, 1, 3, 5, 7\}$ . Then:

$$A \cup B = \{-1, 1, 2, 3, 4, 5, 7\} \quad A \cap B = \{1, 3, 5\}$$

b)  $\{x \in \mathbb{R} \mid x > 7\}$  and  $\{x \in \mathbb{N} \mid x > 5\}$

Let  $C = \{x \in \mathbb{R} \mid x > 7\}$  and  $D = \{x \in \mathbb{N} \mid x > 5\}$ . Then:

$$C \cup D = \{6\} \cup \{x \in \mathbb{R} \mid x \geq 7\} \quad C \cap D = \{x \in \mathbb{N} \mid x > 7\}$$

4. Find the Cartesian product and disjoint union of  $\{1, 2, 3, 4, 5\}$  and  $\{-1, 1, 3, 5, 7\}$ .

Let  $A = \{1, 2, 3, 4, 5\}$  and  $B = \{-1, 1, 3, 5, 7\}$ . Then:

$$\begin{aligned} A \times B = \{ & (1, -1), (1, 1), (1, 3), (1, 5), (1, 7), (2, -1), (2, 1), (2, 3), (2, 5), (2, 7), \\ & (3, -1), (3, 1), (3, 3), (3, 5), (3, 7), (4, -1), (4, 1), (4, 3), (4, 5), (4, 7), \\ & (5, -1), (5, 1), (5, 3), (5, 5), (5, 7) \} \end{aligned}$$

$$A \uplus B = \{(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, -1), (2, 1), (2, 3), (2, 5), (2, 7)\}$$

5. Let  $I = \{2, 3, 4, 5\}$  and for each  $i \in I$ , let  $A_i = \{i, i + 1, i - 1, 2 \cdot i\}$ .

a) List the elements of all sets  $A_i$  for  $i \in I$ .

$$A_2 = \{2, 3, 1, 4\} \quad A_3 = \{3, 4, 2, 6\} \quad A_4 = \{4, 5, 3, 8\} \quad A_5 = \{5, 6, 4, 10\}$$

b) Let  $\{A_i \mid i \in I\}$  stand for  $\{A_2, A_3, A_4, A_5\}$ . Find  $\bigcup\{A_i \mid i \in I\}$  and  $\bigcap\{A_i \mid i \in I\}$ .

$$\bigcup\{A_i \mid i \in I\} = \{1, 2, 3, 4, 5, 6, 8, 10\} \quad \bigcap\{A_i \mid i \in I\} = \{4\}$$

♪ The last three exercises are intended to make you comfortable with these important set-theoretic constructions through concrete examples – make sure you have a good intuition for them going forward.

6. Let  $U$  be a set. For all  $A, B \in \mathcal{P}(U)$ , prove that:

a)  $A^c = B \iff (A \cup B = U \wedge A \cap B = \emptyset)$

( $\Rightarrow$ ) Let  $A, B \in \mathcal{P}(U)$  be sets and assume  $A^c = B$ . We show that  $A \cup B = U$  and  $A \cap B = \emptyset$ . By definition,  $A \cup A^c = \{a \in U \mid x \in A \vee x \in A^c\} = \{x \in U \mid x \in A \vee x \notin A\} = \{x \in U \mid \top\} = U$ , since every element of  $U$  is either in  $A$  or not in  $A$ . Dually,  $A \cap A^c = \{x \in U \mid x \in A \wedge x \in A^c\} = \{x \in U \mid x \in A \wedge x \notin A\} = \{x \in U \mid \perp\} = \emptyset$ , since no element can be both in  $A$  and not in  $A$ .

♪ This direction proves that  $(\cdot)^c$  satisfies the complementation laws on [Slide 300](#).

( $\Leftarrow$ ) Let  $A, B \in \mathcal{P}(U)$  be sets and assume  $A \cup B = U$ ; that is, every  $x \in U$  is in  $A$  or  $B$ . This is logically equivalent to

$$\forall x \in U. x \notin A \implies x \in B$$

so  $A \cup B = U$  implies that  $A^c \subseteq B$ . Similarly, assume  $A \cap B = \emptyset$ ; that is, for every  $x \in U$ , it is not the case that  $x$  is in both  $A$  and  $B$ . This is logically equivalent to

$$\forall x \in U. x \in B \implies x \notin A$$

so  $A \cap B = \emptyset$  implies that  $B \subseteq A^c$ . By the antisymmetry of  $\subseteq$ , we conclude that  $A^c = B$ .

b) **Double complement elimination:**  $(A^c)^c = A$

Presented are two different arguments.

Ⓐ We reason using part (a): to show  $(A^c)^c = A$ , it is sufficient to show that  $A^c \cup A = U$  and  $A^c \cap A = \emptyset$ . Both of these follow from the complementation laws and the commutativity of union and intersection.

Ⓑ We can prove the equality of the sets directly via equational reasoning.

$$\begin{aligned} (A^c)^c &= \{x \in U \mid \neg(x \in A^c)\} \\ &= \{x \in U \mid \neg(x \in \{y \in U \mid y \notin A\})\} \\ &= \{x \in U \mid \neg(x \notin A)\} \\ &= \{x \in U \mid x \in A\} && \text{(double negation elimination)} \\ &= A \end{aligned}$$

♪ This self-inverse property of complementation (and negation) is called *involution*.

c) The de Morgan laws:  $(A \cup B)^c = A^c \cap B^c$  and  $(A \cap B)^c = A^c \cup B^c$

Presented are two different arguments.

Ⓐ We reason using part (a). To show the de Morgan law  $(A \cup B)^c = A^c \cap B^c$ , it is enough to show that

$$(A \cup B) \cup (A^c \cap B^c) = U \quad \text{and} \quad (A \cup B) \cap (A^c \cap B^c) = \emptyset$$

We calculate as follows:

$$\begin{aligned} & (A \cup B) \cup (A^c \cap B^c) \\ &= ((A \cup B) \cup A^c) \cap ((A \cup B) \cup B^c) && (\cup \text{ distributes over } \cap) \\ &= (B \cup (A \cup A^c)) \cap (A \cup (B \cup B^c)) && (\text{commutativity and associativity of } \cup) \\ &= (B \cup U) \cap (A \cup U) && (\text{complementation laws}) \\ &= U \cap U && (U \text{ annihilates } \cup) \\ &= U && (\text{idempotence of } \cap) \end{aligned}$$


$$\begin{aligned} & (A \cup B) \cap (A^c \cap B^c) \\ &= (A \cap (A^c \cap B^c)) \cup (B \cap (A^c \cap B^c)) && (\cap \text{ distributes over } \cup) \\ &= ((A \cap A^c) \cap B^c) \cup ((B \cap B^c) \cap A^c) && (\text{commutativity and associativity of } \cap) \\ &= (\emptyset \cap B^c) \cup (\emptyset \cap A^c) && (\text{complementation laws}) \\ &= \emptyset \cup \emptyset && (\emptyset \text{ annihilates } \cap) \\ &= \emptyset && (\text{idempotence of } \cup) \end{aligned}$$

To show the other de Morgan law  $(A \cap B)^c = A^c \cup B^c$ , one proceeds analogously or derives it from the previous de Morgan law and part (b):

$$\begin{aligned} A^c \cup B^c &= ((A^c \cup B^c)^c)^c && (\text{complement is an involution}) \\ &= ((A^c)^c \cap (B^c)^c)^c && (\text{previous de Morgan law}) \\ &= (A \cap B)^c && (\text{complement is an involution}) \end{aligned}$$

Ⓑ We can work with iff-reasoning, where the crucial third step uses the propositional de Morgan laws  $\neg(P \vee Q) \iff \neg P \wedge \neg Q$  and  $\neg(P \wedge Q) \iff \neg P \vee \neg Q$ .

$$\begin{array}{ll} x \in (A \cup B)^c \iff \neg(x \in A \cup B) & x \in (A \cap B)^c \iff \neg(x \in A \cap B) \\ \iff \neg(x \in A \vee x \in B) & \iff \neg(x \in A \wedge x \in B) \\ \iff \neg(x \in A) \wedge \neg(x \in B) & \iff \neg(x \in A) \vee \neg(x \in B) \\ \iff x \in A^c \wedge x \in B^c & \iff x \in A^c \vee x \in B^c \\ \iff x \in A^c \cap B^c & \iff x \in A^c \cup B^c \end{array}$$

 Many set-theoretic proofs involve establishing the equality of two sets, and there are several ways of formulating such proofs. Two sets  $A$  and  $B$  are equal if they have the same

elements:  $\forall x. x \in A \iff x \in B$ . Separating the bi-implication into two directions gives rise to a derived proof technique: to prove  $A = B$ , it is sufficient to prove  $A \subseteq B$  and  $B \subseteq A$ . These individual subset relations may be established element-wise ( $\forall x \in A. x \in B$  and  $\forall x \in B. x \in A$ ), or via a transitive chain of subset relations. The equality  $A = B$  itself can be shown via equivalence reasoning, either by equating set comprehensions, or using a collection of known equalities (such as the ones proved in this exercise) and “algebraic manipulation” of sets. Finally, a way to combine element-wise and calculational reasoning is via a chain of bi-implications between membership predicates, which often reduces the proof to a purely logical argument, treating “ $x \in A$ ” as an atomic proposition. All of these proof techniques are perfectly appropriate (as long as the nontrivial calculational steps are all justified): one may be easier or harder than another, depending on the problem.

♪ The non-calculational proofs in parts (A) of (b) and (c) may seem rather contrived: they are longer and fiddlier than the alternative proofs, and proceed in a very roundabout way compared to directly calculating with elements. But this is precisely what makes them illuminating: they make no reference whatsoever to notions and constructions specific to sets, like the membership relation or set comprehension. The reasoning is carried out entirely using the abstract properties of unions, intersections, and complementation, such as commutativity, distributivity, annihilation, etc. Thus, the proofs can be directly translated to any setting that supports operators with similar properties; namely order-theoretic structures called *Boolean algebras*<sup>a</sup>. Powersets of a set form a Boolean algebra (see §5.3.2), but so does the familiar set of truth values with conjunction, disjunction and negation. If logical negation is only characterised via the properties  $P \vee \neg P \iff \top$  and  $P \wedge \neg P \iff \perp$ , the proofs above show that negation must also satisfy the property  $\neg(\neg P) \iff P$  and the familiar de Morgan dualities (that we used as a *given* in the alternative proofs in part (c)).

Why bother with all this, you may ask? We can already *obviously* see that  $\neg(\neg P) \iff P$  and the de Morgan laws hold from the truth tables. The guiding principle here (and most of mathematics) is simple: results that have fewer assumptions are stronger than results that have more. The fact that these propositions follow from a small, discrete set of algebraic properties makes them stronger than if we had to rely on notions like truth tables, set comprehensions, etc., which would restrict them to the particular area of mathematics we are working with. Sure, the proofs are more cumbersome than using these “domain-specific” concepts, but they are possible and therefore need not be reproved as long as we are working with a Boolean algebra. Much of mathematics is about proving more abstract and general results than what one actually needs, because relying on fewer assumptions makes them more widely applicable.

<sup>a</sup> You should be familiar with the notion of Boolean algebra as a branch of “normal” algebra which uses truth values instead of numbers – this is what you used in Digital Electronics. Thus it may seem weird to refer to the plural “algebras” in this context. The name clash is unfortunate, but here we are discussing the *algebraic structure* known as “a Boolean algebra”, similarly to how we would discuss “a monoid” or “a field” – and Boolean algebras are the setting in which you can do Boolean algebra (the calculational process). There are other kinds of algebras too, like Heyting algebras and Lindenbaum–Tarski algebras, all with slightly different operations and properties.

## 5.2. Core exercises

1. Prove that for all for all sets  $U$  and subsets  $A, B \subseteq U$ :

$$\text{a) } \forall X. A \subseteq X \wedge B \subseteq X \iff (A \cup B) \subseteq X \quad \text{b) } \forall Y. Y \subseteq A \wedge Y \subseteq B \iff Y \subseteq (A \cap B)$$

a) Let  $X \subseteq U$  be a set.


( $\Rightarrow$ ) Assume that ①  $A \subseteq X$  and ②  $B \subseteq X$ . We need show that for all  $x \in U$ ,  $x \in A \vee x \in B$  implies  $x \in X$ . So, let  $x \in U$  and assume ③  $x \in A \vee x \in B$ . Then, if  $x \in A$  we have  $x \in X$  by assumption ①; and, if  $x \in B$  we also have  $x \in X$ , by assumption ②. Thus, assumption ③ yields  $x \in X$  as required.

( $\Leftarrow$ ) Assume  $A \cup B \subseteq X$ . Then, since  $A \subseteq A \cup B$  and  $B \subseteq A \cup B$ , we have by transitivity of  $\subseteq$  ([Lemma 84](#)) both that  $A \subseteq X$  and  $B \subseteq X$ .

b) Let  $Y \subseteq U$  be a set.

( $\Rightarrow$ ) Assume that ①  $Y \subseteq A$  and ②  $Y \subseteq B$ . We need show that for all  $y \in U$ ,  $y \in Y$  implies  $y \in A \wedge y \in B$ . So, let  $y \in U$  and assume  $y \in Y$ . Then, by assumption ①,  $y \in A$  and, by assumption ②,  $y \in B$ , as required.

( $\Leftarrow$ ) Assume  $Y \subseteq A \cap B$ . Then, since  $A \cap B \subseteq A$  and  $A \cap B \subseteq B$ , we have by transitivity of  $\subseteq$  ([Lemma 84](#)) both that  $Y \subseteq A$  and  $Y \subseteq B$ .

 These properties are also given in [Proposition 86](#), but they are reproduced here to highlight their importance. Once again, these are if-and-only-if characterisations of unions and intersections, which usually hints at an underlying universal property that uniquely describes what it means for something to be a union/intersection. In other words, these properties are the *specification* for the set-theoretic concept of a union/intersection, and the proofs above verify that the specific way we define them (via disjunction/conjunction of membership) satisfies the specification. As [Corollary 87](#) suggests, this gives rise to a proof strategy for showing that a set  $C$  is the union of  $A$  and  $B$ : if  $C$  is a superset of both  $A$  and  $B$ , and it is a subset of any other set  $X$  that is a superset of  $A$  and  $B$ , then  $C$  must equal  $A \cup B$  (and there is a dual pair of conditions for intersections). The benefit of this formulation is that some properties about unions/intersections are easier to approach via this universal property, rather than directly unwrapping the set-theoretic definitions of the operators. Moreover, just like with the complement proofs above, it gives us a way of proving properties with minimal reference to set-theoretic constructions like membership or comprehension, making them more general.

Now, you may have recognised similar formulations from last term, in an entirely different field of mathematics: number theory. Indeed, if we spell out the universal property of intersections:

$$\text{① } A \cap B \subseteq A \wedge A \cap B \subseteq B \quad \text{② } \forall Y \in \mathcal{P}(U). (Y \subseteq A \wedge Y \subseteq B) \implies Y \subseteq A \cap B$$



and compare it with the universal property of greatest common divisors:

$$\textcircled{1} \gcd(m, n) \mid m \wedge \gcd(m, n) \mid n \quad \textcircled{2} \forall d \in \mathbb{Z}^+. (d \mid m \wedge d \mid n) \implies d \mid \gcd(m, n)$$

we can notice a clear and striking similarity. The secret is that both intersections and gcds are instances of a more general construction called a *greatest lower bound*, which is a concept that can be defined (but doesn't necessarily have to exist) in any partially ordered set (poset), i.e. a set with a reflexive, transitive and antisymmetric ordering relation.

Intuitively, a greatest lower bound of two elements of a poset is the largest element that is smaller than both (it is “just below” both of them, but not smaller than necessary). For an abstract poset  $(P, \sqsubseteq)$ , the greatest lower bound of two elements  $x$  and  $y$  in  $P$  is usually denoted  $x \sqcap y$  (also an element of  $P$ , if it exists) and often called their (*binary*) *meet*. It has the properties that it is a lower bound of both  $x$  and  $y$ :

$$\textcircled{1} x \sqcap y \sqsubseteq x \wedge x \sqcap y \sqsubseteq y$$

and it is greater than any lower bound of both  $x$  and  $y$ :

$$\textcircled{2} \forall l \in P. (l \sqsubseteq x \wedge l \sqsubseteq y) \implies l \sqsubseteq x \sqcap y$$

Binary meets (just like all concepts defined via an universal property) are unique, if they exist; an important consequence is that *any* element that satisfies these two properties will be equal to  $x \sqcap y$ . This lets us prove results about  $x \sqcap y$  completely abstractly, without knowing how it is defined in a particular poset.

The properties  $\textcircled{1}$  and  $\textcircled{2}$  of intersections and gcds exactly align with those of the abstract definition of a binary meet: they are the exact same concept, manifested in different partially ordered sets. In the case of intersections, the poset is the powerset  $\mathcal{P}(U)$  with the subset ordering relation, which we proved in §5.1.1 to be a partial order. In the case of gcds, the poset is that of natural numbers, with divisibility as the ordering relation:  $d$  is “less than”  $n$  if  $d \mid n$ . This may seem like a peculiar way of ordering (in particular, 0 becomes the “greatest element” since  $d \mid 0$  for any  $d \in \mathbb{N}$ , and 1 becomes the “least element” since  $1 \mid n$  for any  $n \in \mathbb{N}$ ), but it satisfies all the required properties of being a partial order. In the divisibility poset, “lower bounds” of two numbers are their common divisors, so the “greatest lower bound” is indeed the greatest common divisor.

With this abstract understanding, we can consider binary meets in other known posets. What would the meet of two natural numbers  $m$  and  $n$  be in the standard  $\leq$  ordering? It would be the number that is less than (or equal to) both  $m$  and  $n$ , but not “too small”; the obvious choice would simply be the minimum of the two numbers, the greatest number that is not larger than either  $m$  or  $n$ . Slightly more esoteric is the partial order of Boolean truth values,  $\{\top, \perp\}$ , whose intuitive ordering (reflexive and  $\perp$  less than  $\top$ ) coincides with the Boolean operator of implication. Then, the binary meet of propositions  $P$  and  $Q$  is a proposition that implies both  $P$  and  $Q$ , and is implied by anything that implies both  $P$  and

Q. It is easy to see that the conjunction  $P \wedge Q$  satisfies this characterisation: it clearly implies both  $P$  and  $Q$ , and if  $R \implies P$  and  $R \implies Q$ , we can conclude that  $R \implies P \wedge Q$ .

And of course, by dualising everything, we get two concepts for the price of one: the dual notion of a greatest lower bound is the *least upper bound*, also called the (*binary*) *join* and denoted  $x \sqcup y$ , with universal properties:

$$\textcircled{1} x \sqsubseteq x \sqcup y \wedge y \sqsubseteq x \sqcup y \quad \textcircled{2} \forall u \in P. (x \sqsubseteq u \wedge y \sqsubseteq u) \implies x \sqcup y \sqsubseteq u$$

All the posets described above have binary joins given by the “expected” dual constructions: union, least common multiple, minimum, disjunction.

One important point to highlight is the “unique, if it exists” nature of meets and joins: they are not guaranteed to exist for any pair of elements in a poset. This is different from saying that in any monoid  $(M, \bullet, \varepsilon)$ , we can combine any two elements  $a$  and  $b$  into  $a \bullet b \in M$ . The monoid product  $a \bullet b$  is guaranteed to exist because  $\bullet: M \times M \rightarrow M$  is a binary operator, mapping two elements  $a$  and  $b$  to a new element of  $M$  – it “generates” the element  $a \bullet b$ , and since it is an operator on  $M$ , asking if  $a \bullet b$  exists in  $M$  or not is uninteresting. The symbols  $\sqcap$  and  $\sqcup$  are *not* operators: they are simply used to denote the unique meet/join of two elements of a poset, if it happens to exist. The family of sets  $\mathcal{F} = \{\emptyset, \{1\}, \{2\}\}$  is a poset under subset inclusion, and  $\emptyset$  is indeed the meet (intersection) of  $\{1\}$  and  $\{2\}$ ; but their union  $\{1, 2\}$  is not an element of  $\mathcal{F}$  and there is no other common upper bound that could be called the join, so the elements  $\{1\}$  and  $\{2\}$  in  $\mathcal{F}$  have no binary join.

Having said that, we *can* analyse sets which have binary meets and joins for all pairs of elements; such posets are called *lattices*. It so happens that all of the above examples are lattices, and some are moreover *bounded* lattices with greatest and least elements (with the exception being that  $(\mathbb{N}, \leq)$  has no greatest element). A consequence of this is that the meet and join can indeed be presented as binary operators in a lattice, since  $x \sqcap y$  and  $x \sqcup y$  are guaranteed to exist for any  $x, y \in P$ . The binary operators  $\sqcap: P \times P \rightarrow P$  and  $\sqcup: P \times P \rightarrow P$  satisfy several properties “for free”: they are associative, commutative, and idempotent. These follow – unsurprisingly – from the universal properties of the corresponding (order-theoretic) concepts: for example, to show that  $x \sqcap x = x$ , it is sufficient to show that  $x \sqsubseteq x$  (by reflexivity), and for any  $y \in P$  such that  $y \sqsubseteq x$ ,  $y \sqsubseteq x$  holds (sure). Similar proofs of commutativity and associativity are actually demonstrated in the notes, following the statements of these properties for gcds in [Lemma 63](#). Since the arguments are done purely by universal properties, they can be adapted directly to binary meets in any poset, and then specialised to other lattices like  $(\mathcal{P}(U), \subseteq)$  or  $(\mathbb{N}, \leq)$ .

2. Either prove or disprove that, for all sets  $A$  and  $B$ ,

a)  $A \subseteq B \implies \mathcal{P}(A) \subseteq \mathcal{P}(B)$

Assume  $A \subseteq B$  and let  $X \in \mathcal{P}(A)$ . Then,  $X \subseteq A$  and  $A \subseteq B$ . Hence,  $X \subseteq B$  and so  $X \in \mathcal{P}(B)$ .

b)  $\mathcal{P}(A \cup B) \subseteq \mathcal{P}(A) \cup \mathcal{P}(B)$

One can disprove it by taking two different singleton sets  $A$  and  $B$  and noticing that  $(A \cup B) \in \mathcal{P}(A \cup B)$  while it is not the case that  $(A \cup B) \in \mathcal{P}(A) \cup \mathcal{P}(B)$ . For instance,  $\{1\} \cup \{2\} = \{1, 2\} \in \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ , but  $\{1, 2\} \notin \{\emptyset, \{1\}, \{2\}\}$ .

c)  $\mathcal{P}(A) \cup \mathcal{P}(B) \subseteq \mathcal{P}(A \cup B)$

Assume  $X \in \mathcal{P}(A) \cup \mathcal{P}(B)$ ; that is, either ①  $X \in \mathcal{P}(A)$  or ②  $X \in \mathcal{P}(B)$ .

In case ①,  $X \subseteq A$  and since  $A \subseteq (A \cup B)$  we have  $X \subseteq (A \cup B)$ ; and hence  $X \in \mathcal{P}(A \cup B)$ .


In case ②,  $X \subseteq B$  and since  $B \subseteq (A \cup B)$  we have  $X \subseteq (A \cup B)$ ; and hence  $X \in \mathcal{P}(A \cup B)$ .

d)  $\mathcal{P}(A \cap B) \subseteq \mathcal{P}(A) \cap \mathcal{P}(B)$

Assume  $X \in \mathcal{P}(A \cap B)$ ; that is  $X \subseteq (A \cap B)$  or, equivalently,  $X \subseteq A$  and  $X \subseteq B$ . Hence,  $X \in \mathcal{P}(A)$  and  $X \in \mathcal{P}(B)$ ; so that  $X \in \mathcal{P}(A) \cap \mathcal{P}(B)$ .

e)  $\mathcal{P}(A) \cap \mathcal{P}(B) \subseteq \mathcal{P}(A \cap B)$

Assume  $X \in \mathcal{P}(A) \cap \mathcal{P}(B)$ ; that is,  $X \in \mathcal{P}(A)$  and  $X \in \mathcal{P}(B)$ . Then,  $X \subseteq A$  and  $X \subseteq B$ ; so that  $X \subseteq (A \cap B)$  and hence  $X \in \mathcal{P}(A \cap B)$ .

 Parts (d) and (e) used the universal property of intersections from §5.2.1. We could have formulated the proof for (c) entirely using universal properties: to show  $\mathcal{P}(A) \cup \mathcal{P}(B) \subseteq \mathcal{P}(A \cup B)$ , it is sufficient to show that  $\mathcal{P}(A \cup B)$  is an upper bound of  $\mathcal{P}(A)$  and  $\mathcal{P}(B)$ ; but both follow from the fact that  $A \subseteq A \cup B$ ,  $B \subseteq A \cup B$ , and part (a) which lifts these subset relations to powersets. A UP proof for (d) is very similar: to show  $\mathcal{P}(A \cap B) \subseteq \mathcal{P}(A) \cap \mathcal{P}(B)$ , it is sufficient to show that  $\mathcal{P}(A \cap B)$  is a common subset of  $\mathcal{P}(A)$  and  $\mathcal{P}(B)$ , both of which follow from lifting  $A \cap B \subseteq A$  and  $A \cap B \subseteq B$  to powersets. It is not necessarily the case that a UP proof is shorter or simpler than one done from first principles, but it often allows for higher-level reasoning than going down to element-wise definitions. The rule of thumb should be: if the proof goal is of the form  $A \cup B \subseteq X$  or  $Y \subseteq A \cap B$ , a UP proof may be possible since it is sufficient to show that  $X$  is a common superset or  $Y$  is a common subset.

3. Let  $U$  be a set. For all  $A, B \in \mathcal{P}(U)$  prove that the following statements are equivalent.

$$\text{a) } A \cup B = B \quad \text{b) } A \subseteq B \quad \text{c) } A \cap B = A \quad \text{d) } B^c \subseteq A^c$$

Let  $U$  be a set and consider  $A, B \in \mathcal{P}(U)$ . To prove that the statements are equivalent, it is sufficient that they cyclically imply each other.

(a)  $\Rightarrow$  (b) Assume  $A \cup B = B$ . Then,  $A \subseteq (A \cup B) = B$  and we are done.

(b)  $\Rightarrow$  (c) Assume  $A \subseteq B$ . Since,  $(A \cap B) \subseteq A$  we need only show  $A \subseteq (A \cap B)$  or, by §5.2.1, that  $A \subseteq A$  and  $A \subseteq B$ ; which respectively hold by reflexivity of  $\subseteq$  and assumption.

(c)  $\Rightarrow$  (d) Assume  $(A \cap B) = A$  and let  $x \in U$ . Then,  $x \notin B$  implies  $x \notin (A \cap B) = A$ .

(d)  $\Rightarrow$  (b) Because  $B^c \subseteq A^c$  stands for  $x \notin B \implies x \notin A$  for all  $x \in U$  which is the

contrapositive of  $x \in A \implies x \in B$  for all  $x \in U$ .

**(b)  $\implies$  (a)** Assume  $A \subseteq B$ . Since also  $B \subseteq B$ , by §5.2.1 above, we have  $(A \cup B) \subseteq B$ ; and as  $B \subseteq (A \cup B)$  we are done.

♪ Questions of the form “prove that the following  $n$  statements are equivalent” require one to prove (at least)  $n$  implications that form a cycle; thanks to the transitivity and symmetry of implication, this is sufficient to take care of a bi-implication between any two statements. The order and number of implications proved is not important, as long as there is a way to get from any statement to another. In this question we could have done the chain  $(a) \implies (b) \implies (c) \implies (d) \implies (a)$ , but  $(d)$  is simply a contrapositive of  $(b)$  so the implication  $(d) \implies (b)$  is easier. However, this only results in one outgoing implication from  $(a)$ , so that needed to be “patched” up with an extra (straightforward) proof.

4. For sets  $A, B, C, D$ , prove or disprove at least three of the following statements:

a)  $(A \subseteq C \wedge B \subseteq D) \implies A \times B \subseteq C \times D$

Assume  $A \subseteq C$  and  $B \subseteq D$ , and let  $(a, b)$  be an element of  $A \times B$ . We need to show that  $(a, b)$  is also an element of  $C \times D$ . Since  $a \in A$  and  $A \subseteq C$ , we have that  $a \in C$ ; similarly,  $b \in D$ . Thus,  $(a, b) \in C \times D$ , as required.

♪ We slightly glossed over a few formal steps here, but this is rarely an issue. What we mean by “let  $(a, b)$  be an element of  $A \times B$ ” is that we consider an element  $x \in A \times B$  and use the fact that all elements of  $A \times B$  are pairs; so  $x$  must be of the form  $(a, b)$  for an  $a \in A$  and  $b \in B$ . Such “pattern-matching” is very common in formal proofs and needs not be elaborated on too much, unless the patterns are interesting in their own right or the relationships between the sets are more complex.

b)  $(A \cup C) \times (B \cup D) \subseteq (A \times B) \cup (C \times D)$

This statement is false. As a counterexample, consider  $A = \{1\}$ ,  $D = \{2\}$  and  $B = C = \emptyset$ .  $A \cup C = \{1\}$  and  $B \cup D = \{2\}$ ; the first Cartesian product is thus  $\{(1, 2)\}$ . However,  $A \times B = C \times D = \emptyset$ , so their union is also the empty set, not a superset of  $\{(1, 2)\}$ .

c)  $(A \times C) \cup (B \times D) \subseteq (A \cup B) \times (C \cup D)$

By the universal property of unions it is sufficient to prove that  $A \times C \subseteq (A \cup B) \times (C \cup D)$  and  $B \times D \subseteq (A \cup B) \times (C \cup D)$ . Part (a) implies the former with  $A \subseteq A \cup B$  and  $C \subseteq C \cup D$ , as well as the latter with  $B \subseteq A \cup B$  and  $D \subseteq C \cup D$ .

♪ Again we notice that the question asks us to prove that the union of two sets  $X$  and  $Y$  is below another set  $Z$ , for which it is sufficient to prove that  $Z$  is a common superset of both  $X$  and  $Y$  so the least common superset (a.k.a. the union) is necessarily going to be below it. The requirements can be discharged using property (a), that lets us

“apply” subset relations within components of a Cartesian product. Note how we do not need to refer to elements of the sets at all.

d)  $A \times (B \cup C) \subseteq (A \times B) \cup (A \times C)$

Let  $(a, x)$  be an element of  $A \times (B \cup C)$ . By the definition of unions,  $x$  is either in  $B$  or in  $C$ ; in the former case the tuple  $(a, x)$  is in  $A \times B$ , which is a subset of  $(A \times B) \cup (A \times C)$ ; in the latter,  $(a, x) \in A \times C$  which is again a subset of the union  $(A \times B) \cup (A \times C)$ , as required.

e)  $(A \times B) \cup (A \times D) \subseteq A \times (B \cup D)$

By the universal property of unions it is sufficient to prove that  $A \times B \subseteq A \times (B \cup D)$  and  $A \times D \subseteq A \times (B \cup D)$ . Both follow from property (a) and the fact that  $B \cup D$  is a common superset of both  $B$  and  $D$ .

5. For sets  $A, B, C, D$ , prove or disprove at least three of the following statements:

a)  $(A \subseteq C \wedge B \subseteq D) \implies A \uplus B \subseteq C \uplus D$

Assume  $A \subseteq C$  and  $B \subseteq D$ , and let  $x$  be an element of  $A \uplus B$ . We need to prove that  $x$  is in  $C \uplus D$ . By the definition of disjoint union,  $x$  is either of the form  $(1, a)$  for  $a \in A$ , or  $(2, b)$  for  $b \in B$ . In the first case we use the assumption  $A \subseteq C$  to derive that  $a \in C$ , but then  $(1, a)$  is in  $C \uplus D$ . In the second case we use the assumption  $B \subseteq D$  to derive that  $b \in D$ , so  $(2, b)$  is in  $C \uplus D$ , as required.

b)  $(A \cup B) \uplus C \subseteq (A \uplus C) \cup (B \uplus C)$

Let  $x$  be an element of  $(A \cup B) \uplus C$ . By the definition of unions and disjoint unions, we consider three cases:  $x$  is of the form  $(1, a)$  with  $a \in A$ , or  $(1, b)$  for  $b \in B$ , or  $(2, c)$  for  $c \in C$ . In the first case,  $(1, a)$  is in  $A \uplus C$  and therefore in  $(A \uplus C) \cup (B \uplus C)$ ; similarly, in the second case,  $(1, b)$  is in  $B \uplus C$  so in  $(A \uplus C) \cup (B \uplus C)$ . Finally, in the third case,  $(2, c)$  is in both  $A \uplus C$  and  $B \uplus C$ , so it will certainly be in  $(A \uplus C) \cup (B \uplus C)$ .

c)  $(A \uplus C) \cup (B \uplus C) \subseteq (A \cup B) \uplus C$

By the UP of unions it is sufficient to prove that  $A \uplus C \subseteq (A \cup B) \uplus C$  and  $B \uplus C \subseteq (A \cup B) \uplus C$ . Both follow using part (a) and the fact that  $A \cup B$  is a superset of both  $A$  and  $B$ .

d)  $(A \cap B) \uplus C \subseteq (A \uplus C) \cap (B \uplus C)$

By the UP of intersections it is sufficient to prove that  $(A \cap B) \uplus C$  is in  $A \uplus C$  and in  $B \uplus C$ . Both follow using part (a) and the fact that  $A \cap B$  is a common subset of both  $A$  and  $B$ .

e)  $(A \uplus C) \cap (B \uplus C) \subseteq (A \cap B) \uplus C$

Let  $x$  be an element of  $(A \uplus C) \cap (B \uplus C)$ . By definition of intersections it must be both in  $A \uplus C$  and  $B \uplus C$ , which is possible if it has the same tag: either  $x$  is of the form  $(1, y)$  where  $y$  is both in  $A$  and  $B$ , or of the form  $(2, c)$  with  $c \in C$ . In the first case  $(1, y)$  is the first injection of  $(A \cap B) \uplus C$ , and in the second case  $(2, c)$  is the second injection of  $(A \cap B) \uplus C$ .

6. Prove the following properties of the big unions and intersections of a family of sets  $\mathcal{F} \subseteq \mathcal{P}(A)$ :

$$\text{a) } \forall U \subseteq A. (\forall X \in \mathcal{F}. X \subseteq U) \iff \bigcup \mathcal{F} \subseteq U$$

Let  $U$  be a subset of  $A$ .

( $\Rightarrow$ ) Assume that  $U$  is a superset of every element of  $\mathcal{F}$  and let  $x$  be a member of  $\bigcup \mathcal{F}$ . We need to show that  $x$  is also in  $U$ . By the definition of big unions, there exists a set  $F \in \mathcal{F}$  such that  $x \in F$ ; but since  $U$  is a superset of every set in  $\mathcal{F}$ , we know that  $F \subseteq U$  and therefore that  $x \in U$ .

( $\Leftarrow$ ) Assume  $\bigcup \mathcal{F} \subseteq U$  and let  $X$  be a set in  $\mathcal{F}$ . Since  $\bigcup \mathcal{F}$  is the union of all sets in  $\mathcal{F}$ , we know that  $X \subseteq \bigcup \mathcal{F}$ , and by transitivity with the first assumption we can conclude that  $X \subseteq U$ , as required.

$$\text{b) } \forall L \subseteq A. (\forall X \in \mathcal{F}. L \subseteq X) \iff L \subseteq \bigcap \mathcal{F}$$

Let  $L$  be a subset of  $A$ .

( $\Rightarrow$ ) Assume that  $L$  is a subset of every element of  $\mathcal{F}$  and let  $x$  be a member of  $L$ . We need to show that  $x$  is also in  $\bigcap \mathcal{F}$ , that is, it is a member of every set in  $\mathcal{F}$ . To this end, let  $F$  be an arbitrary element of  $\mathcal{F}$ . By assumption, we know that  $L \subseteq F$ , but then  $x \in L$  must also be in  $F$ . Since  $F$  was arbitrary, this holds for every element of  $\mathcal{F}$ , so indeed  $x \in \bigcap \mathcal{F}$ .

( $\Leftarrow$ ) Assume  $L \subseteq \bigcap \mathcal{F}$  and let  $X$  be a set in  $\mathcal{F}$ . Since  $\bigcap \mathcal{F}$  is the intersection of all sets in  $\mathcal{F}$ , we know that  $\bigcap \mathcal{F} \subseteq X$ , and by transitivity with the first assumption we can conclude that  $L \subseteq X$ , as required.

$\square$  These two propositions generalise the universal properties of unions and intersections. The union of a family of sets  $\mathcal{F}$  is the least common superset of all the sets in the family:

$$\textcircled{1} \forall X \in \mathcal{F}. X \subseteq \bigcup \mathcal{F} \quad \textcircled{2} \forall U \subseteq A. (\forall X \in \mathcal{F}. X \subseteq U) \implies \bigcup \mathcal{F} \subseteq U$$

Dually, the intersection is the greatest common subset of all sets in  $\mathcal{F}$ :

$$\textcircled{1} \forall X \in \mathcal{F}. \bigcap \mathcal{F} \subseteq X \quad \textcircled{2} \forall L \subseteq A. (\forall X \in \mathcal{F}. L \subseteq X) \implies L \subseteq \bigcap \mathcal{F}$$

You should be able to read such properties with relative ease: the first one says that the big union is an upper bound of all elements in  $\mathcal{F}$ , and the second one characterises it as

the smallest such set. As in the binary case, the real power of universal properties comes when proving statements of the form  $\bigcup \mathcal{F} \subseteq U$  or  $L \subseteq \bigcap \mathcal{F}$ , since all one needs to show next is that  $U$  and  $L$  are upper and lower bounds, respectively.

7. Let  $A$  be a set.

a) For a family  $\mathcal{F} \subseteq \mathcal{P}(A)$ , let  $\mathcal{U} \triangleq \{U \subseteq A \mid \forall S \in \mathcal{F}. S \subseteq U\}$ . Prove that  $\bigcup \mathcal{F} = \bigcap \mathcal{U}$ .

The family  $\mathcal{U}$  is the set of upper bounds of  $\mathcal{F}$ , i.e. the family of sets which are all supersets of every set in  $\mathcal{F}$ .

( $\subseteq$ ) By the universal property of big unions (§5.2.6), it is sufficient to prove that  $\forall X \in \mathcal{F}. X \subseteq \bigcap \mathcal{U}$ . By the universal property of intersections, for this it is sufficient to prove that  $\forall X \in \mathcal{F}. \forall U \in \mathcal{U}. X \subseteq U$ , and this holds by the definition of  $\mathcal{U}$  as the set of upper bounds of  $\mathcal{F}$ .

( $\supseteq$ ) We know from the universal property of big unions that  $\bigcup \mathcal{F}$  is an upper bound, so  $\forall S \in \mathcal{F}. S \subseteq \bigcup \mathcal{F}$ . But then  $\bigcup \mathcal{F}$  must be in  $\mathcal{U}$ , the set of upper bounds. By the UP of intersections,  $\bigcap \mathcal{U}$  is a subset of every element of  $\mathcal{U}$ , and in particular,  $\bigcap \mathcal{U} \subseteq \bigcup \mathcal{F}$ , as required.

b) Analogously, define the family  $\mathcal{L} \subseteq \mathcal{P}(A)$  such that  $\bigcap \mathcal{F} = \bigcup \mathcal{L}$ . Also prove this statement.

The family  $\mathcal{L}$  is the set of lower bounds of  $\mathcal{F}$ , that is,

$$\mathcal{L} \triangleq \{L \subseteq A \mid \forall S \in \mathcal{F}. L \subseteq S\}$$

We prove that  $\bigcap \mathcal{F} = \bigcup \mathcal{L}$ .

( $\subseteq$ ) We know from the universal property of big intersections that  $\bigcap \mathcal{F}$  is a lower bound, so  $\forall S \in \mathcal{F}. \bigcap \mathcal{F} \subseteq S$ . But then  $\bigcap \mathcal{F}$  must be in  $\mathcal{L}$ , the set of lower bounds. By the UP of unions,  $\bigcup \mathcal{L}$  is a superset of every element of  $\mathcal{L}$ , and in particular,  $\bigcap \mathcal{F} \subseteq \bigcup \mathcal{L}$ , as required.

( $\supseteq$ ) By the universal property of big unions (§5.2.6), it is sufficient to prove that  $\forall L \in \mathcal{L}. L \subseteq \bigcap \mathcal{F}$ . By the universal property of intersections, for this it is sufficient to prove that  $\forall L \in \mathcal{L}. \forall X \in \mathcal{F}. L \subseteq X$ , and this holds by the definition of  $\mathcal{L}$  as the set of lower bounds of  $\mathcal{F}$ .

♪ One could approach such questions from first principles, by expanding all definitions and reasoning purely by logic. However, it would be rather cumbersome, and the higher-level proof techniques derived from universal properties have a clear advantage. They may take some getting used to, but it's worth the practice.

### 5.3. Optional advanced exercises

1. Prove that for all families of sets  $\mathcal{F}_1$  and  $\mathcal{F}_2$ ,

$$(\bigcup \mathcal{F}_1) \cup (\bigcup \mathcal{F}_2) = \bigcup (\mathcal{F}_1 \cup \mathcal{F}_2)$$

State and prove the analogous property for intersections of non-empty families of sets.

The stated identity for unions is a special case of the associativity law for big unions, so let us just consider the case of intersections; that is: for non-empty collections of sets  $\mathcal{F}_1$  and  $\mathcal{F}_2$ ,

$$\left(\bigcap \mathcal{F}_1\right) \cap \left(\bigcap \mathcal{F}_2\right) = \bigcap (\mathcal{F}_1 \cup \mathcal{F}_2)$$

Indeed, for all  $x$ , we have

$$\begin{aligned} x \in \left(\bigcap \mathcal{F}_1\right) \cap \left(\bigcap \mathcal{F}_2\right) &\iff (x \in \bigcap \mathcal{F}_1) \wedge (x \in \bigcap \mathcal{F}_2) \\ &\iff (\forall X \in \mathcal{F}_1. x \in X) \wedge (\forall X \in \mathcal{F}_2. x \in X) \\ &\iff \forall X. (X \in \mathcal{F}_1 \Rightarrow x \in X) \wedge (X \in \mathcal{F}_2 \Rightarrow x \in X) \\ &\iff \forall X. (X \in \mathcal{F}_1 \vee X \in \mathcal{F}_2) \Rightarrow x \in X \\ &\iff \forall X. X \in (\mathcal{F}_1 \cup \mathcal{F}_2) \Rightarrow x \in X \\ &\iff x \in \bigcap (\mathcal{F}_1 \cup \mathcal{F}_2) \end{aligned}$$

2. For a set  $U$ , prove that  $(\mathcal{P}(U), \subseteq, \cup, \cap, U, \emptyset, (\cdot)^c)$  is a **Boolean algebra**.

Let  $U$  be a set. We have the following:

- $\subseteq: \mathcal{P}(U) \times \mathcal{P}(U) \rightarrow \mathbb{B}$  is a partial order, as shown in §5.1.1.
- Every two sets  $A$  and  $B$  have a union  $A \cup B$  which is their least upper bound, as well as an intersection  $A \cap B$  which is their greatest lower bound (§5.2.1). It follows from the universal properties that both operations are commutative, associative, and idempotent.
- The full set  $U$  is the neutral element of intersection: given any set  $A \subseteq U$ , it is the case that  $A \cap U = A$  by §5.2.3.
- The empty set  $\emptyset$  is the neutral element of union: given any set  $A \subseteq U$ , we know that  $\emptyset \subseteq A$  so §5.2.3 implies that  $A \cup \emptyset = A$ .
- Similarly using §5.2.3 we can deduce that  $U$  is the annihilator for union (since  $A \subseteq U$  implies  $A \cup U = U$ ) and  $\emptyset$  is the annihilator for intersection (since  $\emptyset \subseteq A$  implies  $\emptyset \cap A = \emptyset$ ).
- To show the absorption laws, we let  $A$  and  $B$  be subsets of  $U$  and prove that  $A \cup (A \cap B) = A$ . Let  $x$  be an element of  $A \cup (A \cap B)$ ; by definition, it has to be either in  $A$  or in  $A \cap B$ , i.e. in  $A$  or in both  $A$  and  $B$ . In both cases  $x$  must be in  $A$ . Conversely, assume  $x \in A$ ; then it is clearly in  $A \cup (A \cap B)$ , as required. The second absorption law is similar.
- To show distributivity, we let  $A, B$  and  $C$  be subsets of  $U$  and prove that  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ . Let  $x$  be an element of  $A \cap (B \cup C)$ ; by definition, it has to be both in  $A$  and either in  $B$  or in  $C$ . If it is in  $B$ , then it is also in  $A \cap B$  and hence in  $(A \cap B) \cup (A \cap C)$ . Otherwise, if it is in  $C$ , then it is in  $A \cap C$  and hence in  $(A \cap B) \cup (A \cap C)$ . The other distributive law is similar.

Thus we can conclude that  $(\mathcal{P}(U), \subseteq, \cup, \cap, U, \emptyset, (\cdot)^c)$  is indeed a Boolean algebra.



As the name implies, a Boolean algebra is an algebraic (or order-theoretic) structure that generalises Boolean truth values and operators. As an algebraic structure, it is a carrier set with two idempotent, commutative and associative operators that distribute over each other and are absorptive; two elements that are units for one operator and annihilators for the other; and a unary complement operator. As an order-theoretic structure, it is a complemented, distributive lattice; that is, a poset in which every element has a meet and a join (i.e. a lattice, see §5.3.2) which distribute over each other, and every element has a complement.

As this exercise shows, powersets of a set also form a Boolean algebra – this underlies the intuitive similarity between logical operators (conjunction, disjunction) and set operators (intersection, union). An interesting question to ponder is the status of *implication*: it does not form part of the algebraic structure and is instead defined as  $P \Rightarrow Q \triangleq \neg P \vee Q$ . Set-theoretically the corresponding notion would be  $A^c \cup B$ , i.e. all the elements of the universe except the ones that are exclusively in  $A$  – not a particularly common or useful notion! We can also choose to axiomatise logic in terms of implication, and define negation as  $\neg P \triangleq P \Rightarrow \perp$ . A lattice with least and greatest elements and an appropriately characterised “implication” operator is called a *Heyting algebra*. Every Boolean algebra is a Heyting algebra with the implication defined as above, but not every Heyting algebra is a Boolean algebra – as a consequence, some logical tautologies like double negation elimination  $\neg(\neg P) \Rightarrow P$  or the law of excluded middle  $P \vee \neg P$  do not in general hold in a Heyting algebra. The distinction between Boolean and Heyting algebras is the distinction between *classical* and *intuitionistic* logic, the latter of which is particularly important in computer science and will be covered in much detail in future courses.

## 6. On relations

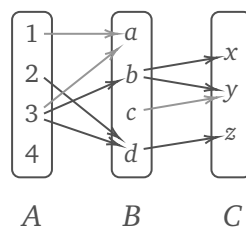
### 6.1. Basic exercises

1. Let  $A = \{1, 2, 3, 4\}$ ,  $B = \{a, b, c, d\}$  and  $C = \{x, y, z\}$ .

Let  $R = \{(1, a), (2, d), (3, a), (3, b), (3, d)\}: A \rightarrow B$

and  $S = \{(b, x), (b, y), (c, y), (d, z)\}: B \rightarrow C$ .

Draw the internal diagrams of the relations. What is the composition  $S \circ R: A \rightarrow C$ ?



The composite  $S \circ R$  is the relation  $\{(2, z), (3, x), (3, y), (3, z)\}$ .

2. Prove that relational composition is associative and has the identity relation as the neutral element.

Let  $R: A \leftrightarrow B$ ,  $S: B \leftrightarrow C$ , and  $T: C \leftrightarrow D$  be three relations. We show that their composition is associative:  $T \circ (S \circ R) = (T \circ S) \circ R$ . Take a pair  $(a, d) \in T \circ (S \circ R)$ ; by the definition of relational composition, there must be a  $c \in C$  such that  $(a, c) \in S \circ R$  and  $c T d$ ; expanding the former, there must be a  $b \in B$  such that  $a R b$  and  $b S c$ . But then  $(b, d) \in T \circ S$  via  $c$ , and  $(a, d) \in (T \circ S) \circ R$  via  $b$ . The converse proof follows analogously from the definition, so we conclude that relational composition is associative.

Let  $R: A \leftrightarrow B$  be a relation. We show that  $\text{id}_B \circ R = R = R \circ \text{id}_A$ . Take a pair  $(a, b) \in \text{id}_B \circ R$ ; there must exist a  $b' \in B$  such that  $a R b'$  and  $b' \text{id}_B b$ , but since the identity relation is the equality, we have that  $b' = b$  and therefore  $(a, b) \in R$ . Conversely, to show that  $(a, b) \in R$  is also in  $\text{id}_B \circ R$ , we observe that  $b$  can be used as the intermediate step in showing that  $(a, b) \in R$  and  $(b, b) \in \text{id}_B$ . The right inverse proof is analogous, so we conclude that the identity relation is the two-sided unit of relational composition.

3. For a relation  $R: A \leftrightarrow B$ , let its *opposite*, or *dual relation*,  $R^{\text{op}}: B \leftrightarrow A$  be defined by:

$$b R^{\text{op}} a \iff a R b$$

For  $R, S: A \leftrightarrow B$  and  $T: B \leftrightarrow C$ , prove that:

a)  $R \subseteq S \implies R^{\text{op}} \subseteq S^{\text{op}}$

Assume  $R \subseteq S$  and show that for all  $b R^{\text{op}} a$ ,  $b S^{\text{op}} a$ . By the definition of opposite relations,  $b R^{\text{op}} a$  if  $a R b$ , but by assumption,  $a S b$  and thus  $b S^{\text{op}} a$ , as required.

b)  $(R \cap S)^{\text{op}} = R^{\text{op}} \cap S^{\text{op}}$

By the previous part and UP of intersections, we have that  $(R \cap S)^{\text{op}} \subseteq R^{\text{op}}$  and  $(R \cap S)^{\text{op}} \subseteq S^{\text{op}}$ , so  $(R \cap S)^{\text{op}} \subseteq R^{\text{op}} \cap S^{\text{op}}$ . Conversely, take a pair  $(b, a)$  in  $R^{\text{op}}$  and  $S^{\text{op}}$ ; then,  $(a, b)$  is both in  $R$  and  $S$  so it is in the intersection and  $(b, a) \in (R \cap S)^{\text{op}}$ .

c)  $(R \cup S)^{\text{op}} = R^{\text{op}} \cup S^{\text{op}}$

For  $(b, a) \in (R \cup S)^{\text{op}}$ , we calculate as follows:

$$\begin{aligned} (b, a) \in (R \cup S)^{\text{op}} &\iff (a, b) \in (R \cup S) \\ &\iff (a R b \vee a S b) \\ &\iff (b R^{\text{op}} a \vee b S^{\text{op}} a) \\ &\iff (b, a) \in R^{\text{op}} \cup S^{\text{op}} \end{aligned}$$

d)  $(T \circ S)^{\text{op}} = S^{\text{op}} \circ T^{\text{op}}$

We calculate as follows:


$$\begin{aligned}
 (T \circ S)^{\text{op}} &= \{(c, a) \mid (c, a) \in (T \circ S)^{\text{op}}\} \\
 &= \{(c, a) \mid (a, c) \in T \circ S\} \\
 &= \{(c, a) \mid \exists b \in B. a S b \wedge b T c\} \\
 &= \{(c, a) \mid \exists b \in B. b S^{\text{op}} a \wedge c T^{\text{op}} b\} \\
 &= \{(c, a) \mid (c, a) \in S^{\text{op}} \circ T^{\text{op}}\} = S^{\text{op}} \circ T^{\text{op}}
 \end{aligned}$$

As before, these questions concern the equality of sets which can be established in several ways; three possibilities (universal properties, bi-implication reasoning and set comprehension reasoning) are demonstrated here.

## 6.2. Core exercises

- Let  $R, R' \subseteq A \times B$  and  $S, S' \subseteq B \times C$  be two pairs of relations and assume  $R \subseteq R'$  and  $S \subseteq S'$ . Prove that  $S \circ R \subseteq S' \circ R'$ .

Assume  $(a, c) \in (S \circ R)$ . Hence, there exists  $b \in B$  such that  $a R b$  and  $b S c$ . Then, since  $(a, b) \in R$  and  $R \subseteq R'$ , we have that  $(a, b) \in R'$ ; similarly,  $(b, c) \in S'$ . By the definition of composition, this implies that  $(a, c) \in S' \circ R'$ , as required.

 A simple, but useful lemma which states that subset relationships can be applied on both operands of relational composition. We have seen similar properties for powersets (§5.2.2(a)), Cartesian products (§5.2.4(a)) and disjoint unions (§5.2.5(a)). As usual, special cases of this property can be derived by expanding only one of the two operands: for example,  $S' \circ R$  and  $S \circ R'$ .

- Let  $\mathcal{F} \subseteq \mathcal{P}(A \times B)$  and  $\mathcal{G} \subseteq \mathcal{P}(B \times C)$  be two collections of relations from  $A$  to  $B$  and from  $B$  to  $C$ , respectively. Prove that

$$(\bigcup \mathcal{G}) \circ (\bigcup \mathcal{F}) = \bigcup \{S \circ R \mid R \in \mathcal{F}, S \in \mathcal{G}\}: A \leftrightarrow C$$

Recall that the notation  $\{S \circ R: A \leftrightarrow C \mid R \in \mathcal{F}, S \in \mathcal{G}\}$  is common syntactic sugar for the formal definition  $\{T \in \mathcal{P}(A \times C) \mid \exists R \in \mathcal{F}. \exists S \in \mathcal{G}. T = S \circ R\}$ . Hence,

$$T \in \{S \circ R \in A \leftrightarrow C \mid R \in \mathcal{F}, S \in \mathcal{G}\} \iff \exists R \in \mathcal{F}. \exists S \in \mathcal{G}. T = S \circ R$$

( $\subseteq$ ) We show:  $(\bigcup \mathcal{G}) \circ (\bigcup \mathcal{F}) \subseteq \bigcup \{S \circ R \mid R \in \mathcal{F}, S \in \mathcal{G}\}$ .

Assume  $(a, c) \in (\bigcup \mathcal{G}) \circ (\bigcup \mathcal{F})$ . Hence, there exists  $b \in B$  such that  $(a, b) \in \bigcup \mathcal{F}$  and  $(b, c) \in \bigcup \mathcal{G}$ . Then, by the definition of big unions, we have  $a R b$  for some  $R \in \mathcal{F}$  and  $b S c$  for some  $S \in \mathcal{G}$  so it follows that  $(a, c) \in S \circ R$  for some  $R \in \mathcal{F}$  and  $S \in \mathcal{G}$ . That is,  $(a, c) \in \bigcup \{S \circ R \mid R \in \mathcal{F}, S \in \mathcal{G}\}$ .

( $\supseteq$ ) By the universal property of unions, we have that  $\bigcup \{S \circ R \mid R \in \mathcal{F}, S \in \mathcal{G}\} \subseteq (\bigcup \mathcal{G}) \circ$

$(\bigcup \mathcal{F})$  if and only if  $S \circ R \subseteq (\bigcup \mathcal{G}) \circ (\bigcup \mathcal{F})$  for all  $R \in \mathcal{F}$  and  $S \in \mathcal{G}$ . This is the case by §6.2.1 and the fact that  $R \subseteq \bigcup \mathcal{F}$  for all  $S \in \mathcal{F}$  and  $S \subseteq \bigcup \mathcal{G}$  for all  $S \in \mathcal{G}$ , since the big unions are upper bounds.

♪ One direction required a direct proof of membership, but the other direction was of the form  $\bigcup \mathcal{U} \subseteq X$  and therefore could be approached via the universal property of big unions as the least upper bound of a family of sets; to show that it is below  $X$ , it is sufficient to show that every element of the family  $\mathcal{U}$  is below  $X$ .

What happens in the case of big intersections?

One direction follows in both cases from the universal property of intersections:

$$(\bigcap \mathcal{G}) \circ (\bigcap \mathcal{F}) \subseteq \bigcap \{S \circ R \mid R \in \mathcal{F}, S \in \mathcal{G}\}$$

However, the other inclusion fails. Consider a pair  $(a, c) \in \bigcap \{S \circ R \mid R \in \mathcal{F}, S \in \mathcal{G}\}$ : it means that for all  $R \in \mathcal{F}$  and  $S \in \mathcal{G}$ , there exists a  $b_{R,S} \in B$  such that  $(a, b_{R,S}) \in R$  and  $(b_{R,S}, c) \in S$ . We need to show that  $(a, c) \in (\bigcap \mathcal{G}) \circ (\bigcap \mathcal{F})$ , that is, there exists a  $b \in B$  such that for all  $R \in \mathcal{F}$ ,  $a R b$ , and for all  $S \in \mathcal{G}$ ,  $b S c$ . Note the order of quantification: our assumption produces an intermediate  $b_{R,S}$  for any choices of  $S$  and  $R$  (and the  $b_{R,S}$ s may be different depending on the choice), while the goal asks for a single  $b \in B$  that acts as an intermediate for every relation in  $\mathcal{F}$  and  $\mathcal{G}$ . Since we won't be able to find such a single  $b$  in general, this direction cannot hold. Abstractly, we only have the implication  $\exists x. \forall y. P(x, y) \implies \forall y. \exists x. P(x, y)$  but not the other direction; this was not an issue with union since existentials can be swapped.

3. Suppose  $R$  is a relation on a set  $A$ . Prove that

a)  $R$  is reflexive iff  $\text{id}_A \subseteq R$

$R$  is reflexive iff for all  $a \in A$ ,  $a R a$ . Equivalently, for all  $a, a' \in A$ , if  $a = a'$  then  $a R a'$ . Since the identity relation is equality, this is equivalent to  $\text{id}_A$  being a subset of  $R$ .

b)  $R$  is symmetric iff  $R = R^{\text{op}}$

$R$  is symmetric iff for all  $a, b \in A$ , if  $a R b$  then  $b R a$ . Equivalently, we can express this as  $a R b$  implying  $a R^{\text{op}} b$ , or  $b R^{\text{op}} a$  implying  $b R a$ . These conditions in turn say that  $R \subseteq R^{\text{op}}$  and  $R^{\text{op}} \subseteq R$ , so  $R = R^{\text{op}}$  is equivalent to  $R$  being symmetric.

c)  $R$  is transitive iff  $R \circ R \subseteq R$

$R$  is transitive iff for all  $a, b, c \in A$ , if  $(a, b) \in R$  and  $(b, c) \in R$  then  $(a, c) \in R$ . We first assume  $R$  is transitive and prove that  $R \circ R \subseteq R$  by taking a pair  $(a, c) \in R \circ R$ . By the definition of relation composition, there exists a  $b \in A$  such that  $a R b$  and  $b R c$ , but  $R$  is transitive, so  $a R c$ . Conversely, assume  $R \circ R \subseteq R$  and suppose  $a R b$  and  $b R c$  for three elements  $a, b, c \in A$ . Then,  $(a, c) \in R \circ R$  via  $b$ , and by assumption,  $a R c$ , proving that  $R$  is transitive.

♪ The calculational proof of this property would depend on the equivalence

$$\begin{aligned} & \forall a, c \in A. (\exists b \in A. aRb \wedge bRc) \implies aRc \\ \iff & \forall a, c \in A. \forall b \in A. (aRb \wedge bRc) \implies aRc \end{aligned}$$

which is precisely an instance of the equivalence of the formulae  $(\exists x. P(x)) \implies Q$  and  $\forall x. (P(x) \implies Q)$  way back from §1.3.2.

d)  $R$  is antisymmetric iff  $R \cap R^{\text{op}} \subseteq \text{id}_A$

$R$  is antisymmetric iff for all  $a, b \in A$ ,  $aRb$  and  $bRa$  implies  $a = b$ . This is equivalent to the statement that  $aRb$  and  $aR^{\text{op}}b$  implies  $a = b$ , that is,  $(a, b) \in R \cap R^{\text{op}}$  implies  $(a, b) \in \text{id}_A$ . This, in turn, is equivalent to  $R \cap R^{\text{op}} \subseteq \text{id}_A$ .

♪ These are sufficient and necessary conditions for establishing properties of relations in terms of set-theoretic operators rather than element-wise proofs. As before, having the ability to reason without “going down to the level of elements” often results in more direct and elegant proofs that capture the algebraic nature of set-level calculations; in addition, not having to introduce a lot of new variable names for elements make such proofs less finicky and error-prone as well.

4. Let  $R$  be an arbitrary relation on a set  $A$ , for example, representing an undirected graph. We are interested in constructing the smallest transitive relation (graph) containing  $R$ , called the *transitive closure* of  $R$ : a relation  $\text{Cl}_t[R]$  that satisfies ①  $R \subseteq \text{Cl}_t[R]$ ; ②  $\text{Cl}_t[R]$  is transitive; and ③  $\text{Cl}_t[R]$  is the smallest such relation.

a) We define the family of relations which are transitive supersets of  $R$ :

$$\mathcal{T}_R \triangleq \{Q: A \leftrightarrow A \mid R \subseteq Q \text{ and } Q \text{ is transitive}\}$$

$R$  is not necessarily going to be an element of this family, as it might not be transitive. However,  $R$  is a *lower bound* for  $\mathcal{T}_R$ , as it is a subset of every element of the family.

Prove that the set  $\bigcap \mathcal{T}_R$  is the transitive closure for  $R$ .

We need to prove that  $\bigcap \mathcal{T}_R$  is the ③ smallest ② transitive relation ① containing  $R$ .

① By the UP of intersections,  $R \subseteq \bigcap \mathcal{T}_R$  holds iff  $R \subseteq Q$  for all  $Q \in \mathcal{T}_R$ ; but by definition of  $\mathcal{T}_R$  we have that  $R$  must be a subset of all its elements.

② To show that  $\bigcap \mathcal{T}_R$  is transitive, it is sufficient to show that  $\bigcap \mathcal{T}_R \circ \bigcap \mathcal{T}_R \subseteq \bigcap \mathcal{T}_R$  by §6.2.3. By the UP of intersections (similar to §6.2.2),  $\bigcap \mathcal{T}_R \circ \bigcap \mathcal{T}_R \subseteq \bigcap \{Q \circ Q \mid Q \in \mathcal{T}_R\}$ , but since all  $Q \in \mathcal{T}_R$  are transitive,  $Q \circ Q \subseteq Q$  and thus  $\bigcap \{Q \circ Q \mid Q \in \mathcal{T}_R\} \subseteq \bigcap \{Q \mid Q \in \mathcal{T}_R\} = \bigcap \mathcal{T}_R$ .

③ To show that  $\bigcap \mathcal{T}_R$  is the smallest transitive superset of  $R$ , we let  $S$  be a transitive relation with  $R \subseteq S$  and prove that  $\bigcap \mathcal{T}_R \subseteq S$ . Since  $S$  is transitive and  $R \subseteq S$ , it must also be an element of  $\mathcal{T}_R$ , and by the UP of intersections,  $\bigcap \mathcal{T}_R$  is a subset of every

element of  $T_R$ , in particular  $S$ .

- b)  $\bigcap \mathcal{T}_R$  is the intersection of an infinite number of relations so it's difficult to compute the transitive closure this way. A better approach is to start with  $R$ , and keep adding the missing connections until we get a transitive graph. This can be done by repeatedly composing  $R$  with itself: after  $n$  compositions, all paths of length  $n$  in the graph represented by  $R$  will have a transitive connection between their endpoints.

Prove that the (at least once) iterated composition  $R^{\circ+} \triangleq R \circ R^{\circ*}$  is the transitive closure for  $R$ , i.e. it coincides with the greatest lower bound of  $\mathcal{T}_R$ :  $R^{\circ+} = \bigcap \mathcal{T}_R$ . *Hint*: show that  $R^{\circ+}$  is both an element and a lower bound of  $\mathcal{T}_R$ .

By the definition of  $R^{\circ*}$  and §6.2.2 (with  $\mathcal{F} = \{R^{\circ k} \mid k \in \mathbb{N}\}$  and  $\mathcal{G} = \{R\}$ ), we have that

$$R^{\circ+} = R \circ R^{\circ*} = R \circ \bigcup \{R^{\circ k} \mid k \in \mathbb{N}\} = \bigcup \{R \circ R^{\circ k} \mid k \in \mathbb{N}\} = \bigcup_{n \in \mathbb{N}^+} R^{\circ n}$$

where  $\mathbb{N}^+$  is the set of positive natural numbers. Again, we show that  $\bigcup_{n \in \mathbb{N}^+} R^{\circ n}$  is the ③ smallest ② transitive relation ① containing  $R$ , where ① and ② amounts to proving that  $R^{\circ+} \in \mathcal{T}_R$  and ③ that it is a lower bound of  $\mathcal{T}_R$ .

① We have that  $R \subseteq \bigcup_{n \in \mathbb{N}^+} R^{\circ n}$  since  $R = R^{\circ 1}$  is an element of the indexed family and big unions are upper bounds.

② To show that  $R^{\circ+}$  is transitive, it is sufficient to show that  $R^{\circ+} \circ R^{\circ+} \subseteq R^{\circ+}$ . By §6.2.2, we have the following:

$$R^{\circ+} \circ R^{\circ+} = \left( \bigcup_{n \in \mathbb{N}^+} R^{\circ n} \right) \circ \left( \bigcup_{m \in \mathbb{N}^+} R^{\circ m} \right) = \bigcup_{n \in \mathbb{N}^+} \bigcup_{m \in \mathbb{N}^+} R^{\circ n} \circ R^{\circ m}$$

To proceed, we prove the following lemma: for all  $k, l \in \mathbb{N}$ ,  $R^{\circ k} \circ R^{\circ l} = R^{\circ(k+l)}$ .

**Base case:**  $k = 0$ . Then,  $R^{\circ 0} \circ R^{\circ l} = \text{id}_A \circ R^{\circ l} = R^{\circ(0+l)}$  since the identity relation is a left unit for composition.

**Inductive step:**  $k + 1$ . Assume the  $\textcircled{H}$ :  $R^{\circ k} \circ R^{\circ l} = R^{\circ(k+l)}$ . By definition of iterated composition,  $R^{\circ(k+1)} \circ R^{\circ l} = (R \circ R^{\circ k}) \circ R^{\circ l}$ , but since relational composition is associative, this equals  $R \circ (R^{\circ k} \circ R^{\circ l})$  which, by the  $\textcircled{H}$ , is  $R \circ R^{\circ(k+l)} = R^{\circ((k+1)+l)}$ , as required.

By this lemma,  $\bigcup_{n \in \mathbb{N}^+} \bigcup_{m \in \mathbb{N}^+} R^{\circ n} \circ R^{\circ m} = \bigcup_{n \in \mathbb{N}^+} \bigcup_{m \in \mathbb{N}^+} R^{\circ(n+m)}$ . Now, to show that  $\bigcup_{n \in \mathbb{N}^+} \bigcup_{m \in \mathbb{N}^+} R^{\circ(n+m)} \subseteq R^{\circ+}$ , we can use the UP of big unions twice and equivalently establish

$$\forall n \in \mathbb{N}^+. \forall m \in \mathbb{N}^+. R^{\circ(n+m)} \subseteq R^{\circ+}$$

but this is the case because  $R^{\circ(n+m)} \in \{R^{\circ k} \mid k \in \mathbb{N}^+\}$  and big unions are upper bounds. Thus, we have shown that  $R^{\circ+} \circ R^{\circ+} \subseteq R^{\circ+}$ , and by §6.2.3, it is transitive.

③ We need to show that  $R^{\circ+}$  is the smallest such relation, i.e. it is a lower bound of

$\mathcal{T}_R$ . By the UP of unions, we equivalently have

$$\bigcup_{n \in \mathbb{N}^+} R^{\circ n} \subseteq \bigcap \mathcal{T}_R \iff \forall n \in \mathbb{N}^+. \forall Q \in \mathcal{T}_R. R^{\circ n} \subseteq Q$$

The latter statement can be proved by induction on  $n$ .

**Base case:**  $n = 1$ . We need to show that for all  $Q \in \mathcal{T}_R$ ,  $R^{\circ 1} = R \subseteq Q$ ; but this is the case since  $R \subseteq Q$  by the definition of  $\mathcal{T}_R$ .

**Inductive step:**  $n = k + 1$ . Assume the  $\textcircled{\text{H}}$ :  $\forall Q \in \mathcal{T}_R. R^{\circ k} \subseteq Q$ . We need to prove that  $\forall Q \in \mathcal{T}_R. R^{\circ(k+1)} \subseteq Q$ . Let  $Q \in \mathcal{T}_R$  be such a relation, and show that  $R^{\circ(k+1)} = R \circ R^{\circ k} \subseteq Q$ . By the induction hypothesis,  $R^{\circ k} \subseteq Q$  and  $R \subseteq Q$  by assumption on  $Q$ , so §6.2.1 implies that

$$R \circ R^{\circ k} \subseteq Q \circ Q \subseteq Q$$

where the last step follows from the fact that  $Q$  is transitive. Thus,  $R^{\circ(k+1)} \subseteq Q$ . By the principle of mathematical induction, we have that  $\forall n \in \mathbb{N}^+. R^{\circ n} \subseteq Q$  for all  $Q \in \mathcal{T}_R$ , so  $R^{\circ+}$  is indeed a lower bound of  $\mathcal{T}_R$ .

Putting everything together, we have that  $R^{\circ+}$  is the transitive closure of  $R$ , as required.

🎵 A rather involved proof with many distinct steps, references to established properties and several proof techniques. Notice, however, that at no point did we have to reason about elements of the relations: we got to the end without ever having to say “take  $(a, a') \in R^{\circ+}$ ”, for example. It would have been possible to get a low-level proof like this, but expanding all definitions and resorting to purely logical reasoning is often lengthier and more error-prone. Gaining the fluency to work with universal properties and recognising common patterns (sufficient conditions for transitivity, operand-wise application of subsets in composition, etc.) is a worthwhile, time-saving skill to learn for discrete mathematics and other mathematical subjects.

🎵 The concept of a *closure* is a common and powerful tool for characterising mathematical constructions. Abstractly, we say that a set  $A$  is *closed* under an  $n$ -ary operation  $f$  if it maps  $n$  elements of  $A$  to an element of  $A$ ; that is, if the operation can be represented as a function  $f : A^n \rightarrow A$ . Familiar examples are addition and multiplication on natural numbers, union and intersection on  $\mathcal{P}(U)$  for a set  $U$ , list concatenation on the set of all lists of some type. However, natural numbers are not closed under subtraction (e.g.  $2 - 5 = -3 \notin \mathbb{N}$ ), odd numbers are not closed under addition (e.g.  $3 + 5 = 8$ ), subsets of  $U$  are not closed under Cartesian product (because the output is a set of pairs in  $U \times U$ , not an element of  $U$ ), etc.

More generally, we can talk about the closure of a set under some property  $P$ : for example  $P(G) \iff G$  is transitive, or  $P(A) \iff A$  is closed under operation  $f$ .

Naturally, we may be interested in taking a set  $A$  and turning it into one that is closed under a particular property  $P$  “with the least amount of effort”. In particular, we don’t

want to do anything if  $A$  already satisfies  $P$ ; but if it doesn't, we only want to add the minimal number of extra elements to make it so, not more. Thus, we want to construct the set  $\text{Cl}_P[A]$  with the following properties: ① it should certainly contain all elements of  $A$ , so  $A \subseteq \text{Cl}_P[A]$ ; ② it should satisfy property  $P$ ; and ③ it should be the smallest superset of  $A$  which satisfies  $P$ . Hopefully you recognise this as a universal construction, defining the smallest set  $\text{Cl}_P[A]$  in the family  $\mathcal{C}_P$  defined as:

$$\mathcal{C}_P \triangleq \{ C \mid A \subseteq C \text{ and } P(C) \}$$

As we saw before, the least element of such a family is exactly the big intersection  $\bigcap \mathcal{C}_P$ , since it is below every closed superset  $C$  of  $A$  by its UP – this is what part (a) shows in the particular case of the transitive closure of a graph. While this proof succeeds, the construction of a closure as a big intersection is inconvenient: it proceeds by overapproximating (potentially quantifying over an infinite number of supersets) and taking the common elements of every overapproximation. In many cases the closure can be built bottom-up, adding elements to the set up until the closure property is satisfied. The exact approach depends on what property one is considering, but often involves repeated phases of adding elements to a set to fix all the current deficiencies, and checking if the new elements gave rise to new holes that need to be fixed. For example, if a graph  $G$  has edges  $(a, b)$  and  $(b, c)$ , its transitive closure will have to include the edge  $(a, c)$ ; however, if  $G$  also has an edge  $(c, d)$ , it can combine with  $(a, c)$  so the edge  $(a, d)$  will be included in the *next* phase. This is repeated until there are no more edges needed to make the graph transitive – for a finite graph, this state will be reached in a finite number of steps. As this question shows, the step of glueing together transitive edges is done via relation composition, and iterating this process a potentially infinite number of times will construct the transitive closure.

## 7. On partial functions

### 7.1. Basic exercises

- Let  $A_2 = \{1, 2\}$  and  $A_3 = \{a, b, c\}$ . List the elements of the sets  $\text{PFun}(A_i, A_j)$  for  $i, j \in \{2, 3\}$ . *Hint:* there may be quite a few, so you can think of ways of characterising all of them without giving an explicit listing.

$\text{PFun}(A_2, A_2)$ . We have 4 possible total functions:  $\{(1, 1), (2, 1)\}$ ,  $\{(1, 1), (2, 2)\}$ ,  $\{(1, 2), (2, 1)\}$ ,  $\{(1, 2), (2, 2)\}$ . All singleton subsets of these are also partial functions, of which there are 4 more:  $\{(1, 1)\}$ ,  $\{(1, 2)\}$ ,  $\{(2, 1)\}$ ,  $\{(2, 2)\}$ . Finally we have the totally undefined function  $\{\}$ , giving the expected number of  $(2 + 1)^2 = 9$  of partial functions.

$\text{PFun}(A_2, A_3)$ . We have 9 possible total functions:  $\{(1, x), (2, y) \mid x, y \in A_3\}$ . The singletons map 1 or 2 to any of  $a, b, c$ , so there are 6 of those:  $\{(1, x) \mid x \in A_3\} \cup \{(2, y) \mid y \in A_3\}$ . With  $\{\}$ , we have  $16 = (3 + 1)^2$  partial functions, as expected.

$\text{PFun}(A_3, A_2)$ . We have 8 possible total functions:  $\{(a, x), (b, y), (c, z) \mid x, y, z \in A_2\}$ .



There are  $3 \cdot 2 \cdot 2 = 12$  partial functions undefined at one argument (where the notation  $\{A_i\}_{i \in I}$  for an indexed family of sets stands for  $\{A_i \mid i \in I\}$ ):

$$\{\{(a, x), (b, y)\}\}_{x, y \in A_2} \cup \{\{(a, x), (c, z)\}\}_{x, z \in A_2} \cup \{\{(b, y), (c, z)\}\}_{y, z \in A_2}$$

There are  $3 \cdot 2 = 6$  partial functions undefined at two arguments:

$$\{\{(a, x)\}\}_{x \in A_2} \cup \{\{(b, y)\}\}_{y \in A_2} \cup \{\{(c, z)\}\}_{z \in A_2}$$

With  $\{\}$ , we have  $27 = (2 + 1)^3$  partial functions, as expected.

$\text{PFun}(A_3, A_3)$ . We have 27 possible total functions:  $\{\{(a, x), (b, y), (c, z)\} \mid x, y, z \in A_3\}$ .

There are  $3 \cdot 3 \cdot 3 = 27$  partial functions undefined at one argument:

$$\{\{(a, x), (b, y)\}\}_{x, y \in A_3} \cup \{\{(a, x), (c, z)\}\}_{x, z \in A_3} \cup \{\{(b, y), (c, z)\}\}_{y, z \in A_3}$$

There are  $3 \cdot 3 = 9$  partial functions undefined at two arguments:

$$\{\{(a, x)\}\}_{x \in A_3} \cup \{\{(b, y)\}\}_{y \in A_3} \cup \{\{(c, z)\}\}_{z \in A_3}$$

With  $\{\}$ , we have  $64 = (3 + 1)^3$  partial functions, as expected.

2. Prove that a relation  $R: A \leftrightarrow B$  is a partial function iff  $R \circ R^{\text{op}} \subseteq \text{id}_B$ .

( $\Rightarrow$ ) Assume  $R: A \leftrightarrow B$  is a partial function: that is, for all  $a \in A$  and  $b_1, b_2 \in B$ , if  $a R b_1$  and  $a R b_2$  then  $b_1 = b_2$ . We need to show that if  $(b_1, b_2) \in R \circ R^{\text{op}}$ ,  $b_1 = b_2$ . By the definition of relational composition and the opposite relation, there exists a  $a \in A$  such that  $a R b_1$  and  $a R b_2$ ; but since  $R$  is functional,  $b_1 = b_2$ .


( $\Leftarrow$ ) Assume  $R \circ R^{\text{op}} \subseteq \text{id}_B$  and take  $a \in A$ ,  $b_1, b_2 \in B$  with  $a R b_1$  and  $a R b_2$ . Then,  $b_1 R^{\text{op}} a$  and therefore  $(b_1, b_2) \in R \circ R^{\text{op}}$  through  $a$ . By assumption, this implies that  $b_1 = b_2$ , as required.

3. Prove that the identity relation is a partial function, and that the composition of partial functions is a partial function.

We show that for all  $a, a_1, a_2 \in A$ , if  $a \text{id}_A a_1$  and  $a \text{id}_A a_2$ ,  $a_1 = a_2$ . Since  $\text{id}_A$  is the equality relation, we have that  $a = a_1$  and  $a = a_2$ , so  $a_1 = a_2$ .

Let  $f: A \rightarrow B$  and  $g: B \rightarrow C$  be two partial functions. To show that  $g \circ f$  is a partial function, it is sufficient to show that  $(g \circ f) \circ (g \circ f)^{\text{op}} \subseteq \text{id}_C$  (§7.1.2). By §6.1.3(d), we have that  $(g \circ f) \circ (g \circ f)^{\text{op}} = g \circ f \circ f^{\text{op}} \circ g^{\text{op}}$ . Since  $f$  is a partial function,  $f \circ f^{\text{op}} \subseteq \text{id}_B$  and  $g \circ g^{\text{op}} \subseteq \text{id}_C$ ; thus, by §6.2.1, we have:

$$g \circ (f \circ f^{\text{op}}) \circ g^{\text{op}} \subseteq g \circ \text{id}_B \circ g^{\text{op}} = g \circ g^{\text{op}} \subseteq \text{id}_C.$$

 We could of course prove the latter by unwrapping the definition of partial functions and composition, or doing case analysis on when the functions are defined. But approaching it via a sufficient condition is quite neat too!

## 7.2. Core exercises

1. Show that  $(\text{PFun}(A, B), \subseteq)$  is a partial order. What is its least element, if it exists?

Any subset of a partial function is itself a partial function, since it may be defined on fewer elements of the domain, but functionality is not violated. The set of partial functions between two sets therefore has the standard subset ordering  $f \subseteq g$  which is reflexive, transitive and antisymmetric as shown in §5.1.1. The least element is the empty set seen as the totally undefined partial function from  $A$  to  $B$ .

2. Let  $\mathcal{F} \subseteq \text{PFun}(A, B)$  be a non-empty collection of partial functions from  $A$  to  $B$ .

- a) Show that  $\bigcap \mathcal{F}$  is a partial function.

By §7.1.2, it is sufficient to show that  $(\bigcap \mathcal{F}) \circ (\bigcap \mathcal{F})^{\text{op}} \subseteq \text{id}_B$ . We calculate as follows:


$$\begin{aligned} (\bigcap \mathcal{F}) \circ (\bigcap \mathcal{F})^{\text{op}} &= (\bigcap \mathcal{F}) \circ (\bigcap \{F^{\text{op}} \mid F \in \mathcal{F}\}) && \text{(by §6.1.3(b))} \\ &\subseteq \bigcap \{F \circ F^{\text{op}} \mid F \in \mathcal{F}\} && \text{(by UP of intersections)} \\ &\subseteq \bigcap \{\text{id}_B \mid F \in \mathcal{F}\} = \text{id}_B && \text{(by §7.1.2 and assumption)} \end{aligned}$$

- b) Show that  $\bigcup \mathcal{F}$  need not be a partial function by defining two partial functions  $f, g : A \rightarrow B$  such that  $f \cup g : A \rightarrow B$  is a non-functional relation.

We can simply have  $f = \{(1, a)\}$  and  $g = \{(1, b)\}$  for  $A = \{1\}$  and  $B = \{a, b\}$ . Both are partial (in fact total) functions, but the union  $\{(1, a), (1, b)\}$  maps 1 to both  $a$  and  $b$ , violating functionality.

- c) Let  $h : A \rightarrow B$  be a partial function. Show that if every element of  $\mathcal{F}$  is below  $h$  then  $\bigcup \mathcal{F}$  is a partial function.

If for all  $f \in \mathcal{F}$ ,  $f \subseteq h$ , then  $h$  is an upper bound of  $\mathcal{F}$  and therefore we have  $\bigcup \mathcal{F} \subseteq h$ . But subsets of partial functions are themselves partial functions, since they cannot have more mappings from any particular element of  $A$  than  $h$ .

 You may wonder why the high-level proof we used for intersections doesn't work for unions. The issue is that the UP of unions only allows the inclusion

$$\bigcup \{F \circ F^{\text{op}} \mid F \in \mathcal{F}\} \subseteq (\bigcup \mathcal{F}) \circ (\bigcup \{F^{\text{op}} \mid F \in \mathcal{F}\})$$

and while a seemingly more general property

$$\bigcup \{F \circ F' \mid F, F' \in \mathcal{F}\} = (\bigcup \mathcal{F}) \circ (\bigcup \mathcal{F})$$

holds in both directions (see §6.2.2), the  $F$  and  $F'$  are independent (since they come from two existential assumptions) and  $F'$  cannot be specialised to  $F^{\text{op}}$ .

## 8. On functions

### 8.1. Basic exercises

1. Let  $A_2 = \{1, 2\}$  and  $A_3 = \{a, b, c\}$ . List the elements of the sets  $\text{Fun}(A_i, A_j)$  for  $i, j \in \{2, 3\}$ .

The total functions have already been listed amongst the partial functions in §7.1.1:

$$\text{Fun}(A_2, A_2) = \{ \{(1, x), (2, y)\} \mid x, y \in A_2 \}$$

$$\text{Fun}(A_2, A_3) = \{ \{(1, x), (2, y)\} \mid x, y \in A_3 \}$$

$$\text{Fun}(A_3, A_2) = \{ \{(a, x), (b, y), (c, z)\} \mid x, y, z \in A_2 \}$$

$$\text{Fun}(A_3, A_3) = \{ \{(a, x), (b, y), (c, z)\} \mid x, y, z \in A_3 \}$$

2. Prove that the identity partial function is a function, and the composition of functions yields a function.

We need to show that for all  $a \in A$  there exists a unique  $a' \in A$  such that  $\text{id}_A(a) = a'$ . Of course,  $a$  is the witness of the existence, and it is unique since sets have no duplicate elements.


Let  $f : A \rightarrow B$  and  $g : B \rightarrow C$  be functions. We show that the composite  $g \circ f$  is also a function, that is, for all  $a \in A$ , there exists a unique  $c \in C$  such that  $(g \circ f)(a) = c$ . By the definition of function composition,  $(g \circ f)(a) = g(f(a))$ , where  $f(a) = b$  for a unique  $b \in B$  and  $g(b) = c$  for a unique  $c \in C$ . Thus, a unique  $c$  does exist, and  $g \circ f$  is a function.

3. Prove or disprove that  $(\text{Fun}(A, B), \subseteq)$  is a partial order.

Unlike partial functions, functions are not closed under taking subsets or supersets: the number of mappings (i.e. the graph) of a function must be equal to the size of the domain (or, more precisely, isomorphic), so we can't add or remove mappings without breaking functionality or totality. We may be tempted to conclude that  $(\text{Fun}(A, B), \subseteq)$  is not a partial order, but we should remember that there is still an ordering on the set even if different functions can't be compared:  $f \subseteq g$  if and only if  $f = g$ . Thus, the subset ordering on functions simply restricts to equality, which is trivially a partial order: we have reflexivity since  $f \subseteq f$ , and antisymmetry and transitivity hold because the hypotheses like  $f \subseteq g$  and  $g \subseteq h$  simply become  $f = g = h$ .

4. Find endofunctions  $f, g : A \rightarrow A$  such that  $f \circ g \neq g \circ f$ .

Let  $f : \mathbb{N} \rightarrow \mathbb{N}$  be the successor function  $n \mapsto n + 1$ , and  $g : \mathbb{N} \rightarrow \mathbb{N}$  be the squaring function  $m \mapsto m^2$ . Then, for all  $n$ ,  $(g \circ f)(n) = (n + 1)^2 = n^2 + 2n + 1$ , but  $(f \circ g)(n) = n^2 + 1$ .

 Many other examples exist of course. This is merely a reminder that function composition (and relational composition in general) is not commutative, and it doesn't have many other properties that we tend to expect from binary operators: for example,  $f \circ g = f \circ h$  does not imply  $g = h$  in general.


## 8.2. Core exercises

1. A relation  $R: A \leftrightarrow B$  is said to be *total* if  $\forall a \in A. \exists b \in B. a R b$ . Prove that this is equivalent to  $\text{id}_A \subseteq R^{\text{op}} \circ R$ . Conclude that a relation  $R: A \leftrightarrow B$  is a function iff  $R \circ R^{\text{op}} \subseteq \text{id}_B$  and  $\text{id}_A \subseteq R^{\text{op}} \circ R$ .

( $\Rightarrow$ ) Assume that  $R: A \leftrightarrow B$  is a total relation, that is, for all  $a \in A$  there exists a  $b \in B$  such that  $a R b$ . We need to show that for all  $(a, a') \in \text{id}_A$ ,  $(a, a') \in R^{\text{op}} \circ R$  – that is, that  $R^{\text{op}} \circ R$  is reflexive. A pair  $(a, a)$  for  $a \in A$  is in  $R^{\text{op}} \circ R$  if there exists a  $b \in B$  such that  $a R b$  and  $b R^{\text{op}} a$ , i.e.  $a R b$ , which is satisfied if there exists a  $b \in B$  such that  $a R b$ . But this follows from the assumption that  $R$  is total.

( $\Leftarrow$ ) Assume that  $R^{\text{op}} \circ R$  is reflexive and show that  $R$  is total. Take  $a \in A$ ; as  $R^{\text{op}} \circ R$  is reflexive,  $(a, a) \in R^{\text{op}} \circ R$ , so there exists a  $b \in B$  such that  $a R b$ . Taking this  $b$  as the witness of existence, we conclude that  $R$  is a total relation.

A total function is both a total relation and a partial function, so a relation  $R$  is total if and only if it satisfies both  $\text{id}_A \subseteq R^{\text{op}} \circ R$  (from above) and  $R \circ R^{\text{op}} \subseteq \text{id}_B$  (from §7.1.2).

 This question establishes that partial functions are in some sense dual to total relations: instead of asking for uniqueness (functionality), they require an existence (totality). Consequently, we can dualise several results from the previous section. For example, we have that the union of total relations is total, but the intersection is not, with the proofs being the duals of the arguments in §7.2.2 (and the proof attempt for intersections failing because their universal property is the “wrong way around”).

2. Let  $\chi: \mathcal{P}(U) \rightarrow (U \Rightarrow [2])$  be the function mapping subsets  $S \subseteq U$  to their characteristic functions  $\chi_S: U \rightarrow [2]$ .

a) Prove that for all  $x \in U$ ,

- $\chi_{A \cup B}(x) = (\chi_A(x) \vee \chi_B(x)) = \max(\chi_A(x), \chi_B(x))$

Let  $x \in U$ . Then,

$$\chi_{A \cup B}(x) \iff x \in (A \cup B) \iff (x \in A) \vee (x \in B) \iff (\chi_A(x) \vee \chi_B(x))$$

and the latter holds iff  $\chi_A(x) = 1$  or  $\chi_B(x) = 1$ , so  $\max(\chi_A(x), \chi_B(x)) = 1$ .

- $\chi_{A \cap B}(x) = (\chi_A(x) \wedge \chi_B(x)) = \min(\chi_A(x), \chi_B(x))$

Let  $x \in U$ . Then,

$$\chi_{A \cap B}(x) \iff x \in (A \cap B) \iff (x \in A) \wedge (x \in B) \iff (\chi_A(x) \wedge \chi_B(x))$$

and the latter holds iff  $\chi_A(x) = 1$  and  $\chi_B(x) = 1$ , so  $\min(\chi_A(x), \chi_B(x)) = 1$ .

- $\chi_{A^c}(x) = \neg(\chi_A(x)) = (1 - \chi_A(x))$

Let  $x \in U$ . Then,

$$\chi_{A^c}(x) \iff x \notin A \iff \neg(x \in A) \iff \neg(\chi_A(x))$$

and the latter holds iff  $\chi_A(x) = 0$ , so  $1 - \chi_A(x) = 1$ .

b) For what construction  $A?B$  on sets  $A$  and  $B$  does it hold that

$$\chi_{A?B}(x) = (\chi_A(x) \oplus \chi_B(x)) = (\chi_A(x) +_2 \chi_B(x))$$

for all  $x \in U$ , where  $\oplus$  is the *exclusive or* operator? Prove your claim.

The element  $x$  must be exactly in one of  $A$  or  $B$ , not their intersection. This leads to the definition

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

which is known as the *symmetric difference* and written  $A \Delta B$ . Then, for  $x \in U$ ,

$$\begin{aligned} \chi_{A\Delta B}(x) &\iff x \in (A \Delta B) \\ &\iff (x \in A \wedge x \notin B) \vee (x \in B \wedge x \notin A) \\ &\iff (\chi_A(x) \wedge \neg\chi_B(x)) \vee (\chi_B(x) \wedge \neg\chi_A(x)) \\ &\iff \chi_A(x) \oplus \chi_B(x) \end{aligned}$$

and the latter doesn't hold iff  $\chi_A(x) = \chi_B(x) = 0$  or  $\chi_A(x) = \chi_B(x) = 1$ . Adding  $\chi_A(x)$  and  $\chi_B(x)$  can give the values of 0, 1 or 2, but we only want the case where the sum is 1 and use the addition modulo 2 to route the other possibilities to 0.

### 8.3. Optional advanced exercise

Consider a set  $A$  together with an element  $a \in A$  and an endofunction  $f : A \rightarrow A$ .

Say that a relation  $R : \mathbb{N} \leftrightarrow A$  is  $(a, f)$ -closed whenever

$$R(0, a) \quad \text{and} \quad \forall n \in \mathbb{N}, x \in A. R(n, x) \implies R(n+1, f(x))$$

Define the relation  $F : \mathbb{N} \leftrightarrow A$  as

$$F \triangleq \bigcap \{ R : \mathbb{N} \leftrightarrow A \mid R \text{ is } (a, f)\text{-closed} \}$$

- Prove that  $F$  is  $(a, f)$ -closed.
- Prove that  $F$  is total, that is:  $\forall n \in \mathbb{N}. \exists y \in A. F(n, y)$ .
- Prove that  $F$  is a function  $\mathbb{N} \rightarrow A$ , that is:  $\forall n \in \mathbb{N}. \exists! y \in A. F(n, y)$ .

*Hint:* Proceed by induction. Observe that, in view of the previous item, to show that  $\exists! y \in A. F(k, y)$  it suffices to exhibit an  $(a, f)$ -closed relation  $R_k$  such that  $\exists! y \in A. R_k(k, y)$ . (Why?) For instance, as the relation  $R_0 = \{(m, y) \in \mathbb{N} \times A \mid m = 0 \implies y = a\}$  is  $(a, f)$ -closed one has that  $F(0, y) \implies R_0(0, y) \implies y = a$ .

- Show that if  $h$  is a function  $\mathbb{N} \rightarrow A$  with  $h(0) = a$  and  $\forall n \in \mathbb{N}. h(n+1) = f(h(n))$  then  $h = F$ .

Thus, for every set  $A$  together with an element  $a \in A$  and an endofunction  $f : A \rightarrow A$  there exists a


unique function  $F: \mathbb{N} \rightarrow A$ , typically said to be *inductively defined*, satisfying the recurrence relation

$$F(n) = \begin{cases} a & \text{for } n = 0 \\ f(F(n-1)) & \text{for } n \geq 1 \end{cases}$$

## 9. On bijections

### 9.1. Basic exercises

1. a) Define a function that has (i) none, (ii) exactly one, and (iii) more than one retraction.
- b) Define a function that has (i) none, (ii) exactly one, and (iii) more than one section.

 The general pattern (for finite sets) is that the domain of sections is smaller than (or equal to) the codomain so elements can be “selected” from a larger set. Conversely, the domain of a retraction is greater than or equal to the codomain, so a group of elements can be “collapsed” into one. The section-retraction condition states that a section at  $a \in A$  selects one of the elements that get mapped to  $a$  by the retraction.

2. Let  $n$  be an integer.

- a) How many sections are there for the absolute-value map  $x \mapsto |x|: [-n..n] \rightarrow [0..n]$ ?

The absolute value function maps two integers  $k$  and  $-k$  to the same natural number  $|k|$  (other than 0), so a section for this map can select either of the two integers. The codomain  $[0..n]$  has size  $n + 1$  but 0 can only be mapped to  $0 \in [-n..n]$ ; for the remaining  $n$  inputs we have 2 choices each, giving us  $2^n$  possible sections.

- b) How many retractions are there for the exponential map  $x \mapsto 2^x: [0..n] \rightarrow [0..2^n]$ ?

The retraction only needs to map the powers of two back to their exponents, leaving  $2^n - n$  naturals in  $[0..2^n]$  that are not in the range of the exponential map and therefore are not constrained by the section-retraction condition. Since each of these can be mapped to any of the  $\#[0..n] = n + 1$  possible inputs, the exponential map has  $(n + 1)^{2^n - n}$  retractions.

3. Give an example of two sets  $A$  and  $B$  and a function  $f: A \rightarrow B$  such that  $f$  has a retraction but no section. Explain how you know that  $f$  has these properties.

See §9.1.1.

4. Prove that the identity function is a bijection and that the composition of bijections is a bijection.

To show that the identity  $\text{id}_A: A \rightarrow A$  is a bijection, it is sufficient to exhibit a two sided inverse, namely  $\text{id}_A$  itself. Since it is the unit of composition, we have  $\text{id}_A \circ \text{id}_A = \text{id}_A$ , which is both the left and right inverse condition.

Let  $f: A \rightarrow B$  and  $g: B \rightarrow C$  be bijections, with respective inverses  $f^{-1}$  and  $g^{-1}$ . We need to show that the composite  $g \circ f: A \rightarrow C$  is a bijection. Consider the function  $f^{-1} \circ g^{-1}: C \rightarrow A$ ,

and calculate using the inverse properties of  $f$  and  $g$ :

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ \text{id}_B \circ f = f^{-1} \circ f = \text{id}_A$$

$$(g \circ f) \circ (f^{-1} \circ g^{-1}) = g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ \text{id}_B \circ g^{-1} = g \circ g^{-1} = \text{id}_C$$

Thus,  $f^{-1} \circ g^{-1}$  is a two-sided inverse of  $g \circ f$ , making it into a bijection.

5. For  $f : A \rightarrow B$ , prove that if there are  $g, h : B \rightarrow A$  such that  $g \circ f = \text{id}_A$  and  $f \circ h = \text{id}_B$  then  $g = h$ . Conclude as a corollary that, whenever it exists, the inverse of a function is unique.

We show that if a map  $f : A \rightarrow B$  has two opposite-sided inverses, they must be equal. Assume  $g, h : B \rightarrow A$  satisfy  $g \circ f = \text{id}_A$  and  $f \circ h = \text{id}_B$ . We consider the composite  $g \circ f \circ h$  and calculate:

$$g \circ (f \circ h) = g \circ \text{id}_B = g \quad (g \circ f) \circ h = \text{id}_A \circ h = h$$

and since composition is associative, we have that  $g = h$ .

Assume a function  $f : A \rightarrow B$  has two inverses  $f_1^{-1}, f_2^{-1} : B \rightarrow A$ . Then, in particular, they satisfy  $f_1^{-1} \circ f = \text{id}_A$  and  $f \circ f_2^{-1} = \text{id}_B$ , so by the first part, we have that  $f_1^{-1} = f_2^{-1}$ .

## 9.2. Core exercises

1. We say that two functions  $s : A \rightarrow B$  and  $r : B \rightarrow A$  are a *section-retraction pair* whenever  $r \circ s = \text{id}_A$ ; and that a function  $e : B \rightarrow B$  is an *idempotent* whenever  $e \circ e = e$ . This question demonstrates that section-retraction pairs and idempotents are closely connected: any section-retraction pair gives rise to an idempotent function, and any idempotent function can be split into a section-retraction pair.

a) Let  $f : C \rightarrow D$  and  $g : D \rightarrow C$  be functions such that  $f \circ g \circ f = f$ .

- (i) Can you conclude that  $f \circ g$  is idempotent? What about  $g \circ f$ ? Justify your answers.

Both are idempotent, since by associativity of  $\circ$  and the assumption we have:

$$(f \circ g) \circ (f \circ g) = (f \circ g \circ f) \circ g = f \circ g$$

$$(g \circ f) \circ (g \circ f) = g \circ (f \circ g \circ f) = g \circ f$$

- (ii) Define a map  $g'$  using  $f$  and  $g$  that satisfies both

$$f \circ g' \circ f = f \quad \text{and} \quad g' \circ f \circ g' = g'$$

Let  $g' = g \circ f \circ g$ . Then:

$$f \circ g' \circ f = f \circ g \circ (f \circ g \circ f) = f \circ g \circ f = f$$

$$g' \circ f \circ g' = g \circ (f \circ g \circ f) \circ g \circ f \circ g = g \circ (f \circ g \circ f) \circ g = g \circ f \circ g = g'$$

♪ Straightforward questions intended to get you used to “the algebra of functions”: calculating with compositions of functions, rather than their values at arguments.

- b) Show that if  $s: A \rightarrow B$  and  $r: B \rightarrow A$  are a section-retraction pair then the composite  $s \circ r: B \rightarrow B$  is idempotent.

Let  $s: A \rightarrow B$  and  $r: B \rightarrow A$  be a section-retraction pair with  $r \circ s = \text{id}_A$ . We show that  $s \circ r: B \rightarrow B$  is idempotent as follows:

$$(s \circ r) \circ (s \circ r) = s \circ (r \circ s) \circ r = s \circ \text{id}_A \circ r = s \circ r$$

where we use assumption along with the associativity of composition and neutrality of the identity function.

- c) Show that for every idempotent  $e: B \rightarrow B$  there exists a set  $A$  (called a *retract* of  $B$ ) and a section-retraction pair  $s: A \rightarrow B$  and  $r: B \rightarrow A$  such that  $s \circ r = e$ .

Let  $e: B \rightarrow B$  be an idempotent function. We need to show that there exists a set  $A$  such that  $e$  can be split into the composition  $e = s \circ r$  where  $s: A \rightarrow B$  and  $r: B \rightarrow A$  form a section-retraction pair.

Take  $A$  to be the subset  $\{e(x) \mid x \in B\} \subseteq B$ , i.e. the direct image of  $B$  under  $e$ . Let  $s: A \rightarrow B$  be the subset injection  $A \hookrightarrow B$ , and  $r: B \rightarrow A$  be  $e$  with its codomain restricted to its range. That is:

$$s(x) = x \quad r(y) = e(y)$$

Now, the composite  $s \circ r$  maps  $x \in B$  to  $e(x) \in A$  which is then injected to  $B$  unchanged, so  $s \circ r = e$ . The reverse composite  $r \circ s$  maps an element  $y: A$  to  $e(y)$ , but by definition of  $A$  there must be an  $x \in B$  such that  $y = e(x)$ , and by the idempotence of  $e$  we have that  $e(y) = e(e(x)) = e(x) = y$ ; thus,  $r \circ s = \text{id}_A$  and the two maps form a section-retraction pair.

♪ This is a rather abstract exercise which establishes a connection between idempotent maps and section-retraction pairs, namely: every sr-pair gives rise to an idempotent map, and every idempotent map can be split into a sr-pair.

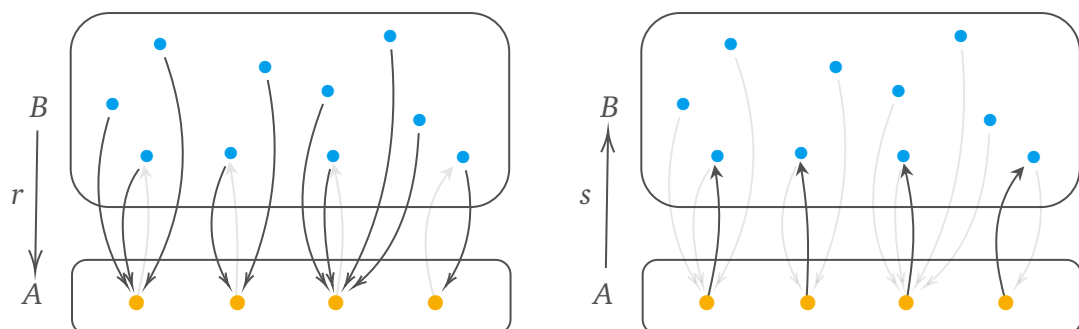
Idempotent maps are functions which do not need to be applied more than once:  $f(f(x)) = f(x)$ , so in general  $f^n(x) = f(x)$  for any natural number  $n$ . Examples are the absolute value function  $|\cdot|: \mathbb{Z} \rightarrow \mathbb{Z}$ , sorting algorithms and other “normalisation” procedures (once something is brought into a standardised, normal form, it should not change if normalised again), mapping a set  $X$  to its closure under some property  $\text{Cl}_p(X)$  (e.g. for an arbitrary relation  $R$ , taking the transitive closure of  $\text{Cl}_t(R)$  should be a no-op), pressing the elevator or road crossing indicator button, etc.

Section-retraction pairs normally capture the idea of sorting a set of elements  $B$  into disjoint groupings labelled by  $A$ : the retraction  $r: B \rightarrow A$  maps an element  $b \in B$  to



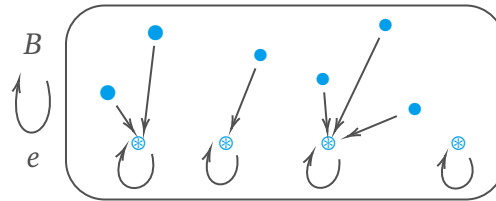
its group label in  $A$ , while the section  $s: A \rightarrow B$  selects a particular element  $s(a) \in B$  labelled by  $a \in A$ . Clearly the group that a particular element of the group belongs to will be the starting group, giving rise to the required one-sided inverse condition. Examples are cities grouped by countries, students grouped by subject/college, products grouped by brands, employees grouped by department, and so on (you can probably find even more examples in the Databases course).

A common characteristic of all of these is that the set of entities is larger than the set of groups, and each group has at least one element (since the section has to select something). This is usually visualised as a vertical internal diagram, where the retraction  $r$  maps several entities in  $B$  to a single element in  $A$ , and the section  $s$  maps an element in  $A$  to one of the elements that is mapped to it by  $r$ . As this representation demonstrates, elements in  $B$  get clustered by which group they belong to, which equips  $B$  with an implicit partitioning (see §10.2.3). The section then selects a “representative element” of each partition. The section-retraction condition  $r \circ s = \text{id}_A$  simply states that the representative elements are in the clusters they represent – certainly desirable! For the example of cities grouped by countries, the representative element of each city cluster may be the capital city (and let’s assume every country has exactly one capital), which of course has to be in the country it is a capital of.



Section-retraction parts only have a one-directional inverse property  $r \circ s = \text{id}_A$ ; nevertheless, the reverse composite  $s \circ r$  – not required to be the identity – cannot be completely arbitrary either. As shown in the exercise, it has to be an idempotent map: once we do one round trip between  $B$  and  $A$ , we are “stuck” no matter how many new round trips we do. Mathematically, we have found the *fixed point* of an endofunction, i.e. the value  $x$  such that  $f(x) = x$ . It is easy to see that every idempotent map  $e: B \rightarrow B$  has a fixed point  $e(x)$  for all  $x \in B$ , since the idempotence condition  $e(e(x)) = e(x)$  is precisely a fixed point equation. Graphically, we can see that following any 2-step path from  $x \in B$  will lead us to the representative element, from which any round trip is merely the identity map. The composite  $s \circ r$  can therefore be seen as a function on  $B$ , representing the mapping of any element in a cluster to its representative; for example, any city to the capital of its country, any student to their college student union president, any employee to their department manager.

Now, we consider a different problem: we start with a set  $B$  and an endomap  $e: B \rightarrow B$  satisfying  $e \circ e = e$ . As before, idempotence clusters elements in  $B$  since one application of  $e$  maps them to a unique representative and any new applications will simply loop on the representatives.



The question is: can we recover the set  $A$  and the section-retraction pair that induces  $e$  just from  $B$  and  $e$ ? While we can't expect to be able to do this exactly – we'd need to figure out the names of colleges only based on the students – we can do the next best thing: find a decomposition which will be *isomorphic* to the original grouping. Looking at the diagram, it should be quite clear which elements act as representatives of the clusters and can therefore be abstractly characterised: all the outputs of the idempotent map  $e$ , i.e. the set of fixed points of  $e$ . Thus, we take the retract  $A$  to be nothing more than the subset  $A \triangleq \{f(x) \mid x \in B\}$ . Intuitively, we exploit the (simplified) fact that the set of capitals/presidents/managers is isomorphic to the set of countries/colleges/departments. Now, we need to find  $s: A \rightarrow B$  and  $r: B \rightarrow A$  satisfying  $r \circ s = \text{id}_A$  and  $s \circ r = e$ . Since  $A$  is a subset of  $B$ , there is a canonical section  $s: A \rightarrow B$  that embeds  $A$  into its superset:  $s(x \in A) = x \in B$ . Conversely, the retraction that maps  $B$  to  $A$  is the idempotent function  $e$  itself, with its codomain restricted to its range:  $r(y \in B) = e(y)$ . The composite  $s \circ r$  is an application of  $e$  followed by an “identity” map, so we clearly have  $s \circ r = e$ . To prove the section-retraction condition, take an  $x \in A$  and consider  $r(s(x)) = r(x) = e(x)$ , which is not exactly what we need; however, we know that  $x \in A$  so it must be of the form  $x = e(y)$  for some  $y \in B$ . Thus,  $r(s(x)) = r(s(e(y))) = e(e(y)) = e(y) = x$ , as required.

## 10. On equivalence relations

### 10.1. Basic exercises

1. Prove that the isomorphism relation  $\cong$  between sets is an equivalence relation.

**Reflexive.** The identity  $\text{id}_A: A \rightarrow A$  is a bijection (§9.1.4), so we have the isomorphism  $A \cong A$  for all sets  $A$ .

**Symmetric.** Assume  $A \cong B$ ; that is, there is a bijection  $f: A \rightarrow B$ . Its inverse  $f^{-1}: B \rightarrow A$  is a bijection too, so we have the isomorphism  $B \cong A$ , as required.

**Transitive.** Assume  $A \cong B$  and  $B \cong C$  with respective bijections  $f$  and  $g$ . Then the composite  $g \circ f: A \rightarrow C$  is a bijection too (§9.1.4) and exhibits the isomorphism  $A \cong C$ .

2. Prove that the identity relation  $\text{id}_A$  on a set  $A$  is an equivalence relation, and that  $A/\text{id}_A \cong A$ .

The identity relation  $\text{id}_A: A \leftrightarrow A$  is equal to the equality relation  $\{(x, y) \in A \times A \mid x = y\}$  which is an equivalence relation.

The quotient set  $A/\text{id}_A$  is the set of equivalence classes of  $A$  under the equality relation:  $A/\text{id}_A = \{[a]_{=} \subseteq A \mid a \in A\}$ . The equivalence class  $[a]_{=}$  contains all elements that are equal to  $a$ , which of course is  $a$  itself since sets have no repeated elements. Hence every equivalence class is the singleton set, and we can construct a bijection  $f: A \rightarrow A/\text{id}_A$  by mapping  $x \in A$  to  $\{x\} \in A/\text{id}_A$ , and the inverse  $f^{-1}$  mapping  $\{y\}$  to  $y$ .

3. Show that, for a positive integer  $m$ , the relation  $\equiv_m$  on  $\mathbb{Z}$  given by

$$x \equiv_m y \iff x \equiv y \pmod{m}$$

is an equivalence relation. What are the equivalence classes of this relation?

We have already proved that congruence is reflexive, transitive and symmetric in §2.1.1, so it is indeed an equivalence relation. The equivalence classes of congruence modulo  $m$  are the congruence classes  $k_m = \{n \in \mathbb{Z} \mid (m \mid k - n)\}$ , and the quotient  $\mathbb{Z}/\equiv_m$  is isomorphic to the set  $\mathbb{Z}_m$  of integers modulo  $m$ .

4. Show that the relation  $\equiv$  on  $\mathbb{Z} \times \mathbb{Z}^+$  given by

$$(a, b) \equiv (x, y) \iff a \cdot y = x \cdot b$$

is an equivalence relation. What are the equivalence classes of this relation?

**Reflexive.** We have to show that  $(a, b) \equiv (a, b)$  for  $a \in \mathbb{Z}$  and  $b \in \mathbb{Z}^+$ ; by definition, this is  $a \cdot b = a \cdot b$ , which is true by reflexivity of equality.

**Symmetric.** Assume  $(a, b) \equiv (x, y)$ ; that is,  $ay = xb$ . By symmetry of equality we have  $x \cdot b = a \cdot y$  which implies  $(x, y) \equiv (a, b)$ , as required.

**Transitive.** Assume  $(a, b) \equiv (x, y)$  and  $(x, y) \equiv (m, n)$ ; then, ①  $a \cdot y = x \cdot b$  and ②  $x \cdot n = m \cdot y$ . We have to show that ③  $a \cdot n = m \cdot b$ . Multiplying both sides of ① by  $n$ , then rearranging and applying ②, we have the following:

$$a \cdot y \cdot n = x \cdot b \cdot n = x \cdot n \cdot b = m \cdot y \cdot b$$

Since  $y \in \mathbb{Z}^+$ , it is nonzero and we can divide both sides of  $a \cdot y \cdot n = m \cdot b \cdot y$  to get ③, as required.

An equivalence class of this relation for a pair  $(a, b)$  contains all pairs  $(x, y)$  such that  $a \cdot y = x \cdot b$ ; in other words,  $\frac{a}{b} = \frac{x}{y}$ . Thus, the relation expresses the equality of fractions, with an equivalence class corresponding to different “representations” of the same fraction, and a representative element for every class being the fraction in lowest terms. The quotient set  $\mathbb{Z} \times \mathbb{Z}^+ / \equiv$  has elements corresponding to rational numbers (represented by an infinite number of distinct, but equivalent fractions of integers) so it is isomorphic to  $\mathbb{Q}$ .

## 10.2. Core exercises

1. Let  $E_1$  and  $E_2$  be two equivalence relations on a set  $A$ . Either prove or disprove the following statements.

- a)  $E_1 \cup E_2$  is an equivalence relation on  $A$ .

The statement is false. Let  $A = \{a, b, c\}$  and consider the equivalence relations  $E_1 = \{(a, a), (a, b), (b, a), (b, b), (c, c)\}$  and  $E_2 = \{(a, a), (b, b), (b, c), (c, b), (c, c)\}$ . Then, the union  $E_1 \cup E_2$  contains the pairs  $(a, b)$  and  $(b, c)$  but not the pair  $(a, c)$ , so the union – while still being reflexive and symmetric – is not transitive.

- b)  $E_1 \cap E_2$  is an equivalence relation on  $A$ .

Let  $E_1$  and  $E_2$  be two equivalence relations on  $A$ . We show that  $E_1 \cap E_2$  is an equivalence relation as well, using the sufficient conditions of §6.2.3.

**Reflexive.** We show that  $\text{id}_A \subseteq E_1 \cap E_2$  or equivalently,  $\text{id}_A \subseteq E_1$  and  $\text{id}_A \subseteq E_2$ , which hold since  $E_1$  and  $E_2$  are reflexive.

**Symmetric.** By §6.1.3(b),  $(E_1 \cap E_2)^{\text{op}} = E_1^{\text{op}} \cap E_2^{\text{op}} = E_1 \cap E_2$ , since both relations are symmetric.

**Transitive.** It is sufficient to show that  $(E_1 \cap E_2) \circ (E_1 \cap E_2) \subseteq E_1 \cap E_2$ , or equivalently,  $(E_1 \cap E_2) \circ (E_1 \cap E_2) \subseteq E_1$  and  $(E_1 \cap E_2) \circ (E_1 \cap E_2) \subseteq E_2$ . Since  $E_1 \cap E_2 \subseteq E_1$ , by §6.2.1 we have  $(E_1 \cap E_2) \circ (E_1 \cap E_2) \subseteq E_1 \circ E_1$ , and by transitivity of  $E_1$ ,  $E_1 \circ E_1 \subseteq E_1$ . The case for  $E_2$  is similar.

🎵 Could we have done this quite easily with element-wise reasoning? Yes. Is this approach far more satisfying? Also yes.

2. For an equivalence relation  $E$  on a set  $A$ , show that  $[a_1]_E = [a_2]_E$  iff  $a_1 E a_2$ , where

$$[a]_E = \{x \in A \mid x E a\}.$$

Let  $E$  be an equivalence relation on  $A$ , and take two elements  $a_1, a_2 \in A$ .

( $\Rightarrow$ ) Assume  $[a_1]_E = [a_2]_E$ ; we need to prove that  $a_1 E a_2$ . By definition of equivalence classes and set equality, all elements  $x \in A$  are related to  $a_1$  if and only if they are related to  $a_2$ :  $x E a_1 \iff x E a_2$ . In particular, for  $x = a_1$ , we have  $a_1 E a_1 \iff a_1 E a_2$ ; but  $E$  is reflexive, so  $a_1 E a_1$  and from this  $a_1 E a_2$  follows.

( $\Leftarrow$ ) Assume  $a_1 E a_2$  and prove that for all  $x \in A$ ,  $x E a_1$  if and only if  $x E a_2$ . If  $x E a_1$ , then by assumption and the transitivity of  $E$ ,  $x E a_2$ . Conversely, if  $x E a_2$ , we can chain this with the opposite assumption  $a_2 E a_1$  to get  $x E a_1$ , as required.

3. For a function  $f : A \rightarrow B$  define a relation  $\equiv_f$  on  $A$  by the rule: for all  $a, a' \in A$ ,

$$a \equiv_f a' \iff f(a) = f(a')$$

- a) Show that for every function  $f : A \rightarrow B$ , the relation  $\equiv_f$  is an equivalence relation on  $A$ .

**Reflexive.** We need to show that for all  $a \in A$ ,  $a \equiv_f a$ , or equivalently,  $f(a) = f(a)$  – but the latter holds by reflexivity.

**Symmetric.** Assume  $a \equiv_f b$ , that is,  $f(a) = f(b)$ . Then  $f(b) = f(a)$ , so  $b \equiv_f a$ , proving that  $\equiv_f$  is symmetric.

**Transitive.** Assume  $a \equiv_f b$  and  $b \equiv_f c$ , that is,  $f(a) = f(b)$  and  $f(b) = f(c)$ . By transitivity of equality,  $f(a) = f(c)$ , so  $a \equiv_f c$ , as required.

- b) Prove that every equivalence relation  $E$  in a set  $A$  is equal to  $\equiv_q$ , where  $q: A \rightarrow A/E$  is the quotient function  $q(a) = [a]_E$ .

Let  $E$  be an equivalence relation on  $A$ . We need to show that for all  $a, b \in A$ ,  $a E b$  if and only if  $a \equiv_q b$ , or, by definition,  $[a]_E = [b]_E$ . But this follows directly from §10.2.2.

- c) Prove that for every surjection  $f: A \rightarrow B$ ,

$$B \cong (A / \equiv_f)$$

Let  $f: A \rightarrow B$  be a surjection. We prove the isomorphism by exhibiting a bijection  $g: B \rightarrow (A / \equiv_f)$  with a two-sided inverse  $g^{-1}: (A / \equiv_f) \rightarrow B$ .

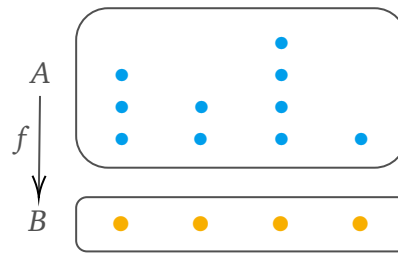
Let  $g$  map a  $b \in B$  to the set  $\{a \in A \mid f(a) = b\}$ ; this is a non-empty set since  $f$  is a surjection, and it is an equivalence class of elements under  $\equiv_f$  because any  $a_1, a_2$  in the set gets mapped to  $b$  by  $f$  and therefore  $f(a_1) = b = f(a_2)$  implies  $a_1 \equiv_f a_2$ .

Let  $g^{-1}$  be a mapping from an equivalence class  $[a]_{\equiv_f}$  of an  $a \in A$  to  $f(a) \in B$ .

We show that for all  $b \in B$ ,  $g^{-1}(g(b)) = b$ . By definition of  $g$ ,  $g(b) = \{a \in A \mid f(a) = b\}$  which is nonempty by the surjectivity of  $f$ ; let  $a$  be one of its representative elements so that  $g(b) = [a]_{\equiv_f}$ . Then,  $g^{-1}([a]_{\equiv_f}) = f(a)$ , but by assumption,  $f(a) = b$ , so we indeed have  $g^{-1} \circ g = \text{id}_B$ .

Conversely, let  $[a]_{\equiv_f} \in (A / \equiv_f)$  be the equivalence class of an element  $a \in A$ . We then have  $g(g^{-1}([a]_{\equiv_f})) = g(f(a)) = \{a' \in A \mid f(a') = f(a)\}$ . But this set is precisely  $\{a' \in A \mid a' \equiv_f a\}$ , the equivalence class  $[a]_{\equiv_f}$  of  $a$ . Thus,  $g \circ g^{-1} = \text{id}_{A/\equiv_f}$ , and the bijection  $g$  exhibits the isomorphism  $B \cong (A / \equiv_f)$ , as required.

🎵 As before, the best way to get an intuition for this question is to draw a diagram. A useful visualisation of functions – similar to the clustering representation in §9.2.1 – is as elements of the domain stacked over the elements of the codomain that they are mapped to, with the individual mapping arrows left implicit (and functionality captured by the fact that a dot in  $A$  can only be over exactly one dot in  $B$ ).



If the function is an injection, each column can at most one element; if it is a surjection (like  $f$  here), each column must have at least one element. The equivalence relation  $\equiv_f$  in this question relates two elements in  $A$  precisely when they are in the same column, and since  $f$  is a surjection, the elements of  $A$  are all partitioned into disjoint, non-empty columns. Since each column is above an element in  $B$ , there is a bijection between the set of partitions (i.e. the set  $A$  quotiented by  $\equiv_f$ ) and  $B$ , exhibited by a mapping between  $b \in B$  and the stack of  $A$  elements that get mapped to  $b$  (sometimes called the *fiber* over  $b$ ).

## 11. On surjections and injections

### 11.1. Basic exercises

1. Give two examples of functions that are surjective, and two examples of functions that are not.

**Surjective.** Absolute value function to the naturals  $|-|: \mathbb{Z} \rightarrow \mathbb{N}$ ; natural log function  $\ln: \mathbb{R}_0^+ \rightarrow \mathbb{R}$ ; first projection function from the Cartesian product of two (nonempty) sets  $\pi_1: A \times B \rightarrow A$ .

**Not surjective.** Integer squaring function on the naturals:  $(-)^2: \mathbb{N} \rightarrow \mathbb{N}$  (only returns perfect squares); constant function  $c_b: A \rightarrow B$  with value  $b \in B$  (always outputs  $b$  if  $B$  is not the singleton set); successor function  $(-) + 1: \mathbb{N} \rightarrow \mathbb{N}$  (0 is not the successor of any number).

2. Give two examples of functions that are injective, and two examples of functions that are not.

**Injective.** The inclusion/injection function  $\iota: S \rightarrow A$  for any subset  $S$  of  $A$ ; exponential function  $x \mapsto e^x: \mathbb{R} \rightarrow \mathbb{R}$ ; [perfect hash function](#).

**Not injective.** Integer squaring function:  $(-)^2: \mathbb{Z} \rightarrow \mathbb{Z}$  (since  $x^2 = (-x)^2$ ); quotient function  $q(a) = [a]_E: A \rightarrow A/E$  for an equivalence relation  $E$  (related elements map to the same equivalence class);  $\sin(x): [0, 2\pi] \rightarrow [-1, 1]$  since  $\sin(0) = \sin(2\pi) = 0$ .

### 11.2. Core exercises

1. Explain and justify the phrase *injections can be undone*.

Every injection (from a non-empty domain) has a retraction which “undoes” its effect. If  $i: A \rightarrow B$  is an injection, every  $b$  in  $B$  is mapped to by at most  $a \in A$ ; thus, a retraction can

be defined as

$$r(b) = \begin{cases} a & \text{if } \exists a \in A. i(a) = b \\ a_0 & \text{otherwise} \end{cases}$$

where  $a_0$  is any element of  $A$ . This is total, since every  $b$  is either mapped to the source  $a$  for which  $i(a) = b$ , or to the fixed  $a_0$ . It is also functional, since there may only be at most one  $a$  for which  $i(a) = b$ . By construction,  $r \circ i = \text{id}_A$ , so the two form a section-retraction pair.

The implication holds in the other direction as well: every section  $s : A \rightarrow B$  (with a retraction  $r : B \rightarrow A$ ) is an injection. To see this, consider  $a, a' \in A$  and assume  $s(a) = s(a')$ . But since  $r \circ s = \text{id}_A$ , we have that  $r(s(a)) = r(s(a'))$  implies  $a = a'$ , so  $s$  must be an injection.

2. Show that  $f : A \rightarrow B$  is a surjection if and only if for all sets  $C$  and functions  $g, h : B \rightarrow C$ ,  $g \circ f = h \circ f$  implies  $g = h$ .

( $\Rightarrow$ ) Let  $f : A \rightarrow B$  be a surjection: for all  $b \in B$  there exists an  $a \in A$  such that  $f(a) = b$ . Furthermore, let  $g, h : B \rightarrow C$  be functions and assume ①  $g \circ f = h \circ f$ . We need to show that  $g = h$ , that is, for all  $b \in B$ ,  $g(b) = h(b)$ . But by assumption any  $b \in B$  is equal to  $f(a)$  for some  $a \in A$ , so the condition is equivalent to  $g(f(a)) = h(f(a))$ , which is just ①.

( $\Leftarrow$ ) We show the contrapositive: if  $f : A \rightarrow B$  is not surjective, then there exists a  $C$  and functions  $g, h : B \rightarrow C$  such that  $g \circ f = h \circ f$  but  $g \neq h$ . If  $f$  is not surjective, there exists a  $b_0 \in B$  such that for all  $a \in A$ ,  $f(a) \neq b_0$ . We can therefore choose two functions  $g$  and  $h$  such that they match on the range of  $f$ , but differ on  $b_0$ . For example, take  $C$  to be  $B$  with a new distinguished element  $\star$  added:  $C = B \cup \{\star\}$ . Let  $g : B \rightarrow B \cup \{\star\}$  be the inclusion  $g(b) = b$ , and let  $h(b_0) = \star$  and  $h(b) = b$  for all  $b \neq b_0$ . Then,  $g \circ f = h \circ f$ , since  $g$  and  $h$  defined to be equal for all elements in the range of  $f$ , but they differ on the element  $b_0$  not “covered” by  $f$ , hence  $g \neq h$ .

♪ The ( $\Leftarrow$ ) direction can be presented as a non-contrapositive argument as well. Let  $f : A \rightarrow B$  be a function and assume for all  $g, h : B \rightarrow C$ , if  $g \circ f = h \circ f$  then  $g = h$ . We need to show that for all  $b \in B$  there exists an  $a \in A$  such that  $f(a) = b$ . Choose  $C = [2] = \{0, 1\}$  and define  $g = \chi_B$  and  $h = \chi_{\vec{f}(A)}$ , where  $\vec{f}(A) \subseteq B$  is also called the range of  $f$ , i.e. the set  $\{f(a) \in B \mid a \in A\}$ . That is,  $g(b) = 1$  for all  $b$ , and  $h(b) = 1$  for all  $b$  in the range of  $f$ , and 0 otherwise. Now, for all  $a \in A$ ,  $g(f(a)) = h(f(a))$ , but by assumption this implies that  $g = h$ . This is only possible if the range of  $f$  is  $B$  itself, i.e.  $f$  is surjective.

**What would be an analogous condition for injections?**

Injectivity is equivalent to left-cancellability:  $f : B \rightarrow C$  is an injection iff for all sets  $A$  and functions  $g, h : A \rightarrow B$ , if  $f \circ g = f \circ h$  then  $g = h$ .

( $\Rightarrow$ ) Assume  $f : B \rightarrow C$  is an injection, and suppose that  $f \circ g = f \circ h$  for some  $g, h : A \rightarrow B$ . We need to show that for all  $a \in A$ ,  $g(a) = h(a)$ . Injectivity means that for all  $b_1, b_2 \in B$ , if

$f(b_1) = f(b_2)$  then  $b_1 = b_2$ . Instantiating this for  $g(a), h(a) \in B$ , and using the assumption  $f \circ g = f \circ h$ , we deduce that  $f(g(a)) = f(h(a))$  implies  $g(a) = h(a)$ . Since  $a \in A$  was arbitrary, we have that  $g = h$ .

( $\Leftarrow$ ) Assume that for all  $A$  and  $g, h: A \rightarrow B$ , ①  $f \circ g = f \circ h$  implies  $g = h$ . We need to show that for all  $b_1, b_2 \in B$ , if  $f(b_1) = f(b_2)$  then  $b_1 = b_2$ . Take  $b_1, b_2 \in B$  and assume that ②  $f(b_1) = f(b_2)$ ; for  $A = \{\ () \}$  the singleton set, define  $g, h: A \rightarrow B$  as  $g() = b_1$ , and  $h() = b_2$ . Then, by ②,  $f(g()) = f(b_1) = f(b_2) = f(h())$ , but then by ①  $g = h$  so  $b_1 = b_2$ .

3. Use the above sufficient condition to show that the identity function is a surjection, and the composition of surjections is a surjection.

**Identity.** We show that  $\text{id}_A: A \rightarrow A$  is a surjection. Let  $X$  be a set and  $g, h: A \rightarrow X$  be two functions, and assume  $g \circ \text{id}_A = h \circ \text{id}_A$ . Since the identity is the unit of composition, we get  $g = h$  immediately, so  $\text{id}_A$  is a surjection by §11.2.2.

**Composition.** Let  $f: A \rightarrow B$  and  $g: B \rightarrow C$  be surjections. We show that  $g \circ f: A \rightarrow C$  is a surjection via §11.2.2. Let  $h, i: C \rightarrow X$  be two functions and assume  $h \circ (g \circ f) = i \circ (g \circ f)$ ; we need to prove that  $h = i$ . Composition is associative, so  $(h \circ g) \circ f = (i \circ g) \circ f$  – and  $f$  is a surjection, so we have that  $h \circ g = i \circ g: B \rightarrow X$ . Similarly,  $g$  is surjective, so  $h = i$ .

♪ As is hopefully apparent now, exercises of the form *[object] is [defined set-theoretic concept] iff [it satisfies set-theoretic property]* are powerful and reusable proof principles:

|                                 |  |
|---------------------------------|--|
| $X$ is the union of $A$ and $B$ | iff $A \subseteq X, B \subseteq X$ , and for all $Y, (A \subseteq Y \wedge B \subseteq Y) \Rightarrow X \subseteq Y$ |
| $R$ is an equivalence relation  | iff $\text{id}_A \subseteq R, R = R^{\text{op}}$ and $R \circ R \subseteq R$   |
| $R$ is a partial function       | iff $R \circ R^{\text{op}} \subseteq \text{id}_B$  |
| $f$ is surjective               | iff for all $g, h: B \rightarrow X, g \circ f = h \circ f \Rightarrow g = h$   |

Thus, to prove that *[object] is [defined set-theoretic concept]*, instead of expanding the set-theoretic definition (which is usually given in terms of individual elements), we can reason via the higher-level set-theoretic properties which may result in more abstract and elegant proofs with preorder or equational reasoning. While perhaps not immediately as intuitive and not necessarily shorter, once you get used to the approach, you will be able to recognise and appreciate opportunities for reasoning via sufficient/universal properties rather than reaching for “let  $x \in A$  and show  $x \in B$ ” right away and making your proof low-level and often harder to follow. That said, for explicitly defined sets and functions, proving surjectivity (for example) from first principles may be more direct – see some exercises in the next section.

## 12. On images


### 12.1. Basic exercises

1. Let  $R_2 = \{(m, n) \mid m = n^2\}: \mathbb{N} \rightarrow \mathbb{Z}$  be the integer square-root relation. What is the direct image of  $\mathbb{N}$  under  $R_2$ ? And what is the inverse image of  $\mathbb{N}$ ?



By the definition of the direct and inverse relational images, we have:

$$\vec{R}_2(\mathbb{N}) = \mathbb{Z} \quad \overleftarrow{R}_2(\mathbb{N}) = \{0\} \cup \{n \in \mathbb{N} \mid n \text{ is not square}\}$$

 This may well be called a “trick question”, since the answer could hardly be more counterintuitive – then again, it follows directly from the definition of inverse relation images so there is not much to argue about!  $R_2$  relates every integer (on the right) with its square (on the left), a natural number:  $R_2 = \{(0, 0), (1, -1), (1, 1), (4, 2), (4, -2), (9, -3), (9, 3), \dots\}$ . The direct image of the natural numbers is therefore  $\mathbb{Z}$  itself, since the square of every integer is in  $\mathbb{N}$ . It may seem intuitively obvious that the inverse image of  $\mathbb{N} \subseteq \mathbb{Z}$  under the square root relation would be the set of square numbers, but this is distinctly *not* the case. Recall the definition of inverse relational images:

$$\overleftarrow{R}(Y \subseteq B) \triangleq \{a \in A \mid \forall b \in B. a R b \Rightarrow b \in Y\}$$

For  $R_2$ , and  $Y = \mathbb{N} \subseteq \mathbb{Z}$ , this becomes:

$$\overleftarrow{R}_2(\mathbb{N}) \triangleq \{m \in \mathbb{N} \mid \forall n \in \mathbb{Z}. m = n^2 \Rightarrow n \in \mathbb{N}\}$$

In other words, if there are any integers that square to an element of  $\overleftarrow{R}_2(\mathbb{N})$ , they all have to be natural numbers. 0 is certainly in the inverse image, since the only number that squares to 0 is 0 itself, and it is in  $\mathbb{N}$ . The problems start with nonzero square numbers like 1, 4, 9, etc.: there are exactly two integers that square to the same perfect square number, namely the square root, and the negative of the square root. Only one of these is a natural number, the other violates  $m = n^2 \Rightarrow n \in \mathbb{N}$  and therefore cannot be an element of the inverse image. Thus, the inverse image of natural numbers under the square-root relation contains no square numbers other than 0. Even worse is that every natural number which is *not* a perfect square (and therefore isn't related to any integers) vacuously satisfies the condition: for any  $n \in \mathbb{Z}$ ,  $2 \neq n^2$  so the hypothesis is never satisfied and the implication holds! As a result, the inverse image contains all the non-square natural numbers and 0.

You may rightly ask: why do we define inverse images in such a way? The answer is simply that this is the most natural way to define it as a dual of the direct image  $\vec{R}(X) \triangleq \{b \in B \mid \exists x \in X. x R b\}$ . Indeed, if we slightly rephrase the condition  $\exists x \in X. x R b$  to separate existence and membership of  $X$ , and compare it to the inverse image definition, we get:

$$\begin{aligned} \vec{R}(X) &\triangleq \{b \in B \mid \exists x \in A. x R b \wedge x \in X\} \\ \overleftarrow{R}(Y) &\triangleq \{a \in A \mid \forall y \in B. a R y \Rightarrow y \in Y\} \end{aligned}$$

As is often the case with mathematics, symmetry and simplicity takes precedence over intuition, and trying to define the inverse image to yield the “expected” results would needlessly complicate the definition. In fact, what we intuitively expect the inverse image of  $\mathbb{N}$  under  $R_2$  to be (the set of perfect squares) is nothing more than the direct image of  $\mathbb{N}$  under the opposite relation  $R_2^{\text{op}}$ .

2. For a relation  $R: A \leftrightarrow B$ , show that:

a)  $\vec{R}(X) = \bigcup_{x \in X} \vec{R}(\{x\})$  for all  $X \subseteq A$


Let  $X$  be a subset of  $A$ . We calculate as follows:

$$\begin{aligned} \vec{R}(X) &= \{b \in B \mid \exists x \in X. x R b\} \\ &= \{b \in B \mid \exists x \in X. \exists y' \in \{x\}. y' R b\} \\ &= \{b \in B \mid \exists x \in X. b \in \vec{R}(\{x\})\} \\ &= \bigcup_{x \in X} \vec{R}(\{x\}) \end{aligned}$$

b)  $\overleftarrow{R}(Y) = \{a \in A \mid \vec{R}(\{a\}) \subseteq Y\}$  for all  $Y \subseteq B$ .

Let  $Y$  be a subset of  $B$ . We calculate as follows:

$$\begin{aligned} \overleftarrow{R}(Y) &= \{a \in A \mid \forall y \in B. a R y \Rightarrow y \in Y\} \\ &= \{a \in A \mid \forall y \in B. (\exists a' \in \{a\}. a' R y) \Rightarrow y \in Y\} \\ &= \{a \in A \mid \forall y \in B. y \in \vec{R}(\{a\}) \Rightarrow y \in Y\} \\ &= \{a \in A \mid \vec{R}(\{a\}) \subseteq Y\} \end{aligned}$$

 This equivalent characterisations of inverse images highlights the requirement that every  $y \in B$  related to an  $a \in A$  has to be in  $Y$ , not just at least one.


## 12.2. Core exercises

1. For  $X \subseteq A$ , prove that the direct image  $\vec{f}(X) \subseteq B$  under an injective function  $f: A \rightarrow B$  is in bijection with  $X$ ; that is,  $X \cong \vec{f}(X)$ .

Let  $f: A \rightarrow B$  be an injective function and let  $X$  be a subset of  $A$ . We show that the direct image of  $X$  under  $f$  is isomorphic to  $X$  by constructing a bijection  $h: X \xrightarrow{\cong} \vec{f}(X)$ . Define  $h$  as

$$h(x \in X) = f(x) \in \vec{f}(X)$$

By construction,  $h$  is a function from  $X$  to  $\vec{f}(X)$  because every output of  $f$  for an input in  $X$  ends up in the direct image. We show that  $h$  is surjective and injective. Take any element  $y \in \vec{f}(X)$ ; by definition, there must exist an element  $x \in X$  such that  $f(x) = h(x) = y$ , which is the condition for surjectivity of  $h$ . Now, take  $x_1, x_2 \in X$  and assume that  $h(x_1) = h(x_2)$ . Then,  $f(x_1) = f(x_2)$ , but  $f$  is injective, so  $x_1 = x_2$  – proving that  $h$  is injective too. As a direct corollary, the range of an injection is isomorphic to the domain:  $\vec{f}(A) \cong A$ .

 This is a situation where proving injectivity and surjectivity is more convenient than trying to precisely formulate an inverse function that maps  $y \in \vec{f}(X)$  to “the element in  $X$  that got uniquely mapped to  $y$ ” and using this to calculate the inverse laws.

2. Prove that for a surjective function  $f: A \rightarrow B$ , the direct image function  $\vec{f}: \mathcal{P}(A) \rightarrow \mathcal{P}(B)$  is surjective.

Assume  $f : A \twoheadrightarrow B$  is a surjection: for all  $b \in B$ , there exists an  $a \in A$  such that  $f(a) = b$ . We need to prove that for any element  $Y \in \mathcal{P}(B)$  there exists an  $X \in \mathcal{P}(A)$  such that  $\vec{f}(X) = Y$ . Thus, take a subset  $Y \subseteq B$ , and let the corresponding subset of  $A$  be the inverse image  $\overleftarrow{f}(Y) \subseteq A$ . We now need to show that  $\vec{f}(\overleftarrow{f}(Y)) = Y$ , for which we calculate:

$$\begin{aligned}\vec{f}(\overleftarrow{f}(Y)) &= \{b \in B \mid \exists a \in \overleftarrow{f}(Y). f(a) = b\} \\ &= \{b \in B \mid \exists a \in A. f(a) \in Y \wedge f(a) = b\} \\ &= \{b \in Y \mid \exists a \in A. f(a) = b\}\end{aligned}$$

but the last set is precisely  $Y$  since  $f$  is surjective and therefore the comprehension condition holds for all  $b \in Y$ . As a direct corollary, the range of a surjection is equal to the codomain:  $\vec{f}(A) = B$ . A bijection is both an injection and a surjection, so  $A \cong \vec{f}(A) = B$ .

3. Show that any function  $f : A \rightarrow B$  can be decomposed into an injection and a surjection: that is, there exists a set  $X$ , a surjection  $s : A \twoheadrightarrow X$  and an injection  $i : X \hookrightarrow B$  such that  $f = i \circ s$ .

Let  $f : A \rightarrow B$  be a function, not necessarily a surjection or injection. Take  $X$  to be the range of  $f$ , that is, the direct image of its domain:  $X = \vec{f}(A) \subseteq B$ . Then, by definition, every element  $b \in \vec{f}(A)$  has an associated element  $a \in A$  such that  $f(a) = b$ , so  $f$  with its codomain restricted to its range is a surjection – hence,  $s(a) = f(a) : A \twoheadrightarrow \vec{f}(A)$ . The range of  $f$  is a subset of the codomain, so we have the canonical inclusion  $i(b) = b : \vec{f}(A) \hookrightarrow B$  which is an injection. For all  $a \in A$ ,  $i(s(a)) = i(f(a)) = f(a)$ , so the composite  $i \circ s$  is indeed equal to  $f$ , as required.

♪ When  $A = B$  (and  $f : A \rightarrow A$  is an endofunction), the construction of course still works. In fact, it gives one half of the idempotent-splitting example §9.2.1, in which an idempotent endofunction  $e : A \rightarrow A$  is split through its range  $\{e(b) \mid b \in B\} = \vec{e}(B)$  into functions  $r$  and  $s$  as  $s \circ r = e$  which, thanks to the idempotence condition, form a section-retraction pair:  $r \circ s = \text{id}_B$ . Sections are always injections, and the constructed retraction is a surjection, matching the result shown in this exercise.

4. For a relation  $R : A \leftrightarrow B$ , prove that

a)  $\vec{R}(\bigcup \mathcal{F}) = \bigcup \{\vec{R}(X) \mid X \in \mathcal{F}\}$  for all  $\mathcal{F} \subseteq \mathcal{P}(A)$

Let  $\mathcal{F} \subseteq \mathcal{P}(A)$  be a family of subsets. We have the following calculation:

$$\begin{aligned}b \in \vec{R}(\bigcup \mathcal{F}) &\iff \exists a \in \bigcup \mathcal{F}. a R b \\ &\iff \exists X \in \mathcal{F}. \exists a \in X. a R b \\ &\iff \exists X \in \mathcal{F}. b \in \vec{R}(X) \\ &\iff \exists Y \in \{\vec{R}(X) \mid X \in \mathcal{F}\}. b \in Y \\ &\iff b \in \bigcup \{\vec{R}(X) \mid X \in \mathcal{F}\}\end{aligned}$$

b)  $\overleftarrow{R}(\bigcap \mathcal{G}) = \bigcap \{\overleftarrow{R}(Y) \mid Y \in \mathcal{G}\}$  for all  $\mathcal{G} \subseteq \mathcal{P}(B)$

Let  $\mathcal{F} \subseteq \mathcal{P}(A)$  be a family of subsets. We have the following calculation:

$$\begin{aligned} a \in \overline{\bigcap \mathcal{G}} &\iff \forall b \in B. aRb \Rightarrow a \in \bigcap \mathcal{G} \\ &\iff \forall b \in B. aRb \Rightarrow \forall Y \in \mathcal{G}. a \in Y \\ &\iff \forall Y \in \mathcal{G}. \forall b \in B. aRb \Rightarrow a \in Y \\ &\iff \forall Y \in \mathcal{G}. a \in \overline{R(Y)} \\ &\iff \forall X \in \{\overline{R(Y)} \mid Y \in \mathcal{G}\}. a \in X \\ &\iff a \in \bigcap \{\overline{R(Y)} \mid Y \in \mathcal{G}\} \end{aligned}$$

5. Show that, by the inverse image, every map  $A \rightarrow B$  induces a *Boolean algebra map*  $\mathcal{P}(B) \rightarrow \mathcal{P}(A)$ . That is, for every function  $f : A \rightarrow B$ , its inverse image preserves set operations:

- $\overline{f(\emptyset)} = \emptyset$

$$a \in \overline{f(\emptyset)} \iff f(a) \in \emptyset \iff \text{false} \iff a \in \emptyset$$

- $\overline{f(B)} = A$

$$a \in \overline{f(B)} \iff f(a) \in B \iff \text{true} \iff a \in A$$

- $\overline{f(X \cup Y)} = \overline{f(X)} \cup \overline{f(Y)}$

$$\begin{aligned} a \in \overline{f(X \cup Y)} &\iff f(a) \in (X \cup Y) \iff f(a) \in X \vee f(a) \in Y \\ &\iff a \in \overline{f(X)} \vee a \in \overline{f(Y)} \iff a \in \overline{f(X)} \cup \overline{f(Y)} \end{aligned}$$

- $\overline{f(X \cap Y)} = \overline{f(X)} \cap \overline{f(Y)}$

$$\begin{aligned} a \in \overline{f(X \cap Y)} &\iff f(a) \in (X \cap Y) \iff f(a) \in X \wedge f(a) \in Y \\ &\iff a \in \overline{f(X)} \wedge a \in \overline{f(Y)} \iff a \in \overline{f(X)} \cap \overline{f(Y)} \end{aligned}$$

- $\overline{f(X^c)} = (\overline{f(X)})^c$

$$a \in \overline{f(X^c)} \iff f(a) \in X^c \iff \neg(f(a) \in X) \iff \neg(a \in \overline{f(X)}) \iff a \in (\overline{f(X)})^c$$

## 13. On countability

### 13.1. Basic exercises

1. Prove that every finite set is countable.

If the set is empty, it is countable by definition. Otherwise, if  $A$  is finite, it has at most  $\#A = n > 0$  elements. Thus, an enumeration  $\mathbb{N} \rightarrow A$  can be constructed by mapping the first  $n$  natural numbers to distinct elements of  $A$  (e.g. by putting them in some order and assigning  $k : [0..n-1]$  to the  $k^{\text{th}}$  element), and the rest of the naturals to a single element  $a_0 \in A$ . The mapping is surjective by construction (the  $k^{\text{th}}$  element of  $A$  is listed at  $k$ ) so it

is an enumeration.

2. Demonstrate that  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  are countable sets.

$\mathbb{N}$  is enumerated by the identity function, which is in particular a surjection.

$\mathbb{Z}$  is enumerated by alternating between positive and negative numbers:  $0, 1, -1, 2, -2, \dots$

Explicitly,  $e: \mathbb{N} \rightarrow \mathbb{Z}$  is the enumeration

$$e(n) \triangleq \begin{cases} \frac{n+1}{2} & \text{if } n \text{ is odd} \\ -\frac{n}{2} & \text{if } n \text{ is even} \end{cases}$$

$\mathbb{Q}$  is enumerable using the traversal of the coordinate plane demonstrated on [Slide 398](#).

### 13.2. Core exercises

1. Let  $A$  be an infinite subset of  $\mathbb{N}$ . Show that  $A \cong \mathbb{N}$ . *Hint:* Adapt the argument shown in the proof of [Proposition 144](#), showing that the map  $\mathbb{N} \rightarrow A$  is both injective and surjective.

Let  $A$  be an infinite subset of  $\mathbb{N}$ . We construct a bijection  $\mathbb{N} \xrightarrow{\cong} A$  to show that they are isomorphic. To this end, define the function  $\mu: \mathbb{N} \rightarrow A$  as follows:

$$\mu(0) \triangleq \min(A) \quad \mu(n+1) \triangleq \min\{k \in A \mid \mu(n) < k\} = \min(A \setminus \{\mu(k) \mid k \leq n\})$$

We will denote the set  $A \setminus \{\mu(k) \mid k \leq n\}$  as  $A_n$ , so that  $\mu(n+1) = \min(A_n)$ .

To show that  $\mu$  is an injection, we need to prove that if  $\mu(m) = \mu(n)$  then  $m = n$ . We equivalently prove the contrapositive: if  $m \neq n$ , then  $\mu(m) \neq \mu(n)$ . Without loss of generality, assume that  $m < n$ ; then,  $\mu(m) \in \{\mu(k) \mid k \leq n-1\}$ , so  $\mu(m) \notin \{\mu(k) \mid k \leq n-1\}^c = A \setminus \{\mu(k) \mid k \leq n-1\}$ . On the other hand,  $\mu(n)$  is an element of the latter set (its minimum), which means that  $\mu(m)$  cannot equal  $\mu(n)$ .

To show that  $\mu$  is a surjection, we let  $a$  be an arbitrary element of  $A$  and show that there is an  $i \in \mathbb{N}$  such that  $\mu(i) = a$ . Consider the set  $\{k \in \mathbb{N} \mid \mu(k) < a\}$  of numbers which get mapped to an element below  $a$  in  $A$ , and let  $N$  be the size of this set (which, by the Pigeonhole Principle, must be at most  $a$ ). Now,  $A_N$  is the subset of  $A$  obtained by removing its  $N$  least elements, and by construction, its least element is  $a$ . But if  $a = \min(A_N)$ , then it is equal to  $\mu(N+1)$  by the definition of  $\mu$ , so we indeed have the natural number  $i = N+1$  such that  $\mu(i) = a$ .

2. For an infinite set  $A$ , prove that the following are equivalent:

- There is a bijection  $\mathbb{N} \xrightarrow{\cong} A$ .
- There is a surjection  $\mathbb{N} \rightarrow A$ .
- There is an injection  $A \rightarrow \mathbb{N}$ .


(a)  $\Rightarrow$  (b), (c) Every bijection  $f: \mathbb{N} \xrightarrow{\cong} A$  has an inverse  $f^{-1}: A \xrightarrow{\cong} \mathbb{N}$  which is itself a bijection. Every bijection is a surjection (giving  $\mathbb{N} \rightarrow A$  from  $f$ ) and an injection (giving  $A \rightarrow \mathbb{N}$  from  $f^{-1}$ ).

**(b)  $\Rightarrow$  (c)** Let  $s: \mathbb{N} \rightarrow A$  be a surjection. We need to construct an injection  $i: A \rightarrow \mathbb{N}$ , assigning a unique numeric code to every element of  $A$ . As  $s$  is a surjection, the inverse images of the singleton subsets of  $A$  (also called the set of *fibres* of elements of  $A$ ) are all non-empty, and as they are subsets of the natural numbers, they must have a minimal element. Thus, define  $i: A \rightarrow \mathbb{N}$  as a function that maps an  $x \in A$  to the smallest natural number that maps to  $x$ :

$$i(x) = \min(\overline{s}(\{x\})) = \min\{n \in \mathbb{N} \mid s(n) = x\}$$

This encodes an element  $x \in A$  by the position of its first occurrence in the enumeration given by  $s$ . We can see that  $i$  is injective as  $s$  acts as its retraction: for any  $x \in S$ ,  $s(i(x)) = s(n)$  where  $n$  is the smallest natural number such that  $s(n) = x$  so clearly  $s(i(x)) = s(n) = x$  and  $s \circ i = \text{id}_A$ , as required.

**(c)  $\Rightarrow$  (a)** Let  $i: A \rightarrow \mathbb{N}$  be an injection. We need to construct a bijection  $A \xrightarrow{\cong} \mathbb{N}$ , or equivalently, show that  $A$  and  $\mathbb{N}$  are isomorphic. By §12.2.1, the direct image of the domain under the injection  $i$  (i.e. the range of  $i$ ) is isomorphic to the domain:  $\overrightarrow{i}(A) \cong A$ . By assumption,  $A$  is infinite, so  $\overrightarrow{i}(A) \subseteq \mathbb{N}$  is infinite as well. But then it is an infinite subset of the natural numbers, and by §13.2.1, it is isomorphic to  $\mathbb{N}$ . Hence we have the chain  $A \cong \overrightarrow{i}(A) \cong \mathbb{N}$ , establishing the bijection  $\mathbb{N} \xrightarrow{\cong} A$ .

 If you look at other resources on countability, you will face several competing, but equivalent (but sometimes not *quite* equivalent) definitions which make translating between various statements and proofs a bit of a chore – especially since the same terms are used by different authors for different purposes. This course uses *enumerable* for sets which have a surjection  $\mathbb{N} \rightarrow A$ , and *countable* for sets which are enumerable or empty (since one can't have a function into the empty set so this needs to be handled as a special case). Other literature (e.g. [Wikipedia](#)) calls sets which are isomorphic to  $\mathbb{N}$  *countably infinite*, and sets which are either finite or countably infinite are called *countable*. The countability condition can be equivalently stated as the set being isomorphic to some subset of the natural numbers, i.e. coming with an injection  $A \rightarrow \mathbb{N}$ . Yet other terms used for the above notions are *at most countable*, *enumerable*, *denumerable*, *equinumerous*, *listable*, etc.

As this exercise shows, the bijection/surjection/injection notions are equivalent when the set  $A$  is infinite, and appropriate connections can be made when the sets are empty or finite. This gives us two equivalent ways of showing that a set is enumerable: either by constructing an enumeration  $\mathbb{N} \rightarrow A$ , or by defining an encoding function  $A \rightarrow \mathbb{N}$  that maps every element of  $A$  to a unique natural number. This is related to a concept called *Gödel encoding* which will be covered in more detail in the IB Computation Theory course.

3. Prove that:

a) Every subset of a countable set is countable.

Assume  $S \subseteq A$  for some sets  $A$ . If  $A$  is finite, so is  $S$  and it is countable. If  $A$  is infinite and  $S$  is finite,  $S$  is countable. If  $S$  is also infinite, we can show that it is enumerable by providing an injection  $S \rightarrow \mathbb{N}$ . But by assumption we have an injection  $f : A \rightarrow \mathbb{N}$  and subsets come with a canonical injective inclusion function  $\iota : S \rightarrow A$ , so the composite  $S \rightarrow A \rightarrow \mathbb{N}$  is an injection.

**b) The product and disjoint union of countable sets is countable.**

Assume  $A$  and  $B$  are countable sets.

If either is empty, the Cartesian product will be empty too and therefore countable. In the general case, assume they are enumerable and there are injections  $f : A \rightarrow \mathbb{N}$  and  $g : B \rightarrow \mathbb{N}$  that uniquely encode the elements of the sets. We define a function  $p : A \times B \rightarrow \mathbb{N}$  as follows:


$$p(a, b) = 2^{f(a)} \cdot 3^{g(b)}$$

By the Fundamental Theorem of Arithmetic, the mapping is injective: the output of this mapping will have a unique prime decomposition, and the number of 2 and 3 factors will give the unique code of elements of  $A$  and  $B$ , respectively. By §13.2.2, the injection  $p$  will imply that  $A \times B$  is enumerable.

If both sets are empty, their disjoint union will be empty and therefore countable. If either is empty, the disjoint union will be isomorphic to the other set, which is countable by assumption. In the general case, assume  $A$  and  $B$  are enumerable and come with injections  $f : A \rightarrow \mathbb{N}$  and  $g : B \rightarrow \mathbb{N}$ . We define the function  $u : A \uplus B \rightarrow \mathbb{N}$  as follows:

$$u(0, a) = 2^{f(a)} \quad u(1, b) = 3^{g(b)}$$

Again, by the Fundamental Theorem of Arithmetic, the prime decomposition of the output of  $u$  will uniquely determine the output of  $f(a)$  or  $g(b)$  which in turn uniquely determine the input  $a$  and  $b$  by assumption. By §13.2.2, the injection  $u$  will imply that  $A \uplus B$  is enumerable.

 Constructing unique encodings using products of primes is a useful alternative to the visually descriptive “diagonal traversal” enumeration which is often quite difficult to define explicitly. The specific choice of encoding can of course vary (e.g. we could have encoded disjoint unions via even and odd numbers) but there is no reason to look for the most “efficient” solution since all we care about is whether the enumeration/encoding is possible or not.

**4. For a set  $A$ , prove that there is no injection  $\mathcal{P}(A) \rightarrow A$ .**

We suppose there is an injection  $f : \mathcal{P}(A) \rightarrow A$  and derive a contradiction. By §11.2.1, the injection  $f$  has a retraction  $r : A \rightarrow \mathcal{P}(A)$  which must be a surjection since it undoes the application of  $f$  on any element of  $A$ . But then  $r$  would be a surjection from a set to its powerset, which is impossible due to [Cantor’s Theorem](#).

### 13.3. Optional advanced exercise

1. Prove that if  $A$  and  $B$  are countable sets then so are  $A^*$ ,  $\mathcal{P}_{\text{fin}}(A)$  and  $\text{PFun}_{\text{fin}}(A, B)$ .

All the results follow from [Proposition 154](#): an enumerable indexed disjoint union of enumerable sets is enumerable. An enumeration-style proof is presented in the notes, but an encoding-style argument is straightforward too: we can encode elements of  $\bigsqcup_{i \in I} A_i$  as

$$d(i \in I, a \in A_i) = p(c(i), c_i(a)) = 2^{c(i)} \cdot 3^{c_i(a)}$$

where  $c : I \rightarrow \mathbb{N}$  is the encoding of the index set, and  $c_i : A_i \rightarrow \mathbb{N}$  is an encoding for every element of the indexed family.

$A^* = \bigsqcup_{n \in \mathbb{N}} A^n$  is the set of finite sequences on  $A$ . If  $A$  is empty, the only finite sequence with elements from  $A$  is the empty sequence, so  $\{()\}$  is finite and countable. In the general case, we know that  $\mathbb{N}$  is enumerable, and  $A^n$  is the iterated binary Cartesian product of enumerable sets and hence is an enumerable set for every  $n \in \mathbb{N}$  (quick inline induction proof:  $A^0 = \emptyset$  is countable;  $A^{k+1} = A^k \times A$  is the Cartesian product of countable  $A^k$  by IH, and countable  $A$  by assumption). By the above proposition, the  $\mathbb{N}$ -indexed disjoint union of countable sets is countable.

$\mathcal{P}_{\text{fin}}(A) = \{S \subseteq A \mid S \text{ is finite}\}$  is the set of finite subsets of  $A$ . If  $A$  is empty,  $\mathcal{P}_{\text{fin}}(A) = \{\emptyset\}$  which is finite so countable. Otherwise, the set  $A$  has an encoding  $c : A \rightarrow \mathbb{N}$  which imposes an ordering on the elements on  $A$  based on the ordering of their code: for  $a, b \in A$ ,  $a \sqsubseteq b$  if  $c(a) \leq c(b)$ . This ordering restricts to every subset of  $A$ , so in particular, finite subsets of  $A$  can be mapped to finite sequences of elements of  $A$  according to the ordering  $\sqsubseteq$ . Then, the set of finite subsets of  $A$  is isomorphic to the set of finite sequences on  $A$ , which is countable for a countable  $A$ .

$\text{PFun}_{\text{fin}}(A, B) = \bigsqcup_{S \in \mathcal{P}_{\text{fin}}(A)} S \Rightarrow B$  is the set of partial functions with a finite domain of definition from  $A$  to  $B$ . If  $A$  or  $B$  are empty, the totally undefined function is the only element of the set so it is countable. Otherwise, the disjoint union is indexed by  $\mathcal{P}_{\text{fin}}(A)$  which is enumerable by the result above. The function space  $S \Rightarrow B$  has a finite domain  $S$ , so a single function  $f : S \rightarrow B$  can be captured as a finite sequence of elements of  $B$  as  $(f(s_1), f(s_2), f(s_3), \dots, f(s_n))$  where  $n = \#S$  and  $s_i$  is the  $i^{\text{th}}$  element of  $S$  in some ordering (which is always possible to define for a finite  $S$ ). Thus,  $S \Rightarrow B \cong B^{\#S}$  which is countable for any countable  $B$ . By Proposition 154, the set  $\text{PFun}_{\text{fin}}(A, B)$  is a countable indexed disjoint union of countable sets and is therefore itself countable.

## 14. On inductive definitions

1. Let  $L$  be the subset of  $\{a, b\}^*$  inductively defined by the axiom  $\frac{u}{\varepsilon}$  and rule  $\frac{u}{aub}$  for  $u \in \{a, b\}^*$ .

a) Use *rule induction* to prove that every string in  $L$  is of the form  $a^n b^n$  for some  $n \in \mathbb{N}$ .



We prove that for every string  $s$  in the set  $L$  inductively defined by the axiom and rule, there exists a natural number  $n$  such that  $s = a^n b^n$ .

**Axiom**  $\frac{\varepsilon}{\varepsilon}$  The string  $s$  must be the empty string  $\varepsilon$ , and for  $n = 0$  we have that  $\varepsilon = a^0 b^0$ .

**Rule**  $\frac{u}{aub}$  Let  $s = aub$  for some string  $u$  and assume the  $\textcircled{\text{IH}}$ : there exists a natural number  $k$  such that  $u = a^k b^k$ . Then,  $s = aub \stackrel{\text{IH}}{=} aa^k b^k b = a^{k+1} b^{k+1}$  so the witness  $n = k + 1$  satisfies the required property.

b) Use *mathematical induction* to prove that for all  $n \in \mathbb{N}$ ,  $a^n b^n \in L$ .

**Base case:**  $n = 0$ . The string  $a^0 b^0$  is the empty string  $\varepsilon$ , which is an element of  $L$  by the defining axiom.

**Inductive step**  $n = k + 1$ . Assume the  $\textcircled{\text{IH}}$ :  $a^k b^k \in L$ . We prove that the string  $a^{k+1} b^{k+1}$  is in  $L$  as well. By definition of string repetition,  $a^{k+1} b^{k+1} = aa^k b^k b$ . The  $\textcircled{\text{IH}}$  states that  $a^k b^k \in L$ , and the rule can be applied to deduce that  $aa^k b^k b \in L$  as well.

c) Conclude that  $L = \{a^n b^n \mid n \in \mathbb{N}\}$ .

In the previous two parts we have shown that every string of  $L$  is of a particular form  $a^n b^n$  for  $n \in \mathbb{N}$ , and that every string of this form is in  $L$ . Thus, we have the subset inclusions  $L \subseteq \{a^n b^n \mid n \in \mathbb{N}\}$  and  $\{a^n b^n \mid n \in \mathbb{N}\} \subseteq L$ , proving that the sets are equal.

d) Suppose we add the string  $a$  to  $L$  to get  $L' = L \cup \{a\}$ . Is  $L'$  closed under the axiom and rule? If not, characterise the strings that would be in the smallest set containing  $L'$  that is closed under the axiom and rule.

The resulting language  $L'$  would not be closed: we can use the rule to generate the strings  $aab$ ,  $aaabb$ ,  $a^{n+1} b^n$ , which are not of the required form and therefore are not already part of the language. The closure of  $L'$  under the rule and axiom would therefore be  $\{a^n b^n \mid n \in \mathbb{N}\} \cup \{a^{n+1} b^n \mid n \in \mathbb{N}\}$ .

2. Suppose  $R: X \leftrightarrow X$  is a binary relation on a set  $X$ . Let  $R^\dagger: X \leftrightarrow X$  be inductively defined by the following axioms and rules:

$$\frac{}{(x, x) \in R^\dagger} \quad (x \in X) \qquad \frac{(x, y) \in R^\dagger}{(x, z) \in R^\dagger} \quad (x \in X \text{ and } y R z)$$

a) Show that  $R^\dagger$  is reflexive and that  $R \subseteq R^\dagger$ .

We show that  $R^\dagger$  is reflexive by giving a derivation of  $(x, x) \in R^\dagger$  for all  $x \in X$ . This is simply the first axiom defining the relation.

Next, we show that for all  $(x, y) \in R$ ,  $(x, y) \in R^\dagger$  by providing a derivation:

$$\frac{\overline{(x, x) \in R^\dagger}}{(x, y) \in R^\dagger} \quad (x \in X \text{ and } x R y)$$

b) Use rule induction to show that  $R^\dagger$  is a subset of

$$S \triangleq \{ (y, z) \in X \times X \mid \forall x \in X. (x, y) \in R^\dagger \implies (x, z) \in R^\dagger \}$$

Deduce that  $R^\dagger$  is transitive.

We show that for all  $(y, z) \in R^\dagger$ , we have that for all  $x \in X$  such that  $(x, y) \in R^\dagger$ ,  $(x, z) \in R^\dagger$  by rule induction.

**Axiom**  $\frac{}{(y, y) \in R^\dagger}$  We clearly have  $(x, y) \in R^\dagger$  implying  $(x, y) \in R^\dagger$ , as required.

**Rule**  $\frac{(x, y) \in R^\dagger}{(x, z) \in R^\dagger}$  with ①  $y R z$ . Assume the ⑩:  $(x, y) \in S$ . To show  $(x, z) \in S$ , let  $w \in X$  be an element and suppose that ②  $(w, x) \in R^\dagger$ ; we prove  $(w, z) \in R^\dagger$  by giving a derivation:

$$\frac{\text{⑩}}{\frac{(w, y) \in R^\dagger}{(w, z) \in R^\dagger}} \quad (w \in X \text{ and } \text{① } y R z)$$

where the ⑩  $(x, y) \in S$  is applied to the assumption ②  $(w, x) \in R^\dagger$  to deduce  $(w, y) \in R^\dagger$ , as required.

To prove that  $R^\dagger$  is transitive, we need to show that  $(x, y), (y, z) \in R^\dagger$  implies  $(x, z) \in R^\dagger$ . Since  $R^\dagger \subseteq S$ , we also have  $(y, z) \in S$ , which, by definition of  $S$  and the assumption  $(x, y) \in R^\dagger$  implies  $(x, z) \in R^\dagger$ .

c) Suppose that  $T : X \rightarrow X$  is a reflexive and transitive binary relation and that  $R \subseteq T$ . Use rule induction to show that  $R^\dagger \subseteq T$ .

We show that for all  $(x, y) \in R^\dagger$ ,  $(x, y) \in T$  by rule induction.

**Axiom**  $\frac{}{(y, y) \in R^\dagger}$  Since  $T$  is reflexive, we have that  $(y, y) \in T$ .

**Rule**  $\frac{(x, y) \in R^\dagger}{(x, z) \in R^\dagger}$  with ①  $y R z$ . Assume the ⑩:  $(x, y) \in T$ . Since  $R \subseteq T$ , we also have  $(y, z) \in T$  from ①; then, since  $T$  is transitive, we deduce  $(x, z) \in T$  using the ⑩, which is what we were meant to prove.

d) Deduce from above that  $R^\dagger$  is equal to  $R^*$ , the reflexive-transitive closure of  $R$ .

In parts (a) and (b) we showed that  $R^\dagger$  is reflexive and transitive; in part (a) we also proved that  $R \subseteq R^\dagger$ . Finally, part (c) established that  $R^\dagger$  is smaller than any other reflexive-transitive superset of  $R$ , which is the universal characterisation of the reflexive-transitive closure of  $R$ .

3. Let  $L$  be a subset of  $\{a, b\}^*$  inductively defined by the axiom and rules (for  $u \in \{a, b\}^*$ ):

$$\frac{}{ab} \qquad \frac{au}{au^2} \qquad \frac{ab^3u}{au}$$

a) Is  $ab^5$  in  $L$ ? Give a derivation, or show that there isn't one.

The string  $ab^5$  is indeed in  $L$ , as witnessed by the following derivation:

$$\frac{\frac{\frac{\frac{\frac{}{ab}}{ab^2}}{ab^4}}{ab^8}}{ab^5}}$$

b) Use rule induction to show that every  $u \in L$  is of the form  $ab^n$  with  $n = 2^k - 3m \geq 0$  for some  $k, m \in \mathbb{N}$ .

**Axiom**  $\frac{}{ab}$  We have that  $ab = ab^1$  and  $1 = 2^k - 3m$  for  $k = m = 0$ .

**Rule**  $\frac{au}{au^2}$  If by the IH we have that  $u = b^{2^l - 3n}$ , then  $u^2 = b^{2^{l+1} - 3 \cdot (2n)}$ , so  $au^2$  is of the required form with  $k = l + 1$  and  $m = 2n$ .

**Rule**  $\frac{ab^3u}{au}$  If by the IH we have that  $b^3u = b^{2^l - 3n}$ , then by removing the first 3  $b$ s we have that  $u = b^{2^l - 3(n+1)}$ ; thus,  $au$  is of the required form with  $k = l$  and  $m = n + 1$ .

c) Is  $ab^3$  in  $L$ ? Give a derivation, or show that there isn't one.

If  $ab^3$  were in  $L$ , by part (b) it must be of the form  $ab^{2^k - 3m}$  for some  $k, m \in \mathbb{N}$ . This is not possible however, since  $2^k - 3m = 3 \iff 2^k = 3(m + 1)$  would require  $3(m + 1)$  to be a power of 2; but the only prime factor of  $2^k$  is 2 so it can't be a multiple of 3.

d) Find an explicit characterisation of the elements of the language as a set comprehension, and prove (along the lines of §14.1) that it coincides with the inductively defined set  $L$ .

We claim that  $L = \{ab^{2^k - 3m} \mid 2^k - 3m \geq 0\}$ . We've already shown the  $\subseteq$  direction, proving that every string in  $L$  is of the appropriate form. We now show that every string of the appropriate form has a derivation; namely, that for all  $k \in \mathbb{N}$ ,

$$\forall m \in \mathbb{N}. 2^k - 3m \geq 0 \implies ab^{2^k - 3m} \in L$$

which we prove by mathematical induction on  $k$ .

**Base case:**  $k = 0$ . The only  $m$  for which the hypothesis  $2^0 - 3m \geq 0$  is satisfied is  $m = 0$ , and for this we have a derivation of  $ab^{2^0 - 3 \cdot 0} = ab \in L$  by the axiom.

**Inductive step:**  $k = l + 1$ . Assume the  $\textcircled{\text{IH}}$ :

$$\forall m \in \mathbb{N}. 2^l - 3m \geq 0 \implies ab^{2^l - 3m} \in L$$

and prove  $\forall m \in \mathbb{N}. 2^{l+1} - 3m \geq 0 \implies ab^{2^{l+1} - 3m} \in L$  by nested mathematical induction on  $m$ .

**Inner base case:**  $m = 0$ . By the  $\textcircled{\text{IH}}$  we have that  $ab^{2^l} \in L$ , and by applying the rule  $\frac{au}{au^2}$  we can derive  $ab^{2^{l+1}} \in L$ .

**Inner inductive step:**  $m = n + 1$ . If, by the nested IH we have that  $ab^{2^{l+1} - 3n} \in L$ , then by applying the rule  $\frac{ab^3u}{au}$  we can derive  $ab^{2^{l+1} - 3n - 3} = ab^{2^{l+1} - 3(n+1)} \in L$ .

## 15. On regular expressions

1. Find regular expressions over  $\{0, 1\}$  that determine the following languages:

a)  $\{u \mid u \text{ contains an even number of 1's}\}$

We should only be able to add 1s in pairs, so we take the regex  $(0|10^*1)^*$ .


b)  $\{u \mid u \text{ contains an odd number of 0's}\}$

After requiring one 0, we ask for an even number of 0s:  $1^*0(1|01^*0)^*$ .

2. Show that  $b^*a(b^*a)^*$  and  $(a|b)^*a$  are equivalent regular expressions, that is, a string matches one iff it matches the other. Your reasoning should be rigorous but can be informal.

First note that any string  $u$  matching  $b^*a(b^*a)^*$  is a concatenation  $u = u_1u_2 \cdots u_n$  of one or more (i.e.  $n \geq 1$ ) strings in  $\{a, b\}^*$  matching  $b^*a$ . Each  $u_i$  ends with an  $a$  and hence (because  $n \geq 1$ ), so does  $u$ . Therefore  $u$  matches  $(a|b)^*a$ .

Conversely, if  $u$  matches  $(a|b)^*a$  it is a string in  $\{a, b\}^*$  ending with an  $a$ : looking at the occurrences of  $a$  in  $u$ , we can express  $u$  as  $u = b^{n_1}ab^{n_2}a \cdots b^{n_k}a$  for some  $k \geq 1$  and some  $n_1, \dots, n_k \geq 0$ ; and hence  $u$  matches  $b^*a(b^*a)^*$ .

 Equivalence of regular expressions is more difficult to establish in general – reasoning by “observation” or pattern analysis like above does not scale to more complicated regexes. The question will be revisited, however, in the second half of the course, using some additional developments that will allow us to check equivalence of regular expressions in finite time.

3. Extend the [concrete syntax](#), [abstract syntax](#), [parsing relation](#) of regular expressions, and the [matching relation](#) between strings and regular expressions with the following constructs:

a)  $r?$ : matches the regex  $r$  zero or one times. For example,  $ab?c$  is matched by  $ac$  and  $abc$ , but not  $abbc$ .

b)  $r^+$ : matches the regex  $r$  one or more times. For example,  $ab^+c$  is matched by  $abc$  and  $abbbbc$ , but not  $ac$ .

We extend the alphabet  $\Sigma'$  with the symbols  $?$  and  $^+$  and the concrete syntax with the following rules:

$$\frac{r}{r?} \qquad \frac{r}{r^+}$$

The abstract syntax is extended with the unary constructors *Opt* and *Plus*, and parsing is modified as follows:

$$\frac{r \sim R}{r? \sim \text{Opt}(R)} \qquad \frac{r \sim R}{r^+ \sim \text{Plus}(R)}$$


Finally, we add the following axioms and rules to the matching relation:

$$\frac{}{(\varepsilon, r?)}, \quad \frac{(u, r)}{(u, r^+)}, \quad \frac{(u, r)}{(u, r?)}, \quad \frac{(u, r) \quad (v, r^+)}{(uv, r^+)}$$

Show that  $(r^+)?$  is equivalent to  $r^*$ . Is that the case for  $(r?)^+$  as well?

The regex  $(r^+)?$  matches either the empty string  $\varepsilon$ , or one or more repetitions of a string matched by  $r$ . Combined, it matches zero or more repetitions of a string matched by  $r$ , which is precisely the meaning of  $r^*$ .

The regex  $(r?)^+$  matches one or more repetitions of either the empty string, or a string matched by  $r$ . In particular, it matches the empty string (if all the repetitions are empty), and any nonzero number of occurrences of  $r$ . Again, this is the same as the meaning of  $r^*$ .

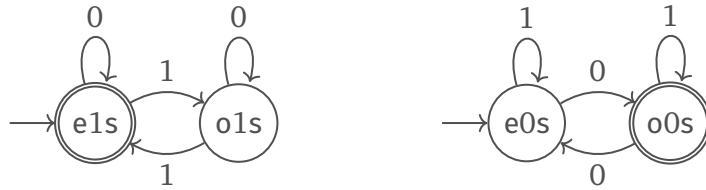
 This question involved adding two new constructs to our regex syntax, and as the last part showed, the system is now “non-orthogonal” in that certain regexes are interderivable. This is not necessarily a problem – many formal systems exhibit this form of redundancy – but it does make the inductively defined syntax larger which may complicate reasoning about the system (for example, every recursive definition or inductive proof on regexes now has two extra cases that would be covered by existing ones). Thus we may also reasonably choose to define  $r?$  and  $s^+$  as “syntactic sugar” (extra notation added for convenience) abbreviating  $r|\varepsilon$  and  $ss^*$ , respectively. These, and other derived operators and patterns form the basis of practical regex engines used widely in text processing applications.

## 16. On finite automata

1. For each of the two languages mentioned in §15.1 (string containing an even number of 1's or an odd number of 0's), find a DFA that accepts exactly that set of strings.

We can construct both with only two states each, corresponding to whether we have seen

an even or odd number of 1's or 0's and making the appropriate state accepting.



2. Given an NFA<sup>ε</sup>  $M = (Q, \Sigma, \Delta, s, F, T)$ , we write  $q \xRightarrow{u} q'$  to mean that there is a path in  $M$  from state  $q$  to state  $q'$  whose non- $\varepsilon$  labels form the string  $u \in \Sigma^*$ . Show that  $L = \left\{ (q, u, q') \mid q \xRightarrow{u} q' \right\}$  is equal to the subset of  $Q \times \Sigma^* \times Q$  inductively defined by the axioms and rules:

$$\frac{}{(q, \varepsilon, q)} \quad \frac{(q, u, q')}{(q, u, q'')} \text{ if } q' \xrightarrow{\varepsilon} q'' \text{ in } M \quad \frac{(q, u, q')}{(q, ua, q'')} \text{ if } q' \xrightarrow{a} q'' \text{ in } M$$

*Hint:* recall the method from §14.1. for showing that a language defined via set comprehension is equal to an inductively defined set: first show that  $L$  is closed under the rules and axioms, then show that every string in  $L$  has a derivation.

( $\subseteq$ ) We show that every element  $(q, u, q')$  of the inductively defined set  $L$  satisfies  $q \xRightarrow{u} q'$  by rule induction.

**Axiom**  $\frac{}{(q, \varepsilon, q)}$  We can always transition from a state to itself without consuming any symbols, so  $q \xRightarrow{\varepsilon} q$  holds vacuously.

**Rule**  $\frac{(q, u, q')}{(q, u, q'')}$  where  $q' \xrightarrow{\varepsilon} q''$  in  $M$ . The IH states that  $q \xRightarrow{u} q'$ . If there is an  $\varepsilon$ -transition from  $q'$  to  $q''$ , we can make one further step without consuming a symbol, so the overall string formed by the non- $\varepsilon$  labels will still be  $u$  – hence,  $q \xRightarrow{u} q''$ , as required.

**Rule**  $\frac{(q, u, q')}{(q, ua, q'')}$  where  $q' \xrightarrow{a} q''$  in  $M$ . The IH states that  $q \xRightarrow{u} q'$ . If there is a transition from  $q'$  to  $q''$  labelled with  $a$ , we can make a further step that extends the recognised string with the symbol  $a$  – hence,  $q \xRightarrow{ua} q''$ , as required.

( $\supseteq$ ) We show that we can derive  $(q, u, q') \in L$  whenever  $q \xRightarrow{u} q'$  in  $M$  by mathematical induction on the length  $|u| \in \mathbb{N}$  of the string  $u$ .

**Base case:**  $|u| = 0$ , so  $u = \varepsilon$ . If the number of non- $\varepsilon$  symbols in the path is 0, all of the steps must have been  $\varepsilon$ -transitions. Such paths can be captured by the axiom, and any number of applications of the first rule to transition between states without consuming an input.

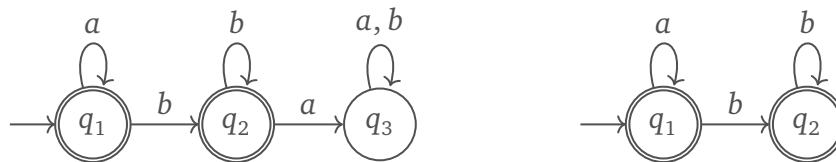
**Inductive step:**  $|u| = k + 1$ , so  $u = va$  for  $a \in \Sigma$  and  $|v| = k$ . Assume the  $\textcircled{\text{H}}$ : we have a path  $q \xRightarrow{v} q'$  and therefore  $(q, v, q') \in L$ . If we have a path labelled  $va$ , there must be a

transition from  $q'$  to  $q''$  labelled by  $a$ ; then, the second rule can be applied to  $(q, v, q') \in L$  and  $q' \xrightarrow{a} q''$  to deduce  $(q, va, q'') \in L$ , as required.

♪ Even though strings are no different from finite lists of symbols of the alphabet, they are formally elements of  $\Sigma^*$ , not a set inductively defined with an axiom for the empty string, and a rule for “consing” a symbol to the string. Performing induction on the length of the string simulates the kind of (structural) induction one would perform on an OCaml-style list.

3. The example of the subset construction given on [Slide 58](#) constructs a DFA with eight states whose language of accepted strings happens to be  $L(a^*b^*)$ . Give an “optimised” DFA with the same language of accepted strings, but fewer states. Give an NFA with even fewer states that does the same job.

The simplified NFA and DFA are as follows:



The main difference is that the DFA needs to handle the symbol  $a$  occurring in state  $q_2$ , which would mean seeing an occurrence of  $a$  after a  $b$  which disqualifies the string from being in  $L(a^*b^*)$ . The usual way of marking this as an invalid input in a DFA is to transition into a state from which it is impossible to reach an accepting state; upon any further input just stays stuck in  $q_3$ .

## 17. On regular languages

1. Why can't the automaton  $Star(M)$  used in [step \(iv\)](#) of the proof of part (a) of Kleene's Theorem be constructed by simply taking  $M$ , making its start state the only accepting state and adding new  $\varepsilon$ -transitions back from each old accepting state to its start state?

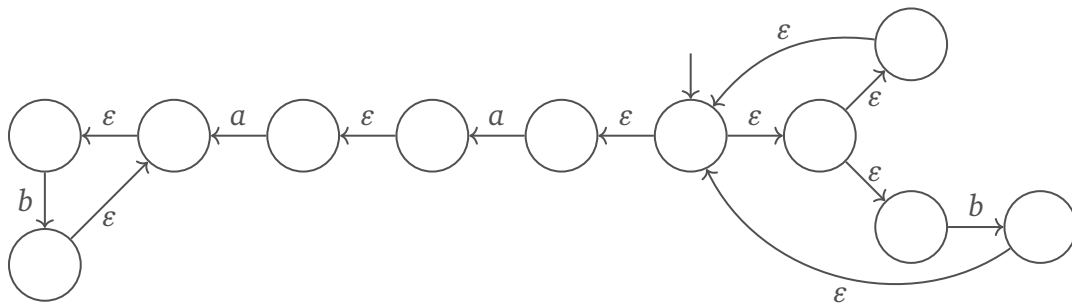
The problem is that we would be meddling with the internals of the automaton in unexpected ways by turning a (potentially) non-accepting start state into an accepting one. If  $M$  has transitions looping back to the start state, we may be able to accept partially recognised strings prematurely. For example, the automaton on the left below recognises the language  $L(a(aa)^*b)$ , but naively adding an  $\varepsilon$ -transition from  $q_3$  to  $q_1$  and making  $q_1$  accepting would result in a machine that accepts not only the expected  $L((a(aa)^*b)^*)$ , but also  $(aa)^*$ . By adding a new start state we ensure that the automaton “commits” to

performing a full repetition by explicitly transitioning into the start state of  $M$ .



2. Construct an  $NFA^\epsilon M$  satisfying  $L(M) = L((\epsilon|b)^*aab^*)$  using Kleene's construction.

Using the entirely algorithmic construction we get the following automaton:



Of course, there is a lot of redundancy, especially the large number of  $\epsilon$ -transitions that could be safely collapsed. The regex is not in its simplest form either, since  $(\epsilon|b)^*$  is equivalent to  $b^*$ . However, these are concerns of implementation efficiency, and Kleene's theorem is a result that the two formalisms are "in principle" equivalent. There are examples of languages that could be concisely expressed as regexes but the size of DFAs recognising the language is exponential in the length of the redex (such as the strings which have a specific symbol in the  $k^{\text{th}}$  last position); conversely, some languages are simple to recognise by a DFA, but the corresponding regexes are enormous (divisibility by 7 requires a DFA of 7 states, and converts to a regex of [10791 characters](#)).

3. Show that any finite set of strings is a regular language.

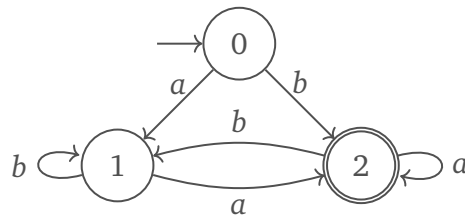
Let  $L = \{u_1, u_2, \dots, u_k\}$  be a finite set of strings. We construct the regular expression  $u_1|u_2|\dots|u_k$  which is clearly matched by all strings in  $L$ . Since such a regex exists, by Kleene's Theorem we conclude that  $L$  is regular.

4. Use the construction given in the proof of part (b) of Kleene's Theorem to find a regular expression for the DFA  $M$  whose state set is  $\{0, 1, 2\}$ , whose start state is 0, whose only accepting state is 2, whose alphabet of input symbols is  $\{a, b\}$ , and whose next-state function is given by the following table.

| $\delta$ | $a$ | $b$ |
|----------|-----|-----|
| 0        | 1   | 2   |
| 1        | 2   | 1   |
| 2        | 2   | 1   |



The DFA specified in the question is as follows:

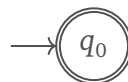


We apply Kleene's regex construction by first removing state 1, then state 2 – other orderings are possible. The recursive cases are not expanded further when the required redex is easily constructed by observation.

$$\begin{aligned}
 r_{0,2}^{\{0,2\}} &= ba^* \\
 r_{0,1}^{\{0,2\}} &= r_{0,1}^{\{0\}} \mid r_{0,2}^{\{0\}}(r_{2,2}^{\{0\}})^* r_{2,1}^{\{0\}} = a \mid ba^*b \\
 r_{1,1}^{\{0,2\}} &= r_{1,1}^{\{0\}} \mid r_{1,2}^{\{0\}}(r_{2,2}^{\{0\}})^* r_{2,1}^{\{0\}} = b \mid aa^*b \\
 r_{1,2}^{\{0,2\}} &= r_{1,2}^{\{0\}} \mid r_{1,2}^{\{0\}}(r_{2,2}^{\{0\}})^* r_{2,2}^{\{0\}} = a \mid aa^*a \\
 r_{0,2}^{\{0,1,2\}} &= r_{0,2}^{\{0,2\}} \mid r_{0,1}^{\{0,2\}}(r_{1,1}^{\{0,2\}})^* r_{1,2}^{\{0,2\}} = ba^* \mid (a \mid ba^*b)(b \mid aa^*b)^*(a \mid aa^*a)
 \end{aligned}$$

5. If  $M = (Q, \Sigma, \Delta, s, F)$  is an NFA, let  $Not(M)$  be the NFA  $(Q, \Sigma, \Delta, s, Q \setminus F)$  obtained from  $M$  by interchanging the role of accepting and nonaccepting states. Give an example of an alphabet  $\Sigma$  and an NFA  $M$  with set of input symbols  $\Sigma$  such that  $\{u \in \Sigma^* \mid u \notin L(M)\}$  is *not* the same as  $L(Not(M))$ .

A simple minimal example is the following automaton  $M$  with alphabet  $\Sigma = \{a\}$ :



We have that  $L(M) = \{\epsilon\}$ , but interchanging the accepting and nonaccepting states would turn  $q_0$  into a nonaccepting state so the language recognised is  $\emptyset$ . However,  $\emptyset \neq \{a\}^* \setminus \{\epsilon\}$ .

6. Let  $r = (a|b)^*ab(a|b)^*$ . Find a regular expression that is equivalent to the complement for  $r$  over the alphabet  $\{a, b\}$  with the property  $L(\sim r) = \{u \in \{a, b\}^* \mid u \notin L(r)\}$ .

The language matching  $r$  consists of all strings that contain  $ab$  as a substring. Thus, the complement of  $L(r)$  is the set of strings that do not contain  $ab$  as a substring, which is only possible if there are no occurrences of  $b$  after the first occurrence of  $a$ . The corresponding regular expression is thus simply  $b^*a^*$ .

7. Given DFAs  $M_i = (Q_i, \Sigma, \delta_i, s_i, F_i)$  for  $i = 1, 2$ , let  $And(M_1, M_2)$  be the DFA

$$(Q_1 \times Q_2, \Sigma, \delta, (s_1, s_2), F_1 \times F_2)$$

where  $\delta: (Q_1 \times Q_2) \times \Sigma \rightarrow (Q_1 \times Q_2)$  is given by

$$\delta((q_1, q_2), a) = (\delta_1(q_1, a), \delta_2(q_2, a))$$

for all  $q_1 \in Q_1, q_2 \in Q_2$  and  $a \in \Sigma$ . Show that  $L(\text{And}(M_1, M_2)) = L(M_1) \cap L(M_2)$ .

We prove the following lemma: for all strings  $u \in \Sigma^*$  and states  $q_1, q_2, q'_1, q'_2 \in Q$ ,

$$(q_1, q_2) \xRightarrow{u} (q'_1, q'_2) \text{ in } \text{And}(M_1, M_2) \iff q_1 \xRightarrow{u} q'_1 \text{ in } M_1 \wedge q_2 \xRightarrow{u} q'_2 \text{ in } M_2$$


This will directly imply the required result for  $q_1, q_2$  the start states  $s_1, s_2$  and  $q'_1, q'_2$  a pair of accepting states in  $F_1, F_2$ . The lemma will be proved by induction on the length of the string  $u$ .

**Base case:**  $|u| = 0$ , so  $u = \varepsilon$ . If  $(q_1, q_2) \xRightarrow{\varepsilon} (q'_1, q'_2)$  in  $\text{And}(M_1, M_2)$ , we must have that  $q_1 = q'_1$  and  $q_2 = q'_2$  because the automaton is deterministic and has no  $\varepsilon$ -transitions. But this is exactly the case when  $q_1 \xRightarrow{\varepsilon} q_1$  and  $q_2 \xRightarrow{\varepsilon} q_2$ .

**Inductive case:**  $|u| = k + 1$ , so  $u = va$  with  $|v| = k$  and  $a \in \Sigma$ . Assume the  $\textcircled{\text{H}}$  for the string  $v$ . If there is a path  $(p_1, q_1) \xRightarrow{va} (p_3, q_3)$  in  $\text{And}(M_1, M_2)$ , there must be two states  $p_2, q_2$  such that  $(p_1, q_1) \xRightarrow{v} (p_2, q_2)$  and  $(p_2, q_2) \xrightarrow{a} (p_3, q_3)$ . By the  $\textcircled{\text{H}}$ , we have that  $p_1 \xRightarrow{v} p_2$  and  $q_1 \xRightarrow{v} q_2$ , and  $\delta((p_2, q_2), a) = (p_3, q_3)$  by definition holds if  $\delta_1(p_2, a) = p_3$  and  $\delta_2(q_2, a) = q_3$ . Combining  $p_1 \xRightarrow{v} p_2$  with  $p_2 \xrightarrow{a} p_3$  and  $q_1 \xRightarrow{v} q_2$  with  $q_2 \xrightarrow{a} q_3$ , we have a path  $p_1 \xRightarrow{va} p_3$  and  $q_1 \xRightarrow{va} q_3$  in  $M_1$  and  $M_2$  respectively, as required.

## 18. On the Pumping Lemma

1. Briefly summarise the proof of the Pumping Lemma in your own words.

 Bookwork exercise, mainly intended to get you to read and understand the proof in order to be able to reproduce it if needed. The core points to remember are:

- The pumping lemma property is a *necessary condition*: regularity of  $L$  implies PLP, but not the other way around.
- The statement of the PLP should be read as a kind of dialogue: what are the objects and constraints you are given (a regular language  $L$ , a word of an appropriate length, and the number of repetitions of the central string), and what are things you have control over (the minimum length of the string, and the decomposition).
- The negation of PLP (used in the contrapositive) is the same dialogue but flipped around. An important consequence that is easy to overlook in informal proofs of non-regularity is that the “opponent” chooses the decomposition: they will try their very best to “catch you out” so the proof must not make any assumptions on how the string is split (other than the constraints stated, which are there precisely to stop the opponent choosing a decomposition for which the expected reasoning doesn’t work).

2. Consider the language  $L \triangleq \{c^m a^n b^n \mid m \geq 1 \wedge n \geq 0\} \cup \{a^m b^n \mid m, n \geq 0\}$ . The notes show that  $L$  has the pumping lemma property. Show that there is no DFA  $M$  which accepts  $L$ .

*Hint:* argue by contradiction. If there were such an  $M$ , consider the DFA  $M'$  with the same states


as  $M$ , with alphabet of input symbols just consisting of  $a$  and  $b$ , with transitions all those of  $M$  which are labelled by  $a$  or  $b$ , with start state  $\delta_M(s_M, c)$  where  $s_M$  is the start state of  $M$ , and with the same accepting states as  $M$ . Show that the language accepted by  $M'$  has to be  $\{a^n b^n \mid n \geq 0\}$  and deduce that no such  $M$  can exist.

We follow the hint and take  $M$  and  $M'$  as given. We show that  $L(M') = \{a^n b^n \mid n \geq 0\}$ .

( $\subseteq$ ) If  $w \in \{a, b\}^*$  is accepted by  $M'$ , then  $cw \in L(M)$  since the start state of  $M'$  is reached with a single  $c$ -transition from the start state of  $M$ . By definition of  $M$ ,  $w$  must be of the form  $a^n b^n$  for some  $n \in \mathbb{N}$ .

( $\supseteq$ ) If  $w = a^n b^n$ , we have that  $ca^n b^n \in L(M)$ , and by the definition of  $M'$ , we have that  $a^n b^n \in L(M')$ .

Thus, from the assumption that the DFA  $M$  exists, we constructed a DFA  $M'$  such that  $L(M') = \{a^n b^n \mid n \geq 0\}$ ; however, we know from the contrapositive statement of the Pumping Lemma that  $\{a^n b^n \mid n \geq 0\}$  is not regular, so our assumption that  $M$  exists was wrong. Consequently, the language  $L$  is not regular.

 The proof presented here is done using a technique called *reduction*: we reduce the question of determining the membership of a string  $u$  in  $\{a^n b^n \mid n \geq 0\}$  to the question of determining the membership of  $cu$  in  $L$ , so if the latter is answerable using a DFA, then so is the former (which leads to a contradiction). The central property of the mapping  $u \mapsto cu$  is that  $u \in \{a^n b^n \mid n \geq 0\}$  if and only if  $cu \in L$ . Reduction proofs will be discussed in more detail in the IB *Computation Theory* and *Complexity Theory* courses.