Teaching Mathematics in the Digital Age with Structured Derivations

Ralph-Johan Back

Four Ferries Publishing

Teaching Mathematics in the Digital Age with Structured Derivations

Ralph-Johan Back

Four Ferries Publishing

Contact information

Ralph-Johan Back, Professor of Computer Science,
Abo Akademi University
Joukahaisenkatu 3 – 5, 50250 Turku, Finland
mail: backrj@abo.fi, web: www.abo.fi/~backrj

Four Ferries Publishing

ISBN 978-952-7147-01-6

Cover picture: John Towner. License: CC0. Source::
https://unsplash.com/search/bike?photo=Hf4Ap1-ef40

Contents

Acknowledgements

My research on structured derivations started as part of a long and fruitful co-operation with Joakim von Wright when working on a monograph on refinement calculus in the 1990s. Much of the later research on structured derivations was carried out at the *Learning and Reasoning Laboratory,* a research laboratory at the IT-departments of Åbo Akademi University and University of Turku, jointly chaired by Tapio Salakoski and me.

Most of the ideas presented in this book have taken form during long and detailed discussions with many people. In particular, I want to thank Stefan Asikainen, Antti Lempinen, Linda Mannila, Mia Peltomäki, Viorel Preoteasa, Tapio Salakoski, Petri Sallasmaa, Petri Salmela, Patrik Sibelius, Anton Vistbacka and Joakim von Wright for the many insightful and critical discussions we have had during the years, about mathematics education in general and about structured derivations in particular. A special thanks goes to the members of the IFIP Working group 2.3 who have given both inspiration for developing the structured derivation method and much useful feedback on the method throughout the years.

I want to thank Linda Mannila and Mia Peltomäki for their long and continued work on piloting the practical use of structured derivations in teaching mathematics at high school level. I also want to thank the numerous high school teachers and students in Finland, Sweden and Estonia who have participated in our pilot studies during the years. They have given us plenty of valuable feedback, which has inspired and motivated the further development of the method.

The development of our series of math textbooks based on structured derivations (*eMath*) has been done as teamwork, with a highly dedicated and skilled authors team. I wish to thank everyone who has been working in this team, Stefan Asikainen, Kim Gustafsson, Kadri Hiob, Terhi Hovi, Topi Hurtig, Matti Hutri, Esbjörn Hägerstedt, Joonatan Jalonen, Tero Kesäläinen, Siiri Künnapas, Antti Lempinen, Marie Linden-Slotte, Saara Mäkinen, Mia Peltomäki, Mari Pöld, Tom Rydell, Eva Rönnqvist, and Lars Wingård. I also want to thank Teemu Rajala, Petri Salmela, and Petri Sallasmaa for their work on implementing en efficient syntax-directed editor for structured derivations, and Johannes Eriksson, Viorel Preoteasa, Anton Vistbacka, and Aman Yadav for building an automatic correctness checker for structured derivations. These efforts have all had a strong influence on the development of structured derivations, making the syntax more streamlined and the proof theory more powerful, as well as showing that the method can be successfully applied in all areas of high school mathematics.

I want to thank Stefan Asikainen for providing the assignments for the book, and Joonatan Jalonen, Marie Linden-Slotte, and Anton Vistbacka for careful reading and commenting on preliminary drafts of the book, and Jenna Lainio for helping with the layout.

The research on structured derivations has been financed by the Academy of Finland, TEKES (the Finnish Funding Agency for Technology and Innovation), the Centennial Foundation of Technology Industries of Finland, the Finnish National Board of Education, the EU Central Baltic Program and the Swedish Cultural Foundation

in Finland. I want to extend my gratitude to each of these organizations, as well as to Abo Akademi University and its IT department. The support given by these organizations has advanced the method from theoretical ideas on how to structure mathematical arguments to a comprehensive set of mathematics course books based on structured derivations, the development of an interactive e-book environment to support these course books, numerous teachers training courses on structured derivations, and extensive piloting of the ideas expressed here in practical mathematics courses at different education levels.

Last but not least, I want to thank my wife, Barbro Back, for many many discussions and her unwavering support throughout the development of the ideas presented in this book.

CHAPTER 1

Introduction

Mathematics education has not changed much in the last centuries, even as the number of students that study mathematics has exploded around the world. A central problem with learning mathematics is that it is cumulative: new topics can only be understood if the student already masters previous topics. A student that does not get a specific idea, like how to add and multiply fractions, will drop off when the class moves on to the next topic, say polynomials and equation solving. The course progresses at a given speed, and once the student has dropped off, it is difficult to get back on track again. Some do it, but many don't, the latter creating an ever growing backlog of things not understood, and therefore having more and more difficulties with the new topics they are faced with.

The way mathematics and mathematical reasoning is presented and taught is part of this problem. Mathematical arguments are traditionally presented in a very concise fashion, as a narrative with special mathematics notation. The assumption is that the reader is a skilled mathematician who can fill in the missing details. This is also an efficient way of communication between people who are well versed in mathematics. However, this presentation is less suitable when the receiving end is not that strong in mathematics, maybe less motivated, and is still learning the stuff. This is the situation in schools today. The overall structure of the mathematical argument is hidden in the narrative. Many students learn to see the logical structure behind the narrative that the teacher presents at the blackboard, or that is presented in the textbook. They will learn to think like mathematicians, feel comfortable with using mathematics, and many will go on to study in fields where mathematics is required. However, the rest will not get it, and will gradually classify themselves as weak in mathematics. They will avoid mathematics in their future studies.

This book describes an alternative way of presenting mathematical arguments, *structured derivations,* that aims at making the reasoning more transparent and easier to understand. A structured derivation shows clearly the overall structure of the argumentation, while at the same time requiring that each step in the derivation is carefully justified. Structured derivations have a well-defined syntax that can be understood, supported and analyzed by computers, while still being quite close to the standard way we reason in mathematics. Hence, we do not have to change the way we think, only how we present our reasoning.

One of our aims with structured derivations is to reintroduce proofs and careful argumentation as the solid basis for mathematics education. Structured derivations provide a universal format for presenting mathematical arguments of different forms: calculations, proofs, derivations, constructions, simplifications, etc. The method does not put any restrictions on the mathematical domain where we reason, nor on the level of detail or on the mathematical rigor of the argumentation. Hence, structured derivations can be used in any area of mathematics, and at any level of education.

Structured derivations are a further development of the *calculational proof style* that was originally proposed by Edsger W. Dijkstra and his colleagues (Wim Feijen, Netty van Gasteren, and Carel Scholten) [16, 36] (see [17] for a nice overview of the calculational proof style). They present a proof in a fixed format, as a sequence of calculation steps, with an explicit justification for each step. The original motivation for this proof style was to provide a simpler and more intuitive way of reasoning about program correctness. Joakim von Wright and I developed the structured derivation format in the late 90s [12, 7], as a tool for studying a particular programming logic, the *refinement calculus* [2, 12, 3]. As it turned out, this approach worked well also for teaching mathematics at high school level [13, 8, 30].

The structured derivations format has since then been gradually enhanced over the years, based on experiences gained from using the method in practice ([4]). Structured derivations in their present form combine *forward proofs*, *backward proofs* and *calculations* (the three main proof paradigms) into a single unified proof method [4]. The format is based on a hierarchical view of a derivation, where the main derivation can be split up into a number of more detailed observations and sub-derivations, which in turn can be further split up into even more detailed observations and sub-derivations, and so on.

Structured derivations support the use of logical notation and explicit logical inference rules in high school mathematics. This is another feature that is inspired by the calculational proof style. Logic is everywhere in school mathematics, but is usually hidden behind informal verbal explanations. Logical notation is used to some extent in mathematics education, but students are not usually taught the rules for how to manipulate logical expressions. Hence, the use of logic remains at the level of informal understanding, and the opportunity to speed up reasoning by symbolic manipulation of logical expressions is missed. The basic logic needed for school mathematics is not particularly difficult, and can be taught in a small number of lectures. The time spent on learning basic logic pays itself back many times over in subsequent studies.

Structured derivations have been designed with computer support in mind. We can build editors that support and enforce the syntax of structured derivations, and we can build computer systems that check whether a given structured derivation is mathematically correct. This requires that the mathematics used in the derivation has been mechanized, i.e., formalized in a way that that is understood by automatic proof checking systems. Large areas of mathematics have in fact already been mechanized and are checkable by proof checking systems like HOL [20], Isabelle [29], and PVS [28].

We have recently finished a completely new textbook series for mathematics based on structured derivations. These textbooks cover all compulsory courses in the present Finnish mathematics curriculum at high school level (grades 10 - 12). The textbooks are available in English, Finnish and Swedish [6]. The books are interactive e-books that come with an integrated editor for creating structured derivations and a facility for checking that a structured derivation is correct (using automatic theorem proving techniques).

The structured derivations method has been developed in a tight feedback loop with empirical studies in mathematics courses, primarily at high school level. This piloting work started already in 2001 with a number of pilot courses based on structured derivations, and has continued throughout the years. The method has also been extensively used in teachers education courses. The results have been very positive and encouraging. Students learn the approach quite quickly, and appreciate the added clarity gained, both when they use the structured derivation format themselves, and when the teacher presents examples in this style [10, 26]. We also see definite performance improvements in the students grades when the method is used [30]. The big advantage of structured derivations is that all the information needed to understand the argumentation is explicit in the derivation. This makes it easier for students to follow the teachers presentation in class, and it helps them to understand a derivation or a proof later, when solving home assignments or studying for an exam. At the same time, the method gives students a template for how to solve their own mathematical problems in a systematic fashion.

Structured derivations show how to present mathematical arguments *in-the-large* (the overall structure of the argument), while logic is needed for the mathematical argumentation *in-the-small* (the detailed justification for each step in the argumentation). These two story lines are intertwined in the book. They support each other, but we have tried to treat them in separate chapters or sections. Our book can therefore also be seen as a practical course on logic, describing how to use logic in standard mathematical reasoning. Both propositional and predicate calculus are covered. Together with the structured derivation format, this provides a solid foundation for building correct, yet easily understandable solutions to mathematical problems.

For those who are mainly interested in the structuring aspect and do not want to go into the details of logic, we recommend to have a look at another of out books, *Structured Derivations: Teaching Mathematical Reasoning in High School* [5]. There the emphasis is on structured derivations as a systematic way of organizing mathematical arguments at high school level, while the logic part is treated more lightly.

The content of the book has been roughly ordered according to educational levels. Structured calculations, described in Chapter 2, can be used already at grades 7 - 9 (in the Finnish school system). Chapters 3 to 14 add progressively more and more features to the structured derivations method. The approach outlined in these chapters is intended to be used for teaching mathematics at high school level (grades 10

- 12 in the Finnish school system). The teacher may choose whether he or she wants to include the logic part in the teaching (Chapters 3, 4, 5, 7, 11), or whether to only focus on the use of structured derivations for presenting mathematical arguments in a systematic fashion (Chapters 2, 6, 8, 9, 10, 12, 13, and 14). Chapters 15 to 18 is on reasoning with quantifiers and the use of mathematical theories, and would therefore be particularly useful for introductory mathematics education at university level. However, the intuitive reasoning that is formalized in Chapters 14 and 15 is already needed in high school, so the teacher should understand this material, even if it is not explicitly taught in class.

The remaining chapters describe the structured derivation method itself in more detail. Chapter 19 summarizes the syntax of structured derivation, while Chapter 20 explains how to check that a structured derivation is mathematically correct. Chapter 21 discusses computer support for structured derivations. Chapter 22 gives some more details on the background of the method. We have also included an appendix on the logical soundness and completeness of the structured derivation method. This is included for completeness, and is mainly intended for researchers who want to develop the structured derivation method further, or use ideas from this method in developing their own approach to mathematics education.

Finally, a few words about what this book is about and what it is not about. The book is not about the logical foundations of mathematics, but rather a practical guide for how to use standard logic in mathematical reasoning. The book is not about any specific mathematical topic, but about how to structure mathematical arguments in general. We assume that the reader is already familiar with high school mathematics, and we choose our examples freely from different areas there. Structured derivations tries to provide a solid framework for presenting mathematical arguments in such a way that the overall structure of the argument is shown explicitly, and the correctness of the argument can be checked systematically, step by step. The framework is designed to make use of the increasingly digital environment that our students now live in, enabling the use of computers for constructing, editing and checking mathematical arguments.

A bicycle would be a good analogue to what we want to achieve. Mathematics is the beautiful surrounding that the student can explore on a bicycle. Logic is the standard collection of ready made parts that are used when building bicycles. Structured derivations is a new (and hopefully improved) model of a bicycle, built from existing parts with a dash of some more recent technologies (like a small electric motor to make the uphill ride easier). The selling argument for this bicycle is that it is easier to use, faster, and more reliable than the standard models. This book can be seen as the user manual for the new bicycle.

Chapter 2

Calculations

The following chapters will show how to present mathematical arguments as structured derivations. We will introduce structured derivations one feature at a time. We start with basic features that are needed in almost any derivation, in this chapter. The subsequent chapters then introduce new features, one by one, and show how they are useful when the argumentation becomes more demanding and complex.

Let us start by introducing *structured calculations*, a simple form of structured derivations. Consider the following basic task: calculate the value of the expression $3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2$. A traditional solution may look like this:

$$\begin{array}{rcl} 3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2 & = & 3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16 \\ & = & 24 + 36 - 32 \\ & = & 60 - 32 \\ & = & 28 \end{array}$$

We rewrite this as a *structured calculation*:

$$\begin{array}{ll} \bullet & 3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2 \\ = & \{\text{calculate the powers}\} \\ & 3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16 \\ = & \{\text{carry out the multiplications}\} \\ & 24 + 36 - 32 \\ = & \{\text{carry out the first addition}\} \\ & 60 - 32 \\ = & \{\text{carry out the subtraction}\} \end{array}$$

$$28$$

□

The solution is the same as above, but now each step is justified explicitly. A justification is written on separate line and is enclosed in curly brackets. The justification explains why equality holds between the expressions on the previous line and the next line. The first step thus says that

$$3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2 = 3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16$$

by the rules for calculating powers.

The calculation is written in two columns: the equality sign is in the first column, while expressions and justifications are written in the second column. The two column format is used throughout in structured derivations. The *bullet* "•" indicates the start of the calculation, and the *square* "□" the end of it.

It is, of course, possible to give explicit justifications in the traditional format too:

$$\begin{array}{lll}
 & 3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2 & \text{calculate the powers} \\
= & 3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16 & \text{perform the multiplications} \\
= & 24 + 36 - 32 & \text{perform the addition} \\
= & 60 - 32 & \text{perform the subtraction} \\
= & 28 &
\end{array}$$

The justification is here written on the same line as the expression. The problem with this format is that it does not allow for longer expressions and/or justifications. Justifications are easy to omit, in particular for steps that seem more or less obvious, so there are usually few or no justifications in the calculation. But there is a problem with selective justifications: what is obvious to someone writing the calculation may not be obvious to the one who tries to understand it. A calculation without justifications is harder to understand. Even if you made the calculation yourself, it may still be hard to check the calculation afterwards, when you have forgotten the reasoning behind the calculation. Even an obvious explanation, such as "perform the addition" in the calculation above, can provide useful information to the inexperienced, it shows that there is no hidden complexity in this step.

The simpler format with few or no justifications rarely causes problems for an experienced mathematician, he/she will quickly see what rules were used in each step and how. But for a student who is just trying to learn something, the lack of justifications is an additional hurdle. It makes it more difficult to follow the reasoning, thus lowering the motivation for learning and decreasing the confidence in understanding the issues at hand. Problems in learning mathematics may often be due to a communication problem like this, rather than too weak motivation, inability to focus, or lack of mathematical ability.

The traditional method of teaching mathematics, used for centuries, is that the teacher writes calculations and arguments on the blackboard, and the students copy these into their notebooks. The teacher justifies each step of the calculation verbally. However, the students will usually only copy what the teacher writes on the blackboard, they do not write down the verbal justifications. The teacher thinks that he/she is giving the whole story, but the students only write down half of it. When students later do their homework, they start by looking at the examples that the teacher presented, trying to reconstruct the justification for each step. The best students will manage this, sharpening their understanding of mathematics in the process. But other students, who lack motivation, or have not done their previous homework properly, will run into problems. They create a backlog of unresolved issues, derivation steps not understood because of some misunderstanding or confusion about the underlying mathematics. The problems pile up as the course progresses, since there are more and more steps in the teacher's examples that the student does not understand. And these problems carry over to the next course. These students will gradually lose confidence in their ability to master mathematics, and classify themselves as bad at math. Mathematics teaching then becomes a process where we filter out the future mathematicians, those who will go on to study sciences, engineering and medicine. The rest are left on their own, they have failed at math and will in future studies avoid all subjects that even smell of mathematics.

Structured derivations try to solve this communication problem. Understanding the calculation afterwards becomes easier when we justify each step explicitly. Even students who were absent, unable, or unwilling to pay attention when the teacher showed the calculation for the first time, can now understand the calculation steps on their own. A step without justification is also easy to spot in a structured derivation, and the empty line indicates that the calculation is incomplete. The format forces both students and teachers to write out explicitly the justification for each step.

There are more reasons for insisting that each step is explicitly justified. The students usually have or will get access to the answers to their assignments. This can give them the impression that getting the correct answer is all that matters. But calculations in real life are carried out precisely when we do not know the answer; there is no reason to calculate if we already know the answer. The only way to convince ourselves and others that we have found the correct answer is then to carefully check that each step of the calculation is correct, i.e., check that the justification for each step is correct and that we have not made any errors in calculating the next step.

The traditional calculation format (with explicit justifications) requires fewer lines than a structured derivation, and it looks more concise. However, the number of symbols in the two calculations are roughly the same. So we do not save any ink or keystrokes with the traditional calculation format, only some paper. On a computer screen, we do not even save this.

Our example calculation is simple and trivial, so the explicit justifications may seem unnecessary. But the calculation may not be that simple and obvious for a student who is learning about powers for the first time. We illustrate the need for explicit justifications with another, less trivial example, where the calculation steps are not as obvious.

Example 1. We want to calculate the tangent of the expression $\frac{17\pi}{3}$. We start by giving a solution in the traditional format, without explicit justifications.

$$\begin{aligned}
\tan\frac{17\pi}{3} &= \tan\left(\frac{6\cdot 2\pi + 5\pi}{3}\right)\\
&= \tan\left(2\cdot 2\pi + \frac{5\pi}{3}\right)\\
&= \tan\frac{5\pi}{3}\\
&= \tan\left(2\pi - \frac{\pi}{3}\right)\\
&= -\tan\frac{\pi}{3}\\
&= -\sqrt{3}
\end{aligned}$$

The same argumentation, now written as a structured derivation:

• $\tan\frac{17\pi}{3}$

= {factor out 2π}

$\tan\left(\frac{6\cdot 2\pi + 5\pi}{3}\right)$

= {write the angle in the form $n\cdot 2\pi + \alpha$}

$\tan\left(2\cdot 2\pi + \frac{5\pi}{3}\right)$

= {we can ignore full circles 2π}

$\tan\frac{5\pi}{3}$

= {the angle is in the fourth quadrant, so we can write it in the form $2\pi - \alpha_0$ where α_0 is between 0° and 90°}

$\tan\left(2\pi - \frac{\pi}{3}\right)$

= {ignore full circles, $\tan\left(-\frac{\pi}{3}\right) = -\tan\left(\frac{\pi}{3}\right)$}

$-\tan\frac{\pi}{3}$

= {this is a 30 - 60 - 90 triangle}

$-\sqrt{3}$

□ ■

(We use a black square to the right on a page to indicate the end of an example, a definition or a theorem).

The explicit justifications makes it easier to understand the argumentation. Writing the justification on a separate line gives us enough room to properly explain each step. The structured calculation format allows both terms and explanations to stretch over two or more lines, without compromising ease of reading.

Calculations like the one above are based on the fact that equality is *transitive.* This means that for arbitrary values $a_1, a_2, \ldots, a_n$: if $a_1 = a_2$, $a_2 = a_3$, $\ldots$, and $a_{n-1} = a_n$, then $a_1 = a_n$.

Example 2. Calculate

$$\lim_{x \to 1} \frac{x-1}{\sqrt{x^2+3}-2}$$

We notice that the denominator is 0, when $x = 1$, so we have to manipulate the expression into a form where this does not happen.

- $\lim_{x \to 1} \dfrac{x-1}{\sqrt{x^2+3}-2}$

= {eliminate the radical from the denominator by expanding with $\sqrt{x^2+3}+2$}

$\lim_{x \to 1} \dfrac{(x-1)(\sqrt{x^2+3}+2)}{(\sqrt{x^2+3}+2)(\sqrt{x^2+3}-2)}$

= {use the rule $(a-b)(a+b) = a^2 - b^2$}

$\lim_{x \to 1} \dfrac{(x-1)(\sqrt{x^2+3}+2)}{(\sqrt{x^2+3})^2 - 2^2}$

= {simplify the denominator}

$\lim_{x \to 1} \dfrac{(x-1)(\sqrt{x^2+3}+2)}{x^2-1}$

= {write the denominator in the form $(x-1)(x+1)$}

$\lim_{x \to 1} \dfrac{(x-1)(\sqrt{x^2+3}+2)}{(x-1)(x+1)}$

= {we cancel out $(x-1)$; this is allowed because $x - 1 \neq 0$ in the neighborhood of $x = 1$}

$\lim_{x \to 1} \dfrac{\sqrt{x^2+3}+2}{x+1}$

= {calculate the limit by substituting $x = 1$}

$\dfrac{\sqrt{1^2+3}+2}{1+1}$

= {calculate the value}

2

□

Transitivity of equality gives the answer:

$$\lim_{x\to 1} \frac{x-1}{\sqrt{x^2+3}-2} = 2$$

■

Example 3. We can also use a calculation to prove a theorem. For instance, the next calculation proves the conjugate rule for binomials: $(a+b)(a-b) = a^2 - b^2$.

• $(a+b)(a-b)$

= {the distributive law for polynomials}

$(a+b)a - (a+b)b$

= {the distributive law for polynomials}

$a^2 + ba - ab - b^2$

= {the second and third terms cancel out}

$a^2 - b^2$

□

Transitivity of equality shows that the theorem is true. ■

Example 4. Equation solving is another area where we usually use calculations. Consider the problem of solving the equation $x + 2 = 2x - 1$.

We solve the equation by transforming it step by step into another, equivalent form where the solution is explicitly shown. Note that we use "≡" to indicate that two equations are equivalent, rather than the more common "⇔". We will treat equivalence in much more detail in Chapter 4.

• $x + 2 = 2x - 1$

≡ {add -2 to both sides of the equation}

$x + 2 - 2 = 2x - 1 - 2$

≡ {simplify both sides}

$x = 2x - 3$

$\equiv$ $\quad$ {add $-2x$ to both sides of the equation}

$\qquad x - 2x = 2x - 3 - 2x$

$\equiv$ $\quad$ {simplify both sides}

$\qquad -x = -3$

$\equiv$ $\quad$ {divide both sides by -1}

$$\frac{-x}{-1} = \frac{-3}{-1}$$

$\equiv$ $\quad$ {simplify}

$\qquad x = 3$

□

Adding the same expression to both sides of an equation does not change the truth of the equation, regardless of the value of x. Similarly, the truth of the equation is preserved when we multiply both sides by the same expression (provided it is not 0). Finally, truth of the equation is also preserved when we replace an arithmetic expression in an equation by another expression with the same value. Since equivalence is also transitive, the calculation shows that

$$(x + 2 = 2x - 1) \equiv (x = 3)$$

In other words, $x = 3$ is the solution to the equation. ■

2.1 Syntax of Structured Calculations

A *structured calculation* is written in a specific way, shown by the template here. We write the calculation in two columns. The bullet starts the calculation, and the square shows where it ends. The initial mathematical expression is written in the second column. On the next line we write a relation in the first column (denoted *rel*, this can, e.g., be "=" or "≤" or "≡") followed by a justification in the second column. A new mathematical expression is then written on the next line, in the second column. We continue in this way, until we have reached the final mathematical expression. We have used color coding for the different syntactic categories in the calculation: red for relations, blue for justifications, and black for expressions.

***calculation*:**

● $\quad$ *expression*

rel $\quad$ *justification*

$\qquad$ *expression*

⋮

rel $\quad$ *justification*

$\qquad$ *expression*

□

Figure 2.1: Structured calculation

The three vertical dots show that we can add 0 or more steps to the first calculation step. Every subsequent calculation step has two lines, a relation and justification line followed by an expression line. The justification explains why the relation shown in the first column holds between the expression on the preceding line and the expression on the next line. The justifications in our examples so far have been simple, just explaining text enclosed in curly brackets. We will later encounter more complex justifications (in Chapter 9).

Below is an example of a structured calculation with 4 steps. On the left we show the general format, and on the right an example of a structured calculation that follow this syntax.

●	*expression*	●	$3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2$
rel	*justification*	$=$	{calculate the powers}
	expression		$3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16$
rel	*justification*	$=$	{perform the multiplications}
	expression		$24 + 36 - 32$
rel	*justification*	$=$	{perform the addition}
	expression		$60 - 32$
rel	*justification*	$=$	{perform the subtraction}
	expression		28
□		□	

A structured calculation is a collection of facts. The calculation on the right says that

$$3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2 \quad = \quad 3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16$$

by calculating the powers, and

$$3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16 \quad = \quad 24 + 36 - 32$$

by performing the multiplications, and

$$24 + 36 - 32 \quad = \quad 60 - 32$$

by performing the additions, and

$$60 - 32 \quad = \quad 28$$

by performing the subtraction. From this collection of facts, we may then conclude (by transitivity) that

$$3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2 \quad = \quad 28$$

2.2 Expressions and Relations

Calculations are performed within the framework of some branch of mathematics, like algebra, geometry, analytic geometry, calculus, etc. The underlying theory will then determines the notation we use, i.e., what kind of expressions and relations we can use in the calculation. For polynomials, we use expressions like $x^2+2x-y+1$ and $\dfrac{x^2-y}{x+2y}$, in analytic geometry we have expressions that describe lines (e.g., equations) and points (coordinates), etc. In addition to standard mathematical expressions, we can also use informal expressions like "the circumference of the circle $+$ 3" or "the base $\cdot$ the height of the triangle", to make the solutions more intuitive. An example of this is the following start of a derivation:

$$\begin{array}{ll}
 & \text{area of the triangle}\\
= & \{\text{area formula}\}\\
 & \dfrac{\text{base } \cdot \text{ height of the triangle}}{2}\\
= & \{\text{the height of the triangle is 3 times longer than the base according to}\\
 & \text{the assumption}\}\\
 & \dfrac{3 \cdot (\text{base of the triangle })^2}{2}\\
\vdots &
\end{array}$$

We are free to use any binary relations between the terms. Typically, we use transitive order relations like $\leq, <, \geq, > \ldots$, equality $=$, and logical relations like implication ($\Rightarrow$) or equivalence ($\equiv$ or $\Leftrightarrow$). We are free to mix different binary relations in the same derivation.

Equality can be combined with any binary relation: if the relation $a \sim b$ holds and $b = c$, then $a \sim c$ also holds. Consider, e.g., the chain

$$a = b < c < d = e = f$$

This says that

$$a = b \text{ and } b < c \text{ and } c < d \text{ and } d = e \text{ and } e = f$$

We conclude that $a < f$, since $<$ is transitive.

We can also use non-transitive relations between the terms. For instance, assume that $a \rightarrow_k b$ says that the distance from a to b is k kilometers. The calculation

$$a \rightarrow_{12} b \rightarrow_8 c \rightarrow_7 d$$

then says that

$$a \rightarrow_{12} b \text{ and } b \rightarrow_8 c \text{ and } c \rightarrow_7 d$$

We conclude that the distance from a to d is $12+8+7=27$ km, if we drive through b and c, i.e.,

$$a \rightarrow_{27} d$$

This is an example of a conclusion we can draw from a calculation that is not based on transitivity.

Inequality is an example of a relation that is not transitive: $a \neq b$ and $b \neq c$ does not imply that $a \neq c$ (counterexample: $0 \neq 1$ and $1 \neq 0$ do not imply that $0 \neq 0$). This restricts the usefulness of multiple inequalities in structured derivations. Similarly, mixing "$\leq$" and "$\geq$" in the same derivation is usually not a good idea: from

$$a \leq b \geq c$$

we can only conclude that $a \leq b$ and $c \leq b$, i.e., that b is the largest of the numbers a, b, c.

2.3 Justifying the Steps

There are basically two different ways of justifying a calculation step. We can justify a calculation step with a mathematical rule, or we can refer to a permissible operation. In the first case, the justification states which rule is used:

$$\begin{array}{ll} & (x+1)(x+y) \\ = & \{\text{the distributive law for polynomials}\} \\ & (x+1)x+(x+1)y \end{array}$$

In the second case, the justification states the operation that we apply:

$$\begin{array}{ll} & (x+1)(x+y) \\ = & \{\text{distribute the first term across the second term}\} \\ & (x+1)x+(x+1)y \end{array}$$

In the second case, we know that distributing the first term across the second term is permissible, since the distributive law holds. An operation is allowed if there is a rule that says that the operation results in a new term that has the desired relation to the original term. In our example, we start from the term $(x+1)(x+y)$ and distribute the first term $(x+1)$ across the second term $(x+y)$. The result is a new expression $(x+1)x+(x+1)y$. The operation is permitted, since distribution preserves equality between terms, i.e.,

$$(x+1)(x+y) = (x+1)x+(x+1)y$$

holds according to the distributive law.

Both ways of explaining a step are useful but they have different characteristics. In the first case, we see the justification as a static observation of why the equality holds. In the second case, we justify why transforming an expression in a specific way is permissible. When the relation is equality, we are allowed to transform the original expression as long as we do not change the value of the expression.

2.4 The Level of Detail

The level of detail in a justification depends on whom we are trying to convince. If the reader is an experienced mathematician, a short and concise explanation may be enough, as in the examples above. If the purpose of the derivation is to illustrate how we use a certain rule, we can be more careful, e.g., by explicitly stating the rule in the justification:

$$\begin{array}{ll} & (x+1)(x+y) \\ = & \{\text{the distributive law for polynomials: } a(b+c) = ab+ac\} \\ & (x+1)x + (x+1)y \end{array}$$

If we want to be even more explicit, we can also say how the rule is applied:

$$\begin{array}{ll} & (x+1)(x+y) \\ = & \{\text{the distributive law for polynomials: } a(b+c) = ab+ac\text{, where } a \text{ is } x+1\text{, } b \text{ is } x \text{ and } c \text{ is } y\} \\ & (x+1)x + (x+1)y \end{array}$$

Mathematical rules are often conditional. For instance, the rule for expanding and reducing fractions state that

$$\frac{a}{b} = \frac{k \cdot a}{k \cdot b}, \text{ when } k \neq 0$$

When a conditional rule is used, the justification has to explain why the condition is satisfied, e.g., like this:

$$\begin{array}{ll} & \dfrac{x^2+3}{x-2} \\ = & \{\text{expand the fraction by } x-1\text{, permitted since } x > 1 \text{ by assumption, so } x-1 \neq 0\} \\ & \dfrac{(x-1)(x^2+3)}{(x-1)(x-2)} \end{array}$$

The justification refers to an assumption (stated elsewhere) that implies that the expansion is permitted.

The the number of steps in a calculation may also vary, depending on the target audience. A teacher can do the parts of the calculation that illustrate new concepts in small and detailed steps, while the parts that are based on earlier material can be done in larger steps. The calculation should, however, always be detailed enough so that an interested reader can check every step of the proof directly, without having to do complicated calculations in head or on paper. This makes it easier to follow the calculation and also prevents a lot of trivial mistakes when constructing the calculation.

Example 5. We want to calculate the value of $2^8 + 2^7$. The following calculation shows the main steps, with explanations. The target audience would be students in secondary schools:

• $2^8 + 2^7$

$=$ $\{$the product rule: $a^m a^n = a^{m+n}$, and $2^1 = 2\}$

$2 \cdot 2^7 + 2^7$

$=$ $\{$factor out $2^7\}$

$(2 + 1) \cdot 2^7$

$=$ $\{$calculate the value, $2^7 = 128\}$

384

□ ■

Example 6. A more detailed calculation may look like this:

• $2^8 + 2^7$

$=$ $\{$the product rule: $a^m a^n = a^{m+n}\}$

$2^1 \cdot 2^7 + 2^7$

$=$ $\{a = 1 \cdot a\}$

$2^1 \cdot 2^7 + 1 \cdot 2^7$

$=$ $\{$the distributive law: $(a + b)c = ac + bc\}$

$(2^1 + 1) \cdot 2^7$

$=$ $\{$power rule: $a^1 = a\}$

$(2 + 1) \cdot 2^7$

$=$ $\{$arithmetics: $1 + 2 = 3\}$

$3 \cdot 2^7$

$=$ $\{$arithmetics: $2^7 = 128\}$

$3 \cdot 128$

$=$ $\{\text{arithmetics: } 3 \cdot 128 = 384\}$

384

□

Here every step is justified by an explicit rule. This is a suitable level of detail when the purpose is to illustrate the rules for manipulating arithmetic expressions in a more axiomatic context. The earlier level of detail was sufficient if the purpose was just to calculate the value of the expression. ■

2.5 Assignments

1. Simplify $a^{x+3} \cdot a^{x-2} \cdot \left(a^{-x-1}\right)^2$ (assume $a \neq 0$).
2. Solve the equation $5x - 2\,(x-1) = 2$.
3. Solve the equation $x^2 + 5x - 24 = 0$.
4. Solve the equation $x^3 - 6\frac{1}{2}x^2 - 3\frac{1}{2}x = 0$.
5. Calculate $\int\limits_0^\pi (\sin(x) + \cos(x))\,\mathrm{d}x$.
6. Calculate $\frac{d}{dx}\left(x^2 \cos(2x)\right)$.
7. Solve the simultaneous equations $y = 2x - 3$ and $5x = -2y + 39$.
8. Prove $(\sin(x) + \cos(x))^2 - 1 = \sin(2x)$.
9. Solve the absolute value equation $|2x - 8| = 16$.
10. Prove that the equation $5 \cdot (2 - x) - 9 = 6 \cdot (3 - x) - (16 - x)$ is not satisfied for any values of the variable x.

Chapter 3

The Logic Behind Calculations

Calculations are easy to understand intuitively, a main reason for why they are so popular. However, there are things going on behind the scenes also in calculations. We explain here the basic logic of calculations, in a less formal manner, saving a more formal treatment to later.

A *logical proposition* is a proposition that is either true or false. We know that the proposition “the Moon is a cheese” is false. The proposition “the earth is flat” has been considered true, but now we consider it false. Some propositions cannot be false. The proposition that a triangle has three sides is, for instance, true by necessity, since it follows from the definition of a triangle. A proposition like “the younger brother turns 10 later than the older brother” is also true by necessity, it follows from the meaning of younger and older.

Similarly, mathematical propositions can be true or false. The proposition

$$2 + 1 = 4 - 1$$

is true, while the proposition

$$2 + 2 = 4 - 2$$

is false. The proposition

$$2 + x = 4 - x$$

is true for some values of x (when $x = 1$) and false for other values (when $x \neq 1$).

A logical proposition has the value T if it is true, and F if it is false. We call T and F *truth values*. We denote the *set of truth values* by $\mathbb{B}$. This set contains only two elements,

$$\mathbb{B} = \{F, T\}$$

The letter B stands for *Boolean algebra*, another name for propositional calculus. George Boole was a 19th century British mathematician and philosopher, who laid the foundations for the propositional calculus. Logic can largely be characterized as the theory about these two truth values and how to determine whether a logical proposition has value T or F.

Example 7. The proposition $2+1=4-1$ is true, so it has the value T, i.e.

$$(2+1=4-1)=T$$

The proposition $2+2=4-2$ is false, i.e.

$$(2+2=4-2)=F$$

The proposition $2+x=4-x$ is true when $x=1$ and else false, i.e.

$$(2+x=4-x)=\begin{cases} T, & \text{when } x=1 \\ F, & \text{when } x\neq 1 \end{cases}$$

The proposition $x+2\leq 2x$ is true when $x\geq 2$ and else false, i.e.

$$(x+2\leq 2x)=\begin{cases} T, & \text{when } x\geq 2 \\ F, & \text{when } x<2 \end{cases}$$

■

We say that the value $x=3$ *satisfies* proposition $x+2\leq 2x$, if the proposition is true (i.e., has the value T) when $x=3$.

We use above the same symbol "=" for equality between real numbers and between truth values. In both cases it describes the same basic concept, that two expressions have the same value. But for more complex logical propositions, introduced later, this can make the expressions difficult to read. Because of this, we prefer to use the symbol "≡" for equality between truth values and call this *equivalence.* However, it is important to understand that this is just another name for equality. The order of calculation is, however, different for equality and equivalence: we calculate equality between arithmetic expressions before we calculate whether logical propositions are equivalent, i.e. the order is "=" before "≡".

Using a special symbol for equivalence means that we do not need parentheses to distinguish equality between arithmetic expressions and equality between truth values. Instead of writing $(2+1=4-1)=T$, we can write $2+1=4-1\equiv T$, and we can write

$$2+x=4-x\equiv\begin{cases} T, & \text{when } x=1 \\ F, & \text{when } x\neq 1 \end{cases}$$

Parentheses can be omitted around the logical proposition, because we know that equality is calculated before equivalence.

The symbol "⇔" is often used for equivalence between logical propositions, instead of "≡". We prefer the symbol "≡", because it emphasizes that equivalence is equality between truth values. We read the proposition $p\equiv q$ as p *is as true as* q (or p and q are *equally true*). Since there are only two truth values, this means that p and q are either both true or are both false. As Example 4 showed, we can carry out calculations with equivalence between equations in the same way as we carry out calculations with equality between algebraic expressions.

3.1 Mathematical Facts and Rules

A logical proposition can be true or false, depending on the values that we assign to the variables in the proposition. A *mathematical fact* is a logical proposition that we can prove is true (i.e., has the value T) for all value of the variables in the proposition. The distribution rule of algebra is an example of a mathematical fact. It says that

$$a(b+c) = ab + ac$$

is true for all real numbers a, b and c. Some rules include conditions that must be satisfied for the proposition to be true. The rule

$$\sqrt{a^2} = a$$

e. g., only holds if a is non-negative, i.e. when $a \geq 0$.

Whether a mathematical statement is to be understood as a logical proposition (which may be true or false) or a mathematical fact (which is always true) is usually clear from the context in mathematical text. We need, however, to make this distinction as clear as possible. Therefore we will introduce a special symbol "$\vdash$" to indicate that a logical proposition is a mathematical facts. The first rule above is, e.g., written more exactly as

$$\vdash a(b+c) = ab + ac$$

where the symbol "$\vdash$" shows that it is a fact rather than a logical expression. This says that the logical expression $a(b+c) = ab + ac$ can be *proved* to be true (i.e., proved to have the value T) for every possible combination of values for the variables a, b and c. Here it is understood that the variables a, b and c only take real numbers as values.

The second rule is written as

$$a \geq 0 \vdash \sqrt{a^2} = a$$

This states that we can prove that $\sqrt{a^2} = a$ is true for every real number a that satisfies the condition $a \geq 0$.

The square root function on real numbers is an example of a *partially defined function*: we know the value of $\sqrt{a}$ when a is a non-negative real number, but we have not defined the value of $\sqrt{a}$ when a is a negative real number. We can interpret this situation in two ways: either we think that the square root function does not have any value for negative arguments, or we think that the square root function does have a value for negative arguments, but that we know nothing about this value. In either case, we say that the value of the square root function is *not defined* (or that it is *undefined*) for negative arguments. Which interpretation we choose does not matter, the main thing is that we should be careful not to use the square root function unless we know that the argument is non-negative.

The general form for a *mathematical fact* (also known as a *sequent*) is

$$A_1, A_2, \ldots, A_m \vdash C$$

where $A_1, A_2, \ldots, A_m$ are logical propositions (the *assumptions* of the rule) and C is another logical proposition (the *consequence* of the rule). This fact says that the consequence C *follows from* the set of assumptions $A_1, A_2, \ldots, A_m$. If propositions $A_1, \ldots, A_m$ and C contain variables, say $x_1, \ldots, x_n$, then the fact says that C can be proved to be true for any combination of values for $x_1, \ldots, x_n$ that satisfies all assumptions $A_1, \ldots, A_m$. How to prove this is one of the central topics of logic, which we will be discussing a lot in the sequel. Notice that we here talk about a *set of assumptions,* not a sequence of list of assumptions. This means that it does not matter in what order we list assumptions in $A_1, A_2, \ldots, A_m \vdash C$, or whether we have duplicate assumptions in the list.

We assumed above that the variables only take real number values. We can make this explicit, by adding assumptions about the domain of variables:

$$\begin{array}{rcl} a \in \mathbb{R},\, b \in \mathbb{R},\, c \in \mathbb{R} & \vdash & a(b+c) = ab + ac \\ a \in \mathbb{R},\, a \geq 0 & \vdash & \sqrt{a^2} = a \end{array}$$

Let us consider the distribution rule in a little bit more detail. It holds for any real values of a, b, and c. This means that it will also be true for any real valued expressions that we substitute for the variables a, b, and c in the law. Consider, e.g., the expressions x^2, $x + y$ and $y - 1$. The distribution law allows us to conclude that

$$x^2 \in \mathbb{R},\, (x{+}y) \in \mathbb{R},\, (y{-}1) \in \mathbb{R} \vdash x^2{\cdot}((x{+}y){+}(y{-}1)) = x^2{\cdot}(x{+}y){+}x^2{\cdot}(y{-}1) \quad (3.1)$$

We get this new fact by choosing x^2 as the value for a, $x + y$ as the value for b, and $y - 1$ as the value of c. This is referred to as an *instance* of the original fact. We *instantiate* a fact by choosing specific expressions for the variables that occur in it. We write an *instantiation* as a sequence of assignment statements,

$$a := x^2,\, b := x + y,\, c := y - 1$$

To use the distribution rule in a calculation, we also need a general proof principle called *modus ponens.* It says that we can prove

$$A_1, \ldots, A_m \vdash C$$

by first proving a collection of lemmas $B_1, \ldots, B_k$, and then prove C under the original assumptions $A_1, \ldots, A_m$ together with the newly proved lemmas $B_1, \ldots, B_k$. In other words, we prove

$$A_1, \ldots, A_m \vdash C$$

by proving

$$\begin{array}{rcl} A_1, \ldots, A_m & \vdash & B_1 \\ A_1, \ldots, A_m, B_1 & \vdash & B_2 \\ & \vdots & \\ A_1, \ldots, A_m, B_1, \ldots, B_{k-1} & \vdash & B_k, \text{ and} \\ A_1, \ldots, A_m, B_1, \ldots, B_k & \vdash & C \end{array}$$

Note that we allow the use of previously proved lemmas $B_1, \ldots, B_{i-1}$ when proving the next lemma B_i. We apply the lemma rule to our example. The basic arithmetic operations are all closed over reals, so we know that

$$\begin{aligned} x \in \mathbb{R},\, y \in \mathbb{R} &\ \vdash\ x^2 \in \mathbb{R} \\ x \in \mathbb{R},\, y \in \mathbb{R} &\ \vdash\ x + y \in \mathbb{R} \\ x \in \mathbb{R},\, y \in \mathbb{R} &\ \vdash\ y - 1 \in \mathbb{R} \end{aligned}$$

Hence, from these rules and the instance of the distribution shown in (3.1), we conclude by modus ponens that

$$x \in \mathbb{R},\, y \in \mathbb{R} \vdash x^2 \cdot ((x + y) + (y - 1)) = x^2 \cdot (x + y) + x^2 \cdot (y - 1)$$

This is a new mathematical fact, which we presumably needed in some calculation.

A *rule* is a mathematical fact that can be applied in a wide variety of circumstances (using instantiation and modus ponens). The distribution and the square root rules above are useful in this way and deserve to be called rules. The last mathematical fact that we have derived above is too specific to be called a rule.

3.2 Properties of Equality

Equality has certain basic properties that are used over and over again: it is reflexive, symmetric and transitive. *Reflexivity* means that each expression is equal to itself, e.g., $2 = 2$ and $x + 1 = x + 1$ are both true. *Symmetry* means that equality holds in both directions: if $2x = 3$, then we also have that $3 = 2x$. Algebraic laws can therefore be read in both directions: if we have proved that $(a + b)(a - b) = a^2 - b^2$, then we also know that $a^2 - b^2 = (a + b)(a - b)$. *Transitivity* again says that we can chain equalities: if we have proved that $2x = 3y^2$ and that $3y^2 = 27$, then we know that $2x = 27$. These rules are listed in Table 3.1.

	$\vdash$	$a = a$	$\{= \textit{is reflexive}\}$
$a = b$	$\vdash$	$b = a$	$\{= \textit{is symmetric}\}$
$a = b,\, b = c$	$\vdash$	$a = c$	$\{= \textit{is transitive}\}$
$e = e'$	$\vdash$	$E(e) = E(e')$	$\{\textit{Leibniz' rule}\}$

Table 3.1: Equality rules

The properties of equality should be compared to other binary relations that do not have all these properties. Consider for instance the less-than relation. It is not reflexive ($2 < 2$ is not true), and it is not symmetric ($2x < 3$ does not mean that $3 < 2x$). The relation is, however, transitive: if we have proved that $2x < 3y^2$ and that $3y^2 < 27$, then we know that $2x < 27$. Another example is the not-equal relation. It is not reflexive ($2 \neq 2$ is not true), it is symmetric (if $2x \neq 3$ is true, the so is $3 \neq 2x$), but it is not transitive ($0 \neq 1$ and $1 \neq 0$ does not mean that $0 \neq 0$). The less-than-or-equal relation "$\leq$" is again both reflexive and transitive, but it is not symmetric.

In addition to these basic properties, we have one more property that is central to all calculations. This is really a property of functions. Assume that f is a function, and that a and b are both in the domain of this function. Then

$$a = b \vdash f(a) = f(b)$$

This is the essential property of functions: a function has a unique value for each element in its domain.

Assume that we know that $x^2 = y + 3$. Then the function rule allows us to deduce that $x^2 + 1 = (y + 3) + 1$. We see this by considering the function $f(a) = a + 1$. The assumption $x^2 = y + 3$ and the fact that f is a function gives us that

$$x^2 = y + 3 \vdash x^2 + 1 = (y + 3) + 1$$

The same rule allows us to further deduce that

$$\frac{x^2 + 1}{x - 2} = \frac{(y + 1) + 1}{x - 2}$$

In general, the function rule allows us to replace any subexpression in a larger expression with another, equal expression, without changing the value of the expression. Let us formulate this as a general principle.

Let $E(e)$ denote an expression with subexpression e. Then we have the following general rule:

$$e = e' \vdash E(e) = E(e') \ \{\text{Leibniz' rule}\}$$

Here $E(e')$ denotes the expression that we get from $E(e)$ when we replace e with e'. The name of the rule comes from the 17th century mathematician and philosopher Gottfried Leibniz, who formulated this principle explicitly in his studies over the laws of thought. This rule is used very frequently (and almost always implicitly) in calculations. Its main use is in simplifying an expression, by replacing a subexpression with a simpler expression that is equal to the original one.

3.3 Correctness of calculation steps

We are now ready to look more carefully on how the steps in a calculation are justified. We will look at the calculation in Example 6. How do we justify the steps in this calculation with the rules that we have presented above? Let us consider the first step,

• $2^8 + 2^7$

= {the product rule: $a^m a^n = a^{m+n}$}

$2^1 \cdot 2^7 + 2^7$

The justification is based on the mathematical rule

$$a \in \mathbb{R}^+,\ m \in \mathbb{N},\ n \in \mathbb{N} \vdash a^m \cdot a^n = a^{m+n}$$

which we may assume has been proved. We first observe that $\vdash 2 \in \mathbb{R}^+$, $\vdash 1 \in \mathbb{N}$, and $\vdash 7 \in \mathbb{N}$. This allows us to instantiate the rule with $a := 2$, $m := 1$, $n := 7$, Hence, we have that

$$\vdash 2^1 \cdot 2^7 = 2^{1+7}$$

The laws of arithmetic gives us that $\vdash 1 + 7 = 8$. Leibniz rule then gives us that

$$\vdash 2^{1+7} = 2^8$$

Transitivity then gives us that

$$\vdash 2^1 \cdot 2^7 = 2^8$$

Symmetry of equality gives us that

$$\vdash 2^8 = 2^1 \cdot 2^7$$

Finally, applying Leibniz' rule once more gives us that

$$\vdash 2^8 + 2^7 = 2^1 \cdot 2^7 + 2^7$$

We have thus proved that the first calculation step is correct, based on the basic logical rules we gave above. The other calculation steps are proved in a similar manner. We repeat the whole calculation below, with more detailed justifications for each step.

Example 8. A more detailed justification for each of the calculation steps in Example 6 is as follows. Each step is justified by some arithmetic or algebraic law.

- $2^8 + 2^7$

= {Rule $a \in \mathbb{R}^+, m, n \in \mathbb{N} \vdash a^m \cdot a^n = a^{m+n}$; instantiate with $a := 2$, $m := 1$, $n := 7$; modus ponens with $2 \in \mathbb{R}$, $1 \in \mathbb{N}$, $7 \in \mathbb{N}$; Leibniz rule with $1 + 7 = 8$; symmetry; Leibniz' rule with $2^8 = 2^1 \cdot 2^7$ (as described above)}

$2^1 \cdot 2^7 + 2^7$

= {Rule $a \in \mathbb{R}^+ \vdash a = 1 \cdot a$; instantiate with $a := 2^7$; modus ponens with $2^7 \in \mathbb{R}$; Leibniz' rule with $2^7 = 1 \cdot 2^7$}

$2^1 \cdot 2^7 + 1 \cdot 2^7$

= {Rule $a, b, c \in \mathbb{R} \vdash (a + b) \cdot c = a \cdot c + b \cdot c$; instantiate with $a := 2^1$, $b := 1$, and $c := 2^7$; modus ponens with $2^1, 1, 2^7 \in \mathbb{R}$; symmetry}

$(2^1 + 1) \cdot 2^7$

= {Rule $a \in \mathbb{R}^+ \vdash a^1 = a$; instantiate with $a := 2$; modus ponens with $2 \in \mathbb{R}$; Leibniz' rule with $2^1 = 2$}

$(2 + 1) \cdot 2^7$

$=$ $\{\text{Rule} \vdash 1 + 2 = 3; \text{ Leibniz' rule}\}$

$3 \cdot 2^7$

$=$ $\{\text{Rule} \vdash 2^7 = 128; \text{ Leibniz' rule}\}$

$3 \cdot 128$

$=$ $\{\text{Rule} \vdash 3 \cdot 128 = 384\}$

384

□

Transitivity gives the desired result for the calculation,

$$2^8 + 2^7 = 384$$

The use of instantiation, modus ponens, reflexivity, symmetry and transitivity, together with Leibniz' rule, is so familiar that we do not recognize these rules when we do calculations. The implicit use of these rules is good, it speeds up the calculations, and makes the presentation more compact. However, it is useful to understand the more detailed reason for why a calculation gives the result it gives.

3.4 What Does a Calculation Say

• t_0

$\sim_1$ *justification*$_1$

t_1

$\sim_2$ *justification*$_2$

t_2

$\vdots$

t_{n-1}

$\sim_n$ *justification*$_n$

t_n

□

Figure 3.1: Structured calculation

Consider the structured calculation on the left. Here $t_0, t_1, \ldots, t_n$ are mathematical expressions, and $\sim_1, \ldots, \sim_n$ are binary relations, $n \geq 1$. This calculation says that

$$\begin{array}{ll} \vdash & t_0 \sim_1 t_1, \text{ by } \textit{justification}_1 \\ \vdash & t_1 \sim_2 t_2, \text{ by } \textit{justification}_2 \\ \vdots & \\ \vdash & t_{n-1} \sim_n t_n, \text{ by } \textit{justification}_n \end{array}$$

The calculation is thus just a sequence of calculation steps with an explanation for why the indicated relationship holds between the two expressions in the step.

We construct a calculation because we want to derive some useful conclusion from it. If we use the same transitive relation $\sim$ (together with r equality) in each step, then we can conclude that $t_0 \sim t_n$. However, is is also possible to mix different relations in a calculation, and come to other kinds of conclusions from the calculation, as explained above. We will expand on this point later, in Chapter 6.

Our definition of correctness means that a calculation step is considered wrong even if the mathematical fact $t_{i-1} \sim_i t_i$ happens to be true, but the argument for it, *justification$_i$,* is wrong. We consider a calculation to be a proof that we write in order to convince ourself and others that a certain fact is true. The proof will not be convincing if there are steps without valid justifications. This highlights the important distinction between something being true and us knowing that something is true. In mathematics, we do not have any other way of knowing that a proposition is true than by proving it. If our proof has holes in it, then we know nothing about the proposition. Only when the whole proof is correct do we know that the proposition is true.

Chapter 4

Logical Calculations

We have described general arithmetic and algebraic calculations above. We now expand the treatment of calculations to logical calculations, i.e., calculations involving logical propositions.

Logic is everywhere in high school mathematics, and it should be taken seriously. We therefore believe that the standard logical notation should be introduced already at high school level, and students should be taught how to reason about logical expressions. Learning to reason about logical propositions is needed in high school mathematics anyway, mathematics just becomes harder if the basic concepts of logic are not properly identified and explained.

Many common functions, such as square roots and absolute values, are defined using logical expressions (side conditions that must be satisfied, definitions by cases, etc.). Manipulation of such functions requires knowledge about the basic rules of logic. The alternative, now prevailing in high school mathematics, is to use natural language to express logical relationships and rely on the student's intuition when it comes to simplifying the logical connectives in a calculation.

Explicit use of logic enhances the students' ability to express problems in a mathematical language. There are many nuances and side conditions that are ignored when a problem is described only in natural language, and the likelihood of getting the solution wrong increases. Formulating the problem using explicit logical connectives shows clearly what should be done, and it also shows what kind of logical rules can be used in the problem. This is particularly useful for students when solving more complex word problems.

The argument against teaching explicit logical notation and logical rules in high school is either that it does not fit into the current curriculum, or that it is too difficult for the students. The first argument is easy to counter (but perhaps more difficult to implement in practice): the curriculum has been changed before, and we have been able to introduce new fields of mathematics when they were considered important to society, such as probability theory and statistics. The second argument is more difficult to give direct answers to, because it is an empirical statement. We have made a number of experiments in schools, where we have taught basic logic already at first grade in high school, and then freely used them later when solving equations and other problems [26, 30, 9, 1, 4, 8]. The students have not considered logic particularly difficult, but rather as an interesting new topic of mathematics. They do not seem to have any serious problems with using logic in practice, in particular when we restrict ourselves to propositional calculus. Predicate calculus is somewhat more complex, and seems to be best taught at higher grades in high school or university level.

4.1 Equations and Equivalence

An equation is a prime example of a logical proposition. The equation

$$2x + 3 = 10x - 5$$

states that x has a value that satisfies the equality, i.e. the equation says that $2x+3$ has the same value as $10x - 5$. This proposition is true for some values of x and false for other values. The proposition is true when $x = 1$, since

$$2 \cdot 1 + 3 = 10 \cdot 1 - 5$$

but false when $x = 2$, since

$$2 \cdot 2 + 3 \neq 1 \cdot 2 - 5$$

We say that a value of x is a *solution* (or *root*) of the equation if the equation is true for this value of x. The equation above has only one solution, $x = 1$.

	$\vdash$	$p \equiv p$	$\{\equiv$ *is reflexive*$\}$
$p \equiv q$	$\vdash$	$q \equiv p$	$\{\equiv$ *is symmetric*$\}$
$p \equiv q,\ q \equiv r$	$\vdash$	$p \equiv r$	$\{\equiv$ *is transitive*$\}$
$e = e'$	$\vdash$	$P(e) \equiv P(e')$	$\{$*Leibniz' rule*$\}$

Table 4.1: Properties of equivalence

Equivalence is equality of truth values. Hence is obeys the same rules as equality: it is *reflexive*, *symmetric* and *transitive*. Table 4.1 repeats these properties for equivalence, together with Leibniz' rule for equivalence. Leibniz' rule says that we can replace an expressions e in a logical proposition $P(e)$ with another, equal expression e'. The resulting logical proposition $P(e')$ will then be equivalent to the original, i.e., $P(e) \equiv P(e')$ [1].

Equivalence between equations is a particularly important notion. Two equations are equivalent, if the same value of x always gives the same truth value for both equations. We have, e.g., that $2x + 3 = 10x - 5$ is equivalent to $2x = 10x - 5 - 3$. We write this as

$$(2x + 3 = 10x - 5) \ \equiv \ (2x = 10x - 5 - 3)$$

Since equality takes precedence over equivalence, we are free to omit the parenthesis, so this is the same as

$$2x + 3 = 10x - 5 \ \equiv \ 2x = 10x - 5 - 3$$

[1] Leibniz rule covers both the case when e and e' are algebraic expressions and $e = e'$, and when e and e' are logical expressions, and $e \equiv e'$.

The definition of equivalence implies that two equations are equivalent exactly when they have the same set of solutions.

Solving an equation means that we try to find an equation that is equivalent to the original equation and which shows explicitly the values of the unknown variables. For the example above, $x = 1$ is the (only) solution to the equation, since

$$2x + 3 = 10x - 5 \;\equiv\; x = 1$$

Equation $x = 1$ is a solution, since it directly shows which value of x satisfies the original equation.

We solve an equation by converting it step-by-step to a solution. Two important rules when we solve an equation are the following:

1. If we have an equation $s = t$ and add an expression e to both sides, we get a new equation that is equivalent to the original equation:

$$\vdash s = t \;\equiv\; s + e = t + e$$

2. Similarly, we get an equivalent equation by multiplying both sides of the equation by an expression e that is different from 0:

$$e \neq 0 \vdash s = t \;\equiv\; s \cdot e = t \cdot e$$

In addition, we use Leibniz' rule to simplify subexpressions that arise in the calculation.

Example 9. Solve equation $2x + 3 = 10x - 5$. The following shows a very careful derivation of the solution, where every step is explicitly justified:

$\bullet$ $2x + 3 = 10x - 5$

$\equiv$ {add $-10x$ to both sides of the equation (rule 1)}

$2x + 3 - 10x = 10x - 5 - 10x$

$\equiv$ {rearrange terms}

$3 + 2x - 10x = 10x - 10x - 5$

$\equiv$ {simplify (Leibniz' rule), using $2x - 10x = -8x$; here $P(e)$ is $3 + e = 10x - 10x - 5$, e is $2x - 10x$ and e' is $-8x$ }

$3 - 8x = 10x - 10x - 5$

$\equiv$ {simplify (Leibniz' rule), $10x - 10x = 0$}

$3 - 8x = 0 - 5$

$\equiv$ {simplify (Leibniz' rule), $0 - 5 = -5$}

$3 - 8x = -5$

$\equiv$ {add -3 to both sides (rule 1)}

$3 - 8x - 3 = -5 - 3$

$\equiv$ {reorder and simplify, $3 - 3 = 0$}

$-8x = -5 - 3$

$\equiv$ {reorder and simplify, $-5 - 3 = -8$}

$-8x = -8$

$\equiv$ {multiply both sides by $-\frac{1}{8}$ (rule 2, $-\frac{1}{8} \neq 0$)}

$-\frac{1}{8} \cdot (-8x) = -\frac{1}{8} \cdot (-8)$

$\equiv$ {simplify, $-\frac{1}{8} \cdot (-8) = 1$}

$x = 1$

□

■

We usually combine rule 1 and 2 with simplification, to shorten the calculation, and talk directly about subtraction and division. Our solution is then much more concise:

• $2x + 3 = 10x - 5$

$\equiv$ {subtract $10x$ from both sides and simplify}

$3 - 8x = -5$

$\equiv$ {subtract 3 from both sides and simplify}

$-8x = -8$

$\equiv$ {divide both sides by -8 and simplify}

$x = 1$

□

A first-degree equation usually has exactly one solution. But not always. Sometimes it has no solution and sometimes every value of x can be a solution. An example of an equation with no solution is shown below.

Example 10. Solve the equation

$$2x + 3 = 2x$$

The solution is as follows:

$\bullet \quad 2x + 3 = 2x$

$\equiv \quad$ {subtract $2x$ from both sides}

$\quad 3 = 0$

$\equiv \quad$ $\{3 = 0$ is not true$\}$

$\quad F$

□

The solution to the equation shows that

$$2x + 3 = 2x \equiv F$$

In other words, the equation has the truth value F for every value of x, i.e. the equation is always false (same as never true). This means that the equation does not have a solution. ■

Next, we show an example of an equation where every value of x is a solution.

Example 11. Solve the equation

$$2x = 2(x + 1) - 2$$

We solve this as follows:

$\bullet \quad 2x = 2(x + 1) - 2$

$\equiv \quad$ {simplify the RHS}

$\quad 2x = 2x$

$\equiv \quad$ {holds for every value of x}

$\quad T$

□

We get that

$$2x = 2(x + 1) - 2 \equiv T$$

Thus the equation is true for every value of x, i.e., the equation is always true. ■

A first-degree equation can thus be

- always false, i.e., have no solution, or
- equivalent to an equation of the form $x = c$ for some value of c, i.e., it has exactly one solution, or
- always true, i.e., every value of x is a solution.

The techniques for solving equations can also be used for solving inequalities, as shown by the following example.

Example 12. Simplify the condition $2x + 1 \leq 4x + 3$.

$$\begin{array}{ll} \bullet & 2x + 1 \leq 4x + 3 \\ \equiv & \{\text{subtract 3 from both sides and simplify}\} \\ & 2x - 2 \leq 4x \\ \equiv & \{\text{subtract } 2x \text{ from both sides and simplify}\} \\ & -2 \leq 2x \\ \equiv & \{\text{divide both sides by 2 and simplify}\} \\ & -1 \leq x \end{array}$$

□

The result,

$$2x + 1 \leq 4x + 3 \equiv -1 \leq x$$

now follows by transitivity.

The truth value of an inequality does not change if we add the same arbitrary number to both sides, or if we divide both sides of the inequality by a positive number. We see this directly in the example: if $2x + 1 \leq 4x + 3$ is true for a given value of x (e.g. $x = 1$), then $-1 \leq x$ is also true for the same value, and if $2x + 1 \leq 4x + 3$ is false for a value of x (e.g. $x = -2$), then $-1 \leq x$ is false as well.

Similarly, the truth value does not change if we simplify a subexpression in an inequality: $2x + 1 - 3 \leq 4x + 3 - 3$ has the same truth value as $2x - 2 \leq 4x$. This is an application of Leibniz' rule for propositions. However, multiplying both sides of an inequality with a negative value will change the direction of the inequality, so here is a difference to equation solving. ■

4.2 Logical Expressions

We often combine logical propositions in everyday language. We say, e.g., that "it is raining *and* blowing", that "the car is white *or* black", that "the door can be opened *if* the key is in the door", that "it is *not* blowing", etc. We can construct complicated logical propositions in this way:

> "The cargo is thoroughly attached to the trailer *if* it is tied by two straps *and* it does not protrude from the sides of the trailer, *neither* to the left *nor* to the right, *and, if* it protrudes behind the trailer, *then* it has a piece of white cloth attached to its end to warn drivers following behind."

Just like there are arithmetic operations like addition, subtraction, multiplication, and division on real numbers, there are *logical operations* on truth values. These operations are called *logical connectives.* A logical connective is an operation on truth values that gives a new truth value as a result.

Basic logic (usually referred to as *proposition calculus*), has five logical operations/connectives. Let p and q be two logical propositions. Then we can form new logical propositions from these with the following operations:

- *negation*, "not p", denoted $\neg p$,
- *conjunction,* "p and q", denoted $p \wedge q$,
- *disjunction*, "p or q" denoted $p \vee q$,
- *implication,* "if p then q", denoted $p \Rightarrow q$, and
- *equivalen*ce, $p = q$, denoted $p \equiv q$.

We can use these logical connectives to construct more complicated logical propositions, i.e. *logical expressions.* We can perform calculations with logical expressions in the same way (but with slightly different rules) as on arithmetic and algebraic expressions.

The following table shows how to construct logical expressions from basic propositions. We start from a collection of basic propositions: "it is raining", "it is blowing", "the car is white", "the car is black", "the key is in the door", "the door can be opened". We rephrase the propositions above using logical connectives in propositional calculus:

Natural language	**Propositional calculus**
it is *not* blowing	$\neg$ it is blowing
it is blowing *and* raining	it is raining $\wedge$ it is blowing
the car is white *or* black	the car is white $\vee$ the car is black
if the key is in the door, *then* the door can be opened	the key is in the door $\Rightarrow$ the door can be opened
the door can be opened, *if* the key is in the door	the door can be opened $\Leftarrow$ the key is in the door

The last example shows that implication can be used in both directions (similar to how we can write $a \leq b$ and $b \geq a$). Forward implication $p \Rightarrow q$ in natural language is written "*if p then q"* while backward implication $q \Leftarrow p$ is written as "*q if p".*

We summarize here the intuitive meaning of all connectives. Let p and q be two logical propositions. Then we define the connectives as follows:

- *negation:* $\neg p$ is true exactly when p is false

- *conjunction:* $p \wedge q$ is true exactly when both p and q are true
- *disjunction:* $p \vee q$ is true exactly when either p or q (or both) are true
- *implication:* $p \Rightarrow q$ is true exactly when p is false or q is true
- *equivalence* $p \equiv q$ is true exactly when p and q are both true or both are false.

The definition of implication requires an explanation (the other connectives should be obvious). If p is true, then q must be true for the implication $p \Rightarrow q$ to be true. The implication $p \Rightarrow q$ is also true if p is false. We understand the implication $p \Rightarrow q$ as "if p then q", with the tacit understanding that the implication is also true when p is false. We call this a *material implication*, as opposed to the so-called *causal implication.* Material implication does not require there to be any connection between p and q, while causal implication assumes that p in one way or another is the reason for q.

An example of material implication is that

$$2 > 7 \Rightarrow \text{the Moon is a cube}$$

This implication is true, since $2 > 7$ is false. The proposition

$$2 < 7 \Rightarrow \text{the Moon is a cube}$$

on the other hand, is false, since $2 < 7$ is true, but the Moon is not a cube. Finally, the proposition

$$2 < 7 \Rightarrow \text{the Moon is a spehere}$$

is true, since $2 < 7$ and the Moon is actually a sphere.

Example 13. We can make the rule for loading a trailer above more precise by writing it as a logical proposition. The basic logical propositions in the sentence are the following:

C = the cargo is thoroughly attached to the trailer
S = the cargo is tied with two straps
L = the cargo protrudes to the left
R = the cargo protrudes to the right
B = the cargo protrudes behind the trailer
W = a piece of white cloth is attached to the cargo

We can then give the same description in the form of the following logical proposition:

$$C \Leftarrow S \wedge \neg L \wedge \neg R \wedge (B \Rightarrow W)$$

We have here used the converse implication $q \Leftarrow p$. ■

p	$\neg p$
T	F
F	T

p	q	$p \wedge q$	$p \vee q$	$p \Rightarrow q$	$p \equiv q$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

Table 4.2: Truth tables

4.3 Truth Tables

Since there are only two truth values, T and F, we can easily define each connective by enumerating its argument values for the different combinations of T and F. This is called a *truth table.* The truth table for negation is shown in Table 4.2, on the left. The first row states that when $p \equiv T$, then $\neg p \equiv F$. The second row states that when $p \equiv F$, then $\neg p \equiv T$. This corresponds to the intuitive definition we gave above.

The other connectives all have two arguments, so we can summarize them with the other truth table in Table 4.2, on the right. This table give the value of every connective for every combination of truth values of p and q. The third row of the table shows, e.g., that when $p \equiv F$ and $q \equiv T$, then $(p \wedge q) \equiv F$, $(p \vee q) \equiv T$, $(p \Rightarrow q) \equiv T$ and $(p \equiv q) \equiv F$. This corresponds to the definitions we gave earlier.

Mathematical propositions are logical propositions. An example is the logical proposition

$$(x + y \leq 3) \wedge (\neg(x = 2))$$

We have here combined two logical propositions by conjunction. The proposition states that $x + y \leq 3$ and $x \neq 2$.

We can calculate the truth value of this proposition whenever the values of x and y are given. Consider, e.g., the situation when $x = 2$ and $y = 0$. We calculate the truth value of the entire proposition as follows:

• $(x + y \leq 3) \wedge (\neg(x = 2))$

$\equiv$ $\{$insert the values of the variables, $x = 2$ and $y = 0\}$

$(2 + 0 \leq 3) \wedge (\neg(2 = 2))$

$\equiv$ $\{$calculate the truth value of the two propositions: $2+0 \leq 3 \equiv T$ and $2 = 2 \equiv T\}$

$T \wedge (\neg T)$

$\equiv$ $\{$the definition of negation, from the truth table$\}$

$T \wedge F$

$\equiv$ {the definition of conjunction, from the truth table}

F

□

This shows that $(x + y \leq 3) \wedge (\neg(x = 2))$ is false when $x = 2$ and $y = 0$. Similarly, we can see that the proposition is true when e.g. $x = 1$ and $y = 2$, and false again when $x = 1$ and $y = 4$.

4.4 Evaluating Logical Expressions

We can get away with fewer parentheses in logical propositions by fixing a preference order for connectives. The order of operation for logical propositions is the following (unless changed by parentheses):

1. Calculate the values of arithmetic expressions, in the usual way
2. Calculate the truth values for equalities, inequalities, and other relations
3. Calculate the truth value of the negations
4. Calculate the truth value of conjunctions and disjunctions
5. Calculate the truth value of implications and equivalences.

Note that there is no specific order between conjunctions and disjunctions, nor is there one between implication and equivalence. This means that we cannot write an expression like $p \wedge q \vee r$. We must use parentheses to state what we mean exactly: do we mean $p \wedge (q \vee r)$ or $(p \wedge q) \vee r$. Similarly, we have to use parentheses to distinguish between $(p \Rightarrow q) \equiv r$ and $p \Rightarrow (q \equiv r)$.

Example 14. Calculate the truth value of the logical proposition

$$x + y \leq 3 \vee \neg x = 2 \Rightarrow y \leq 1 \wedge x \geq 3$$

when $x = 2$ and $y = 0$.

• $x + y \leq 3 \vee \neg x = 2 \Rightarrow y \leq 1 \wedge x \geq 3$

$\equiv$ {insert the values of the variable from the assumptions, $x = 2$ and $y = 0$}

$2 + 0 \leq 3 \vee \neg 2 = 2 \Rightarrow 0 \leq 1 \wedge 2 \geq 3$

$\equiv$ {calculate the values of the arithmetic expressions, $2 + 0 = 2$, step 1}

$2 \leq 3 \vee \neg 2 = 2 \Rightarrow 0 \leq 1 \wedge 2 \geq 3$

$\equiv$ {calculate the truth value of the individual equalities and inequalities, $2 \leq 3 \equiv T$, $2 = 2 \equiv T$, $0 \leq 1 \equiv T$ and $2 \geq 3 \equiv F$, step 2}

$T \vee \neg T \Rightarrow T \wedge F$

$\equiv$ {calculate the negation, $(\neg T) \equiv F$, from the truth table, step 3}

$T \vee F \Rightarrow T \wedge F$

$\equiv$ {calculate the disjunction, $(T \vee F) \equiv T$, from the truth table, step 4}

$T \Rightarrow T \wedge F$

$\equiv$ {calculate the conjunction, $(T \wedge F) \equiv F$, from the truth table, step 4}

$T \Rightarrow F$

$\equiv$ {calculate the implication, $(T \Rightarrow F) \equiv F$, from the truth table, step 5}

F

□

In other words, the proposition is false when $x = 2$ and $y = 0$. ■

4.5 Logical Theorems and Proofs

In the same way as we can have mathematical theorems, we can have purely logical theorems. An example of a logical theorem is the following:

$$\vdash p \wedge T \equiv p$$

Here p is a logical proposition (and hence has value F or T, depending on the values of the variables in p). The rule states that $p \wedge T \equiv p$ for every truth value of p, i.e. both when $p = T$ and when $p = F$. Intuitively, the rule says that eliminating T in a conjunction does not change the truth value of the conjunction.

We can easily see that this rule holds using the truth table for conjunction. Consider the cases $p = T$ and $p = F$ separately. For the case $p = T$, we have

• $p \wedge T$

$\equiv$ {the assumption $p = T$}

$T \wedge T$

$\equiv$ {the definition of conjunction}

T

$\equiv$ {the assumption $p = T$}

p

□

For case $p = F$, we have

• $\quad p \wedge T$

$\equiv \quad$ {the assumption $p = F$}

$\quad F \wedge T$

$\equiv \quad$ {the definition of conjunction}

$\quad F$

$\equiv \quad$ {the assumption $p = F$}

$\quad p$

□

We can summarize this proof in a truth table, where we first calculate the value of $p \wedge T$ for both cases, $p = T$ and $p = F$, and then we calculate the value of $p \wedge T \equiv p$ in both cases:

p	$p \wedge T$	$p \wedge T \equiv p$
F	F	T
T	T	T

The rule is true, since the last column is true for every combination of values of p.

Example 15. Check that disjunction is commutative, i.e. that

$$\vdash p \vee q \equiv q \vee p$$

We check the rule by constructing its truth table:

p	q	$p \vee q$	$q \vee p$	$p \vee q \equiv q \vee p$
F	F	F	F	T
F	T	T	T	T
T	F	T	T	T
T	T	T	T	T

The rule is true, since the last column is true for every combination of values of p and q. ■

Example 16. Check that double negations can be omitted, i.e. that

$$\vdash \neg(\neg p) \equiv p$$

The truth table looks as follows:

p	$\neg p$	$\neg(\neg p)$	$\neg(\neg p) \equiv p$
F	T	F	T
T	F	T	T

This shows that the rule holds. ■

Example 17. Check that the rule for rewriting an implication as a disjunction holds, i.e. that

$$\vdash (p \Rightarrow q) \equiv (\neg p \vee q)$$

Here is the truth table:

p	q	$p \Rightarrow q$	$\neg p$	$\neg p \vee q$	$(p \Rightarrow q) \equiv (\neg p \vee q)$
F	F	T	T	T	T
F	T	T	T	T	T
T	F	F	F	F	T
T	T	T	F	T	T

We see that the rule holds here as well. ■

Instead of determining the truth value of a logical rule directly with a truth table, we can prove the rule with a calculation, based on logical rules that we already know are true.

Example 18. Show that the rule for *contraposition* holds, i.e. that

$$\vdash (p \Rightarrow q) \equiv (\neg q \Rightarrow \neg p)$$

We prove the equivalence with a calculation.

$\bullet$ $\quad p \Rightarrow q$

$\equiv$ $\quad$ {write the implication as a disjunction}

$\quad \neg p \vee q$

$\equiv$ $\quad$ {double negation}

$\quad \neg p \vee \neg(\neg q)$

$\equiv$ $\quad$ {disjunction is commutative}

$\quad \neg(\neg q) \vee \neg p$

$\equiv$ $\quad$ {write the disjunction as an implication}

$\quad \neg q \Rightarrow \neg p$

□ ■

We have two different ways of reasoning about logical expressions (in propositional calculus): algebraic manipulation with the rules of logic, and calculating truth values with truth tables. The latter is more mechanical, but in practice it does not scale up very well to larger and more complex situations. Algebraic and logical manipulation is more efficient, once one has learnt the basic rules. When we go over to quantifiers, this is also the only way, since there are no (finite) truth tables for quantifiers.

4.6 Assignments

1. The proposition $2 = 2$ is a) true, b) false, c) contingent (i.e., depends on the value of x)?
2. Is the proposition $x + 5 = 3$ a) true, b) false c) contingent?
3. Determine whether the mathematical proposition $3 - x = 5 - x$ is true, false, or contingent.
4. Consider the equation $ax+b = cx+d$. Under what conditions on the constants a, b, c and d is the equation always false? a) $a = c$ and $b \neq d$, b) $a = c$ and $b = d$, c) $a \neq c$ and $b = d$, d) $a \neq c$ and $b \neq d$.
5. x^3+1+2x is a) false, b) contingent c) true d) none of the previous alternatives.
6. Assume that p and q are logical expressions. The combination $p \wedge q$ a) is true, b) is false, c) is contingent d) does not have a truth value?
7. The statement "$x + 4 = x + 3$ is false" a) does not have a truth value, b) is false, c) is contingent d) is true?
8. Prove that the equation $2x + 3 = -2(4 - x)$ is false.
9. Determine the truth value of the statement "the equation $2x - 3 = 5x - 3$ is contingent"
10. Simplify the statement $(x > y) \wedge (\neg (x > y) \vee (2 \neq 3))$
11. Determine the truth value of the statement $(x \neq y \vee x > y) \wedge y \neq 1$, when $x = 132$ and $y = -2$.
12. Are the two formulae $p \Rightarrow (q \wedge \neg q)$ and $\neg p$ logically equivalent?
13. Determine the truth value of the statement $\neg (\neg p \Rightarrow (q \wedge \neg r))$, when p, q and r are the logical statements $x < 5$, $y > 10$ and $y > x$, for the case when $x = 7$ and $y = 12$.
14. Simplify the expression $((x < y) \wedge \neg (y \leq x)) \vee (x > y \wedge x = y)$.
15. a) Calculate the truth value of $3 = 2 \Rightarrow (7 \neq 9 \Rightarrow (3 = 2 \wedge 7 \neq 9))$. b) Prove that $x = y \Rightarrow (z \neq v \Rightarrow (x = y \wedge z \neq v))$ will always hold for any choice of integers x, y, z and v. c) Will it always hold if you replace $x = y$ and $z \neq v$ with other random statements?

16. Simplify the expression $\big(((5x = 5x \vee 2 = 3) \Rightarrow ((x < 5 \wedge x - 3 < 0) \wedge x > 2)) \vee \big(x > 4 \wedge x^2 < 25\big)\big)$ $\wedge\ (x \neq 4.5 \vee 7 = y)$, when $y = 2$.

Chapter 5

Solving Equations with Logic

Equation solving is a special case of the more general problem of simplifying a logical proposition p into another, equivalent but simpler logical proposition. In other words, finding a simpler logical proposition q such that $p \equiv q$. We want q to show, as explicitly as possible, which values of the variables satisfy p. The two extreme cases are propositions that are satisfied for every value of the variable, i.e. $p \equiv T$, and propositions that are not satisfied for any values of the variable, i.e. $p \equiv F$. Between these extremes we have propositions that are true for some values of the variables and false for other values.

We will show below how the different logical connectives appear in a natural way when we solve equations of the kind considered in high school. At the same time, we will introduce the basic logical rules that are needed for manipulating logical expressions. These rules are already used intuitively when solving equations, but we make them easier to use by formulating the rules explicitly. Equation solving is then more mechanical, and it is easier to check that the solution is correct.

The logical rules for connectives are important, they form the basis for all mathematical reasoning. In most cases, the rules are also obvious, since they are part of our common sense, of the way we understand mathematical concepts. Sometimes a little bit of reflection is needed to realize that a particular logical rule must be true. But once we have convinced ourselves of these rules, they quickly become part of the toolkit that we use when solving mathematical problems.

5.1 Equations and Conjunction

Conjunctions are common when solving equations. We give here two different examples, simultaneous equations and equations involving square roots.

Consider a pair of simultaneous equations,

$$\begin{cases} 2x + 3 = y - x \\ 5y + 4 = x \end{cases}$$

The simultaneous equation says that both equations have to hold for x and y. Hence, we can write the simultaneous equations as a conjunction,

$$2x + 3 = y - x \wedge 5y + 4 = x$$

This pair of equations is true for x and y when both $2x + 3 = y - x$ and $5y + 4 = x$ are satisfied.

Example 19. We solve this pair of simultaneous equations with the following calculation:

• $2x + 3 = y - x \wedge 5y + 4 = x$

$\equiv$ {substitute the value of x given in the second equation for the value of x in the first equation}

$2 \cdot (5y + 4) + 3 = y - (5y + 4) \wedge 5y + 4 = x$

$\equiv$ {simplify the first equation}

$10y + 11 = -4y - 4 \wedge 5y + 4 = x$

$\equiv$ {rearrange terms in the first equation}

$14y = -15 \wedge 5y + 4 = x$

$\equiv$ {solve the first equation}

$y = -\frac{15}{14} \wedge 5y + 4 = x$

$\equiv$ {insert the value of y into the second equation}

$y = -\frac{15}{14} \wedge 5 \cdot (-\frac{15}{14}) + 4 = x$

$\equiv$ {solve the second equation}

$y = -\frac{15}{14} \wedge x = -\frac{75}{14} + 4$

$\equiv$ {simplify the equations}

$y = -1\frac{1}{14} \wedge x = -1\frac{5}{14}$

□

The solution to the simultaneous equations is thus $x = -1\frac{5}{14}$ and $y = -1\frac{1}{14}$. The solution is given as a conjunction of the form

$$x = -1\frac{5}{14} \wedge y = -1\frac{1}{14}$$

which explicitly shows the values of x and y that satisfy the original equation. ■

The rules for conjunction are given in Table 5.1. Conjunction is both *commutative* and *associative*, in the same way as addition and multiplication. This means that changing the order of a conjunction does not change its truth value, and that we may group the various parts of a conjunction freely. Furthermore, conjunction is *idempotent*, i.e., we are free to omit repetitions in a conjunction (a proposition does not become more true by repeating it).

$\vdash$	$p \wedge q \equiv q \wedge p$	$\{\wedge$ *is commutative*$\}$
$\vdash$	$p \wedge (q \wedge r) \equiv (p \wedge q) \wedge r$	$\{\wedge$ *is associative*$\}$
$\vdash$	$p \wedge p \equiv p$	$\{\wedge$ *is idempotent*$\}$
$\vdash$	$p \wedge T \equiv p$	$\{\wedge$ *with truth*$\}$
$\vdash$	$p \wedge F \equiv F$	$\{\wedge$ *with falsity*$\}$

Table 5.1: Conjunction rules

An important special case is simplification of a conjunction when we know the truth value of one argument. The definition of conjunction implies directly that a false proposition makes the entire conjunction false, while a true proposition does not affect the truth value of the conjunction. This is shown by the last two rules in Table 5.1.

Our next example is about square roots in equations. We use the following definition of a square root in the calculation:

$$\sqrt{a} = b \;\equiv\; a = b^2 \wedge b \geq 0$$

This means that the square root of a number is always positive.

Example 20. The task is to solve the equation $\sqrt{x^2 + 9} = 5$.

• $\sqrt{x + 9} = 5$

$\equiv$ {definition of the square root}

$x + 9 = 5^2 \;\wedge\; 5 \geq 0$

$\equiv$ {simplify, $5 \geq 0$ is true}

$x = 16 \;\wedge\; T$

$\equiv$ {a true proposition can be omitted from a conjunction (the rule for truth: $p \wedge T \equiv p$)}

$x = 16$

□ ■

We illustrate the conjunction rules further with the following example.

Example 21. Simplify the logical proposition $p \wedge (T \wedge p \wedge F) \wedge q$.

• $p \wedge (T \wedge p \wedge F) \wedge q$

$\equiv$ {associativity, group the propositions inside the parenthesis}

$p \wedge ((T \wedge p) \wedge F) \wedge q$

$\equiv$ {$\wedge$ is commutative}

$p \wedge ((p \wedge T) \wedge F) \wedge q$

$\equiv$ {$\wedge$ with truth}

$p \wedge (p \wedge F) \wedge q$

$\equiv$ {$\wedge$ with falsity}

$p \wedge F \wedge q$

$\equiv$ {associativity, grouping}

$(p \wedge F) \wedge q$

$\equiv$ {$\wedge$ with falsity}

$F \wedge q$

$\equiv$ {$\wedge$ is commutative}

$q \wedge F$

$\equiv$ {$\wedge$ with falsity}

F

□

Example 22. The calculation above is quite a bit longer than what is needed in practice. A more concise calculation uses associativity and commutativity implicitly, in the same way as we use these rules in algebraic simplifications.

- $p \wedge (T \wedge p \wedge F) \wedge q$

$\equiv$ {F makes the proposition in parenthesis false}

$p \wedge F \wedge q$

$\equiv$ {F makes the whole conjunction false}

F

□

5.2 Equations and Disjunction

In the same way as conjunction, disjunction is also *commutative, associative* and *idempotent.* Table 5.2 shows the basic rules for disjunction.

$\vdash$	$p \vee q \equiv q \vee p$	$\{\vee$ *is commutative*$\}$
$\vdash$	$p \vee (q \vee r) \equiv (p \vee q) \vee r$	$\{\vee$ *is associative*$\}$
$\vdash$	$p \vee p \equiv p$	$\{\vee$ *is idempotent*$\}$
$\vdash$	$p \vee T \equiv T$	$\{\vee$ *with truth*$\}$
$\vdash$	$p \vee F \equiv p$	$\{\vee$ *with falsity*$\}$

Table 5.2: Disjunction rules

Disjunctions occur naturally when we solve second- or higher-degree equations. The disjunction gets into the calculations because of the *zero-product property,*

$$\vdash (a \cdot b = 0) \equiv (a = 0 \vee b = 0)$$

The rule states that a product is zero exactly when at least one factor is zero.

A second-degree equation can have one or two solutions, or always be true or always be false. If the equation has two solutions, $x = x_0$ and $x = x_1$, then we can describe the solutions as a disjunction,

$$x = x_0 \vee x = x_1$$

While we described the solution to the simultaneous equations above as a conjunction, we describe the solution to a second-degree equation as a disjunction.

Example 23. Solve the equation $x^2 - 2x - 3 = 0$.

• $x^2 - 2x - 3 = 0$

$\equiv$ {rewrite equation to prepare for factorization}

$x^2 - 3x + x - 3 = 0$

$\equiv$ {factor out x in first two terms}

$x(x - 3) + x - 3 = 0$

$\equiv$ {factor out $x - 3$}

$(x + 1)(x - 3) = 0$

$\equiv$ {zero-product rule}

$x + 1 = 0 \vee x - 3 = 0$

$\equiv$ {solve equations}

$x = -1 \vee x = 3$

□

This shows that $x^2 - 2x - 3 = 0$ has the same truth value as $x = -1 \vee x = 3$, for every value of x. Thus $x = -1 \vee x = 3$ gives the two solutions to the equation. ■

The next example shows how to use associativity and idempotence when we solve equations.

Example 24. Solve the equation $x^3 - x^2 - x + 1 = 0$.

$\bullet$ $x^3 - x^2 - x + 1 = 0$

$\equiv$ {rearrange the terms}

$x^3 - x - x^2 + 1 = 0$

$\equiv$ {factor out a common factor of the subexpressions}

$x(x^2 - 1) - (x^2 - 1) = 0$

$\equiv$ {factor out a common factor of the entire expression}

$(x - 1)(x^2 - 1) = 0$

$\equiv$ {the zero-product property}

$x - 1 = 0 \vee x^2 - 1 = 0$

$\equiv$ {add 1 to both sides in the left-hand equation}

$x = 1 \vee x^2 - 1 = 0$

$\equiv$ {add -1 to both sides of the right-hand equation}

$x = 1 \vee x^2 = 1$

$\equiv$ {solve the right equation, $x^2 = 1 \equiv x = 1 \vee x = -1$}

$x = 1 \vee (x = 1 \vee x = -1)$

$\equiv$ {disjunction is associative}

$(x = 1 \vee x = 1) \vee x = -1$

$\equiv$ {disjunction is idempotent}

$x = 1 \vee x = -1$

□

In this case the equation has two solutions, $x = 1 \vee x = -1$. ■

We have similar rules for disjunction with truth and falsity as we have for conjunction, as shown in Table 5.2. These rules say that the truth value of a disjunction does not change if we add a false proposition, while the entire disjunction becomes true if we add a true proposition. In the next example we show how to use the the rule of disjunction with falsity.

Example 25. Solve the equation $x^3 - x^2 + x - 1 = 0$.

$\bullet$ $\quad x^3 - x^2 + x - 1 = 0$

$\equiv$ $\quad$ {rearrange the terms}

$\quad x^3 + x - x^2 - 1 = 0$

$\equiv$ $\quad$ {factor out a common factor of the subexpressions}

$\quad x(x^2 + 1) - (x^2 + 1) = 0$

$\equiv$ $\quad$ {factor out a common factor of the entire expressions}

$\quad (x - 1)(x^2 + 1) = 0$

$\equiv$ $\quad$ {the zero-product property}

$\quad x - 1 = 0 \lor x^2 + 1 = 0$

$\equiv$ $\quad$ {add 1 to both sides of the left-hand equation}

$\quad x = 1 \lor x^2 + 1 = 0$

$\equiv$ $\quad$ {add -1 to both sides of the right-hand equation}

$\quad x = 1 \lor x^2 = -1$

$\equiv$ $\quad$ {a square is never negative, i.e. $(x^2 = -1) \equiv F$}

$\quad x = 1 \lor F$

$\equiv$ $\quad$ {$\lor$ with falsity}

$\quad x = 1$

□

In this case, the equation has only one solution, $x = 1$. ■

5.3 Equations with Conjunction and Disjunction

Disjunction and conjunction are both commutative, like addition and multiplication. For addition and multiplication we also have the property that multiplication distributes over addition: $a \cdot (b + c) = a \cdot b + a \cdot c$. Conjunction and disjunction have the same property: conjunction distributes over disjunction. In addition, disjunction also distributes over conjunction, as shown in Table 5.3. Connectives thus have stronger properties than the arithmetic operations: multiplication distributes over addition, but addition does not distribute over multiplication (generally, $a + (b \cdot c) \neq (a + b) \cdot (a + c)$).

	$\vdash$	$p \wedge (q \vee r) \equiv (p \wedge q) \vee (p \wedge r)$	$\{\wedge$ *distributes over* $\vee\}$
	$\vdash$	$p \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r)$	$\{\vee$ *distributes over* $\wedge\}$
$p \Rightarrow q$	$\vdash$	$p \wedge q \equiv p$	$\{$*redundancy with* $\wedge\}$
$p \Rightarrow q$	$\vdash$	$p \vee q \equiv q$	$\{$*redundancy with* $\vee\}$

Table 5.3: Distribution and redundancy rules

Example 26. Simplify the logical proposition $p \wedge (q \vee p \vee F) \wedge q$. The rules for associativity and commutativity of conjunction and disjunction are used implicitly, in the same way as in normal algebraic derivations.

• $p \wedge (q \vee p \vee F) \wedge q$

$\equiv$ {false propositions can be omitted from a disjunction: $p \vee F \equiv p$}

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

$\equiv$ {conjunction distributes over disjunction: $p \wedge (q \vee p) \equiv (p \wedge q) \vee (p \wedge p)$}

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

$\equiv$ {conjunction is idempotent: $p \wedge p \equiv p$}

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

$\equiv$ {conjunction distributes over disjunction}

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

$\equiv$ {conjunction is idempotent: $q \wedge q \equiv q$}

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

$\equiv$ {disjunction is idempotent}

$p \wedge q$

□

In other words, we have shown that $p \wedge (q \vee p \vee F) \wedge q \equiv p \wedge q$. ■

The rules for conjunction and disjunction are very useful for solving equations involving absolute values. We define absolute values by the following formula

$$|a| = b \equiv ((a \geq 0 \wedge a = b) \vee (a < 0 \wedge a = -b))$$

In other words, $|a| = b$ when either $a \geq 0$ and $a = b$, or $a < 0$ and $a = -b$.

Example 27. Solve the equation $|2x-1| = x+1$.

- $|2x-1| = x+1$

$\equiv$ {definition of absolute values}

$(2x-1 \geq 0 \wedge 2x-1 = x+1) \vee (2x-1 < 0 \wedge 2x-1 = -(x+1))$

$\equiv$ {solve the inequalities}

$\big(x \geq \frac{1}{2} \wedge 2x-1 = x+1\big) \vee \big(x < \frac{1}{2} \wedge 2x-1 = -(x+1)\big)$

$\equiv$ {solve the equations}

$\big(x \geq \frac{1}{2} \wedge x = 2\big) \vee (x < \frac{1}{2} \wedge x = 0)$

$\equiv$ {the condition $x \geq \frac{1}{2}$ in the left disjunct is redundant, since it follows from the stronger condition $x = 2$}

$\big(x = 2 \vee \big(x < \frac{1}{2} \wedge x = 0\big)\big)$

$\equiv$ {the condition $x < \frac{1}{2}$ in the right disjunct is redundant, since it follows from the stronger condition $x = 0$}

$x = 2 \vee x = 0$

□ ■

The last two steps in the derivation use the redundancy rules of Table 5.3. The first of these rules says that we can omit a redundant proposition from a conjunction. A proposition is *redundant*, if it follows from the other propositions in the conjunction. The rule says that $p \Rightarrow q \vdash p \wedge q \equiv p$.

Consider an implication $p \Rightarrow q$. We refer here to p as the *stronger* proposition and q as the *weaker* proposition, because q follows from p. Thus, the redundancy rule for conjunction says that we can omit the weaker proposition from a conjunction (since it is part of the stronger proposition, so to speak).

In the second to last step, we have $x = 2 \Rightarrow x \geq \frac{1}{2}$ (if $x = 2$ it is obvious that $x \geq \frac{1}{2}$, since $2 \geq \frac{1}{2}$). We then have that

$$x \geq \frac{1}{2} \wedge x = 2 \equiv x = 2$$

i.e., we can omit the weaker proposition $x \geq \frac{1}{2}$ from the conjunction). In the last step we have $x = 0 \Rightarrow x < \frac{1}{2}$, so we can omit the weaker proposition $x < \frac{1}{2}$ from the conjunction:

$$x < \frac{1}{2} \wedge x = 0 \equiv x = 0$$

There is a corresponding rule for disjunction, where we omit the stronger proposition: $p \Rightarrow q \vdash p \vee q \equiv q$. In other words, in a conjunction we can omit the weaker proposition, while in a disjunction we can omit the stronger proposition. [1]

When we solve an equation involving absolute values, we get two different cases, depending on whether the expression within the absolute value signs is negative or non-negative. Usually we would handle these two cases with calculations performed somewhere else on the paper. Logic allows us to treat this as one expression with two disjuncts. When solving simultaneous equations by substitution, we also get two separate logical propositions, but in this case they are joined together as a conjunction. With absolute values, we have a disjunction of two propositions, with simultaneous equations we have a conjunction. This distinction is central to the calculation, but it remains obscure when we solve equations in the traditional way. Logical connectives allows calculations that branch out into sub-calculations, where the relation between the different branches is explicitly stated with connectives. Calculations from different branches can be later combined into a single result, and we can continue from there with further calculations.

Example 28. In the following example we solve two different subexpressions simultaneously in a nested derivation. The problem is to solve the inequality $(1+x)^2 \leq 1$.

$$\begin{array}{lll}
\bullet & (1+x)^2 \leq 1 & \\
\equiv & \{\text{the binomial rule } (a+b)^2 = a^2+2ab+b^2\} & \\
 & 1+2x+x^2 \leq 1 & \\
\equiv & \{\text{subtract 1 from both sides}\} & \\
 & 2x+x^2 \leq 0 & \\
\equiv & \{\text{the distributive law: } ab+ac = a\,(b+c)\} & \\
 & x(2+x) \leq 0 & \\
\equiv & \{\text{write the inequality in alternative form: } (a \leq 0) \equiv (a = 0 \vee a < 0) & \\
 & (x(2+x) = 0) \vee (x(2+x) < 0) & \\
\equiv & \{\text{solve both disjunctions separately}\} & \\
 & \bullet & x(2+x) = 0 \\
 & \equiv & \{\text{the zero-product property: } (ab = 0) \equiv (a = 0 \vee b = 0)\} \\
 & & x = 0 \vee 2+x = 0 \\
 & \equiv & \{\text{subtract 2 from both sides in the right-hand conjunct}\} \\
 & & x = 0 \vee x = -2 \\
 & \bullet & x(2+x) < 0
\end{array}$$

[1] In fact, this applies in the other direction as well, the more general rules are

$$\begin{array}{llll}
\vdash & (p \Rightarrow q) & \equiv & (p \wedge q \equiv p) \\
\vdash & (p \Rightarrow q) & \equiv & (p \vee q \equiv q)
\end{array}$$

$\equiv$ {the product is negative only if one of the factors is negative and the other factor is positive: $(ab < 0) \equiv (a < 0 \wedge b > 0) \vee (a > 0 \wedge b < 0)$}

$(x < 0 \wedge 2 + x > 0) \vee (x > 0 \wedge 2 + x < 0)$

$\equiv$ {simplify both disjuncts}

$(x < 0 \wedge x > -2) \vee (x > 0 \wedge x < -2)$

$\equiv$ {the right-hand disjunct is false for every value of x}

$(-2 < x < 0) \vee F$

$\equiv$ {$p \vee F \equiv p$}

$-2 < x < 0$

... $(x = 0 \vee x = -2) \vee (-2 < x < 0)$

$\equiv$ {combine the conditions}

$-2 \leq x \leq 0$

□

The answer to the task is thus

$$(1 + x)^2 \leq 1 \quad \equiv \quad -2 \leq x \leq 0$$

■

5.4 Equations with Negations

Negation has three fundamental properties: a proposition and its negation cannot simultaneously be true, a proposition or its negation is always true, and a double negation is equivalent to the original proposition. These rules are shown in Table 5.4.

$\vdash p \wedge \neg p \equiv F$	{*contradiction*}
$\vdash p \vee \neg p \equiv T$	{*law of excluded middle*}
$\vdash \neg(\neg p) \equiv p$	{*double negation*}
$\vdash \neg(p \wedge q) \equiv \neg p \vee \neg q$	{*De Morgan's law for* $\wedge$}
$\vdash \neg(p \vee q) \equiv \neg p \wedge \neg q$	{*De Morgan's law for* $\vee$}

Table 5.4: Rules for negation

Arguing with negation comes in handy, e. g., when we solve equations involving rational expressions. We use below the following principle for solving an equation involving rational expressions ($P(x)$ and $Q(x)$ are polynomials in x):

$$\frac{P(x)}{Q(x)} = 0 \equiv P(x) = 0 \wedge Q(x) \neq 0$$

Example 29. Solve the equation

$$\frac{(x + 1)(x - 2)}{x - 2} = 0$$

We calculate the solution as follows.

• $\dfrac{x^2 - x - 2}{x - 2} = 0$

≡ {we search for the zeros of the numerator for which the denominator is different from zero}

$x^2 - x - 2 = 0 \wedge x - 2 \neq 0$

≡ {factor the first expression}

$(x + 1)(x - 2) = 0 \wedge x - 2 \neq 0$

≡ {the zero-product property}

$(x + 1 = 0 \vee x - 2 = 0) \wedge x \neq 2$

≡ {solve the equations in the parenthesis}

$(x = -1 \vee x = 2) \wedge x \neq 2$

≡ {the distributive law: $(p \vee q) \wedge r \equiv (p \wedge q) \vee (p \wedge r)$}

$(x = -1 \wedge x \neq 2) \vee (x = 2 \wedge x \neq 2)$

≡ {contradiction: $p \wedge \neg p \equiv F$}

$(x = -1 \wedge x \neq 2) \vee F$

≡ {we can omit a false proposition from a disjunction: $p \vee F \equiv p$}

$x = -1 \wedge x \neq 2$

≡ {redundancy: $x = -1 \Rightarrow x \neq 2$, so the second proposition can be omitted}

$x = -1$

□ ■

One of the most useful rules for negation are the de Morgan's laws, which show how negation distributes over conjunction and disjunction. These are also shown in Table 5.4. Note that conjunction becomes disjunction and disjunction becomes conjunction when the negation is distributed.

Example 30. Find the domain of the rational expression

$$\frac{x - 1}{x^2 - 1}$$

• Find out when $\dfrac{x - 1}{x^2 - 1}$ is well-defined

$\Vdash$ $\dfrac{x - 1}{x^2 - 1}$ is well-defined

≡ {a rational expression is well-defined when the denominator is not 0}

$x^2 - 1 \neq 0$

$\equiv$ {rewrite with explicit negation}

$\neg(x^2 - 1 = 0)$

$\equiv$ {factor the expression in the parenthesis}

$\neg((x+1)(x-1) = 0)$

$\equiv$ {the zero-product property}

$\neg(x = -1 \vee x = 1)$

$\equiv$ {de Morgans law: $\neg(p \vee q) = (\neg p \wedge \neg q)$}

$\neg(x = -1) \wedge \neg(x = 1)$

$\equiv$ {write as inequalities}

$x \neq -1 \wedge x \neq 1$

□

Thus, the expression $\dfrac{x-1}{x^2-1}$ is well-defined exactly when $x \neq -1$ and $x \neq 1$. ■

5.5 Case Analysis

	$\vdash$	$p \equiv (q \wedge p) \vee (\neg q \wedge p)$	{*case analysis*}
$q_1 \vee \ldots \vee q_n$	$\vdash$	$p \equiv (q_1 \wedge p) \vee \ldots \vee (q_n \wedge p)$	{*general case analysis*}

Table 5.5: Case analysis rules

Another very useful logical rule is *case analysis.* This means that we prove that some property is true by considering all the possible different cases that can occur, one at a time. The rule for case analysis is shown in Table 5.5.

Case analysis is an example of a logical rule that can be derived from other, more basic logical rules. We have the following proof of this rule:

• p

$\equiv$ {conjunction with T}

$T \wedge p$

$\equiv$ {excluded middle, $T \equiv q \vee \neg q$}

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

$\equiv$ {distribution}

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

□

Solving equations involving absolute values, like in Example 27, is a good example of when one needs case analysis in equation solving.

5.6 Assignments

1. Solve the set of simultaneous equations $y = 2x - 3$ and $5x = -2y + 39$.
2. Solve the set of simultaneous equations $\frac{x}{\sqrt{3}} + \frac{y}{\sqrt{2}} = \frac{1}{\sqrt{2}}$ and $6\sqrt{2}y - 3\sqrt{2} = -2\sqrt{3}x$
3. Solve the inequality $(x+2)(x-3) < 0$.
4. Prove that $(p \vee q) \wedge (r \vee p) \wedge (\neg q \vee \neg r \vee p) \equiv p$
5. Prove that $(p \wedge q) \vee (\neg p \wedge q) \vee (p \wedge \neg q) \vee (\neg p \wedge \neg q) \equiv T$
6. Prove that $p \vee q \equiv (p \wedge q) \vee \neg(p \equiv q)$.
7. Prove that the statement
$$(p \wedge p \wedge q) \vee \neg(\neg p \vee \neg\neg q) \vee \neg(\neg p \Rightarrow \neg q) \vee (\neg p \wedge \neg q\ )$$
is equivalent to true.
8. Is it possible to choose a value of a in such a manner that the simultaneous equations $x - \frac{y}{2} = -4$ and $2x = y - a$ are contingent (i.e. neither identically false or identically true)?
9. Solve the set of simultaneous equations
$$5x + 3y + z = 15 \ \wedge \ 2x - 2y + 2z = 6 \ \wedge \ -x + 4y + 5z = 27$$
10. Alice needs a 25% solution of hydrochloric acid. She has a 15% solution and 30% solution available. How much of these should she use to make 12 liters of 25% solution?
11. Prove that $p \wedge (\neg q \Rightarrow \neg p) \equiv q \wedge p \wedge (\neg q \Rightarrow \neg p)$.
12. Prove that $(p \wedge (p \vee q)) \Rightarrow (p \vee q)$ is always valid.
13. A set consists of an even number of consecutive natural numbers. Prove that the average of the numbers is not a natural number.
14. Solve the equation $5x - 2(x-1) = 2$.
15. Solve the equation $x^2 + 5x - 24 = 0$.

16. Solve the equation $x^3 - 6\frac{1}{2}x^2 - 3\frac{1}{2}x = 0$.

17. Solve the simultaneous equations $y = 2x - 3$ and $5x = -2y + 39$.

18. Solve the absolute value equation $|2x - 8| = 16$.

19. Prove that the equation $5 \cdot (2 - x) - 9 = 6 \cdot (3 - x) - (16 - x)$ is not satisfied for any values of the variable x.

CHAPTER 6

Basic Structured Tasks

A structured calculation describes the solution to a mathematical problem. A *structured task* extends structured calculations: we write down both the problem that we are solving and the solution to the problem. The solution will typically involve some kind of structured calculation.

A structured task starts with a request, stating what we are supposed to do. This request is really a question that we are asked to answer. This is followed by a calculation to find an answer to the question. We rewrite our introductory example as a structured task.

Example 31. Calculate the value of the expression $3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2$.

• What is the value of expression $3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2$

$\Vdash$ {The answer follows by transitivity}

$3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2$

$=$ {calculate the powers}

$3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16$

$=$ {perform the multiplications}

$24 + 36 - 32$

$=$ {perform the addition}

$60 - 32$

$=$ {perform the subtraction}

28

□ The value is 28 ■

We write the question after the bullet, in the second column. We write the answer after the square, also in the second column. The justification after the "$\Vdash$" symbol explains why the answer is correct. Here we justify the answer with the fact that equality is transitive. The special symbols in a structured derivation can be given more intuitive names in the context of a task: "•" stands for *"task"*, "□" stands for *"answer"*, and "$\Vdash$" stands for *"conclusion"*.

We are often asked to solve a task under some given *assumptions*. The assumptions are listed after the question, writing "-" in the first column for each assumption. Alternatively, we can use numbers or lower case letters in parentheses, like in (a), (b), (c),. . . , to identify assumptions, so that we can refer to specific assumptions in the justifications. The question and the assumptions together form the *problem* that we are to solve. The following is an example of a task with assumptions.

Example 32. Calculate the value of the expression $3 \cdot x^3 + 4 \cdot y^2 - 2 \cdot 4^2$, when $x = 2$ and $y = 3$.

• What is the value of the expression $3 \cdot x^3 + 4 \cdot y^2 - 2 \cdot 4^2$, when

- $x = 2$, and

- $y = 3$

$\Vdash$ {The answer follows by transitivity}

$3 \cdot x^3 + 4 \cdot y^2 - 2 \cdot 4^2$

= {insert the values x and y from the assumptions}

$3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2$

= {calculate the powers}

$3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16$

= {perform the multiplications}

$24 + 36 - 32$

= {perform the addition}

$60 - 32$

= {perform the subtraction}

28

□ The value is 28

Here the conclusion is that the answer 28 follows from the calculation steps by transitivity. ■

The assumptions in a structured task may not always list all the assumptions that we actually make when solving a problem. Some assumptions may remain implicit, usually because they are taken for granted in the area of mathematics that we are working with. The value range of variables is often left implicit: x and y usually range over real numbers in high school mathematics, and i, j, and k stand for integer numbers.

One of the things that mathematicians really hate is to write down redundant information. They prefer brevity and elegance over precise details. The idea is that “an intelligent reader” can fill in the missing details. On the other hand, the basic idea of structured derivations is to make this implicit, hidden information explicit in mathematical arguments. There is a clear conflict between these two goals. We will solve this problem (to some degree) by allowing *default* information in a derivation: a derivation with missing information stands for a complete derivation where the missing information is provided by defaults.

The previous task is an example where we could be more brief. We could omit the answer after the “□” symbol, because it just repeats the last line of the calculation. We may also omit the justification for why this answer is correct, because it is the standard one, i.e., transitivity. In other words,

- the *default answer* is the last line of the calculation, and
- the *default justification* for the answer is transitivity.

We use these defaults in the sequel whenever they are applicable, to make the derivation more concise. Using these defaults in the example above gives us a slightly more compressed derivation.

Example 33. Solving same problem as above, but using defaults.

• What is the value of the expression $3 \cdot x^3 + 4 \cdot y^2 - 2 \cdot 4^2$, when

- $x = 2$, and

- $y = 3$

$\Vdash$ $3 \cdot x^3 + 4 \cdot y^2 - 2 \cdot 4^2$

$=$ {insert the values x and y from the assumptions}

$3 \cdot 2^3 + 4 \cdot 3^2 - 2 \cdot 4^2$

$=$ {calculate the powers}

$3 \cdot 8 + 4 \cdot 9 - 2 \cdot 16$

$=$ {perform the multiplications}

$24 + 36 - 32$

$=$ $\{$perform the addition$\}$

$60 - 32$

$=$ $\{$perform the subtraction$\}$

28

□

Note that we move the first line of the calculation up one step, to the place where the justification of the answer is written, to save space. ■

Example 34. Simplify $\cos(x + \frac{\pi}{3})$ when $\sin x = \cos x$.

• Simplify $\cos(x + \frac{\pi}{3})$, when

- $\sin x = \cos x$

$\Vdash$ $\cos(x + \frac{\pi}{3})$

$=$ $\{$the angle sum formula: $\cos(a + b) = \cos a \cdot \cos b - \sin a \cdot \sin b\}$

$\cos x \cdot \cos \frac{\pi}{3} - \sin x \cdot \sin \frac{\pi}{3}$

$=$ $\{$insert values: $\sin \frac{\pi}{3} = \frac{\sqrt{3}}{2}$ and $\cos \frac{\pi}{3} = \frac{1}{2}\}$

$\frac{1}{2} \cos x - \frac{\sqrt{3}}{2} \sin x$

$=$ $\{$by the assumption$\}$

$\frac{1}{2} \cos x - \frac{\sqrt{3}}{2} \cos x$

$=$ $\{$factor out $\cos x\}$

$\frac{1-\sqrt{3}}{2} \cos x$

□

The assumption actually implies that $\cos x = \pm\frac{1}{\sqrt{2}}$, so we could carry out the simplification even further, to get the answer $\pm\frac{1-\sqrt{3}}{2\sqrt{2}}$. We will later show how to add this kind of observations to solving tasks. ■

Example 35. Derive a formula for the derivative of the product of two functions f and g, when both f and g are differentiable.

The product of functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$, denote $fg : \mathbb{R} \to \mathbb{R}$, is defined by

$$(fg)(x) = f(x) \cdot g(x)$$

for every $x \in \mathbb{R}$.

• Calculate $(fg)'(x)$, when

(a) f and g are differentiable

(b) $(fg)(x) = f(x) \cdot g(x)$, for every $x \in \mathbb{R}$

⊩ $(fg)'(x)$

= {the definition of derivative}

$\lim_{h\to 0} \frac{(fg)(x+h)-(fg)(x)}{h}$

= {the definition of product function}

$\lim_{h\to 0} \frac{f(x+h)g(x+h)-f(x)g(x)}{h}$

= {we can add $f(x)g(x+h) - g(x+h)f(x)$ to the numerator, since the value of this expression is 0}

$\lim_{h\to 0} \frac{f(x+h)g(x+h)-f(x)g(x)+f(x)g(x+h)-g(x+h)f(x)}{h}$

= {regroup the numerator}

$\lim_{h\to 0} \frac{f(x+h)g(x+h)-g(x+h)f(x)+f(x)g(x+h)-f(x)g(x)}{h}$

= {the first two terms have the common factor $g(x+h)$ and the last two terms have the common factor $f(x)$}

$\lim_{h\to 0} \frac{g(x+h)(f(x+h)-f(x))+f(x)(g(x+h)-g(x))}{h}$

= {we split the expression into two separate sums}

$\lim_{h\to 0} \left(\frac{g(x+h)(f(x+h)-f(x))}{h} + \frac{f(x)(g(x+h)-g(x))}{h}\right)$

= {we calculate the limits separately for the two terms}

$\lim_{h\to 0} \frac{g(x+h)(f(x+h)-f(x))}{h} + \lim_{h\to 0} \frac{f(x)(g(x+h)-g(x))}{h}$

= {rewrite the expressions}

$\lim_{h\to 0} g(x+h)\frac{f(x+h)-f(x)}{h} + \lim_{h\to 0} f(x)\frac{g(x+h)-g(x)}{h}$

= {when $h \to 0$, the fractional expressions approach $f'(x)$ and $g'(x)$, and $g(x+h) \to g(x)$}

$g(x)f'(x) + f(x)g'(x)$

□ ■

6.1 A More Verbal Format

Structured tasks contain special symbols that identify the different parts of the task: “•”, “⊩”, “–”, “{”, “}”, and “□”. Formalizing the notation for tasks is similar to what has happened in mathematics in general: the “+” symbol was introduced as a shorthand for writing “the sum of... ” and the “=” symbol was introduced as a shorthand for saying that two expressions have the same value. Introducing special symbols for important mathematical concepts shortens the writing, but also paves the way for a universal mathematical language and provides an unambiguous interpretation of these symbols. In our case, this means that a task looks the same in every language and that a task has an exact logical meaning.

When teaching structured derivations for the first time, it may be easier for students to understand a structured task by first using traditional words instead of symbols.

Example 36. We can use a more verbose notation for structured tasks, e.g., like this:

Task:	Check whether the statement $(1+a)(1+b)(1+c) \geq 1+a+b+c$ is always true
Assumption:	$a, b, c \geq 0$
Conclusion:	{The answer follows from the calculation below}
	$(1+a)(1+b)(1+c)$
$=$	{expand the last two parentheses}
	$(1+a)(1+b+c+bc)$
$=$	{expand the remaining two parentheses}
	$1+b+c+bc+a+ab+ac+abc$
$\geq$	{$ab+ac+bc+abc$ is non-negative, since $a, b, c \geq 0$, according to the assumption}
	$1+a+b+c$
Answer:	The statement is always true ■

The verbal form is more intuitive, but at the same time it gives the impression that the task is an informal description of a problem and its solution, and that it is open to different interpretations. This, however, is not the case, a structured derivation has a logical meaning that is as exact as, e.g., the value of an arithmetic expression. We will stick to the more concise symbolic notation for tasks in the rest of the book, but teachers are free to use the verbal form when they feel that it makes it easier to understand the task and its solution. Table 6.1 shows the two presentation formats side by side. The only difference is that the symbols have been replaced by keywords.

We can also have an intermediate form, where the main steps of the calculation are described in more verbose notation, while the details are described symbolically. We will give some examples of this later on.

•	Check whether the statement		**Task**: Check whether the statement
	$(1+a)(1+b)(1+c) \geq 1+a+b+c$		$(1+a)(1+b)(1+c) \geq 1+a+b+c$
	is always true, when		is always true
–	$a, b, c \geq 0$		**Assumption**: $a, b, c \geq 0$
⊩	{The answer follows from the calculation below}		**Conclusion**: {The answer follows from the calculation below}
	$(1+a)(1+b)(1+c)$		$(1+a)(1+b)(1+c)$
$=$	{expand the last two parentheses}	$=$	{expand the last two parentheses}
	$(1+a)(1+b+c+bc)$		$(1+a)(1+b+c+bc)$
$=$	{expand the remaining two parentheses}	$=$	{expand the remaining two parentheses}
	$1+b+c+bc+a+ab+ac+abc$		$1+b+c+bc+a+ab+ac+abc$
$\geq$	{$ab+ac+bc+abc$ is non-negative, since $a, b, c \geq 0$, according to the assumption}	$\geq$	{$ab+ac+bc+abc$ is non-negative, since $a, b, c \geq 0$, according to the assumption}
	$1+a+b+c$		$1+a+b+c$
□	The statement is always true		**Answer**: The statement is always true

Table 6.1: Symbolic and verbose formats

6.2 Questions and Answers

A task starts with a question: what values will satisfy some given condition. Examples of questions are:

> What values of the variable x satisfy the condition $x^2 + 2x + 1 = 0$ (solving an equation)?
>
> What values of the variables x and y satisfy the conditions that $2x+y = 3$ and $3x - y = 4$ (simultaneous equations)?
>
> What is a simpler form for the expression $s = \dfrac{x^2 - 1}{x + 1}$ (simplification)?
>
> What is the value of $a = \sin(2\pi)^2$ (calculating a value)?

Tasks are usually not phrased as questions but rather as requests: "Solve the equation…", "Simplify the expression…", "Calculate the value of…". But behind a request is a question that we want to answer, i.e., the request is to find the answer to a given question. The answer is then a logical proposition that shows which values satisfy the conditions stated in the question.

There are two kinds of basic questions that we may ask in a task. Sometimes we want to find *all* values that satisfy a given condition (e.g., when solving an equation), while in other cases we are looking for *some* value that satisfies the condition (e.g., in simplification). It is also important to specify the domain where we search for acceptable values. We also need to state whether we are looking for values for a single variable, or whether we are looking for values for two or more variables at the same time.

The general form for a *some-question* is

$$?\, x_1 \in U_1, \ldots, x_m \in U_m :\; Q(x_1, \ldots, x_m)$$

and for an *all-question*

$$!\, x_1 \in U_1, \ldots, x_m \in U_m :\; Q(x_1, \ldots, x_m)$$

Here $x_1, \ldots, x_m$ are the variables for which we want to find suitable values, and $U_1, \ldots, U_m$ are the value domains for these variables. This means that we are looking for values $x_1 \in U_1$, $x_2 \in U_2$ and so on. A list of the form $x_1 \in U_1, \ldots, x_m \in U$ is known as a *declaration.* The purpose of a declaration is to introduce new names for variables (and constants) and associate a value domain with each of these new names. The logical proposition $Q(x_1, \ldots, x_m)$ describes the conditions that the values for $x_1, \ldots, x_m$ must satisfy.

The answer to a some-question will be of the form $(x_1, \ldots, x_m) = (t_1, \ldots, t_m)$. The answer to an all-question will be another logical proposition $R(x_1, \ldots, x_m)$, from which it is easy to see which values satisfy the original proposition.

Example 37. Some example questions:

1. Solve the equation $x^2 + 2x + 1 = 0$:

   $$!\, x \in \mathbb{R} :\; x^2 + 2x + 1 = 0$$

   In other words, find **all** values $x \in \mathbb{R}$ that satisfy the equation $x^2 + 2x + 1 = 0$. The answer could be of the form "$x = e_1 \vee x = e_2$" .

2. Solve the equation pair

   $$\begin{cases} 2x + y = 3 \\ 3x - y = 4 \end{cases}$$

   This can be expressed as the question

   $$!\, x \in \mathbb{R},\; y \in \mathbb{R} :\; 2x + y = 3 \,\wedge\, 3x - y = 4$$

   In other words, we want to find **all** value combinations (x, y) that satisfy the two equations, $2x + y = 3$ and $3x - y = 4$. The answer could be of the form "$x = e_1 \wedge y = e_2$".

3. Simplify the expression $\dfrac{x^2 - 1}{x + 1}$. This can be expressed as the question

   $$?\, s \in \mathbb{R} :\; s = \frac{x^2 - 1}{x + 1}$$

Here we want to find some value (or expression) s that is equal to the original expression, but simpler in some way. What it means to be simpler depends on the context, and the rules are not always that explicit. The answer would be of the form "$s = e$".

4. Solve the equation $\frac{1}{2} = \sin x^2$. This can be expressed as the question

$$!\, x \in \mathbb{R} : \frac{1}{2} = \sin x^2$$

The equation has an infinite number of solutions x, because the sine function is periodic. We are looking for an expression that characterizes all these solutions. The answer would then be a logical proposition that shows all possible values of x that satisfy the equation.

5. Check whether $x^2 \geq x$ is always true. This can be expressed as the question

$$?p \in \mathbb{B} : p \equiv x^2 \geq x$$

Here we are looking for a simplification p of the logical expression $x^2 \geq x$. If we can simplify the expression to T, i.e., $p \equiv T$, then the logical expression is always true. If we can simplify the expression to F, i.e., $p \equiv F$, then the logical expression is always false. Otherwise, the logical expression will be true for some values of x and false for other values.

6. Prove that $(x + y)^2 = x^2 + 2xy + y^2$. Here we are not looking for any values for variables, we just want to prove that the equation holds for any values of x and y. This is similar to the previous question, i.e.,

$$?p \in \mathbb{B} : p \equiv (x + y)^2 = x^2 + 2xy + y^2$$

except that also have filled in the answer, $p \equiv T$. Our task is then to add a proof that shows that this answer is correct. We will usually write this as a question with an empty variable list,

$$?\ (x + y)^2 = x^2 + 2xy + y^2$$

■

We often place additional constraints on what answers are acceptable. A solution to an equation may, e.g., only allow conjunctions and disjunctions of propositions where x occurs only on the left-hand side of an equality. The rules for what is an acceptable answers to a question are often not very explicit, and depend on what we need the answer for. Simplification of an expression is a good example of this. If the task is to simplify the expression $10 \cdot \sqrt{\frac{12}{5}}$, then the answer we are looking for is probably $4 \cdot \sqrt{15}$. This is the simplest form of the original expression, because it has no fraction and the expression under the square root cannot be simplified further. The answer is correct, because $10 \cdot \sqrt{\frac{12}{5}} = 4 \cdot \sqrt{15}$. However, from a purely logical point of view, the answer $10 \cdot \sqrt{\frac{12}{5}}$ is equally correct, because $10 \cdot \sqrt{\frac{12}{5}} = 10 \cdot \sqrt{\frac{12}{5}}$

is also true. In other words, even if an answer is correct in the sense that we have defined above, it may still not be the answer we are looking for, because we place some additional (extra-logical) requirements on the answer we want to have.

We will in the sequel follow standard mathematical practice and express the task as a request rather than as a question. However, it is important to understand that the task is really a question, and that the purpose of the task is to find an answer to this question. With a little experience, it becomes quite easy to see the implied question in a request, and thus what kind of answer we should be looking for.

6.3 Proof Tasks

A *proof task* is the simplest form of a task, exemplified in case 6 of Example 37. In this case, we do not need to determine any variable values, we only want to prove that a given proposition is true.

Example 38. Example 36 shows a proof task. The task is to prove a theorem:

- Show that $(1+a)(1+b)(1+c) \geq 1+a+b+c$, when
  - a, b, $c \geq 0$

The solution (the proof) is given below:

$\Vdash$ {Transitivity of equality}

$(1+a)(1+b)(1+c)$

$=$ {expand the last two parentheses}

$(1+a)(1+b+c+bc)$

$=$ {expand the remaining two parentheses}

$1+b+c+bc+a+ab+ac+abc$

$\geq$ {$ab + ac + bc + abc$ is non-negative, since a, b, and c are all non-negative, according to the assumption}

$1+a+b+c$

□ ■

We can interpret the answer square here as the completion of the proof ("Quad erat demonstrandum", "Which had to be proven", abbreviated Q.E.D.)

6.4 Calculation Tasks

We described structured calculations in Chapter 2. A structured calculation (a *calculation task*) is a special case of a task where only the calculation of the task is written out explicitly. The the question, the answer and the justification for the answer are left implicit, and are given by defaults, as explained earlier. The default answer is the last line of the calculation, and the default justification is transitivity. The default question is to find a simplified version of the expression on the first line of the derivation. There are no explicit assumptions in a calculation task. This means that the possible assumptions we need must be clear from the context of the task. In spite of this, calculation tasks are very useful. Our previous trigonometry example, written as a calculation task, is shown below.

Example 39. Simplify $\cos(x + \frac{\pi}{3})$ when $\sin x = \cos x$. We solve the task directly by calculating:

$\bullet$ $\quad \cos(x + \frac{\pi}{3})$

$=$ $\quad \{\text{the angle sum formula: } \cos(a+b) = \cos a \cdot \cos b - \sin a \cdot \sin b\}$

$\quad \cos x \cos \frac{\pi}{3} - \sin x \sin \frac{\pi}{3}$

$\vdots$

$=$ $\quad \{\text{factor out } \cos x\}$

$\quad \frac{1-\sqrt{3}}{2} \cos x$

□

This calculation is carried out in a context where $\sin x = \cos x$ is known to hold. The default task is "simplify the expression on the first line of the calculation, i.e., $\cos(x + \frac{\pi}{3})$". The default answer is given by the expressions on the last line, i.e., $\frac{1-\sqrt{3}}{2} \cos x$. The default justification is transitivity. ■

6.5 Assignments

1. Solve the equation $(x - 5) - x = \frac{9}{x}$.
2. Solve the equation $\frac{x+7}{x} - 9 = \frac{x-3}{2x}$.
3. Solve the equation $(x + 1)^2 - 3 = \frac{3x^2}{x+1}$.
4. Solve the equation $x \ln(x) - x = 0$, $x > 0$.
5. Is $x = -\frac{5}{\sqrt{2}}$ a solution to the equation $\sqrt{3}x - \sqrt{2}x = 5 + 9^{\frac{1}{4}}x$?

6. Prove that the value of the expression $\dfrac{\log\left(b^2\right) - \log\left(\sqrt[5]{b^2}\right)}{\log\left(\sqrt{b}\right) - \log\left(b^{\frac{1}{3}}\right)}$ is independent of the value of the parameter b, where $b > 0$ and $b \neq 1$.

7. Determine the limit $\lim\limits_{x\to 3} \dfrac{\ln\left(x^3\right) - \ln(27)}{\ln(x) - \ln(3)}$.

8. Prove that the sequence $a_n = \dfrac{n!}{n^n}$ is strictly decreasing, when $n = 1, 2, 3, \ldots$

9. Define the *double factorial* for even numbers as $(2n)!! = 2n \cdot (2n-2) \cdot \cdots \cdot 2,$, when $n = 1, 2, 3, ...$ and $0!! = 1$. Rewrite the double factorial $(2n)!!$ in terms of the regular factorial $n!$.

Chapter 7

Proofs as Logical Calculations

The previous chapter has demonstrated how to solve an equation by simplifying it step by step to another, logically equivalent proposition, which directly shows the values that satisfy the equation. This is in fact a general method for determining the truth value of an arbitrary logical proposition, and for calculating the values of the variables that make the proposition true.

$\vdash$	$p \equiv (p \equiv T)$	$\{\equiv$ *and truth*$\}$
$\vdash$	$\neg p \equiv (p \equiv F)$	$\{\equiv$ *and falsity*$\}$

Table 7.1: Truth rules for equivalence

A logical proposition p is true if it has the value T for every value of the variables in p. This gives us simple and straightforward way to prove that a logical proposition p is true: we show that $p \equiv T$ holds (Table 7.1). We do it by a calculation of the form

$$
\begin{array}{ll}
\bullet & p \\
\equiv & \{\text{justification}_1\} \\
 & p_1 \\
\equiv & \{\text{justification}_2\} \\
 & p_2 \\
\vdots & \\
 & p_{n-1} \\
\equiv & \{\text{justification}_n\} \\
 & T \\
\Box &
\end{array}
$$

Transitivity then gives us that $p \equiv T$, i.e. that p is true.

Example 40. Assume that the vectors $\bar{a}$ and $\bar{b}$ have the same length. Show that this means that the vectors $\bar{a}+\bar{b}$ and $\bar{a}-\bar{b}$ are orthogonal.

• Show that $\bar{a}+\bar{b}$ and $\bar{a}-\bar{b}$ are orthogonal, when

- $|\bar{a}| = |\bar{b}|$

$\Vdash$ {We show that the proposition holds by showing that it is equivalent to T}

the vectors $\bar{a}+\bar{b}$ and $\bar{a}-\bar{b}$ are orthogonal

$\equiv$ {two vectors are orthogonal when their dot product is zero}

$(\bar{a}+\bar{b}) \cdot (\bar{a}-\bar{b}) = 0$

$\equiv$ {the sum and product rules for vectors}

$\bar{a} \cdot \bar{a} - \bar{b} \cdot \bar{b} = 0$

$\equiv$ {the relation between dot products and magnitudes for vectors, $\bar{a} \cdot \bar{a} = |\bar{a}|^2$}

$|\bar{a}|^2 - |\bar{b}|^2 = 0$

$\equiv$ {assumption}

$0 = 0$

$\equiv$ {always true}

T

□ ■

We can use the same method to show that a logical proposition is not true. The easiest way to disprove a proposition is to give a counterexample. Often we cannot directly identify a counterexample, but we can find one with a little bit of calculation.

Example 41. Check whether the proposition $x < x^2$ is true.

• $x < x^2$

$\equiv$ {subtract x^2 from both sides}

$x - x^2 < 0$

$\equiv$ {the distributive law for arithmetic expressions: $ab + ac = a(b + c)$}

$x(1 - x) < 0$

$\equiv$ {a product is negative if, and only if, one factor is negative and the other is positive: $(ab < 0) \equiv (a < 0 \ \wedge \ b > 0) \vee (a > 0 \ \wedge \ b < 0)$}

$(x < 0 \ \wedge \ 1 - x > 0) \vee (x > 0 \ \wedge \ 1 - x < 0)$

$\equiv$ {solve the inequalities in the disjuncts}

$(x < 0 \ \wedge x < 1) \vee (x > 0 \ \wedge x > 1)$

$\equiv$ $\{x < 0 \Rightarrow x < 1$ and $x > 1 \Rightarrow x > 0$, omit redundant propositions from the conjunctions}

$x < 0 \vee x > 1$

$\not\equiv$ {the proposition is false for e.g. $x = 0.5$}

T

□ ■

This shows that the proposition is not true for every value of x. Note that in the last step we do not have $x < 0 \vee 1 < x \equiv F$ either, i.e., the proposition is also not false for every value of x. Instead we have a proposition that is true for some values of x and false for other values. The second to last line of the same derivation gives the exact condition for when the proposition is true:

$$x < x^2 \quad \equiv \quad x < 0 \vee x > 1$$

This is the strength of proofs by logical manipulations. We calculate the value of a logical proposition using equivalences, which leads to the exact condition for when a proposition is true. This is the same situation as we have encountered earlier, when solving first-degree equations.

Example 42. When is $\int_a^{a+1}(2x+3)\ dx \geq \frac{1}{2}$. We simplify the logical proposition so that we get a simple condition for a that is equivalent to the the original condition.

• Determine the values of a for which $\int_a^{a+1}(2x+3)\ dx \geq \dfrac{1}{2}$.

$\Vdash$ $\int_a^{a+1}(2x+3)\ dx \geq \dfrac{1}{2}$

$\equiv$ {the integration formulas $\int_a^b cx^n\ dx = \left[\dfrac{cx^{n+1}}{n+1}\right]_a^b$ and $\int (f(x) + g(x))\, dx = \int f(x)\, dx + \int g(x)\, dx$}

$[x^2 + 3x]_a^{a+1} \geq \dfrac{1}{2}$

$\equiv$ $\{[F(x)]_a^b = F(b) - F(a)\}$

$\big((a+1)^2 + 3(a+1)\big) - (a^2 + 3a) \geq \dfrac{1}{2}$

$\equiv$ {calculate the squares and expand the parentheses}

$$a^2 + 2a + 1 + 3a + 3 - a^2 - 3a \geq \frac{1}{2}$$

$\equiv$ {add together similar terms}

$$2a + 4 \geq \frac{1}{2}$$

$\equiv$ {solve for a}

$$a \geq -\frac{7}{4}$$

□

We have shown that the original inequality holds exactly when $a \geq -\frac{7}{4}$. ■

Solving equations is thus a special case of a more general method for solving problems, where we simplify a logical proposition to find out for which values of the variables the proposition is true. Each simplification step leads to a new logical proposition that is equivalent to the previous proposition. We continue until we have the proposition in a form that we are satisfied with. This new form of the proposition is equivalent to the original proposition by transitivity.

The truth value of a logical proposition $p(x)$ depends on the value of the variable x. We have three basic cases:

- The proposition is always false: $p(x) \equiv F$, i.e., there is no value of x for which the proposition is true. The set of solutions is thus empty, i.e., $\{x|p(x)\} = \emptyset$.
- The proposition is true for some values of x and false for other values: $p(x) \not\equiv F$ and $p(x) \not\equiv T$. The set of solutions is not empty, but it is not full either, i.e., $\{x|p(x)\} \neq \emptyset$ and $\{x|p(x)\} \neq \mathbb{R}$.
- The proposition is always true: $p(x) \equiv T$, i.e., the proposition is true for every value of x. The set of solutions is full, i.e., $\{x|p(x)\} = \mathbb{R}$.

The advantage of this technique is that we do not need to decide in advance whether we want to prove that the proposition is true or that it is false. We simplify the logical proposition as much as possible with a sequence of equivalences. If we can simplify the proposition to T, then it is always true, if we can simplify the proposition to F, it is always false. Otherwise, we have derived an equivalent logical expression that shows exactly when the proposition is true.

7.1 Implication

We have described the general rules of logical equivalence, conjunction, disjunction, and negation above. These rules are very useful when we want to simplify and manipulate logical propositions. We will now treat the last of the core logical connectives, implication, and give the general rules for manipulating expressions with implication.

Logical consequence is basically the same as implication. This means that we can write the rule for the square root of a square for real numbers as

$$a \in \mathbb{R},\, a \geq 0 \vdash \sqrt{a^2} = a$$

or as

$$\vdash a \in \mathbb{R} \wedge a \geq 0 \Rightarrow \sqrt{a^2} = a$$

Generally

$$A_1, A_2, \ldots, A_m \vdash C$$

and

$$A_1, A_2, \ldots, A_k \,\vdash A_{k+1} \wedge \;\ldots \wedge A_m \Rightarrow C$$

have the same meaning. In practice it is often important to make a clear distinction between what a theorem proposes and the assumptions that are made in the theorem, so we usually prefer the first form.

Implication is a partial order relation, i.e., it is *reflexive, transitive* and *antisymmetric.* These rules are shown in Table 7.2. Equivalence is symmetric but implication is antisymmetric*:* if $p \Rightarrow q$ and $q \Rightarrow p$, then $p \equiv q$. This is similar to the order relation between real numbers: if $a \leq b$ and $b \leq a$, then $a = b$. Equivalence can, in fact, be described as mutual implication, as shown in the table. This explains why the notation $p \Leftrightarrow q$ (a shorthand for $(p \Rightarrow q) \wedge (p \Leftarrow q)$) is often used instead of $p \equiv q$.

	$\vdash$	$p \Rightarrow p$	$\{\Rightarrow$ *is reflexive*$\}$
$p \Rightarrow q,\, q \Rightarrow r$	$\vdash$	$p \Rightarrow r$	$\{\Rightarrow$ *is transitive*$\}$
$p \Rightarrow q,\, q \Rightarrow p$	$\vdash$	$p \equiv q$	$\{\Rightarrow$ *is antisymmetric*$\}$
	$\vdash$	$p \equiv (T \Rightarrow p)$	$\{$*prove* $p\}$
	$\vdash$	$\neg p \equiv (p \Rightarrow F)$	$\{$*prove* $\neg p\}$
	$\vdash$	$(p \equiv q) \equiv (p \Rightarrow q) \wedge (q \Rightarrow p)$	$\{\equiv$ *is mutual implication*$\}$
	$\vdash$	$(p \Rightarrow q) \equiv \neg p \vee q$	$\{\Rightarrow$ *is material implication*$\}$

Table 7.2: Implication rules

We have shown earlier that we can prove a logical proposition p by showing that $p \equiv T$. In fact, it is sufficient to show that

$$T \Rightarrow p$$

This will establish that p is true (Table 7.2). We can easily see this using the rules we have given earlier:

• $\quad p \equiv T$

$\equiv \quad$ {equivalence is mutual implication}

$\quad (p \Rightarrow T) \wedge (T \Rightarrow p)$

$\equiv$ {every proposition implies T, i.e. $(p \Rightarrow T) \equiv T$}

$T \wedge (T \Rightarrow p)$

$\equiv$ {we can omit T from a conjunction}

$T \Rightarrow p$

□ ■

The following example shows a typical situation where implication is sufficient to prove a theorem.

Example 43. Show that $pab < a^2 + b^2$, when $a > 0$, $b > 0$ and $0 < p < 2$.

• Show that $pab < a^2 + b^2$, when

- $a > 0$, $b > 0$,

- $0 < p < 2$

$\Vdash$ {We prove the proposition p by showing that $T \Rightarrow p$}

T

$\equiv$ {a square is always non-negative}

$0 \leq (a - b)^2$

$\equiv$ {calculate the square}

$0 \leq a^2 - 2ab + b^2$

$\equiv$ {add $2ab$ to both sides}

$2ab \leq a^2 - 2ab + b^2 + 2ab$

$\equiv$ {simplify}

$2ab \leq a^2 + b^2$

$\Rightarrow$ {$pab < 2ab$ according to the assumption}

$pab < a^2 + b^2$

□ ■

Correspondingly, we can show that a proposition is always false by showing that $p \Rightarrow F$.

In many situations, we do not know whether the proposition p is true, our task is to find out whether the proposition is true or false. We then have two basic strategies to choose from:

- We can first try to prove that p is true, i.e., show that $T \Rightarrow p$. If we succeed in this, we know that p is true. I we fail, we can try to prove that p is false. The easiest way to do this is to find values of the variables in p that make p false. If we succeed in this, we know that the proposition is false. If we do not find values that refute the proposition but we also fail to prove the proposition, we do not know anything about its truth value.

- Alternatively, we can try to simplify p into another equivalent proposition. If we can show that $p \equiv T$, then we know that the proposition is true, and if we can show that $p \not\equiv T$, then we know that the proposition is false. If neither approach is successful, then we still have a characterization of when p is true.

We often have a situation where establishing implication between logical expressions is enough. We can compare implications to order relations: we can prove that $a \leq b$, but we might not be able to prove that $a = b$. The same is true in set theory, proving that $A \subseteq B$ holds may suffice when we cannot prove that $A = B$, for two sets A and B.

A typical example is calculating the value of a function f in a given point x_0. For the function $f(x) = \sqrt{x+1}$ we have $f(3) = 2$, i.e.

$$f(x) = \sqrt{x+1} \text{ for every } x \in \mathbb{R} \Rightarrow f(3) = 2$$

The implication does not, however, hold in the other direction, $f(x) = \sqrt{x+1}$ is not the only function f for which $f(3) = 2$ (e.g. the function $f(x) = x - 1$ gives the same result for $x = 3$).

Example 44. Calculate the derivative of the function $f(x) = \sqrt{x^2+1}$ in the point 1.

• Calculate $f'(1)$, when

- $f(x) = \sqrt{x^2+1}$ for $x \in \mathbb{R}$

$\Vdash$ {We start the calculation from the definition of the function}

$f(x) = \sqrt{x^2+1}$ for $x \in \mathbb{R}$

$\Rightarrow$ {differentiate the expression using the rules $\dfrac{d}{dx}x^n = nx^{n\text{-}1}$ and the chain rule $\dfrac{d}{dx}f^n = nf^{n'1} \cdot \dfrac{d}{dx}f$}

$f'(x) = \dfrac{1}{2\sqrt{x^2+1}} \cdot 2x$ for $x \in \mathbb{R}$

$\equiv$ {simplify}

$f'(x) = \dfrac{x}{\sqrt{x^2+1}}$ for $x \in \mathbb{R}$

$\Rightarrow$ {insert $x = 1$}

$$f'(1) = \frac{1}{\sqrt{1^2+1}}$$

$\equiv$ {calculate the result}

$$f'(1) = \frac{1}{\sqrt{2}}$$

□ ■

We use implication twice here. In both cases the implication holds in one direction, but not in the other direction. The first implication

$$f(x) = \sqrt{x^2+1} \text{ for } x \in \mathbb{R} \;\Rightarrow f'(x) = \frac{1}{2\sqrt{x^2+1}} \cdot 2x, \text{ for } x \in \mathbb{R}$$

holds, but implication in the other direction does not hold, so the first two terms are not equivalent. This is because there are many functions with the same derivative (they differ by a constant term). The second implication calculates the value of the function in a particular point. Since there are many different functions with the same value in the point 1, the reverse implication does not hold here either.

$\vdash$	$p \wedge q \Rightarrow p$	{*conjunction is stronger*}
$\vdash$	$p \Rightarrow p \vee q$	{*disjunction is weaker*}
$\vdash$	$p \Rightarrow T$	{*T is weakest*}
$\vdash$	$F \Rightarrow p$	{*F is strongest*}

Table 7.3: Strengthening and weakening

As we already explained, p in an implication $p \Rightarrow q$ is the *stronger* proposition and q is the *weaker* proposition, since q follows from p. A proposition becomes stronger if we add another proposition with conjunction, while the proposition becomes weaker if we add another proposition with disjunction. Table 7.2 shows these two rules, together with two other basic properties of implication: F implies any proposition p, and every proposition p implies T.

7.2 Solving Equations with Implications

We use equivalence to describe the relationship between two equations. This gives us every solution to an equation. We could also use implication between steps when we solve equations. Implication is transitive, so a calculation would then show that the first equation implies the last equation in the calculation. However, using implication may introduce false solutions, as the following example shows.

Example 45. Consider the equation

$$\sqrt{2x^2+1} = 2x + 1$$

We use the following rule when we solve the equation:

$$\sqrt{a} = b \Rightarrow a = b^2, \text{ when } a \geq 0$$

This rule does not hold in the other direction, i.e. $a = b^2 \not\Rightarrow \sqrt{a} = b$ (check with e.g. the values $a = 4$ and $b = -2$).

We can solve this equation in the following way:

$\bullet$ $\sqrt{2x^2+1} = 2x+1$

$\Rightarrow$ {square both sides}

$2x^2+1 = 4x^2+4x+1$

$\equiv$ {transfer terms, simplify}

$2x^2+4x = 0$

$\equiv$ {factor}

$x(2x+4) = 0$

$\equiv$ {the zero-product property}

$x = 0 \vee x = -2$

□

This proves that

$$\sqrt{2x^2+1} = 2x+1 \Rightarrow x = 0 \vee x = -2$$

If both $x = 0$ and $x = -2$ are solutions to the equation, we should have implication in the other direction as well. A calculation shows that $x = 0$ is a solution to the equation: $\sqrt{2\cdot 0^2+1} = 2\cdot 0+1$. However, $\sqrt{2\cdot(-2)^2+1} \neq 2\cdot(-2)+1$, so $x = -2$ is not a solution. Thus, we have proven that

$$\sqrt{2x^2+1} = 2x+1 \equiv x = 0$$

■

The fact that using implication when solving equations may introduce false solutions is easily seen, by considering an arbitrary equation. Consider, e.g., the equation $x+1 = 2$. From the rules for implication, we know that T is the weakest proposition, i.e., $x+1 = 2 \Rightarrow T$. Any value for x satisfies the proposition T trivially. However, all value except $x = 1$ are false solutions.

The advantage of using equivalence when solving an equation, compared to implication, is that we immediately find every solution to the equation with no false solutions. Implication is, however, important when we give approximate value for the solutions.

Example 46. A sum of money that is deposited in a bank account accrues compound interest so that after 10 years it has grown by a factor 1.5. What is the annual interest rate as a percentage? The withholding tax is not considered. Give the answer with an accuracy of one hundredth of a percent.

$\bullet$ Calculate the annual interest p as a percentage, when

- the initial amount of money is b, and

- the amount grows to $1.5b$ in 10 years

$\Vdash$ T

$\equiv$ {use the rule $a_n = a_1 \cdot q^{n-1}$, where a_n is the value of the deposit after n years, a_1 is the initial capital, and q is the growth rate}

$1.5b = bq^{10-1}$

$\equiv$ {divide both sides by b, switch the sides}

$q^9 = 1.5$

$\equiv$ {solve for q}

$q = \sqrt[9]{1.5}$

$\equiv$ {the growth rate q is 1 + the annual interest rate}

$p + 1 = \sqrt[9]{1.5}$

$\equiv$ {solve for p}

$p = \sqrt[9]{1.5} - 1$

$\Rightarrow$ {calculate and round}

$p \approx 0.0413797$

$\equiv$ {convert to a percentage}

$p \approx 4.13797\,\%$

$\Rightarrow$ {round to hundredths of a percent}

$p \approx 4.14\,\%$

□

In this case, we use implication when we go from exact calculations to approximate values. The implication does not apply in the other direction, since there are many exact values with the same approximate value, and the same applies to rounding.

In this case, we could also argue that there is a difference between the mathematical answer and the solution to the original problem. We could give the answer $p = \sqrt[9]{1.5} - 1$ to the mathematical question posed in the task. Then we interpret the answer by calculating an approximate value of p, $p \approx 4.14\,\%$, and give this as a solution to the original problem. ■

7.3 The Isis Problem

We finally show a slightly more demanding example of solving equations with logical manipulations. We consider an ancient Egyptian problem, the so-called *Isis problem*. The problem is probably at least 3 000 years old. The original solution was verbal, but we show here how to formulate and solve the task using logical manipulations.

We will allow mixing lines from a structured task with text that explain how we solve the task. This is useful for longer tasks with more complex solutions. The informal text can explain the strategy we use for solving the assignment, or the intuition behind the observations, or show us how to interpret the answer. In addition to text, we can also have figures, illustrations, tables, graphs, etc. Mixing prose and displayed formulas is standard in mathematical for longer arguments, and is also useful for structured derivations.

Example 47. (*The Isis problem)* Assume that we have a rectangular field, with integer side lengths. The question is how many different fields have an area that is equal to its perimeter (i.e., the numerical value of the area is equal to the numerical value of the perimeter).

We start by specifying the problem and the assumptions we can make. We exclude the trivial solution where both sides of the field have the length 0 ($0 \cdot 0 = 2 \cdot 0 + 2 \cdot 0$), and observe that we have no solution if one side is 0 and the other side is greater than 0 ($0 \cdot m = 0 \neq 2 \cdot 0 + 2 \cdot m$, when $m > 0$). We are left with the condition that both sides of the field must be positive natural numbers. The problem is then the following:

- • For which different values of m and n is the area mn equal to the perimeter $2m + 2n$, when
- - m and n are positive integers.

We start from the condition that m and n must satisfy, and try to simplify the condition to a form where we can directly see all values for m and n that satisfy the condition. We have the following calculation:

$\Vdash$ $mn = 2m + 2n$

$\equiv$ {regroup}

$mn - 2m = 2n$

$\equiv$ {factor out m}

$m(n - 2) = 2n$

We now want to express m as a function of n. This is easy, provided that we do not divide by 0. We therefore study the two cases $n - 2 = 0$ and $n - 2 \neq 0$ separately, using case analysis:

$m(n-2) = 2n$

$\equiv$ {analyze the two cases $n = 2$ and $n \neq 2$ separately; the logical rule for case analysis $p \equiv (q \wedge p) \vee (\neg q \wedge p)$}

$(n = 2 \wedge m(n-2) = 2n) \vee (n \neq 2 \wedge m(n-2) = 2n)$

$\equiv$ {check the first alternative by inserting the value of n into the second condition}

$(n = 2 \wedge 0 = 4) \vee (n \neq 2 \wedge m(n-2) = 2n)$

$\equiv$ {the first alternative is contradictory and can thus be omitted from the disjunction}

$n \neq 2 \wedge m(n-2) = 2n$

$\equiv$ {divide by $n-2$, permitted, since $n-2 \neq 0$}

$$n \neq 2 \wedge m = \frac{2n}{n-2}$$

We have now expressed m as a function of n, and we know that $n \neq 2$. We also know that both m and n are integers, but that m is expressed as a fraction. This means that there can be values of n for which m is not an integer but a fraction. We simplify the expression for m further, to get a clearer view of when m is an integer and when it is a proper fraction. We get that

$$n \neq 2 \wedge m = \frac{2n}{n-2}$$

$\equiv$ {rewrite the numerator $2n$ as $2(n-2)+4$}

$$n \neq 2 \wedge m = \frac{2(n-2)+4}{n-2}$$

$\equiv$ {simplify}

$$n \neq 2 \wedge m = 2 + \frac{4}{n-2}$$

We have now expressed m as an integer part plus a fraction. We notice that if $n \geq 7$, then $n - 2 > 4$. From this follows that $m = 2 + \frac{4}{n-2}$ is a proper fraction, which contradicts the assumption that m is an integer. Thus, it must be that $n \leq 6$. We now have that:

$$n \neq 2 \wedge m = 2 + \frac{4}{n-2}$$

$\equiv$ {if $n \geq 7$ then m is not an integer; thus $m \leq 6$; we add this information to the conjunction as redundant information}

$$n \leq 6 \wedge n \neq 2 \wedge m = 2 + \frac{4}{n-2}$$

We now have the condition $n \leq 6 \wedge n \neq 2$. Note that we have used the rule for redundant information in a conjunction in the other direction: instead of omitting a proposition from a conjunction because it follows from one of the other propositions, we add redundant information to a conjunction.

Together with the assumption that $n \geq 1$ we now only have 5 values of n that satisfy the condition: $n = 1, 3, 4, 5, 6$. Thus we can rewrite the condition as a disjunction:

$$n \leq 6 \wedge n \neq 2 \wedge m = 2 + \frac{4}{n-2}$$

$\equiv$ {rewrite $n \leq 6 \wedge n \neq 2$ as a disjunction, use the assumption that n is a non-negative integer}

$$(n = 1 \vee n = 3 \vee n = 4 \vee n = 5 \vee n = 6) \wedge m = 2 + \frac{4}{n-2}$$

$\equiv$ {distribute the conjunction over the disjunction; logical rule for distribution}

$$(n = 1 \wedge m = 2 + \frac{4}{n-2}) \vee (n = 3 \wedge m = 2 + \frac{4}{n-2}) \vee$$

$$(n = 4 \wedge m = 2 + \frac{4}{n-2}) \vee (n = 5 \wedge m = 2 + \frac{4}{n-2}) \vee$$

$$(n = 6 \wedge m = 2 + \frac{4}{n-2})$$

$\equiv$ {simplify by inserting the values of n}

$$(n = 1 \wedge m = -2) \vee (n = 3 \wedge m = 6) \vee$$

$$(n = 4 \wedge m = 4) \vee (n = 5 \wedge m = 2 + \frac{4}{3}) \vee$$

$$(n = 6 \wedge m = 3)$$

$\equiv$ {according to an assumption m is a natural number, so alternatives $n = 1$ and $n = 5$ are contradictory and can be omitted by the rule for falsity }

$$(n = 3 \wedge m = 6) \vee (n = 4 \wedge m = 4) \vee (n = 6 \wedge m = 3)$$

□

The answer is that there are two different fields that satisfy the condition: a square field with the side 4 and a rectangular field with the sides 3 and 6.

■

7.4 Assignments

1. Prove that if a and b are odd integers then $a \cdot b$ is also an odd integer.
2. Assume that $a, b, c > 0$. Prove that if $a^2 + b^2 = c^2$, then $a + b \leq c\sqrt{2}$.
3. Check if the formula

   $$((p \Rightarrow \neg q) \vee (\neg r \Rightarrow q)) \wedge \neg ((\neg p \wedge q) \wedge (p \vee \neg q))$$

   is valid.
4. Prove that implication is transitive, i.e., prove that

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

5. Prove that implication is antisymmetric, i.e., prove that

   $$((p \Rightarrow q) \wedge (q \Rightarrow p)) \Rightarrow (p \equiv q)$$

6. Prove that implication not symmetric, i.e. that $(p \Rightarrow q) \Rightarrow (q \Rightarrow p)$ is not a valid formula.
7. Let us define the operation $\circ$(Latex: \circ) over the real numbers by $x \circ y = xy + x + y$. Prove that this operation is associative, i.e., that $a \circ (b \circ c) = (a \circ b) \circ c$.

Chapter 8

Observations

Basic tasks are for situations where we just state the problem and then solve it with a single calculation. For more complex problems, it can be difficult to see how to solve the problem directly. Rather, we have to construct a solution one step at a time. Starting from the assumptions, we make a series of *observations*, until we have enough information to solve the main problem, with or without a final calculation. We have three different kinds of observations, *(mathematical) facts* that follow from the assumptions, *definitions* for introducing new notations and concepts, and *declarations* that just introduce new constants.

8.1 Facts

A *fact* consists of two parts: the justification that explains why the fact follows from the assumptions and earlier observations, and the fact itself. A "+" symbol, or a number in square brackets ([1], [2],...), identifies this as a fact. The justification is written on a line of its own, before the fact itself. The general format for a fact is shown below on the left, an example of a fact is shown on the right.

fact

+ *justification*

proposition

+ {by assumption (b)}

$2 \leq x$

The example states that $2 \leq x$ follows directly from some assumption (b). Note that we are here talking about a *mathematical* fact, i.e., something we know to be true because of the assumptions and previous observations. A mathematical fact says nothing directly about the real world, it only states consequences of the assumptions and definitions in use.

Example 48. Nadja and Peter each rent a car for one day. Nadja pays 50 € per day, plus 0.40 € per km. Peter rents a car from another company that charges 70

€ per day and 0.30 € per km. How many kilometers should Nadja and Peter drive so that they pay the same rent for their cars.

First we identify the question and the assumptions of the problem:

- How many kilometers x should Nadja and Peter drive, to make their cost of renting the cars equal, when

(a) Nadja pays 50 € per day and 0.40 € per km.

(b) Peter pays 70 € per day and 0.30 € per km.

(c) Nadja and Peter rent their cars for one day

Next we observe some facts that follow directly from the assumptions:

[1] {Assumptions (a) and (c) show what Nadja pays}

Nadja pays $50 + 0.40 \cdot x$ euros to drive x km

[2] {Assumptions (b) and (c) show what Peter pays}

Peter pays $70 + 0.30 \cdot x$ euros to drive x km.

We now calculate the answer by writing an equation and solving it:

$\Vdash$ {We solve the equation below for x}

Nadja pays as much as Peter for driving x kilometers

$\equiv$ {observation [1] and [2]}

$50 + 0.40 \cdot x = 70 + 0.30 \cdot x$

$\equiv$ {regroup the terms}

$0.40 \cdot x - 0.30 \cdot x = 70 - 50$

$\equiv$ {simplify}

$0.10 \cdot x = 20$

$\equiv$ {divide by 0.10}

$x = \dfrac{20}{0.10}$

$\equiv$ {calculate}

$x = 200$

□ Peter and Nadja should both drive 200 km

The task has thus been solved: Nadja and Peter pay the same rent if they both drive 200 km. ■

Example 49. A dart board has a radius of 20 cm, and it is divided into ten rings of uniform width, numbered from 1 to 10 (starting from the outside). Gabriel hits the dart board so that the distance r (in cm) from the center of the board is distributed according to the density function

$$f(r) = \begin{cases} \dfrac{3}{16000}(400 - r^2), & \text{when } 0 \le r \le 20 \\ 0, & \text{else} \end{cases}$$

Calculate the probability that Gabriel scores a 9 or a 10.

We start by formulating the problem.

- Calculate the probability that the dart hits 9 or 10, when

(a) the radius of the dart board is 20 cm

(b) each ring has the same width, and

(c) the density function is

$$f(r) = \begin{cases} \dfrac{3}{16000}(400 - r^2), & \text{when } 0 \le r \le 20 \\ 0, & \text{else} \end{cases}$$

where r is the distance from the center of the board to the dart

The dart board looks like this [1]:

[1] Picture by Kallerna

The figure allows us to make the following observation.

[1] {From assumptions (a) and (b)}

the dart hits 9 or 10 if, and only if, $0 \le r \le 4$.

We are now ready to calculate the solution to the problem.

$\Vdash$ P(Gabriel hits 9 or 10)

$=$ {observation [1]}

$P(0 \le r \le 4)$

$=$ $\{P(A) = \int_c^d f(x)\ dx$, the interval of integration is given by [1]}

$\int_0^4 f(r)\ dr$

$=$ $\{f(r)$ is given by assumption (c) }

$$\int_0^4 \frac{3}{16000}(400 - r^2)\ dr$$

$=$ {integrate using formula$\int (f(x) + g(x))\, dx = \int f(x)\, dx + \int g(x)\, dx$}

$$\int_0^4 \frac{3}{16000} \cdot 400\ dr - \int_0^4 \frac{3}{16000} \cdot r^2\ dr$$

$=$ {integrate using formula $\int_a^b\ cx^n\ dx = \left[c \cdot \frac{x^{n+1}}{n+1}\right]_a^b$}

$$\left[\frac{3 \cdot 400}{16000} r\right]_0^4 - \left[\frac{3r^3}{3 \cdot 16000}\right]_0^4$$

$=$ {simplify}

$$\left[\frac{3}{40} r\right]_0^4 - \left[\frac{r^3}{16000}\right]_0^4$$

$=$ $\{[F(x)]_a^b = F(b) - F(a)\}$

$$\frac{3}{40} \cdot 4 - \frac{3}{40} \cdot 0 - \frac{64}{16000} + \frac{0}{16000}$$

$\approx$ {calculate an approximative value}

0.3

□ P(Gabriel hits 9 or 10) ≈ 0.3. ■

8.2 Definitions

It is often useful to introduce new notations in proofs and derivations, e.g., to simplify a calculation by introducing a name for a complex subexpression. We do this using *definitions*. Below is the general format for a definition. On the right, we have an example of a definition:

definition

+ *declaration*

justification

proposition

+ Define $c \in \mathbb{R}$

$\{a \neq 0$ by assumption, so c is well-defined $\}$

$c = \dfrac{e^a - 1}{a}$

The definition is written in three lines. The first line *declares* the name of the constant and its value domain. The declaration must at the very least state the names of the constants that are defined, and their value ranges. The second line is a justification that explains why the constant is well-defined. The third line gives the condition that defines the constant. Note that the name of the constant must be new, we are not allowed to reuse an existing name for the definition.

The name of the constant is c in the example, and the value domain is $\mathbb{R}$. The definition condition is

$$c = \frac{e^a - 1}{a}$$

The justification explains why c is well-defined, i.e., that there is a value in $\mathbb{R}$ that satisfies the definition condition. In this case, the constant c is well-defined when $a \neq 0$. We can then use c freely in the rest of the derivation, and replace c by its definition $\dfrac{e^a - 1}{a}$ whenever needed.

Sometimes we need to introduce two or more constants at the same time. We say that a real number a is rational, if it can be written as a fraction $\dfrac{p}{q}$, where p and q are two integers. A definition that introduces p and q is shown below:

+ Define $p, q \in \mathbb{Z}$

$\{a$ is a rational number$\}$

$a = \dfrac{p}{q}$

Note that this is an *implicit definition* of the constants p and q, it does not stipulate a unique value for these two numbers. For $a = \frac{1}{3}$, we could choose $p = 1$ and $q = 3$, but we could as well choose $p = 6$ and $q = 18$ or $p = 201$ and $q = 603$.

We explicitly write “Define” in front of the declaration of the constant names that are defined. This is optional, the syntax allows us to leave out this word, writing just + $p, q \in \mathbb{Z}$. However, the derivations seems to be easier to follow when we are more explicit in definitions.

Example 50. Three siblings inherit 12 000 € in total. The inheritance is to be split among the sibling in the ratio 5:3:2. How large a share will each sibling get?

• How large are the shares A, B and C of the inheritance, when

(a) the inheritance is 12 000, and

(b) A, B and C split the inheritance in the ratio 5:3:2

[1] {Assumptions (a) and (b) give}

$A + B + C = 12\,000$

We introduce a constant a that allows us to express the shares of each heir.

[2] Define $a \in \mathbb{R}$

{a well-defined because of assumption (b)}

$A = 5a \wedge B = 3a \wedge C = 2a$

$\Vdash$ $A + B + C = 12\,000 \wedge A = 5a \wedge B = 3a \wedge C = 2a$

$\equiv$ {substitute the values of A, B and C}

$5a + 3a + 2a = 12\,000 \wedge A = 5a \wedge B = 3a \wedge C = 2a$

$\equiv$ {solve the equation}

$a = 1\,200 \wedge A = 5a \wedge B = 3a \wedge C = 2a$

$\Rightarrow$ {property of the conjunction}

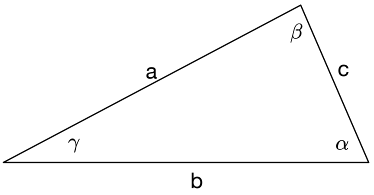
$A = 6\,000 \wedge B = 3\,600 \wedge C = 2\,400$

□ A gets 6 000, B gets 3 600 and C gets 2 400

The last step uses implication instead of equivalence. This is because we omit the condition for a in the last expression, as we do not need it. The implication does not hold in the opposite direction, the last expression does not say anything about the value of a so the two propositions are not equivalent. ■

8.3 Declarations

A definition introduces a new name for a constant, and states it properties. Sometimes it is easier to describe a situation by first introducing a number of new constant names, and then separately listing the assumptions that we make about these constants. This is particularly useful when we building a mathematical model for a situation that we want to analyze in more detail. Consider as an example the geometric figure below.



This figure introduces a triangle with sides a, b, c and angles α, β, γ. We can start modeling this situation by first introducing the constants and their value ranges with *declarations*:

$+ \qquad a, b, c \in \mathbb{R}^+$

$+ \qquad \alpha, \beta, \gamma \in \mathbb{R}^+$

Then we list the assumptions that we make about these constants and how they are related to each other:

- a, b, c are the (lengths of the) sides of a triangle

- α is the angle opposite a, β is the angle opposite b, γ is the angle opposite c

A declaration of the form

$+ \qquad c_1, \ldots, c_m \in A$

can be seen as a shorthand for a definition of the form

$+ \qquad c_1, \ldots, c_m \in A$

$\{A$ is non-empty$\}$

T

In other words, a declaration defines new constants with names $c_1, \ldots, c_m$ and values in A. The justification must show that A is non-empty, to guarantee that the constants have well-defined values.. This is all we know about the values of the constants $c_1, \ldots, c_m$. Any further information is given by the assumptions that follow these declarations. When using value ranges like $\mathbb{N}$, $\mathbb{Q}$, $\mathbb{R}$, etc., that we know are non-empty, a simple constant declaration is sufficient.

A declaration introduces new constants to describe our model. Instead of explicit declarations, we could just assume that the constants that we need already exist in the environment, and our task is just to state what properties we assume of these constants. We could then introduce the geometry problem with a list of assumptions:

- $a, b, c \in \mathbb{R}^+$
- $\alpha, \beta, \gamma \in \mathbb{R}^+$
- a, b, c are the (lengths of the) sides of a triangle
- α is the angle opposite a, β is the angle opposite b, γ is the angle opposite c

From the point of solving the problem, these two approaches are equivalent. Analyzing the properties of the model will proceeds in the same way in both approaches. We usually use assumptions rather than explicit declarations in our example solutions, as it is closer to the way mathematics is usually carried out in schools.

There is, however, a situation where the use of declarations becomes important. This is when we want to build a new theory for some interesting phenomenon or situation. We should then clearly delineate those concepts that are specific to the new theory from the concepts that we may assume already known in the mathematical background for the theory. This includes the constants that we *postulate* for our theory. We will discuss this in more detail in Chapter 18.

8.4 Solving Problems as Tasks

We describe the solution to a mathematical problem in the form of a structured task. But this does not mean that we have to construct the solution step-by-step in the same order as the different components of the task are enumerated in the final solution. We often approach a problem in ways that later turn out to be unsuccessful, we sometimes change the question or the assumptions, and we may make them more precise. We can make irrelevant observations, or identify additional assumptions at a later stage that we did not notice earlier, etc. We need to write the components of a structured task in a particular order to avoid circular reasoning, but this order does not have to be the one we follow when we work out the problem and look for a solution.

Solving a math problem is to some extent analogous to solving programing problems. First we need to find a strategy for solving the problem, and then work out the details. Some of the details are straightforward, while others can be very tricky. A structured task is comparable to a program, it is the format we use to write down the final solution to the problem. While we are working on the problem, a structured task functions as repository for facts and information that we discover during the process. This allows us to work systematically, step by step, towards a final solution of the problem. We show below how one can use structured derivations as a support structure when solving a problem

A computer based editor for structured derivation is very useful here, since we can then easily make changes to the task, add some components, remove others as wrong or unnecessary, and copy expressions from one line to another. This is more cumbersome when we work with pen and paper: scribblings, deletions and corrections in the text make the text cluttered, so that we may need to rewrite the final solution in order to hand in a clean text. .

Example 51. Prove that $m^2 - n^2 \geq 3$, when m and n are positive natural numbers and $m > n$.

Step 1 We start by working out the problem: what should we do (the question) and which assumptions are we allowed to make. In this case, the problem is as follows:

- Prove that $m^2 - n^2 \geq 3$, when

(a) $m \in N$, $m > 0$

(b) $n \in \mathbb{N}$, $n > 0$

(c) $m > n$

Step 2 We want to prove that the statement is true. We start by simplifying the expression $m^2 - n^2$. We immediately notice that we can use the conjugate rule here. After adding this step, the task looks as follows (red text indicate what is new):

- Prove that $m^2 - n^2 \geq 3$, when

(a) $m \in \mathbb{N}$, $m > 0$

(b) $n \in \mathbb{N}$, $n > 0$

(c) $m > n$

$\Vdash$ $\{\}$

$m^2 - n^2$

$=$ {by the conjugate rule}

$(m - n)(m + n)$

The curly bracket for justifying the answer is still empty, since we have not solved the problem yet.

Step 3 We now notice that we can use the monotonicity of a product, i.e., that $ab \geq ab'$, if $a \geq 0$ and $b \geq b'$. We can use this rule to prove $(m-n)(m+n) \geq (m-n) \cdot 3$, if we can show that $m - n \geq 0$ and $m + n \geq 3$. We show this by adding two facts before the calculation. The task now look as follows, with the new facts and the new step shown in red:

- Prove that $m^2 - n^2 \geq 3$, when

(a) $m \in \mathbb{N}$, $m > 0$

(b) $n \in N$, $n > 0$

(c) $m > n$

[1] $\{$(c) implies that $m - n > 0$, so $m - n \geq 1\}$

$m - n \geq 1$

[2] $\{$(a) - (c) imply that $n \geq 1$ and $m \geq n + 1 \geq 2$, so $m + n \geq 3\}$

$m + n \geq 3$

$\Vdash$ $\{\}$

$m^2 - n^2$

$=$ $\{$by the conjugate rule$\}$

$(m - n)(m + n)$

$\geq$ $\{$the product is monotonic: $ab \geq ab'$, when $a \geq 0$ and $b \geq b'$, [1] and [2]$\}$

$(m - n) \cdot 3$

Step 4 We complete the task by applying the same rule once more, now for expression $(m - n) \cdot 3$, to get the final solution.

- Prove that $m^2 - n^2 \geq 3$, when

(a) $m \in \mathbb{N}$, $m > 0$

(b) $n \in \mathbb{N}$, $n > 0$

(c) $m > n$

[1] $\{$from (c) follows that $m - n > 0$, so $m - n \geq 1\}$

$m - n \geq 1$

[2] $\{$from (a) - (c) follows that $n \geq 1$ and $m \geq n + 1 \geq 2$, so $m + n \geq 3\}$

$m + n \geq 3$

$\Vdash$ $\{$The statement follows from the transitivity of $\geq\}$

$m^2 - n^2$

$=$ {by the conjugate rule}

$(m - n)(m + n)$

$\geq$ {the product is monotonic: $ab \geq ab'$, when $a \geq 0$ and $b \geq b'$, [1] and [2]}

$(m - n) \cdot 3$

$\geq$ {the product is monotonic, observation [1]}

$1 \cdot 3$

$=$ {arithmetics}

3

□ ■

The example shows that the task was not constructed in the same order as it is written down in the final solution. We have inserted observations at the beginning when we needed them. We have also added justifications later, when we saw what they should be.

Chapter 9

Nested tasks

We also have another way of breaking up a larger task into smaller, more manageable tasks: *subtasks* or *nested tasks.* We will start by first introducing nested calculations, before moving on to the nesting of general tasks.

9.1 Nested Calculations

Structured calculations require that each step in the calculation is explicitly justified. In many cases, it is sufficient to write a comment in curly brackets as a justification, as we have done in the examples we have described until now. There are, however, many situations where a simple explanation is not sufficient, but we would really need to do another calculation in order to see that the step is correct. We refer to such sub-calculations as *nested calculation*s. Consider as an example the following calculation problem and its solution.

Example 52. Calculate the value of the expression $2+(3\cdot 2^3+4\cdot 3^2)\cdot 2\cdot 4^2-2\cdot 5^2$

• $2+(3\cdot 2^3+4\cdot 3^2)\cdot 2\cdot 4^2-2\cdot 5^2$

= {calculate the powers in the parenthesis}

$2+(3\cdot 8+4\cdot 9)\cdot 2\cdot 4^2-2\cdot 5^2$

= {multiply in the parenthesis}

$2+(24+36)\cdot 2\cdot 4^2-2\cdot 5^2$

= {add in the parenthesis}

$2+60\cdot 2\cdot 4^2-2\cdot 5^2$

= {calculate the powers in the entire expression}

$2+60\cdot 2\cdot 16-2\cdot 25$

= {multiply}

$2 + 1920 - 50$

$=$ {add and subtract}

1872

□ ■

The expression in parentheses has to be evaluated first, before the main calculation. The problem with this calculation is that we have to copy the part of the expression that lies outside the parenthesis from one line to another when we manipulate the expression inside the parenthesis. This is redundant, since this part does not change from one line to the next, and it is error prone, particularly if we do this by hand. For long and complex expressions, it also becomes difficult to see which part of the expression is being manipulated from one step to another.

Nested calculations solve this problem. A nested calculation is a structured calculation that is carried out as part of a larger structured task. We can add a nested calculation to any justification, to give a more detailed explanation for the derivation step. A nested calculation is a separate calculation that supports the justification. The nested calculation is indented one step to the right. The next example shows the previous calculation written with a nested calculation

Example 53. Example of a structured task with a nested calculation.

• $2 + (3 \cdot 2^3 + 4 \cdot 3^2) \cdot 2 \cdot 4^2 - 2 \cdot 5^2$

$=$ {calculate the value inside the parenthesis}

> • $3 \cdot 2^3 + 4 \cdot 3^2$
>
> $=$ {calculate the powers}
>
> $3 \cdot 8 + 4 \cdot 9$
>
> $=$ {multiply}
>
> $24 + 36$
>
> $=$ {add}
>
> 60
>
> □

... $2 + 60 \cdot 2 \cdot 4^2 - 2 \cdot 5^2$

$=$ {calculate the powers in the entire expression}

$2 + 60 \cdot 2 \cdot 16 - 2 \cdot 25$

$=$ {multiply}

$2 + 1920 - 50$

$=$ {add and subtract}

1872

□ ■

The three dots in the left column after the nested calculation shows where the main calculation continues. This will give us more lines in the derivation, but we have to type fewer characters, since we do not have to copy expressions that remain unchanged from one line to the next. The nested calculation also shows clearly which part of the expression is being manipulated.

Using a computer to write structured derivations can again be quite useful here. An *outlining editor*, i.e., an editor that can selectively show and hide indented text, is particularly useful. We can then hide the nested calculation when we want to concentrate on the overall solution, and show the nested calculation again when we want to check the details. Hiding the nested calculation in the previous calculation gives us the following derivation:

• $2 + (3 \cdot 2^3 + 4 \cdot 3^2) \cdot 2 \cdot 4^2 - 2 \cdot 5^2$

$=$ {calculate the parenthesis}

... $2 + 60 \cdot 2 \cdot 4^2 - 2 \cdot 5^2$

$=$ {calculate the powers in the entire expression}

$2 + 60 \cdot 2 \cdot 16 - 2 \cdot 25$

$=$ {multiply}

$2 + 1920 - 50$

$=$ {add and subtract}

1872

□

The three dots now show that the first justification contains a hidden nested calculation. ■

The following example shows a more substantial case of how to organize and simplify the calculation of arithmetic expressions.

Example 54. Simplify the expression $\sqrt{7 + 2\sqrt{11}} + \sqrt{7 - 2\sqrt{11}}$. Our approach is to square the expression, then simplify it and finally take the square root of the simplified expression.

• Simplify the expression $\sqrt{7 + 2\sqrt{11}} + \sqrt{7 - 2\sqrt{11}}$

$\Vdash$ $\sqrt{7 + 2\sqrt{11}} + \sqrt{7 - 2\sqrt{11}}$

= {we square the expression, simplify it and then insert the square root of the simplified expression}

- $(\sqrt{7+2\sqrt{11}}+\sqrt{7-2\sqrt{11}})^2$

  = {the square rule}

  $7+2\sqrt{11}+2\cdot\sqrt{7+2\sqrt{11}}\cdot\sqrt{7-2\sqrt{11}}+7-2\sqrt{11}$

  = {simplify}

  $14+2\cdot\sqrt{7+2\sqrt{11}}\cdot\sqrt{7-2\sqrt{11}}$

  = {focus on the second subexpression $2\cdot\sqrt{7+2\sqrt{11}}\cdot\sqrt{7-2\sqrt{11}}$}

  - $2\cdot\sqrt{7+2\sqrt{11}}\cdot\sqrt{7-2\sqrt{11}}$

    = {the product of two radicals}

    $2\cdot\sqrt{(7+2\sqrt{11})\cdot(7-2\sqrt{11})}$

    = {the conjugate rule}

    $2\sqrt{49-4\cdot 11}$

    = {simplify}

    $2\sqrt{5}$

  □

  ... $14+2\sqrt{5}$

□

... $\sqrt{14+2\sqrt{5}}$

□ The simplified expression is $\sqrt{14+2\sqrt{5}}$. ■

We use two nested derivations here, one inside the other. The original problem is to simplify a sum of square roots. Instead of doing this directly, we simplify the square of this expression in a nested derivation. The square root of the simplified expression is then the solution to our original problem. Inside the first nested derivation, we then carry out a separate nested derivation, where we simplify a part of the complex square expression. Focusing on a part of the expression in the nested derivation makes it easier to see what we manipulate in the derivation. There will also be fewer errors, as we avoid copying long expressions from line to line.

9.2 Nested tasks

We can use nesting for arbitrary tasks, not only for calculation tasks. We rewrite our earlier number theory example to use *nested tasks.*

Example 55. Prove that $m^2-n^2\geq 3$, when m and n are positive integers, and $m>n$.

- Prove that $m^2-n^2\geq 3$, when

(a) m is a positive integer,

(b) n is a positive integer, and

(c) $m > n$

$\Vdash$ $m^2 - n^2$

$=$ {by the conjugate rule}

$(m - n)(m + n)$

$\geq$ {the product is monotonic: $ab' \geq ab$, when $a \geq 0$ and $b' \geq b$}

• Show that $m - n \geq 0$

$\Vdash$ $m - n$

$>$ {assumption (c)}

$n - n$

$=$ {arithmetics}

0

□

• Show that $m + n \geq 3$

$\Vdash$ $m + n$

$\geq$ {assumption (c)}

$n + 1 + n$

$\geq$ {assumption (b)}

$1 + 1 + 1$

$=$ {arithmetics}

3

□

... $(m - n) \cdot 3$

$\geq$ {the product is monotonic, $m - n > 0$ according to (c), and $3 \geq 0$}

$1 \cdot 3$

$=$ {arithmetics}

3

□

The second calculation step uses two nested tasks, to show that the two constraints for applying the rule are both satisfied. We have the same situation in the third step, but now it is easier to see that the constraints are satisfied, so we handle this without nested derivations. ■

9.3 Inheritance

The use of nested tasks can greatly simplify a structured derivation, by dividing the overall problem into smaller and more manageable subproblems. The nested tasks are solved in the context of the justification that they support. This means that all assumptions, facts and definitions that are available at the point where the justification is written are also available in the nested tasks for that justification. In other words, the nested tasks *inherit* the assumptions, facts and definitions of their justification. Therefore, we need not repeat these in the nested tasks.

This inheritance is clearly shown in Example 55. The first nested task uses assumption (c), which is given on the outer level. The second nested task uses both assumptions (c) and (b) from the outer level. Generally, a nested task can refer to any preceding assumptions, observations and definitions at an outer level compared to the nested task and written before this task. We are not allowed to refer to assumptions or observations that are made after the nested task, regardless of what level they are on. This restriction prevents circular reasoning when solving a task.

9.4 General Syntax for Tasks

The general format for structured tasks is shown in Table 9.1. The syntax for observations covers both facts and definitions (the declaration is empty for a fact):

This template shows that tasks are *recursive* in nature:

- Each step in a task comes with a justification.
- A justification can contain zero or more (nested) tasks.

A nested tasks inside a justification may have its own justifications, which again may contain nested tasks, and so on. The recursion ends with a justification that does not introduce any new nested tasks.

We emphasize the recursive structure of tasks by coloring tasks and justifications blue. The other components of a structured derivation are colored in other colors, depending on their type: questions, answers, and declarations are colored green, logical propositions are colored magenta, relations are colored red, justification explanations are colored cyan, while expressions and declarations are colored black.

We illustrate the general syntax for tasks with an example that we have given earlier (Example 51). This task has 3 assumptions, 2 observations and 4 calculation steps. We have chosen to identify the assumptions by letters and the observations by numbers, so that we can refer to them individually in the derivation.

***task*:**

●	*question*
-	*assumption*
⋮	
-	*assumption*
+	*declaration*
	justification
	proposition
⋮	
+	*declaration*
	justification
	proposition
⊩	*justification*
	expression
rel	*justification*
	expression
⋮	
rel	*justification*
	expression
□	*answer*

***justification*:**

{*explanation*}

task

⋮

task

Table 9.1: General syntax for tasks

•	*question*	•	Prove that $m^2 - n^2 \geq 3$, when
(a)	*assumption*	(a)	$m \in \mathbb{N}$, $m > 0$
(b)	*assumption*	(b)	$n \in \mathbb{N}$, $n > 0$
(c)	*assumption*	(c)	$m > n$
[1]	*justification*	[1]	{ (c) implies that $m - n > 0$, so $m - n \geq 1$}
	proposition		$m - n \geq 1$
[2]	*justification*	[2]	{ (a) - (c) imply $n \geq 1$ and $m \geq n + 1 \geq 2$}
	proposition		$m + n \geq 3$
$\Vdash$	*justification*	$\Vdash$	{ follows from the transitivity of $\leq$}
	expression		$m^2 - n^2$
rel	*justification*	$=$	{by the conjugate rule}
	expression		$(m - n)(m + n)$
rel	*justification*	$\geq$	{product is monotonic, observation [1] and [2]}
	expression		$(m - n) \cdot 3$
rel	*justification*	$\geq$	{ product is monotonic, observation [1]}
	expression		$1 \cdot 3$
rel	*justification*	$=$	{arithmetics}
	expression		3
□	*answer*	□	the proposition is true

Note that *justifications* on the left are colored blue, because they stand for general justifications, which may have nested derivations. In the example on the right, each justification is of the simple form *{explanation}*, and is hence colored cyan.

9.5 Assignments

1. Solve the simultaneous equations

$$x = 5 \;\wedge\; y = x + 12 \;\wedge\; z = 2x + y - 2z$$

2. Solve the simultaneous equations

$$2z + y - 2x = 6\frac{1}{3} \;\wedge\; 10y - x = \frac{z}{18} \;\wedge\; z + 2 = 2x + 2\frac{1}{3} - y$$

3. Solve the absolute value equation $|2x-8| = 3x-5$.
4. Solve the inequality $2x^2+20x+32>0$.
5. Solve the absolute value equation $\left|x^3-x^2+2x-2\right| = 4x^2+8$.
6. The points $A=(-4,\ -6,\ 0)$ and $B=(5,\ 3,\ 1)$ are in the same plane. Can the vector $\bar{n} = -49\bar{\imath}+51\bar{\jmath}-18\bar{k}$ be a normal vector to the plane?
7. A line passes through the point $(4,\ 2,\ 1)$and is directed along the vector $2\bar{\imath}+3\bar{\jmath}-\frac{5}{3}\bar{k}$. Determine if the line intersects the xy-plane and if so, where it intersects it.
8. When appending the digits 91 to the end of a specific integer we get the original number multiplied by 107. What is the original number?

Chapter 10

Problem Solving Paradigms

There are essentially three main paradigms for solving mathematical problems: *calculations, forward derivations*, and *backward derivations*. We will show how each of these paradigms can be formulated as a structured task, and then show how a structured task allows us to combine all three paradigms in a single derivation.

We will illustrate each paradigm by proving the same simple theorem, each time following a different paradigm. We conclude by giving a proof that combines these paradigms in a single structured task. The problem we consider is the following:

> *Prove that $k^2 + k$ is an even number when k is a natural number.*

10.1 Calculations

Calculation is a central tool in all of mathematics, in particular in science and engineering. A standard calculation does not contain observations or nested tasks, and the conclusion is usually implicit.

Example 56. We prove our example theorem with a calculation. We do this by calculating the truth value of the logical proposition "$k^2 + k$ is even", and find that it is true.

• Show that $k^2 + k$ is even, when

- $k \in \mathbb{N}$

$\Vdash$ $k^2 + k$ is even

$\equiv$ {write the expression as a product}

$k(k + 1)$ is even

$\equiv$ {a product is even if and only if one of the factors is even}

k is even or $k + 1$ is even

$\equiv$ {every second natural number is even, so either k or $k + 1$ must be even}

T

□

The calculation shows that the original proposition is equivalent to T for any natural number, which means that the proposition is true. ■

A structured task that is solved with a calculation has the following general form:

● *question*

- *assumption*

$\vdots$

- *assumption*

$\Vdash$ {*conclusion*}

expression

rel {*explanation*}

expression

$\vdots$

rel {*explanation*}

expression

□ *answer*

10.2 Forward Derivations

A forward derivation is a proof that starts from given facts (the assumptions) and then adds one observation after the other, until we reach an observation that proves the theorem that we are interested in. Each observation is shown to follow from the assumptions and previous observations. This proof method was introduced by the greek mathematicians, and resulted in, e.g., the impressive treatment of Euclidean geometry.

There are no calculations in a forward derivation, and the arguments are simple, i.e., there are no nested tasks. The assumptions and observations are numbered, so that that we can refer to them when justifying the observations.

Example 57. We prove our example theorem with a forward derivation.

• Show that $k^2 + k$ is even, when

- $k \in \mathbb{N}$

[1] {Each natural number is either even or odd}

k is even or k is odd

[2] {A natural number is odd iff the next number is even, observation [1]}

k is even or $k + 1$ is even

[3] {A product of two natural numbers is even if one of the numbers is even, observation [2]}

$k \cdot (k + 1)$ is even

[4] {Distribution rule: $k \cdot (k + 1) = k^2 + k$, observation [3]}

$k^2 + k$ is even

$\Vdash$ {The theorem follows from observation [4]}

□ ■

A structured task that is solved with forward derivations looks as follows:

• *question*

(a) *assumption*

⋮

(m) *assumption*

[1] {*explanation*}

proposition

⋮

[n] {*explanation*}

proposition

$\Vdash$ {*conclusion*}

□ *answer*

We have for simplicity assumed here that all observations are facts, but we can also introduce new constants in a forward derivation with definitions.

10.3 Backward Derivations

A task solved with backward derivations does not include any calculations nor any forward derivation. Instead, we use justifications with nested tasks. The basic idea is that solving the main problem is reduced to solving a number of simpler subproblems (nested tasks).

Example 58. We prove our example theorem with backward derivation. We will do this in stages. The problem that we want to solve is the following:

• Show that $k^2 + k$ is even, when

- k is a natural number

We first reduce this task to two simpler tasks (marked with red) using case analysis.

• Show that $k^2 + k$ is even, when

- k is a natural number

⊩ {Case analysis, consider the cases that k is even and k is odd separately}

> • Show that $k^2 + k$ is even, when
>
> \- k is an even number
>
> • Show that $k^2 + k$ is even, when
>
> \- k is an odd number

□

The original problem has been reduced to two smaller problems: showing that the original statement is true when k is even, and showing that the statement is true when k is odd. As k must be either even or odd, it is sufficient to prove that the theorem holds in both these cases.

We now complete this proof with arguments that show that the two new theorems are true. We prove these two theorems without any further reductions. The new parts are again written in red.

• Show that $k^2 + k$ is even, when

- k is a natural number

⊩ {Case analysis, consider the two cases that k is even and that k is odd separately}

> • Show that $k^2 + k$ is even, when
>
> \- k is an even number
>
> ⊩ {$k^2 + k$ can be written as $k(k + 1)$; k is even, so $k(k + 1)$ is even}

□

- • Show that $k^2 + k$ is even, when
- - k is an odd number
- $\Vdash$ $\{k^2 + k$ can be written as $k(k+1)$; k is odd, so $k+1$ is even, so $k(k+1)$ is even$\}$

□

□

The recursion stops at the first level of nesting, because the nested tasks are proved directly, without introducing any new nested tasks. ■

A task that we solve with backward derivations has the following general form:

task:

- • *question*
- - *assumption*

$\vdots$

- - *assumption*
- $\Vdash$ $\{$*conclusion*$\}$

task

$\vdots$

task

□ *answer*

We have here substituted the definition of justification directly in the task. A task now only states the question and the assumptions, together with a justification for why the answer is correct. This justification is, however, based on solving a number of other, nested tasks. These new tasks are then either solved directly, or reduced to further subtasks.

10.4 Combining Paradigms

The classical proof paradigms each have their strengths. Algebraic and numeric problems are best solved with calculations, complex mathematical problems are usually solved with forward derivations, while proof strategies, where we break up a larger problem into smaller, more manageable problems, are often based on backward derivations. Backward derivations are often the fastest way to find a proof, and are therefore standard in computer based theorem proving.

Structured tasks allow us to combine all these proof paradigms in a single general paradigm for mathematical argumentation. We can, e.g., start by reducing the original problem to a number of simpler subproblems. We can then use calculations to solve some of the subproblems, observations for some other subproblems, and use further reductions for the remaining problems. Or, we can solve the original problem by combining observations, calculations and reductions in a single task. In essence, this means that we use the proof paradigm that is best suited for the problem and subproblem at hand.

Example 59. We prove that $k^2 + k$ is an even number for any natural number k, by combining backward proofs with calculations that refine the previous proof.

$\bullet$ Show that $k^2 + k$ is even, when

- k is a natural number

$\Vdash$ {Case analysis, consider the two cases, k is even and k is odd}

    $\bullet$ Show that $k^2 + k$ is even, when

    \- k is an even number

    $\Vdash$ $k^2 + k$ is even

    $\equiv$ {write as product}

        $k(k + 1)$ is even

    $\Leftarrow$ {a product is even if one of the factors is even}

        k is even

    $\equiv$ {assumption}

        T

    □

    $\bullet$ Show that $k^2 + k$ is even, when

    \- k is an odd number

    $\Vdash$ $k^2 + k$ is even

    $\equiv$ {write as product}

        $k(k + 1)$ is even

    $\Leftarrow$ {a product is even if one of the factors is even}

        $k + 1$ is even

    $\equiv$ {number theory}

        k is odd

    $\equiv$ {assumption}

T

□

□

Note the use of backward implication in both nested tasks. We have in both cases showed that the desired result (that $k^2 + k$ is even) follows from the assumption made in the subtask. ■

Structured tasks thus combine the three main proof paradigms in a single new proof format. Observations constitute a forward derivation, while justifications describe backward derivations. Thus, we can describe the general syntax of structured tasks as follows:

***task*:**

● *question*

- *assumptions*

forward derivation

⊩ ***justification***

calculation

□ *answer*

The proof paradigms are:

forward derivation:

+ *declaration*

justification

proposition

⋮

+ *declaration*

justification

proposition

justification

{*explanation*}

task

⋮

task

calculation:

expression

rel ***justification***

expression

⋮

rel ***justification***

expression

The conclusion sign "$\Vdash$" separates the three proof paradigms from each other. Note that justifications occur in each step of a forward derivation and a calculation, not only for the conclusion. In a forward derivation, the declaration of an observation is empty (and may thus be omitted) for a fact.

10.5 Examples

We end with a few examples that illustrate the power of combining proof paradigms in problem solving with structured derivations. Our first example is taken from analytic geometry.

Example 60. Find the point on the parabola $y = x^2 - 2x - 3$ where its tangent has the direction angle 45°.

•   Find the point (x, y) on the parabola f, where

(a) $f(x) = x^2 - 2x - 3$ for all $x \in \mathbb{R}$, and

(b)   the tangent of the parabola at point (x, y) has direction angle $\alpha = 45°$

[1]   {Find the first derivative at point x}

    •   the tangent of the parabola at point (x, y) has direction angle 45°

    $\equiv$ {the slope is $\tan \alpha$}

        the tangent of the parabola at point (x, y) has slope $\tan 45°$

    $\equiv$ $\{\tan 45° = 1\}$

        the tangent of the parabola at point (x, y) has slope 1

    $\equiv$ {the first derivative gives the slope}

        $f'(x) = 1$

    □

... $f'(x) = 1$

[2]   {Find x}

    • $f'(x) = 1$

    $\equiv$ {assumption (a), the derivative of f at the point x is $f'(x) = 2x - 2$}

        $2x - 2 = 1$

    $\equiv$ {solve for x}

        $x = \dfrac{3}{2}$

    □

... $x = \dfrac{3}{2}$

$\Vdash$ (x, y)

$=$ {observation [2]}

$(\frac{3}{2}, y)$

$=$ {assumption (a) and observation [2]}

$(\frac{3}{2}, (\frac{3}{2})^2 - 2(\frac{3}{2}) - 3)$

$=$ {calculating}

$(\frac{3}{2}, -\frac{15}{4})$

$\square$ $(x, y) = (\frac{3}{2}, -\frac{15}{4})$ ■

We may also use the verbose format for general tasks, in particular when they have observations. The observations are part of the solution, i.e, they are steps in the derivation that lead us to the final conclusion. Table 10.1 shows the verbose format next to the symbolic presentation.

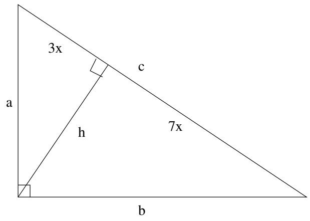
• Find the point (x, y) on the parabola f, where	**Task**: Find the point (x, y) on the parabola f, where
(a) $f(x) = x^2 - 2x - 3$ for all $x \in \mathbb{R}$, and	(a) $f(x) = x^2 - 2x - 3$ for all $x \in \mathbb{R}$, and
(b) the tangent of the parabola at point (x, y) has direction angle $\alpha = 45^\circ$	(b) the tangent of the parabola at point (x, y) has direction angle $\alpha = 45^\circ$
[1] {Find the first derivative at point x}	**Step** 1: {Find the first derivative at point x}
• the tangent of the parabola at point (x, y) has direction angle 45°	• the tangent of the parabola at point (x, y) has direction angle 45°
≡ {the slope is $\tan \alpha$}	≡ {the slope is $\tan \alpha$}
the tangent of the parabola at point (x, y) has slope $\tan 45^\circ$	the tangent of the parabola at point (x, y) has slope $\tan 45^\circ$
≡ {$\tan 45^\circ = 1$}	≡ {$\tan 45^\circ = 1$}
the tangent of the parabola at point (x, y) has slope 1	the tangent of the parabola at point (x, y) has slope 1
≡ {the first derivative gives the slope}	≡ {the first derivative gives the slope}
$f'(x) = 1$	$f'(x) = 1$
□	□
... $f'(x) = 1$	**Thus:** $f'(x) = 1$
[2] {Find x}	**Step** 2: {Find x}
• $f'(x) = 1$	• $f'(x) = 1$
≡ {assumption (a), the derivative of f at the point x is $f'(x) = 2x{-}2$}	≡ {assumption (a), the derivative of f at the point x is $f'(x) = 2x - 2$}
$2x - 2 = 1$	$2x - 2 = 1$
≡ {solve for x}	≡ {solve for x}
$x = \frac{3}{2}$	$x = \frac{3}{2}$
□	□
... $x = \frac{3}{2}$	**Thus**: $x = \frac{3}{2}$
⊩ {Determine second coordinate}	**Calculate:** {Determine second coordinate}
(x, y)	(x, y)
= {observation [2]}	= {observation [2]}
$(\frac{3}{2}, y)$	$(\frac{3}{2}, y)$
= {assumption (a) and observation [2]}	= {assumption (a) and observation [2]}
$(\frac{3}{2}, (\frac{3}{2})^2 - 2(\frac{3}{2}) - 3)$	$(\frac{3}{2}, (\frac{3}{2})^2 - 2(\frac{3}{2}) - 3)$
= {calculating}	= {calculating}
$(\frac{3}{2}, -\frac{15}{4})$	$(\frac{3}{2}, -\frac{15}{4})$
□ $(x, y) = (\frac{3}{2}, -\frac{15}{4})$	**Answer:** $(x, y) = (\frac{3}{2}, -\frac{15}{4})$

Table 10.1: Symbolic and verbose format

Now consider a problem in geometry.

Example 61. The length of two sides of a triangle are 5 and 11. The height to the third side splits that side in ratio 3 : 7. Calculate the length of the unknown side.

We draw a figure to illustrate the problem. We have labeled the sides (a, b and c) and the height (h).



Let us formulate the problem:

- Calculate c, when

- 5, 11 and c are (the lengths of the) sides of a triangle

- h is the height against the side c, and

- h splits c in a ratio 3 : 7

We introduce x as one tenth of c, so that we can describe the two parts of c as $3x$ and $7x$.

[1] Define $x \in \mathbb{R}$

{x is a tenth of c}

$c = 10x$

We can use the Pythagorean theorem for two different right triangles:

[2] {Pythagorean theorem for the triangle with sides h, $3x$ and 5}

$h^2 + 9x^2 = 25$

[3] {Pythagorean theorem for the triangle with sides h, $7x$ and 11}

$h^2 + 49x^2 = 121$

We can now solve x from observations [2] and [3]

[4] {Calculate x}

$\bullet$ [2] and [3]

$\equiv$ {write down the observations}

$h^2 + 9x^2 = 25$ and $h^2 + 49x^2 = 121$

$\Rightarrow$ {subtract the first equation from the second and simplify}

$40x^2 = 96$

$\equiv$ {divide both sides by 40 and simplify}

$x^2 = \frac{12}{5}$

$\equiv$ {take the square root of both sides, note that $x > 0$}

$x = \sqrt{\frac{12}{5}}$

□

$\ldots$ $x = \sqrt{\frac{12}{5}}$

We are now ready to calculate c.

$\Vdash$ c

$=$ {definition [1]}

$10x$

$=$ {observation [4]}

$10 \cdot \sqrt{\frac{12}{5}}$

$=$ {extend by 5 under the root}

$10 \cdot \sqrt{\frac{60}{25}}$

$=$ {extend by 5 under the root and simplify}

$10 \cdot \dfrac{2 \cdot \sqrt{15}}{5}$

$=$ {simplify}

$4 \cdot \sqrt{15}$

□ $c = 4 \cdot \sqrt{15}$ ■

The next example shows how to solve a problem involving series.

Example 62. Calculate the sum of the geometric series

$$a + ar + ar^2 + \ldots + ar^{n-1}$$

for $n \geq 1$, when $r \neq 1$ and $r \neq 0$.

We start by formulating the problem:

- Calculate $a + ar + ar^2 + \ldots + ar^{n-1}$, when
- $r \in R$, $r \neq 0$ and $r \neq 1$, and
- $n \in N$, $n \geq 1$

We introduce an auxiliary constant, s, that turns out to be very useful for solving the problem.

[1] Define $s \in \mathbb{R}$

{Since $r \neq 0$, the expression r^{n-1} is defined for $n = 1$ (note that 0^0 is undefined), hence s is defined for $n \geq 1$}

$s = 1 + r + r^2 + \ldots + r^{n-1}$

Next, we make two observations about s. Both observations make use of nested calculations.

[2] {Calculate $s - rs$}

● $s - rs$

= {the definition of s, calculate rs}

$1 + r + r^2 + \ldots + r^{n-1} - (r + r^2 + r^3 + \ldots + r^n)$

= {simplify}

$1 - r^n$

□

... $s - rs = 1 - r^n$

[3] {Find s by solving the equation in [2]}

● $s - rs = 1 - r^n$

≡ {factor out s}

$s(1 - r) = 1 - r^n$

≡ {divide by $1 - r$, allowed because $r \neq 1$ by assumption}

$s = \dfrac{1 - r^n}{1 - r}$

□

... $s = \dfrac{1 - r^n}{1 - r}$

Finally, we are ready to solve the original problem

$\Vdash$ $a + ar + ar^2 + \ldots + ar^{n-1}$

$=$ $\{$factor out $a\}$

$a \cdot (1 + r + r^2 + \ldots + r^{n-1})$

$=$ $\{$definition [1]$\}$

as

$=$ $\{$observation [3]$\}$

$a\dfrac{1 - r^n}{1 - r}$

□ $a + ar + ar^2 + \ldots + ar^{n-1} = a\dfrac{1 - r^n}{1 - r}$ ■

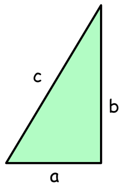
Our final example shows how to solve a problem in geometry that involves the use of figures and geometric constructions.

Example 63. We give here one of the classical proofs of the Pythagorean Theorem. The proof is based on a sequence of geometric constructions. The geometric constructions are shown here in a succession of figures, where the initial situation (the right triangle) is colored green, the first extension is colored orange and the second extension is colored blue.

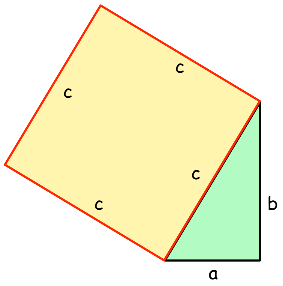
- (*The Pythagorean Theorem*) Show that $a^2 + b^2 = c^2$, where

– c is the hypotenuse of a right triangle and a and b are the legs of the triangle.

We thus have the following initial situation:



Our first step is to draw a square on the hypotenuse.



[1] Define $A_{small} \in \mathbb{R}$

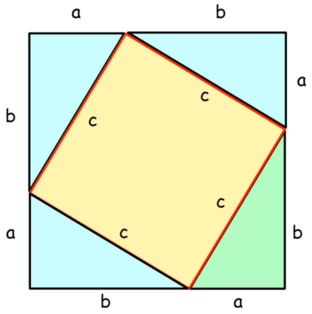
$\{c > 0$, so the small square is well-defined$\}$

A_{small} is the area of the square drawn on the hypothenuse

[2] {Area of square}

$A_{small} = c^2$

Then we draw three copies of the original right triangles around the square, so that the legs a and b are co-linear and coincide at each corner. The four triangles and the small square then form a larger square



[3] Define $A_{large} \in \mathbb{R}$

{the large square is well-defined by the construction}

A_{large} is the area of the square formed by the small square and the four right triangles.

[4] {Area of a square}

$A_{large} = (a + b)^2$

[5] Define $A_{triangle} \in \mathbb{R}$

{area of the original right triangle}

$A_{triangle}$ is the area of the original triangle

[6] {Area of right triangle}

$$A_{triangle} = \frac{a \cdot b}{2}$$

[7] {Calculate the area of the small square A_{small} using the areas of the large square and the triangle}

- A_{small}

= {the small square is the difference between the large square and the four triangles}

$A_{large} - 4 \cdot A_{triangle}$

= {inserting A_{large} and $A_{triangle}$ from the observations [4] and [6]}

$(a+b)^2 - 4 \cdot \frac{a \cdot b}{2}$

= {using the square rule and simplifying the last term}

$a^2 + 2 \cdot a \cdot b + b^2 - 2 \cdot a \cdot b$

= {simplify}

$a^2 + b^2$

□

... $A_{small} = a^2 + b^2$

⊩ {Observations [2] and [7] show that $a^2 + b^2 = c^2$}

□ ■

10.6 Assignments

1. Prove that there exists three consecutive natural numbers whose sum is 171.
2. Prove that $x^3 - x$ is divisible by 3, when $x \in \mathbb{N}$.
3. Determine when the expression $\sqrt{x^2-1} + \frac{\sqrt{10-x^2}}{\sqrt{x-2}}$ is defined.
4. Prove the logarithm rule $\log_a \frac{x}{y} = \log_a x - \log_a y$.
5. A line passes through the points $(5,\ 2,\ -1)$ and $(6,\ 4,\ -3)$. At which point does the line intersect the xz-plane?
6. Prove that the (generally false) formula $(x+y)^3 = x^3 + y^3$ only holds if $y = 0$, $x = 0$, both of the aforementioned or $x = -y$.
7. Provide an example of such numbers x and y that satisfy the (generally false) formula $(x-y)^3 = x^3 - y^3$, but not the formula from the previous task.
8. Prove that lg(25) is not a rational number.

Chapter 11

Proof Strategies

A *proof strategy* is a method for dividing a proof into smaller, more manageable parts. Mathematicians, logicians and philosophers have thought about different proof strategies for a long time. This work has resulted in a basic collection of *inference rules*, which summarize the most common proof strategies used in mathematical reasoning. We may consider these inference rules as *laws of thought*, or as an expression of *common sense*. Most people accept these inference rules as obviously correct. They form the basic toolkit for mathematicians, the mathematician's hammer, saw, screwdriver and wrench.

Case analysis is a typical example of an inference rule. We use case analysis to split a proof of a theorem into two subproofs, depending on whether a specific case A holds or not. The template for using case analysis is shown in Table 11.1.

•	C	
-	Φ	
$\Vdash$	$\{$*Case analysis*$\}$	
	•	C
	-	A
	•	C
	-	$\neg A$
□		

Table 11.1: Case analysis template

The template shows that we can prove that C follows from some set of assumptions $\Phi = A_1, \ldots, A_m$, by proving two subtask:

- prove that C follows from Φ under the additional assumption A, and
- prove that C follows from Φ under the additional assumption $\neg A$.

Note that Φ is inherited by the subtasks in the nested tasks, so we do not need to repeat the assumptions Φ for the subtasks in the template. Any logical proposition A can be used to split up the proof into two cases. The idea here is that either A is true or A is false. If C is true in both cases, then it must be always true. Example 58 shows how case analysis is used to solve a mathematical problem.

Traditionally, the case analysis inference rule would be written as follows:

$$\frac{\Phi, A \vdash C \qquad \Phi, \neg A \vdash C}{\Phi \vdash C} \; \{\textit{case analysis}\}$$

This rule allows us to construct a new mathematical fact $\Phi \vdash C$ (the *conclusion*) from mathematical facts that we already have proved, in this case from $\Phi, A \vdash C$ and $\Phi, \neg A \vdash C$ (the *hypothesis* or the *premises*). The conclusion is written below the horizontal bar and the assumptions above the bar. We write Φ, A for the set of assumptions $\Phi \cup \{A\}$. The name of the inference rule is indicated explicitly inside curly parenthesis.

We read this rule as follows: $\Phi \vdash C$, if both $\Phi, A \vdash C$ and $\Phi, \neg A \vdash C$. In words: C follows from Φ, if C follows from Φ and A, and C also follows from Φ and $\neg A$. The traditional notation for inference rules does not assume that the premises inherit the assumptions of the conclusion, so Φ has to be repeated in the premises.

The general form of an inference rule in classical notation is

$$\frac{\Phi_1 \vdash C_1, \ \ldots \qquad \Phi_n \vdash C_n}{\Phi \vdash C} \ \{R\}$$

where $n \geq 0$. Here $\Phi \vdash C$ is the *conclusion* of rule R, while $\Phi_1 \vdash C_1, \ldots, \Phi_n \vdash C_n$ are the *premises* of R. An inference rule without any premises (i.e., $n = 0$) is called an *axiom*.

A structured derivation template for a general inference rule looks as follows:

$$\begin{array}{lll} \bullet & C & \\ \text{-} & \Phi & \\ \Vdash & \{R\} & \\ & \bullet & C_1 \\ & \text{-} & \Phi_1 \\ & \vdots & \\ & \bullet & C_n \\ & \text{-} & \Phi_n \\ \square & & \end{array}$$

where $n \geq 0$. Subtasks inherit the assumptions of the main task, so rule R written in classical notation, is

$$\frac{\Phi, \Phi_1 \vdash C_1, \ \ldots \qquad \Phi, \Phi_n \vdash C_n}{\Phi \vdash C} \ \{R\}$$

The structured derivation symbol $\Vdash$ thus corresponds to the horizontal bar in the classical notation. The classical notation shows explicitly that the assumptions Φ of the conclusion are also assumptions for all the premises.

The classical notation for inference rules is more expressive, because it allows inference rules where premises do not inherit the assumptions of the conclusion. However, it turns out that we do not need this expressiveness in our framework. We base our inference rules on natural deduction, where the basic inference rules all can be expressed with inheritance, as shown in Table 11.2.

We have shown how to apply an inference rule in a backward proof without any observation or calculation steps. The inference rules can, however, just as well be used to prove that an observation is correct or to prove that a calculation step is correct. We illustrate this with the case analysis rule. To the left, we apply this rule in a reduction step, in the middle, we apply the rule to prove an observation and to the right we apply the rule to a calculation step (the proposition C is of the form $t \sim t'$ in a calculation). The conclusion of the rule is shown in blue, and the premises in red, to highlight the overall structure of the derivation.

Reduction:	Observation:	Calculation step:
• C - Φ $\Vdash$ {*Case analysis*} • C - A • C - $\neg A$ □	- Φ ⋮ + {*Case analysis*} • C - A • C - $\neg A$ ... C ⋮	- Φ ⋮ t $\sim$ {*Case analysis*} • $t \sim t'$ - A • $t \sim t'$ - $\neg A$ ... t' ⋮

The proof strategy templates show that the premises are proved as subtasks. In general, it may not be necessary to prove all premises as subtasks. Some of the premises may have been proved earlier as observations, in which case it is sufficient to refer to the observation in the explanation for the proof step. Some premises may be general mathematical theorems that we know are true from the context in which we carry out our proof (like the logical rules that we have described in the previous chapters). For these, it is sufficient to mention the names of the rules. In some cases, the proof of a premise is so simple that we can argue for it directly in the explanation. The remaining premises then need to be proved as subtasks.

Introduction rules :			
$T-$intro	$\dfrac{}{\Phi \vdash T}$	$\{Prove\, T\}$	
$\wedge$-intro	$\dfrac{\Phi \vdash A \quad \Phi \vdash B}{\Phi \vdash A \wedge B}$	$\{Prove\, \wedge\}$	
$\vee$-intro 1	$\dfrac{\Phi \vdash A}{\Phi \vdash A \vee B}$	$\{Prove\, \vee\}$	
$\vee -$ intro 2	$\dfrac{\Phi \vdash B}{\Phi \vdash A \vee B}$	$\{Prove\, \vee\}$	
$\Rightarrow$-intro	$\dfrac{\Phi, A \vdash B}{\Phi \vdash A \Rightarrow B}$	$\{Prove\ \ \Rightarrow\}$	
$\neg -$ intro	$\dfrac{\Phi, A \vdash F}{\Phi \vdash \neg A}$	$\{Prove\ \ \neg\}$	
Eliminination rules			
RAA	$\dfrac{\Phi, \neg A \vdash F}{\Phi \vdash A}$	$\{RAA\}$	$-$
Assumption	$\dfrac{A \in \Phi}{\Phi \vdash A}$	$\{Assumption\}$	
$\wedge$-elim 1	$\dfrac{\Phi \vdash A \wedge B}{\Phi \vdash A}$	$\{Use\, \wedge\}$	
$\wedge -$ elim 2	$\dfrac{\Phi \vdash A \wedge B}{\Phi \vdash B}$	$\{Use\, \wedge\}$	
$\vee$-elim	$\dfrac{\Phi \vdash A \vee B \quad \Phi, A \vdash C \quad \Phi, B \vdash C}{\Phi \vdash C}$	$\{Use\ \ \vee\}$	
$\Rightarrow$-elim	$\dfrac{\Phi \vdash A \Rightarrow B \quad \Phi \vdash A}{\Phi \vdash B}$	$\{Use\ \ \Rightarrow\}$	
$F -$ elim	$\dfrac{\Phi \vdash F}{\Phi \vdash A}$	$\{Use\ \ F\ \}$	
$\neg$-elim	$\dfrac{\Phi \vdash \neg A \quad \Phi \vdash A}{\Phi \vdash F}$	$\{Use\ \ \neg\}$	

Table 11.2: Natural deduction rules for connectives

11.1 Natural Deduction Rules

Gerhard Gentzen constructed in the 1940s a proof system for first order logic that he referred to as *natural deduction* [19, 32]. The name was chosen to indicate that the inference rules in this proof system correspond very closely to how mathematicians carry out proofs in practice. Natural deduction has become a de facto standard for presenting and analyzing mathematical proofs in logic. Gentzen divided the inference rules into two categories, *introduction rules* and *elimination rules.* Each basic logical connective has one or two introduction rules and one or two elimination rules, shown in Table 11.2. Equivalence is not seen as a basic connective, it is defined in terms of mutual implication, so there are no introduction or elimination rules for equivalence. We base our proof strategies for structured derivations on a *sequent calculus* formalization of natural deduction, as this corresponds directly to the structured derivation format that we use in this book (see the web pages [18] by Frade for a nice overview of logic with natural deduction).

The difference between introduction and elimination rules is that in the former, the logical connective occurs in the conclusion of the rule, while in elimination rules the connective is in the premisses. Introduction rules are useful when we want to prove a logical formula with a specific main connective, so they are often useful for backward/reduction proofs. Elimination rules are again useful when we want to see what conclusions we can derive from a logical formula with a certain main connective, so they are useful in forward proofs. Both introduction and elimination rules are needed for proving mathematical statements. It is also possible to use introduction rules in forward derivations, e.g., when proving an observation, and elimination rules in reduction proofs (the case analysis rule is, e.g., classified as an elimination rule, but it is often used in a reduction proof, as we showed above).

Besides the basic inference rules of natural deduction, there are also a few rules that are called *structural rules.* Because we consider the assumptions $A_1, \ldots, A_m$ in a sequent $A_1, \ldots, A_m \vdash C$ as forming a set rather than a list, we only have one structural rule, called *weakening:*

$$\frac{\Phi \vdash C}{\Phi, A \vdash C}\ \{\mathit{weakening}\}$$

This rule says that we are always allowed to add extra assumptions to a mathematical fact that we have proved, the fact remains valid. This is the only rule where the premise does not inherit the assumptions of the conclusion. However, the weakening rule is *adequate* in natural deduction. This means that any mathematical fact that can be proved with natural deduction using weakening can also be proved with natural deduction without weakening. In other words, the weakening rule is useful, but not strictly necessary.

We will in the following sections present a collection of common proof strategies, in the form of proof strategy templates. These are all based on the natural deduction rules.

11.2 Axioms

•	T	•	A
-	Φ	-	Φ
$\Vdash$	$\{$*Prove* $T\}$	$\Vdash$	$\{$*Use assumption*, $A \in \Phi\}$
□		□	

Table 11.3: Axioms

There are two axioms in natural deduction, *truth* and *assumption*. The first axiom says that T is provable from any set of assumptions, i.e., that T is always true. The second says that any assumption A that occurs in Φ is directly provable. The structured derivation templates for these two axioms are shown in Table 11.3.

11.3 Proof Strategies for Conjunction

•	$A \wedge B$	•	$A \wedge B$	•	A
-	Φ	-	Φ	-	Φ
$\Vdash$	$\{$*Prove* $\wedge\}$	$\Vdash$	$\{$*Prove* $\wedge$ $(2)\}$	$\Vdash$	$\{$*Use* $\wedge\}$
	• A		• A		• $A \wedge B$
	• B		• B	□	
□			- A		
		□			

Table 11.4: Proof strategies for conjunction

The strategies for proving a conjunction are shown in Table 11.4 (the first two inference rules). The first rule is the $\wedge$-intro rule from natural deduction: we prove that $A \wedge B$ is true by proving A and B separately. The second rule is somewhat more powerful: we first prove A, and then we prove B under the additional assumption that A is true. This rule is easily seen to be valid: it is ok to assume that A is true when proving B, because we have just proved A.

The remaining rule is one of the two $\wedge$ - elim rules of natural deduction. It says that we can deduce A from the assumption $A \wedge B$ (a similar template can be given for the other elimination rule). This rule is more useful in a forward proof, where the premises would first be proved in an observation, and the conclusion then being based on this observation, as shown in the template below.

• A

- Φ

⋮

[i] $\{$*justification*$\}$

$A \wedge B$

⋮

⊩ {*Use* $\wedge$, observation [i]}

□

11.4 Proof Strategies for Disjunction

The proof strategies for disjunction are shown in Table 11.5. The first introduction rules is one of the $\vee$ - intro rules. The second rule is a stronger version of the introduction rules: to prove $A \vee B$, it is sufficient to prove that B holds under the assumption that A is false. If A is true, then $A \vee B$ follows directly, so we only need to prove the case when A is false.

• $A \vee B$

- Φ

⊩ {*Prove* $\vee$}

  • A

□

• $A \vee B$ - Φ $\Vdash$ $\{Prove \vee\}$ • A □	• $A \vee B$ - Φ $\Vdash$ $\{Prove \vee (2)\}$ • B - $\neg A$ □	• C - Φ $\Vdash$ $\{Use \vee\}$ • $A \vee B$ • C - A • C - B □

Table 11.5: Proof strategies for disjunction

The third rule is the ∨ - elim rule, which we usually call the case rule. We can derive a general rule for case analysis from this basic rule. The general rule allows us to consider a number of different cases at the same time:

• C

- Φ

$\Vdash$ $\{Case\ analysis\}$

    • $A_1 \vee \ldots \vee A_n$
    • C
    - A_1
    ⋮
    • C
    - A_n

□

The general case rule requires that we prove that C is true for each case A_1,..., A_n. In addition, we need to prove that one of these cases is always true under the assumptions Φ, i.e. that $A_1 \vee \ldots \vee A_n$ follows from Φ. The following example shows how to use the general case analysis rule. The corresponding template entities are indicated in red.

Example 64. We prove that

$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$$

• Show that $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$, when — C

- $a, b \in \mathbb{R}$ and $b \neq 0$ —Φ

⊩ {Case analysis, four difference cases, depending on whether a and b are negative or non-negative.}

  • Show that $(a \geq 0 \wedge b \geq 0) \vee (a \geq 0 \wedge b < 0) \vee (a < 0 \wedge b \geq 0) \vee (a < 0 \wedge b < 0)$ —$A_1 \vee A_2 \vee A_3 \vee A_4$

  ⊩ {obvious}

  □

  • Show that $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$ —C

  - $a \geq 0 \wedge b \geq 0$ —A_1

  ⊩ $\left|\frac{a}{b}\right|$

  = {from the assumption follows that $\frac{a}{b} \geq 0$}

  $\frac{a}{b}$

  = {assumption}

  $\frac{|a|}{|b|}$

  □

  • Show that $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$ —C

  - $a \geq 0 \wedge b < 0$ — A_2

  ⊩ $\left|\frac{a}{b}\right|$

  = {from the assumption follows that $\frac{a}{b} < 0$}

  $-\frac{a}{b}$

  = {arithmetics}

  $\frac{a}{-b}$

  = {assumption}

  $\frac{|a|}{|b|}$

□

• Show that $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$ — C

- $a < 0 \wedge b \geq 0$ — A_3

• Show that $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$ — C

- $a < 0 \wedge b < 0$ — A_4

□

We treat the last two cases in the same way.

11.5 Proof Strategies for Negation

• $\neg A$	• A	• F
- Φ	- Φ	- Φ
$\Vdash$ $\{Prove\ \neg\}$	$\Vdash$ $\{RAA\}$	$\Vdash$ $\{Prove\ F\}$
• F	• F	• A
- A	- $\neg A$	• $\neg A$
□	□	□

Table 11.6: Proof strategies for negation

The proof strategies for negation are shown in Table 11.6. The first rule is the $\neg$ intro rule from natural deduction: to prove $\neg A$, it is sufficient to prove that A leads to a contradiction (i.e., that F is provable from A). The second rule is the *proof by contradiction* rule, also known as *reductio ad absurdum*: we prove that a proposition A is true by proving that the assumption $\neg A$ is contradictory. This rule follows from the first rule in classical logic, when we assume that $\neg(\neg A) \equiv A$. The intuitionistic approach to mathematics does not accept proofs by contradiction as real proofs, but in classical mathematics (that most mathematicians follow) this is no problem. Still, even classical mathematicians tend to prefer constructive proofs of a proposition over proofs by contradiction.

To prove a contradiction, we have to prove that F is true (under the given assumptions), so we need an inference rule for this too. This is provided by the third rule in the table. If we can prove that both B and $\neg B$ are true, then we have proved

$B \wedge \neg B$ (by the inference rule for conjunction). Since $B \wedge \neg B \equiv F$, we have proven F.

Proving F may feel funny, since we know that F is not true. What is important here is to understand that we do not prove that F is true in some absolute sense, but we show that F follows from some set $A_1, \ldots, A_n$ of assumptions (i.e., that $A_1, \ldots, A_n \vdash F$). This is possible only if the assumptions themselves are contradictory, i.e., $A_1 \wedge \ldots \wedge A_n \equiv F$.

Example 65. A classic example of using reductio ad absurdum is the proof that $\sqrt{2}$ is an irrational number. The proof derives a contradiction from the negation of the proposition, which is that $\sqrt{2}$ is rational and can be written as a fraction $\frac{a}{b}$. The following is the first step of the proof:

• $\sqrt{2}$ is an irrational number. — A

⊩ {Reductio ad absurdum}

  • Prove F — F

  - $\sqrt{2}$ is a rational number — $\neg A$

□

Note that Φ is empty here, i.e. there are no assumptions.

In the next step we add the proof of the subtask, i.e. we prove that F follows from the assumption that $\sqrt{2}$ is a rational number.

• $\sqrt{2}$ is an irrational number.

⊩ {Reductio ad absurdum}

  • Prove F

  - $\sqrt{2}$ is a rational number

  + $a, b \in \mathbb{N}$

  {We introduce the integers a and b to describe $\sqrt{2}$ as a rational number. By the assumption, two such numbers always exist}

  $\sqrt{2} = \dfrac{a}{b}$

  ⊩ $\sqrt{2} = \dfrac{a}{b}$

  ⇒ {square both sides}

  $2 = \dfrac{a^2}{b^2}$

  ≡ {multiply both sides by b^2}

  $2b^2 = a^2$

$\equiv$ {the prime factor 2 appear an odd number of times on the LHS and an even number of times on the RHS, which is impossible}

F

□

□ ■

11.6 Proof Strategies for Implication

• $A \Rightarrow B$ - Φ $\Vdash$ {*Prove* $\Rightarrow$} • B - A □	• $A \Rightarrow B$ - Φ $\Vdash$ {*Indirect proof*} • $\neg A$ - $\neg B$ □	• B - Φ $\Vdash$ {*Use* $\Rightarrow$} • $A \Rightarrow B$ • A □	• B - Φ $\Vdash$ {*Lemma rule*} • A • B - A □

Table 11.7: Proof strategies for implication

Most mathematical theorems are implications: $A \Rightarrow B$ says that B is true whenever A is true. The strategies for proving an implication are shown in Table 11.7. The most common way of showing implication $A \Rightarrow B$ is to show that that B holds under the assumptions A. This is referred to as a *direct proof*: we start from the assumptions A and derive the conclusion B. This is the $\Rightarrow$- intro rule, also known as the *deduction theorem*.

Another way to prove implication is with an *indirect proof* (second rule in the table). Instead of proving that $A \Rightarrow B$ holds, we assume $\neg B$ and prove $\neg A$. This proof rule has the traditional name *modus tollendo tonens*. It is based on the equivalence $(A \Rightarrow B) \equiv (\neg B \Rightarrow \neg A)$.

Example 66. We give an example of an indirect proof. We want to prove that if $n^3 + 5$ is an odd number, then n is an even number, for every integer n.

• Show that $odd(n^3 + 5) \Rightarrow even(n)$ — $A \Rightarrow B$

- $n \in \mathbb{Z}$ — Φ

$\Vdash$ {Indirect proof, $\neg even(n) \equiv odd(n)$}

• Show that $even(n^3 + 5)$ — $\neg A$

- $odd(n)$ — $\neg B$

+ $k \in Z$

  {k is well defined, since n is odd}

  $n = 2k + 1$

+ {Calculate $n^3 + 5$}

  • $n^3 + 5$

  = {definition}

  $(2k + 1)^3 + 5$

  = {expand the parenthesis}

  $8k^3 + 12k^2 + 6k + 6$

  = {factor out 2}

  $2(4k^3 + 6k^2 + 3k + 3)$

  □

... $n^3 + 5 = 2(4k^3 + 6k^2 + 3k + 3)$

⊩ {$n^3 + 5$ is even, since we can write it with a factor 2}

□

□ ■

Implication is also transitive, so we can prove an implication with a calculation, similarly to how we prove equivalence.

• $A_0 \Rightarrow A_n$

- Φ

⊩ {$\Rightarrow$ *is transitive*}

  A_0

$\Rightarrow$ {$justification_1$}

  A_1

$\vdots$

  A_{n-1}

$\Rightarrow$ {$justification_n$}

  A_n

□

A special case of proving implication with transitivity is proving that the proposition A is true by proving $T \Rightarrow A$. We discussed this way of proving propositions earlier, when we discussed how to construct proofs as logic calculations.

The third rule in the table is the $\Rightarrow -elim$ rule, also known as *modus ponens*. This rule allows us to conclude that B is true, when $A \Rightarrow B$ is true and A is also true. The most common use of modus ponens in practical mathematical proofs is when we want to apply a mathematical theorem that we have proven earlier. The theorem is often conditional, of the form $A \Rightarrow B$, where A enumerates the assumptions that are necessary for the proposition B to be true. We then apply the rule by showing that the assumptions are satisfied, so that we may use B.

Combining modus ponens and direct proofs gives us a very useful strategy for proving mathematical theorems: the use of *lemmas*. This is the fourth rule in the table above. We want to prove a proposition B. In order to do this, we first prove another proposition A (a *lemma*), and then prove the original proposition B using the new proposition A as an additional assumption.

We can prove the lemma A in a nested derivation, but more often we prove the lemma as an observation, which we then use in the proof of B. The use of observations in proofs is, in fact, just an application of the lemma rule. This is shown in the figure below.

• B

- Φ

$+$ $\{justification\}$

A

$\Vdash$ $\{Lemma\ rule\}$

    • B

    - A

□

11.7 Proof Strategies for Equivalence

Natural deduction does not have special rules for equivalence, as equivalence is defined in terms of the other connectives. However, equivalence is a central connective in practice, and we need special rules for proving properties about equivalences. Table 11.8 shows some of the proof strategies for dealing with connectives. Note that equivalence in these inference rules occurs both in the conclusion and in the premises, so the rules cannot be classified as either introduction or elimination rules. Rather, they can be seen as *congruence* rules, i.e., showing under what conditions equivalence is preserved in an inference.

• $A \equiv C$	• $A \equiv B$
- Φ	- Φ
$\Vdash$ $\{\equiv$ *is transitive*$\}$	$\Vdash$ $\{$*Mutual implication*$\}$
• $A \equiv B$	• $A \Rightarrow B$
• $B \equiv C$	• $B \Rightarrow A$
□	□

Table 11.8: Proof strategies for equivalence

The most common way to prove a proposition of the form $A \equiv B$ is to use the first rule in the table, which says that equivalence is a transitive relation. The transitivity rule does not only apply to equivalence, it applies to every transitive relation ($=$, $\leq$, $<$, etc.).

Another common way to prove an equivalence $A \equiv B$ is shown in the second rule: we prove implication in both directions, $A \Rightarrow B$ and $B \Rightarrow A$.

Another approach is to use the context when proving equivalence, as shown in Table 11.9. The first rule shows that a proposition of the form $C \wedge A \equiv C \wedge B$ holds if we can prove that $A \equiv B$ holds under the additional assumption C. We can easily see that this rule is valid. If C is false, it is trivial to see that the equivalence holds ($C \wedge A \equiv F \wedge A \equiv F \equiv F \wedge B \equiv C \wedge B$). Thus we only need to show that $A \equiv B$ holds when C is true.

• $C \wedge A \equiv C \wedge B$	• $C \vee A \equiv C \vee B$
- Φ	- Φ
$\Vdash$ $\{\wedge$ *context for* $\equiv\}$	$\Vdash$ $\{\vee$ *context for* $\equiv\}$
• $A \equiv B$	• $A \equiv B$
- C	- $\neg C$
□	□

Table 11.9: Proving equivalence in context

There is a corresponding rule for disjunction, but there the additional assumption is $\neg C$. The proof of this rule is analogous to the previous one: the equivalence holds trivially if C is true, so we just need to check the equivalence when C is false.

Finally, we have *Leibniz' rule.* It states that if we have a proposition P that contains a subexpression a (we write this as $P[a]$), and we substitute b for a, the new expression $P[b]$ is equivalent to the original expression, if $a = b$. In other words, we can always replace a term in a logical expressions with another, equal term. Leibniz rule is shown in Table 11.10. To the right we give the same rule for the case when the subexpression is a logic expression.

•	$P[a] \equiv P[b]$	•	$P[A] \equiv P[B]$
-	Φ	-	Φ
$\Vdash$	{*Leibniz' rule for* $=$}	$\Vdash$	{*Leibniz' rule for* $\equiv$}
	• $a = b$		• $A \equiv B$
□		□	

Table 11.10: Leibniz' rule

Leibniz' rule is the rule most commonly used in calculations. We can illustrate the rule by the following calculation step (the subexpression is written in green):

$x^2 + 2(x + y) + y^2 \leq 0$ — $P[a]$

$\equiv$ {Leibniz' rule}

• $2(x + y)$ — a

$=$ {the distributive law}

$2x + 2y$ — b

□

... $x^2 + 2x + 2y + y^2 \leq 0$ — $P[b]$

We usually don't mention Leibniz' rule in this kind of situations, we just write the result directly and only refer to the reason for why $a = b$:

$x^2 + 2(x + y) + y^2 \leq 0$

$\equiv$ {the distributive law}

$x^2 + 2x + 2y + y^2 \leq 0$

In most cases, this is enough for the reader to understand what we have done, and which general rule we have used.

Example 67. Let us look anew at Example 1, where we computed $\tan \frac{17\pi}{3}$. We rewrite the justifications for the derivation so that we explicitly mention each time we use Leibniz's rule.

$\bullet \qquad \tan \dfrac{17\pi}{3}$

$= \qquad \{$Leibniz's rule: $\dfrac{17\pi}{3} = \dfrac{6 \cdot 2\pi + 5\pi}{3}\}$

$\qquad \tan \left(\dfrac{6 \cdot 2\pi + 5\pi}{3} \right)$

$= \qquad \{$Leibniz's rule: $\dfrac{6 \cdot 2\pi + 5\pi}{3} = 2 \cdot 2\pi + \dfrac{5\pi}{3}\}$

$\qquad \tan \left(2 \cdot 2\pi + \dfrac{5\pi}{3} \right)$

$= \qquad \{$we can ignore full circles $2 \cdot 2\pi\}$

$\qquad \tan \dfrac{5\pi}{3}$

$= \qquad \{$Leibniz's rule: $5\pi = 2\pi - \dfrac{\pi}{3}\}$

$\qquad \tan \left(2\pi - \dfrac{\pi}{3} \right)$

$= \qquad \{$the tangent is negative in the 4th quadrant and $\frac{\pi}{3}$ is a special right triangle$\}$

$\qquad -\tan \dfrac{\pi}{3}$

$= \qquad \{$the tangent of a 30 - 60 - 90 triangle$\}$

$\qquad -\sqrt{3}$

□

■

11.8 Assignments

1. Prove the modus ponens rule, i.e., that if p and $p \Rightarrow q$ are true, then q is true.
2. Prove that if $\neg p \Rightarrow q$, $q \Rightarrow p$ and $p \Rightarrow (q \wedge r)$ are true, then $p \wedge q \wedge r$ is also true.
3. Prove $p \Rightarrow (q \wedge r) \vdash (\neg q \vee \neg r) \Rightarrow \neg p$
4. Prove that the formula $p \Rightarrow q$ is a logical consequence of

$$(t \Rightarrow q) \wedge (\neg r \Rightarrow \neg s) \wedge (p \Rightarrow u) \wedge (\neg t \Rightarrow \neg r) \wedge (u \Rightarrow s)$$

5. Prove that it isn't necessary for $p \Rightarrow q$ to hold in order to show that $(p \Rightarrow r) \Rightarrow (p \wedge q \Rightarrow r)$.

6. Prove the absolute value formula $a \leq |a|$ by using the definition

$$a = |b| \Leftrightarrow (a = b \wedge b \geq 0) \vee (a = -b \wedge b < 0)$$

7. Determine whether it is true that if

$$(((q \wedge s) \Rightarrow r) \wedge ((r \wedge p) \Rightarrow s) \wedge q)$$

   then $r \Rightarrow (s \vee q)$.

8. Prove $|x|^2 = x^2$ without using any other rule for absolute values than the definition

$$|a| = b \Leftrightarrow (a = b \wedge a \geq 0) \vee (-a = b \wedge a < 0)$$

   (or the equivalent $(a = b \wedge -a = b) \wedge b \geq 0$).

Chapter 12

Stepwise Refinement of Derivations

We have shown above how to use different proof strategies when solving mathematical problems. A reasonable question is how structured tasks compare to ordinary, more or less informal proofs of mathematical properties. On one hand, the proofs seem to become longer and more detailed when presented as structured derivations. On the other hand, they are more precise, because the structured derivation format requires that each proof step is justified explicitly, and the logical structure of the proof is shown explicitly, rather than being wrapped up in prose. What is the balance, when is it useful to use a structured derivation rather than an informal proof. There is probably not one right answer to this, as different people have different preferences, skills and senses of esthetics. Some people want the proof to be very precise and exact, while others prefer elegant proofs, where the basic ideas are conveyed in a compressed, but still understandable way.

We will propose here an alternative reason for using structured derivation, in addition to just solving a problem directly in this format. The alternative application is to work out a given proof in full detail, in order to understand it better. We call this *stepwise refinement* of a proof, because the process resembles the established method of stepwise refinement of programs. In short, this method starts from a standard mathematical proof presentation, and turns this into a structured derivation, step by step. The following example shows how to do this in practice. We write the part of the original proof that we have not yet analyzed more carefully in blue, while the new part is written in read. The rest of the task is written in black.

12.1 Initial version

We illustrate stepwise refinement of proofs by the example we already presented earlier, that $\sqrt{2}$ is an irrational number. The informal proof is as follows.

Theorem: $\sqrt{2}$ is an irrational number.

Proof: We prove that $\sqrt{2}$ is irrational by showing that the counter assumption leads to a contradiction. Assume therefore that $\sqrt{2}$ is a rational number. This means that we can write $\sqrt{2}$ as a quotient of two natural numbers p and q, $\sqrt{2} = \frac{p}{q}$. Squaring this gives us that $2 = \frac{p^2}{q^2}$, so $p^2 = 2q^2$. There is a unique factorization of p and q, with each having a specific number of factors 2. Then p^2 must have twice as many factors 2 as p , and similarly for q^2 and q. This means that p^2 has an even number of factors 2, while $2 \cdot q^2$ has an odd number of factors 2. This is a contradiction, so $\sqrt{2}$ must be irrational. □

12.2 Second version

We next write this as a structured task, where we have identified the task, but have left the proof itself unchanged, as the justification for the theorem.

• $\sqrt{2}$ is an irrational number

$\Vdash$ {We show that the counter assumption leads to a contradiction. Assume therefore that $\sqrt{2}$ is a rational number. This means that we can write $\sqrt{2}$ as a quotient of two natural numbers p and q, $\sqrt{2} = \frac{p}{q}$. Squaring this gives us that $2 = \frac{p^2}{q^2}$, so $p^2 = 2q^2$. There is a unique factorization of p and q, with each having a specific number of factors 2. Then p^2 must have twice as many factors 2 as p , and similarly for q^2 and q. This means that p^2 has an even number of factors 2, while $2 \cdot q^2$ has an odd number of factors 2. This is a contradiction, so $\sqrt{2}$ must be irrational.}

□

12.3 Third version

Next, we notice that the proof starts by applying the reductio ad absurdum principle. We lift this out on the structured task level, keeping the rest of the proof as informal justifications in curly brackets.

• $\sqrt{2}$ is an irrational number

⊩ {Reductio ad absurdum}

  • F

  - $\sqrt{2}$ is a rational number

  ⊩ {The assumption means that we can write $\sqrt{2}$ as a quotient of two natural numbers p and q, $\sqrt{2} = \frac{p}{q}$. Squaring this gives us that $2 = \frac{p^2}{q^2}$, so $p^2 = 2q^2$. There is a unique factorization of p and q, with each having a specific number of factors 2. Then p^2 must have twice as many factors 2 as p , and similarly for q^2 and q. This means that p^2 has an even number of factors 2, while $2 \cdot q^2$ has an odd number of factors 2. This is a contradiction}

  □

□

12.4 Fourth version

Our next version is to introduce the constants p and q by definitions, as is done in the informal proof.

• $\sqrt{2}$ is an irrational number

⊩ {Reductio ad absurdum}

  • F

  - $\sqrt{2}$ is a rational number

  + $p, q \in \mathbb{N}$

  {p and q are well defined, because any rational number can be written as a quotient of two natural numbers}

  $\sqrt{2} = \frac{p}{q}$

  ⊩ {Squaring gives us that $2 = \frac{p^2}{q^2}$, so $p^2 = 2q^2$. There is a unique factorization of p and q, with each having a specific number of factors 2. Then p^2 must have twice as many factors 2 as p , and similarly for q^2 and q. This means that p^2 has an even number of factors 2, while $2 \cdot q^2$ has an odd number of factors 2. This is a contradiction.}

  □

□

12.5 Fifth version

We are now ready to start calculating the truth value of the counterassumption. This gives us the following task:

• $\sqrt{2}$ is an irrational number

$\Vdash$ {Reductio ad absurdum}

> • F
>
> \- $\sqrt{2}$ is a rational number
>
> $+$ $p, q \in \mathbb{N}$
>
> {p and q are well defined, because any rational number can be written as a quotient of two natural numbers}
>
> $\sqrt{2} = \dfrac{p}{q}$
>
> $\Vdash$ $\sqrt{2} = \dfrac{p}{q}$
>
> $\equiv$ {squaring}
>
> $p^2 = 2q^2$
>
> $\equiv$ {There is a unique factorization of p and q, with each having a specific number of factors 2. Then p^2 must have twice as many factors 2 as p , and similarly for q^2 and q. This means that p^2 has an even number of factors 2, while $2 \cdot q^2$ has an odd number of factors 2. This is a contradiction}
>
> F
>
> □

□

12.6 Sixth version

Finally, we will make the last reasoning more explicit, by working out the exact way in which the unique factorization theorem is applied in this proof. This gives us the final proof.

• $\sqrt{2}$ is an irrational number

⊩ {Reductio ad absurdum}

  • F

  - $\sqrt{2}$ is a rational number

  + $p, q \in \mathbb{N}$

    {p and q are well defined, because any rational number can be written as a quotient of two natural numbers}

    $\sqrt{2} = \frac{p}{q}$

  + $k, r \in \mathbb{N}$ {Unique factorization theorem, k factors 2 in p}

    $p = 2^k \cdot r \wedge \neg even(r)$

  + $m, s \in \mathbb{N}$ {Unique factorization theorem, m factors 2 in q}

    $q = 2^m \cdot s \wedge \neg even(s)$

  + {Squaring}

    $p^2 = 2^{2k} \cdot r^2 \wedge \neg even(r^2\}$

  + {Squaring}

    $q^2 = 2^{2m} \cdot s^2 \wedge \neg even(s^2\}$

  ⊩ $\sqrt{2} = \frac{p}{q}$

  ≡ {squaring}

    $p^2 = 2q^2$

  ≡ {observations}

    $2^{2k} \cdot r^2 = 2 \cdot 2^{2m} \cdot s^2$

  ≡ {observations}

    $2^{2k} \cdot r^2 = 2^{2m+1} \cdot s^2$

  ≡ {The left hand term has an even number of factors 2, while the right hand term has an odd number of factors 2. This contradicts the unique factorization theorem, which says that each number has a unique factorization, and hence also a unique number of factors 2}

    F

  □

□

This completes our derivation.

12.7 What Have We Gained

We started with an informal description of the proof, and have reformulated this as a structured task. The motivation for this exercise was to clarify for ourselves what exactly is the structure of the mathematical argument that establishes that $\sqrt{2}$ is an irrational number, and trying to check that there are no errors in the proof. The final proof is in fact more detailed than the original, because the last step makes the argument for how the unique factorization theorem is applied more precise.

It might seem that presenting a proof as a structured derivation makes it much longer. However, if we compare the original informal proof with the fifth version (which is closest to the original presentation), then we see that the original proof has 534 non-blank characters, while the fifth version has 448 non-blank characters. The structured task is actually shorter than the original proof formulation. The original proof looks more condensed, because it has only 9 lines of text, whereas the structured task has 17 lines. The longer format is the price we pay for making the structure of the proof explicit. This is the same kind of price we pay in programming, when we indicate the program structure with formatting. For longer proofs (and programs), this explicit aid to understanding the structure is more or less indispensable.

Chapter 13

Word Problems

We have in previous chapters showed how to formulate and solve mathematical tasks. Let us now take a step back, and think about how to solve problems that arise in the real world, using mathematics. In school mathematics, this kind of problems are known as *word problems.* We will in this chapter look at how to formulate word problems as tasks, and how to interpret the answers as solutions to the original problem.

13.1 Word Problems as Tasks

We can identify four distinct steps when solving a real-world problem:

1. We start by *analyzing* the *informal description* of the problem: what are the relevant quantities in the problem context, what is the question, and what assumptions are we allowed to make. We then *formulate* the informal problem as a *mathematical problem.*

2. We then *solve* the mathematical problem to get a *mathematical answers* to the problem.

3. Next, we *interpret* the answer in the informal context of the original problem.

4. Finally we *evaluate* the solution, to see whether it is a correct or at least a reasonable solution to the original problem.

We summarize this process in Figure 13.1 taken from a Pisa study [31]. We can reiterate this process a number of times, as long as our evaluation shows that the answer is incorrect or implausible.

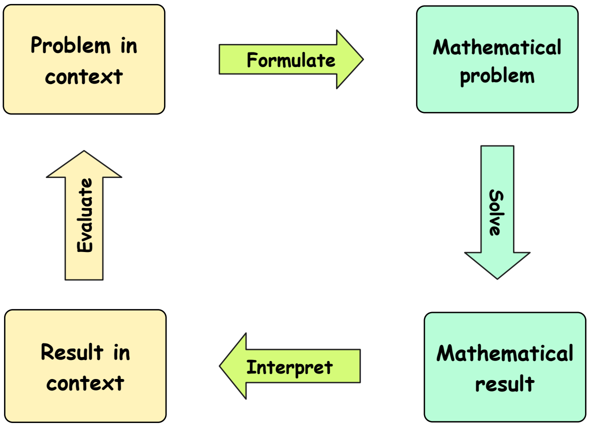


Figure 13.1: Solving real-world problems

We rephrase this model in our framework as shown in Figure 13.2. The informal problem statement is first formulated as the problem of a task, with a question and assumptions. We then complete the task with observations and calculations that lead to an answer to the task question. Having found a mathematical answer, we interpret it in the context of the original problem, to get a solution to the problem. Finally, we try to check whether the solution is correct. If correctness is too difficult to check, we may at least try to see whether the solution is reasonable.

A structured task does not, by itself, have any direct links to the problems in the real world that we are trying to solve. In practice, it is, however, important that we can show how the mathematical problem statement is related to the informal problem statement. We describe this relationship with *comments* that we add to a derivation. A comment is an arbitrary text which can be added to the end on any line in a derivation (in the second column), and is preceded with "—". The comment can continue into the next line (in the second column). We can also have a line with only a comment. The comments have no relevance for the solution of the mathematical problem, i.e., we can omit all comments without changing the meaning of the derivation. We see comments as something that is outside the actual syntax for structured derivations. A comment explains what the question, assumptions or answer of the task mean in the original problem context.

Example 68. A holiday package to Madeira consists of hotel and travel expenses. The cost of the hotel had decreased by 5% since last year, while the travel expenses have increased by 18%. The price of the entire package is still the same as last year. Calculate how many percentages of the total price of last year's package went to travel expenses.

We start by formulating the problem mathematically. First we identify which entities appear in the problem. We introduce symbols for last year's hotel expenses (x) and last year's travel expenses (y), as well as this years hotel expenses (x') and this years travel expenses (y'). Then we can specify the task and the assumptions. The

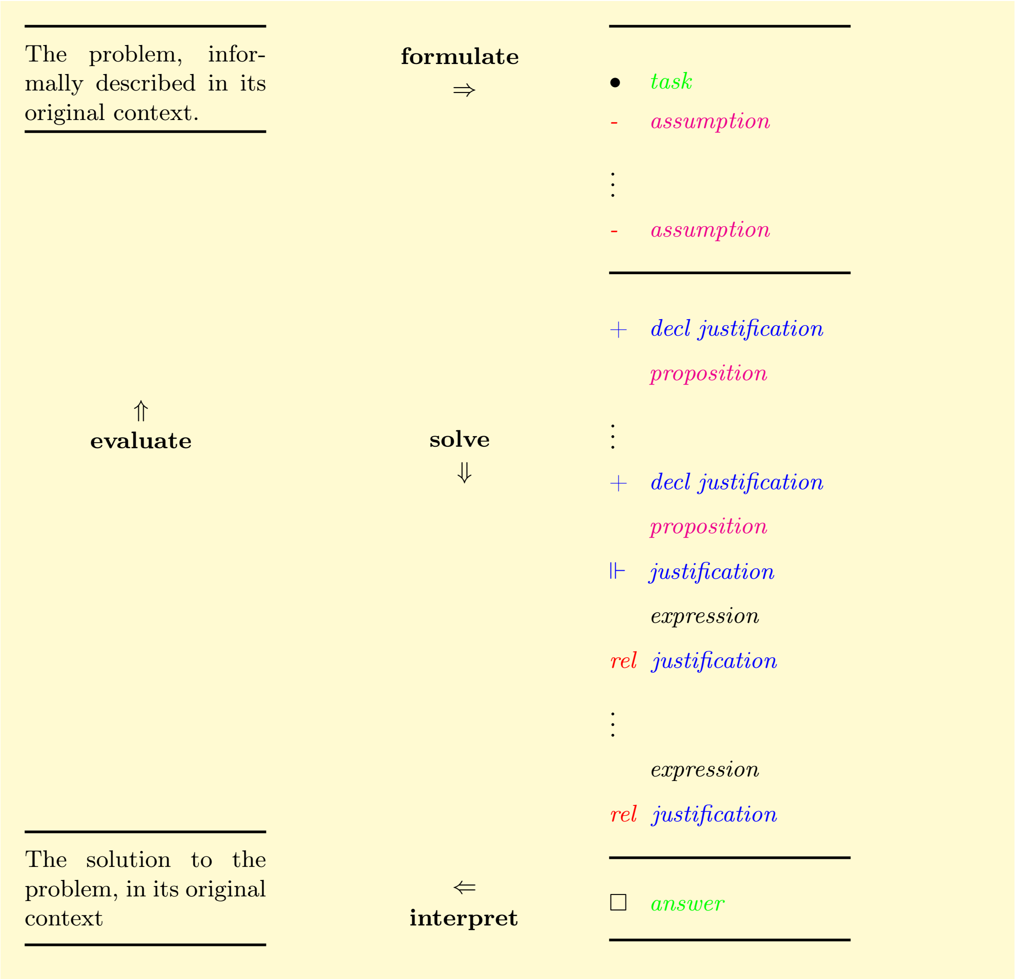


Figure 13.2: Modeling with a structured task

meaning of these entities are explained in comments. For clarity, we have colored all comments blue. We now have the following mathematical formulation of the problem:

- • How many percent is y of $x + y$ — how many percent is last year's travel expenses of the entire holiday package last year
- - $x \in \mathbb{R}^+$ — last year's hotel expenses
- - $y \in \mathbb{R}$ — last year's travel expenses
- - $x' \in \mathbb{R}^+$ — this year's hotel expenses
- - $y' \in \mathbb{R}^+$ — this year's travel expenses

(a) x' is 5 % less than x — this year's hotel expenses are 5 % less than last year

(b) y' is 18 % greater than y — this year's travel expenses are 18 % greater than last year

(c) $x' + y' = x + y$ — the price of the entire holiday package this year is the same as last year

The comments give the connection between the mathematical formulation of the problem and the original informal problem formulation. The quantities that we observe are introduced as variable names, with assumptions about their value domain. These assumptions are not numbered, since we can refer to them directly by name.

Our first observation shows that we can describe this year's hotel and travel expenses using last year's hotel and travel expenses.

[1] {We describe this year's hotel and travel expenses using last year's hotel and travel expenses, assumptions (a) and (b)}

$x' = 0.95x$ and $y' = 1.18y$

The next step is to use assumption (c), which says that the total cost this year is the same as last year.

[2] {Calculate y using assumption (c)}

- • $x + y = x' + y'$
- $\equiv$ {observation [1]}
- $x + y = 0.95x + 1.18y$
- $\equiv$ {subtract x from both sides}
- $y = -0.05x + 1.18y$
- $\equiv$ {subtract $1.18y$ from both sides}
- $-0.18y = -0.05x$
- $\equiv$ {divide both sides by -0.18}
- $y = \frac{0.05}{0.18}x$
- □

… $y = \frac{5}{18}x$

We can now solve the original problem, i.e., calculate how many percent of the entire holiday package went towards travel expenses last year:

$\Vdash$ $\dfrac{y}{x+y}$

$=$ {observation [2]}

$\frac{\frac{5}{18}x}{x+\frac{5}{18}x}$

= {cancel out x}

$\frac{\frac{5}{18}}{1+\frac{5}{18}}$

= {simplify}

$\frac{5}{23}$

≈ {calculate an approximate value}

0.2173913

≈ {round and write as a percentage}

21.7%

□ 21.7% — last years travel expenses were 21.7% of the entire holiday package

We have now an answer to the mathematical problem, 21.7%, which is also the solution to the original problem: last year's travel expenses were 21.7% of the entire holiday package.

Let us finally check if the answer is correct, or at least reasonable. We calculate the share of travel expenses this year.

• travel expenses share of total package this year

= {assumptions}

$\frac{y'}{x'+y'}$

= {assumption (c), observation [1]}

$\frac{1.18y}{x+y}$

= {observation [2]}

$\frac{1.18y}{\frac{18}{5}y+y}$

= {simplify}

$\frac{5 \cdot 1.18}{23}$

≈ {calculate}

25.6%

□

Since the total cost of the package is the same as last year, we see that the increase in travel expenses is

$$\frac{25.6 - 21.7}{21.7} \approx 18\%$$

Thus, our answer seems to be correct. ■

13.2 Is the Solution Correct

How do we know that our solution to the real-world problem is correct? And what does it mean to have a correct solution to a real-world problem? It turns out that there are two different notions of correctness that are relevant here.

- *Correctness of the mathematical solution*: Is the answer we obtain in the structured derivation the correct answer to the question given in the task? In other words: *have we solved the problem right?*
- *Correctness of the problem solution*: Is the solutions that we extracted from the answer the correct solution to original problem? In other words: *have we solved the right problem?*

These are two different things. Consider what can go wrong when we solve a word problem:

- The mathematical formulation of the informal problem may be wrong. We may have misunderstood what to do, or we have formulated an assumption incorrectly, omitted an assumption or added an assumption that is not included in the original problem. This means that we have solved the wrong problem. The informal problem formulation could also be ambiguous, so it may not even be possible to know exactly what the problem is. We can try to avoid errors in formulating the mathematical problem by carefully comparing the original problem text with the mathematical problem.
- The mathematical answer may be wrong. The problem has been formulated correctly, but we have made one or more errors in deriving the answer. In other words, we have a wrong solution to the mathematical problem. This is something that we can avoid by carefully going through the derivation itself, checking the justification of each step and verifying that the step is correct.
- The solution may have been interpreted wrongly. We have formulated the problem correctly, and the mathematical answer is also correct, but we have misinterpreted the answer. This means that we give an incorrect solution to the original task, even though everything was correct almost to the end.

- Finally, we may have evaluated the solution incorrectly. Maybe our solution is correct, but we do not realize it, or maybe it is incorrect, but we manage to convince ourselves that the solution is correct.

So there is a variety of things to take into account when giving a mathematical solution to a real world problem. And many of these errors can occur outside the realm of mathematics, in the formulation of the problem, the interpretation of the answer, or the plausibility check.

13.3 Assignments

1. The friends Amin, Anne, Ada and Arthur are eating pizza. Anne eats half a pizza and Amin eats two thirds of the amount that Anne eats. Ada is not as fond of the pizza, so she only eats half of what Amin ate. Arthur on the other hand is quite hungry, consuming three times the amount that Anne ate. How much pizza did the four friends devour in total?

2. An artist sketches a human body and contemplates the proportions of the body. She remembers from her time at the Academy of Arts that the head is $\frac{2}{15}$ of the entire body and the distance from the nose to the crown is $\frac{1}{2}$ of the height of the head (from the crown to the chin). Moreover the distance between the mouth and the nose should be $\frac{1}{2}$ of the distance between the chin and the nose. Help the artist calculate, what is the distance between the mouth and the crown as a fraction of the height of the body.

3. Anne in mixing some juice for herself. She should mix the juice at a ratio of $1:4$, but she makes a minor mistake. She mixes 1.5 dl concentrate with water and gets 7 dl of mixed juice. At what proportions did Anna mix the juice?

4. You are determined to prove to the world that the number of times that a deck of cards can be ordered in is not really that large. Armed with the folly of youth you set forth going through the different ways that a deck of cards can be ordered in. Assume (in a manner patently hostile to reality) that you can arrange a pack of cards at a rate of one per second and that you never need to eat, sleep or drink. Would you have covered all of the permutations after a year?

5. A company that specializes in hot chili sauces wants to create a sauce with a specific strength, but the company only has two pepper mixtures. One mixture is 15% weaker than the desired strength, while the other mixture it 20% stronger than the desired strength. At what ratio should these mixtures be used to get the desired strength?

6. A shop owner raises the price of his doughnuts by 13%. This results in a 13% drop in sales. Did the increase in price pay off?

7. In 1883, the volcano Krakatoa had an eruption that released an amount of energy equivalent to 150 megatons of dynamite, i.e. $6.3 \cdot 10^{17}$ joules. In

1994, the collision between the comet Shoemaker-Levy and Jupiter released an amount of energy equivalent to 6 teratons of dynamite. How much energy was released during the collision of the Shoemaker-Levy?

Chapter 14

Structured Derivations

A task starts with a specific problem, and then continues to build a solution to this problem. The solution is carried out in some specific context, which lists the facts that we may use in our solution. However, there are often situations where we do not just have to formulate and solve a task, but where we must first build the context for the task. A *structured derivation* describes both the context for one or more tasks, as well as the solutions to these tasks.

A structured derivation is essentially a *mathematical model* of some situation, together with an analysis of that model. The general way to build a mathematical model goes approximately as follows:

- We start from the specific situation that we want to analyze. This could be a real-world problem, or a purely theoretical problem in some domain of science. We identify the quantities that we need to measure, as well as those that we want to determine. We denote these quantities with constant names, and determine their value ranges.
- Next, we identify the constraints that these quantities satisfy and describe how they are related to each other.
- We then formulate the questions that we want to answer about the model. These are formulated as tasks to be solved.
- In order to answer the questions, we may need to first derive some basic facts about the situation, based on what has been defined so far.
- We may also need to define some new concepts, in order to make it easier to formulate constraints and questions about the model.
- We are then ready to determine the answers to the questions posed in the tasks.
- The answers to these tasks are then interpreted as statements/facts about the real-world situation that we are modeling.

14.1 Generalizing Tasks to Structured Derivations

A structured derivation is essentially a sequence of *derivation steps,* of the form

derivation step

⋮

derivation step

where each derivation step is either

- an assumption,
- an observation (a declaration, fact, or a definition), or
- a task.

Assumptions, observations and tasks can thus be freely intermixed in a structured derivation.

A structured derivation allows us to work both on solving some specific tasks and on creating the proper context for these tasks. A structured derivation is the traditional way that a mathematician works on a problem. They first try to identify a specific problem to solve and formulate the problem in mathematical terms. Then they notice that more specific background assumptions are needed to formulate the problem, and that some new concepts have to be introduced by definitions. They then concentrate on solving the problem. Once they have solved the original problem, they may notice that there are other interesting questions that can also be solved in this same context. These may in turn require some additional assumptions and definitions, and so on.

The mathematical development unfolds as a novel, with a plot and some highlights. The difference, compared to a novel, is that each derivation step must be carefully checked for correctness, because a single incorrect observation, definition, or unjustified assumption can spoil the whole story. We also have to be careful not be caught in circular arguments (hence the linear format for the derivation).

A structured derivation gives us more freedom than a structured task:

- We are free to introduce definitions of concepts before we formulate assumptions or tasks that make use of these definitions.
- We can have any number of tasks based on the same set of assumptions, observations and definitions.
- We do not need to introduce all assumptions at once, we can introduce them one by one when they are needed.

Structured derivations are useful for problems with multiple questions. The next example shows a typical case of this: we first define a new concept, and then we ask a number of questions about this concept.

Example 69. The series $a_0, a_1, a_2, \ldots$ is defined by

$$a_n = \frac{n}{2n+1}$$

for $n = 0, 1, 2, 3, \ldots$. Show that (A) $0 < a_n < \frac{1}{2}$ when $n \geq 1$, that (B) $a_{n+1} > a_n$ when $n \geq 0$ and (C) calculate $\lim_{n\to\infty} a_n$.

We solve this task with a general structured derivation. Note that instead of bullets, we indicate the tasks with capital letters, here A, B, and C.

$+$ $\quad a : \mathbb{N} \to \mathbb{R}$

{The function a describes a series, where we denote $a_i = a(i)$, $i = 0, 1, 2, \ldots$. The series is well defined, since $2n + 1 > 0$ when $n = 0, 1, 2, \ldots$ }

$a_n = \dfrac{n}{2n+1}$ when $n = 0, 1, 2, 3, \ldots$

A. Show that $0 < a_n < \dfrac{1}{2}$, when

- $\quad n \in \mathbb{N}$, $n \geq 1$

$\Vdash$ $\quad 0 < a_n < \dfrac{1}{2}$

$\equiv$ $\quad$ {use the definition of a_n}

$0 < \dfrac{n}{2n+1} < \dfrac{1}{2}$

$\equiv$ $\quad$ {multiply both sides by $2n + 1$, write as a conjunction}

$0 < n \wedge n < \dfrac{2n+1}{2}$

$\equiv$ $\quad$ {simplify}

$0 < n \wedge 2n < 2n + 1$

$\equiv$ $\quad$ {$n \geq 1$ by the assumption, so the first proposition is true; the second proposition is always true}

T

□

B. Show that $a_{n+1} > a_n$, when

- $n \in \mathbb{N}$

$\Vdash$ $a_{n+1} > a_n$

$\equiv$ {use the definition of a_n}

$$\frac{n+1}{2(n+1)+1} > \frac{n}{2n+1}$$

$\equiv$ {simplify}

$$\frac{n+1}{2n+3} > \frac{n}{2n+1}$$

$\equiv$ {multiply by $(2n+3)(2n+1)$, which, by the assumption, is positive}

$(2n+1)(n+1) > (2n+3)n$

$\equiv$ {simplify}

$2n^2 + 3n + 1 > 2n^2 + 3n$

$\equiv$ {subtract $2n^2 + 3n$ from both sides}

$1 > 0$

$\equiv$ {arithmetics}

T

□

C. Calculate $\lim_{n\to\infty} a_n$

$\Vdash$ $\lim_{n\to\infty} a_n$

$=$ {the definition}

$$\lim_{n\to\infty} \frac{n}{2n+1}$$

$=$ {reduce by n }

$$\lim_{n\to\infty} \frac{\frac{n}{n}}{\frac{2n}{n} + \frac{1}{n}}$$

$=$ {simplify}

$$\lim_{n\to\infty} \frac{1}{2 + \frac{1}{n}}$$

$=$ $\{\frac{1}{n} \to 0$ when $n \to \infty\}$

$$\frac{1}{2}$$

□ $\lim_{n\to\infty} a_n = \frac{1}{2}$ ■

Structured derivations generalize all the previously defined constructs. A structured task is, e.g., a special case of a structured derivation, where there is only one derivation step, a task. Similarly, an assumption, a simple fact, a definition, or a sequence of these, is each also a special case of a structured derivation.

14.2 Modeling with Structured Derivations

We showed earlier how to formulate and solve a word problem as a mathematical task. We will here show how to formulate and solve a word problem as a structured derivation. This means that we first construct a mathematical model for the word problem, before formulating the question we want to answer. We solve the example problem below with a structured derivation, because we need some preliminary definitions before formulating the problem that we want to solve. These definitions could also be given outside the derivation, as background information, but we include them here in the derivation, to show how structured derivations can be used to build the context for a problem solution.

Example 70. (FNME, Autumn 2002) Since the year 1960 the travel time of the fastest train connection between Helsinki and Lappeenranta has decreased by 37%. Calculate by how many percent the average speed has increased. Assume that the length of the railroad has not changed.

Analyzing the problem statement, we see that we will need notations for the length of the railroad, for the time the trip used to take, and for the time it now takes, in order to express the assumptions in the problem. We declare these as constants in our derivation:

+ $s \in \mathbb{R}$ — the length of the track

+ $t \in \mathbb{R}$ — the original travel time

+ $t' \in \mathbb{R}$ — the current travel time

There is no need to introduce a symbol for the current length of the track, since we may assume that it is unchanged.

We can then formulate the following assumption:

(a) t' is 37% less than t — the current travel time is 37 % less than the original travel time

All variables must be positive real numbers, for the task to be meaningful. We can therefore add assumption (b).

(b) $t > 0$, $t' > 0$ and $s > 0$ — follows from the formulation of the problem

In order to formulate the question in the task, we also need to introduce symbols for the original speed (v), the current speed (v'), and the increase in speed (p). We must show that these new notations are well-defined.

[1] $v \in \mathbb{R}$ — the original speed

{Definition of speed, can be used, since $t > 0$ according to (b)}

$$v = \frac{s}{t}$$

[2] $v' \in \mathbb{R}$ — the current speed

{Definition of speed, can be used, since $t' > 0$ according to (b)}

$$v' = \frac{s}{t'}$$

[3] $p \in \mathbb{R}$ — increase in speed

{Definition of speed increase, can be used since $s > 0$, from which it follows that $v > 0$}

$$p = \frac{v' - v}{v}$$

We can now write down what we are supposed to do, as a task.

- Calculate p — the increase in the speed of the trip between Helsinki and Lappeenranta

We begin solving this task by writing assumption (a) more precisely:

[4] {Calculate the current travel time, based on (a)}

• $t - t' = 0.37 \cdot t$

$\equiv$ {subtract t from both sids}

$-t' = 0.37 \cdot t - t$

$\equiv$ {simplify}

$-t' = -0.63 \cdot t$

$\equiv$ {divide by -1}

$t' = 0.63 \cdot t$

□

... $t' = 0.63 \cdot t$

The solution is determined by the following calculation:

$\Vdash$ p

$=$ {observation [3]}

$$\frac{v' - v}{v}$$

$=$ {observation [1] and [2]}

$$\frac{\frac{s}{t'} - \frac{s}{t}}{\frac{s}{t}}$$

$=$ {simplify}

$$\frac{\frac{s}{t'}}{\frac{s}{t}} - 1$$

$=$ {simplify the fractions}

$$\frac{s \cdot t}{s \cdot t'} - 1$$

$=$ {simplify}

$$\frac{t}{t'} - 1$$

$=$ {$t' = 0.63 \cdot t$ according to observation [4], cancel out t}

$$\frac{1}{0.63} - 1$$

$\approx$ {calculate an approximate value of the expression}

0.59

$=$ {write as a percentage}

59%

□ $p \approx 59\%$ — the increase in speed

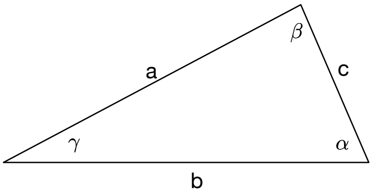
Thus, the answer is that the speed for the fastest connection between Helsinki and Lappeenranta has increased with 59 % since the 1960.

The solution to the modeling problems above is quite long, considering the rather simple calculations involved. A trained person can get the answers much faster and with less effort, by just writing down the relevant equations and solving them directly. The purpose of the two examples above is to show how to derive a solution where every step is carefully justified. At the same time, this provides a checklist of all the information that actually needs to go into the problem formulation and the solution. In practice, much of this information is left implicit. However, if you want to be very careful and certain that the calculation is correct, or if you are teaching problem solving skills to students with little prior experience in this, then it might be a good idea to spell out all steps explicitly. First teach the students how to do it properly, before teaching them the shortcuts.

The following sections gives further and more advanced examples of how to use structured derivations in modeling.

14.3 Example from Geometry

Example 71. (FNME, Autumn 2002). One of the angles in a triangle is α, and its opposite side has the length 5; another angle is 2α and its opposite side has the length 8. Calculate the exact length of the third side of the triangle, and calculate α with accuracy within a tenth of a degree.



Let us start by listing the facts that are given in the assignment, giving names to the important entities at the same time:

(a) The geometric figure is a triangle, with sides a, b and c, and opposing angles α, β and γ

(b) $a = 5$

(c) $b = 8$

(d) $\beta = 2\alpha$

(We have chosen here not to introduce the entities in the model explicitly by declarations, because they are implicitly declared in the figure.)

Let us check that we have all assumptions and assignments written down correctly. For this, we write down the informal problem statement once again, and mark the text fragments with the corresponding entities in the derivation. We mark the assumptions blue and the tasks magenta.

> One of the angles in a triangle is α (assumption a), and its opposite side has the length 5 (assumption b); another angle is 2α (assumption d) and its opposite side has the length 8 (assumption c). Calculate the exact length of the third side of the triangle (task B) , and calculate α with accuracy within a tenth of a degree (task A).

We can see that all assumptions from the informal problem statement have been taken into account in the derivation, and that there are no extra entities (assumptions or tasks) in the derivation. We have also marked the tasks that we need to solve in the problem statement.

We will first calculate the angle α.

A. Calculate the angle α

$\Vdash$ {Two of the the angles and the lengths of two of the sides are known in the triangle, so we can use the law of sines, $\dfrac{a}{\sin(\alpha)} = \dfrac{b}{\sin(\beta)} = \dfrac{c}{\sin(\gamma)}$ and fill in the values from the assumptions}

$$\frac{5}{\sin(\alpha)} = \frac{8}{\sin(2\alpha)}$$

$\equiv$ {multiply both sides by $\sin(\alpha)$ and $\sin(2\alpha)$; the sufficient restriction $0° < \alpha,\ 2\alpha < 180°$ follows from (a) which says that the figure is a triangle}

$$5\sin(2\alpha) = 8\sin(\alpha)$$

$\equiv$ {$\sin(2\alpha) = 2\sin(\alpha)\cos(\alpha)$}

$$5 \cdot 2\sin(\alpha)\cos(\alpha) = 8\sin(\alpha)$$

$\equiv$ {divide by $\sin(\alpha)$; this is allowed because $\sin(\alpha) \neq 0$ ($\alpha \neq 0°$ and $\alpha \neq 180°$ in a triangle)}

$$10\cos\alpha = 8$$

$\equiv$ {divide both sides by 10 and simplify}

$$\cos(\alpha) = \frac{4}{5}$$

□ Angle α is such that $\cos(\alpha) = \dfrac{4}{5}$

[1] {From the solution to task A, calculating the approximate value for α and rounding it off in accordance with the condition}

$$\alpha \approx 36.9°$$

Having calculated the angle α, our next task is to calculate the length of the third side in the triangle.

B. Calculate the length of c

$\Vdash$ {Two sides and one angle are known in a triangle, so the law of cosines $a^2 = b^2 + c^2 - 2bc \cdot \cos\alpha$ can be used to find the third side}

$$5^2 = 8^2 + c^2 - 2 \cdot 8 \cdot c \cdot \cos(\alpha)$$

$\equiv$ {task A}

$$5^2 = 8^2 + c^2 - 2 \cdot 8 \cdot c \cdot \frac{4}{5}$$

$\equiv$ {write it in the form $ax^2 + bx + c = 0$}

$$c^2 - \frac{64}{5}c + 39 = 0$$

$\equiv$ {solve the equation with the quadratic formula}

$$c = \frac{-(-\frac{64}{5}) \pm \sqrt{(-\frac{64}{5})^2 - 4 \cdot 1 \cdot 39}}{2 \cdot 1}$$

$\equiv$ {simplify}

$$c = \frac{64}{10} \pm \frac{\sqrt{\frac{196}{25}}}{2}$$

$\equiv$ $\{\sqrt{\frac{a}{b}} = \frac{\sqrt{a}}{\sqrt{b}}\}$

$$c = \frac{64}{10} \pm \frac{(\frac{\sqrt{196}}{\sqrt{25}})}{2}$$

$\equiv$ {compute the square roots and simplify}

$$c = \frac{64}{10} \pm \frac{14}{10}$$

$\equiv$ {write as disjunction}

$$c = 7.8 \ \vee \ c = 5$$

$\equiv$ {the answer $c = 5$ is false, as the triangle would then be an isosceles with the angles α, α, and 2α, but according to [1], $4\alpha \approx 147.6°$ and $147.6 \neq 180°$. Therefore the figure would not be a triangle, so $c = 5 \equiv F$}

$$c = 7.8 \ \vee F$$

$\equiv$ $\{p \vee F \equiv p\}$

$$c = 7.8$$

□ Third side $c = 7.8$ ■

14.4 Example from Probability Theory

Example 72. (FNME, Autumn 2002). Lena and Sarah toss a coin to decide which one of them will get to ride a horse first. Lena tosses the coin first and is allowed to ride first if she gets a head. If she gets a tail, Sarah will toss the coin and will ride first if she gets a head. If Sarah also gets a tail, the turn to toss the coin goes to Lena again. They continue in this manner until one of them gets a head. What is the probability the Lena is allowed to ride first? What is the probability that Sarah is allowed to ride first?

Let us again start with what we know about the problem.

(a) The probability of the event "head" is $\frac{1}{2}$

(b) The probability of the event "tail" is $\frac{1}{2}$

(c) A "head" lets one ride first, and the coin is tossed until a "head" is acquired

(d) The girls takes turns in tossing the coin

(e) Lena gets the first throw

We then determine the probabilities involving Lena.

[1] $q \in [0, 1]$

{q is the probability that a person lands a head/tail after she has landed a tail. This event can only take place if the second person has landed a tail (probability $\dfrac{1}{2}$), thus giving the first person a chance to throw again (probability $\dfrac{1}{2}$ regardless of whether she lands a head or a tail). Hence, $q = \dfrac{1}{2} \cdot \dfrac{1}{2} = \dfrac{1}{4}$.}

$q = \frac{1}{4}$

[2] {The total probability that Lena is allowed to ride first is given by the expression $\dfrac{1}{2} + \dfrac{1}{2} \cdot q + \dfrac{1}{2} \cdot q^2 + ... + \dfrac{1}{2} \cdot q^n + \ldots$, $n \in \mathbb{N}$, (the probability that she gets to ride on the first turn is $\dfrac{1}{2}$, $\dfrac{1}{2}q$ on the second turn and so forth)}

$$P(\text{“Lena rides first”}) = \frac{1}{2} + \frac{1}{2} \cdot q + \frac{1}{2} \cdot q^2 + ... + \frac{1}{2} \cdot q^n + \ldots,\ n \in \mathbb{N}$$

We now consider the situation from Lena's point of view.

A. Calculate the probability that Lena is allowed to ride first

⊩ $P(\text{“Lena rides first”})$

= {observation [2]}

$\frac{1}{2} + \frac{1}{2} \cdot q + \frac{1}{2} \cdot q^2 + ... + \frac{1}{2} \cdot q^n + \dots$

= {This series is a infinite geometric series. Since $q = \frac{1}{4} < 1$ according to [1], the value of the series is given by $\frac{a}{1-q}$ when $|q| < 1$. Here the first term a is $\frac{1}{2}$}

$\frac{\frac{1}{2}}{1 - \frac{1}{4}}$

= {calculate}

$\frac{2}{3}$

□ P("Lena rides first") $= \frac{2}{3}$

From this we can immediately infer the probability that Sarah rides first.

[3] {Sarah rides with probability $1 - P$("Lena rides first")}

P("Sarah rides first") $= \frac{1}{3}$

The situation involves an unbounded number of throws, so we might be a little bit uncertain about the last observation. We therefore check the answer for Sarah by doing a similar calculation that we did above for Lena, but now from Sarah's point of view.

[5] {Lena gets a head with her first toss with the probability $\frac{1}{2}$. If she lands a tail instead, Sarah will have the probability $\frac{1}{2}$ to land a head.}

Sarah is allowed to ride with her first toss with probability $\frac{1}{4}$.

[6] {The total probability that Sarah is allowed to ride first is given by the expression $\frac{1}{4} + \frac{1}{4} \cdot q + \frac{1}{4} \cdot q^2 + ... + \frac{1}{4} \cdot q^n + \dots$, $n \in \mathbb{N}$.}

P("Sarah rides first") $= \frac{1}{4} + \frac{1}{4} \cdot q + \frac{1}{4} \cdot q^2 + ... + \frac{1}{4} \cdot q^n + \dots, n \in \mathbb{N}$.

B. Calculate the probability that Sarah is allowed ride first.

⊩ P("Sarah rides first")

= {observation [6]}

$\frac{1}{4} + \frac{1}{4} \cdot q + \frac{1}{4} \cdot q^2 + ... + \frac{1}{4} \cdot q^n + \dots$

$=$ {The series above is a infinite geometric series with value $\dfrac{a}{1-q}$ when $|q| < 1$. The first term a is given by $\frac{1}{4}$}

$\dfrac{\frac{1}{4}}{1 - \frac{1}{4}}$

$=$ {calculate}

$\dfrac{1}{3}$

□ $P(\text{“Sarah rides first”}) = \dfrac{1}{3}$

This shows that the original answers were correct: Lena rides first with probability $\frac{2}{3}$ and Sarah rides first with probability $\frac{1}{3}$. ■

14.5 Example from Mechanics

Example 73. Sergeant Riley shoots a cannon ball straight upwards with a canon placed on the ground. The initial velocity of the cannon ball is 80 m/s, and air friction is assumed to be negligible. On what height above ground is the cannonball when 6.0 seconds has passed. Is the cannon ball still going upwards at this point of time?

Let us start by declaring the the entities of the model explicitly:

+ $y \in \mathbb{R}^+ \to \mathbb{R}$ — the height of the ball, as a function of time

+ $y_0 \in \mathbb{R}$ — initial height of the ball

+ $v_0 \in \mathbb{R}$ — initial velocity

+ $a \in \mathbb{R}$ — uniform acceleration

We then list what we know.

(a) $v_0 = 80$ m/s — the initial velocity of the cannon ball

(b) 6.0 seconds have passed

(c) The cannon ball is shot from ground level

(d) Air resistance is negligible

Next we make some initial observations.

[1] {Law of mechanics, the height of ball y as a function of time t}

$y(t) = y_0 + v_0 t + \frac{1}{2}at^2$, for $t \geq 0$

[2] {From (c), the cannon is on the ground}

$y_0 = 0$

[3] {The uniform acceleration due to Earths gravitation}

$a = g = -9.81\,\mathrm{m/s^2}$

A. Calculate height of ball when 6.0 seconds have passed.

$\Vdash$ [1]

$\Rightarrow$ {compute height for time $t = 6.0$ s}

$y(6) = y_0 + v_0 \cdot 6.0\,\mathrm{s} + \frac{1}{2}a \cdot (6.0\,\mathrm{s})^2,$

$\Rightarrow$ {insert initial velocity v_0 from (a), initial height y_0 from [2] and acceleration a from [3]}

$y(6) = 0 + 80\,\mathrm{m/s} \cdot 6.0s - \frac{1}{2} \cdot (-9.81\,\mathrm{m/s^2}) \cdot (6.0\,\mathrm{s})^2$

$\equiv$ {calculate}

$y(6) = 303.42\,\mathrm{m}$

□ The height of the ball is 303.42 m after 6.0 seconds

Next we determine whether the ball is still on its way up after 6.0 seconds.

B. Is the ball still rising after 6.0 seconds?

$\Vdash$ ball is still rising at time 6.0

$\equiv$ {the ball is still rising when the velocity is positive}

velocity is positive at time 6.0

$\equiv$ {velocity is time derivative of distance}

- $\frac{d}{dt}y(t)$

$=$ {[1]}

$\frac{d}{dt}(y_0 + v_0 t + \frac{1}{2}at^2)$

$=$ {calculate derivative}

$v_0 + at$

□

$\ldots$ $v_0 + a \cdot 6.0\,\mathrm{s} > 0$

$\equiv$ {insert values from (a) and [3]}

$80\,\mathrm{m/s} + (-9.81\,\mathrm{m/s^2}) \cdot 6.0\,\mathrm{s} > 0$

$\equiv$ {calculate}

$21.14\,\mathrm{m/s}{>}0$

$\equiv$ {true fact}

T

□ Yes, the ball is still rising at time 6.0 seconds. ■

14.6 Example from Nuclear Physics

Example 74. Calculate the missing mass in the fusion reaction between deuterium and tritium, ${}^2_1H + {}^3_1H \rightarrow {}^4_2He +^1_0 n$. Determine whether energy is consumed or released in this reaction. Give the answer with the precision of four decimals. The interior of a certain star converts every minute at least 45 billion tons of hydrogen to helium (compare to the 36 billion tons converted by our own sun). Assume that all this energy is due to the reaction above. Calculate the power of the star.

We start by listing the facts that we will be using in our calculations.

(a) We study the reaction ${}^2_1H + {}^3_1H \rightarrow {}^4_2He +^1_0 n$

[1] {The mass of the atom nucleus is the mass of the atom/isotope minus the mass of the electrons, where a is the number of electrons in the isotope}

$m(nucleus) = m(isotope) - a \cdot m(e)$

A. Calculate the missing mass Δm.

$\Vdash$ Δm

$=$ {the missing mass Δm is given by the difference between the mass of the initial reactants and the mass of the product of the reaction; use [1] to calculate the mass of each nucleus}

$m({}^2_1H) - m(e) + m({}^3_1H) - m(e) - (m\left({}^4_2He\right) - 2m(e) + m({}^1_0n))$

$=$ {the masses of the electrons cancel each other}

$m({}^2_1H) + m({}^3_1H) - (m\left({}^4_2He\right) + m({}^1_0n))$

$=$ {insert the values for the isotope masses: $m({}^2_1H) = 2.0141018\,\mathrm{u}$, $m({}^3_1H) = 3.0160493\,\mathrm{u}$, $m\left({}^4_2He\right) = 4.0026033\,\mathrm{u}$ and $m({}^1_0n) = 1.0086650\,\mathrm{u}$ and calculate the result}

$0.0188828\,\mathrm{u}$

□ $\Delta m = 0.0188828\,\mathrm{u}$

B. Calculate the energy E that is released in the reaction (a)

$\Vdash$ E

$=$ {the reaction energy is the missing mass converted into energy}

$\Delta m \cdot c^2$

$=$ {A}

$0.0188828\,\mathrm{u} \cdot c^2$

$=$ $\{1\mathrm{u} = 1.6605402 \cdot 10^{-27}\,\mathrm{kg}\}$

$0.0188828 \cdot 1.6605402 \cdot 10^{-27}\,\mathrm{kg} \cdot c^2$

$=$ $\{c = 299792458\,{}^{\mathrm{m}}\!/\!_{\mathrm{s}}\}$

$0.0188828 \cdot 1.6605402 \cdot 10^{-27}\,\mathrm{kg} \cdot (299792458\,{}^{\mathrm{m}}\!/\!_{\mathrm{s}})^2$

$=$ {calculate}

$2.81810514617 \cdot 10^{-12}\,\mathrm{J}$

□ $E = 2.81810514617 \cdot 10^{-12}\,\mathrm{J}$

C. Calculate E in electron volts, with two decimal precision

$\Vdash$ E

$=$ {convert to electron volt}

$17.5892212017\,\mathrm{MeV}$

$\approx$ {round off to two decimals}

$17.59\,\mathrm{MeV}$

□ $E = 17.59\,\mathrm{MeV}$

(b) The star converts 45 billion tons of hydrogen to helium each minute

(c) The reaction (a) is the only one going on in the sun

[2] {The mass of the hydrogen nuclei that participate in the reaction is their combined mass}

- mass of hydrogen nuclei in the reaction

  $=$ {the hydrogen nuclei are ${}^2_1H$ and ${}^3_1H$, according to (a)}

  $m({}^2_1H) + m({}^3_1H)$

  $=$ {insert values for the masses}

  $2.0141018\,\mathrm{u} + 3.0160493\,\mathrm{u}$

  $=$ {add and convert to kilogram, $1\,\mathrm{u} = 1.6605402 \cdot 10^{-27}\,\mathrm{kg}$}

$8.3527681 \cdot 10^{-27}\,\mathrm{kg}$

□

... The combined mass of the hydrogen nuclei is $8.3527681 \cdot 10^{-27}\,\mathrm{kg}$

[3] $n \in \mathbb{R}$

{The number of reactions n taking place each minute is the total mass of hydrogen being converted divided by the the mass needed for one reaction}

$$n = \frac{45 \cdot 10^{12}\mathrm{kg}}{8.3527681 \cdot 10^{-27}\mathrm{kg}}$$

D. Calculate the power P of the star

⊩ P

= {the power is work divided by time}

W/t

= {work is here total amount of energy released in one minute}

nE/t

= {[3]}

$$\frac{45 \cdot 10^{12}\mathrm{kg}}{8.3527681 \cdot 10^{-27}\mathrm{kg}} \cdot E/t$$

= {E is given by (B) and $t = 60\,\mathrm{s}$}

$$\frac{45 \cdot 10^{12}\mathrm{kg}}{8.3527681 \cdot 10^{-27}\mathrm{kg}} \cdot 2.81810514617 \cdot 10^{-12}\mathrm{J}/60\,\mathrm{s}$$

= {calculate}

$2.53039331498 \cdot 10^{26}\,\mathrm{W}$

≈ {approximate}

$2.5 \cdot 10^{26}\,\mathrm{W}$

□ The power of the sun is approximately $2.5 \cdot 10^{26}\,\mathrm{W}$ (in reality it would be higher due to the other reactions taking place).

There are a number of calculation steps above, so we need to summarize the results and check that we are actually answering the original questions. Based on these calculations, we get the following answers:

1. The missing mass in the reaction is $\Delta m = 0.0188828\,\mathrm{u}$ (A)
2. Energy is released in the reaction, the released energy is $E = 17.59\,\mathrm{MeV}$ (C)
3. The power of the star is $2.5 \cdot 10^{26}\,\mathrm{W}$ (D)

■

14.7 Assignments

1. Solve the equation $|x^n| = x^{n+1}$, where n is a positive integer.
2. Determine for which values of x the expression $\ln\left(\left(x^2+x\right)\left(\frac{x+1}{x-1}\right)\right)$ is defined.
3. A space age ruler, who has way too much money, time and cubist passions, wants the Universe to contain at least one cubic object. For this reason, he has the trans-Neptunian dwarf planet Sedna transformed into a cube. a) Calculate the side length of the cube, when Sedna's radius is 995 km and the ruler has access to technology that is advanced enough to use the entire mass of Sedna to create the cube. b)What if the ruler had preferred more traditional methods and transformed the spherical Sedna into a cube by cutting off excess matter. Calculate the volume of Cube-Sedna and the percentage of the original planet that was wasted in the transformation. You may assume that Sedna is spherical and that its mass is uniformly distributed.
4. The barrel of an old cannon is cylindrical. The volume of the cannonballs they used is $8.5\,\mathrm{dm}^3$ and eight of them fit into the barrel so that the last one precisely reaches the muzzle of the barrel. Patrick the Pirate gets overexcited when he loads the cannon with gunpowder. The result is that only a single cannonball barely fits inside the muzzle. a) Calculate the volume of the barrel. b) How much gunpowder did Patrick put into the cannon, when none of it is in front of the cannonball and there is only one cannonball in the barrel?
5. Four letters are randomly picked out of the eleven-letter word UNFORTUNATE. Determine the probability that a) only one of the letters is a consonant, b) only one of the letters is a vowel, and c) you can spell TUNA with the letters that are picked out.
6. Let the function $f\ :\mathbb{R}\to\mathbb{R}$ be $f(x) = x^4+5x+2$. a) Differentiate $f(x)$. b) Determine $f'(2)$. c) Determine the smallest value of $f(x)$.
7. May is designing a sturdy table inspired by mathematical curves. She uses the solid of revolution generated by rotating the curve $y = 0.2x^2 - 3x + 15$ around the x-axis on the interval [1, 20]. Determine a) the volume of the of the solid of revolution (in volume units), b) how much wood it takes to create the table, if it is 95 cm high, c) how much does the table weight if it is made out of oak with a density of $700\,{}^{\mathrm{kg}}/{}_{\mathrm{m}^3}$, and d) how much varnish will it take to coat every surface of the table when you use around 0.12 liters of varnish to cover a square meter.

Chapter 15

Quantifiers

We expand in this chapter the treatment of logic from propositional calculus to *predicate calculus*. Predicate calculus adds two new ways to construct logical propositions, *universal quantification* of the form $(\forall x \in A : p(x))$ ("proposition $p(x)$ is true for every $x \in A$") and *existential quantification* of the form $(\exists x \in A : p(x))$ ("proposition $p(x)$ is true for some $x \in A$"), where $p(x)$ is a logical proposition about the variable x.

Many mathematical propositions require quantifiers to be formulated exactly. The proposition

$$(\forall x \in \mathbb{R} : f(x) \leq a)$$

states that the function f is bounded from above by a, i.e., $f(x) \leq a$ is true for every value of x. The proposition

$$(\forall x \in \mathbb{R} : f(x) \leq f(y))$$

states that the function f has a maximum in y. The proposition

$$(\forall x, y \in \mathbb{R} : x \leq y \Rightarrow f(x) \leq f(y))$$

states that the function f is monotonic, i.e., if $x \leq y$, then $f(x) \leq f(y)$, for every value of x and y.

Divisibility is a good example of a concept where we need existential quantifiers. The number 18 is divisible by 6, since the division 18/6 comes out even (with the result 3). This means that there is a number (namely 3) that multiplied by 6 equals 18. We define the notation $m \,|\, n$ (read as "m divides n" or "m is a factor of n" or "n is divisible by m" as

$$m \,|\, n \;\equiv\; (\exists k \in \mathbb{N} : k \cdot m = n)$$

In words: $m \,|\, n$ is true if, and only if, there is a natural number k, such that $k \cdot m = n$.

The existential quantifier is often used implicitly. One example is provided by solutions to trigonometric equations. These are often periodic, and can be described using the existential quantifier. The equation

$$\sin x = 1$$

holds for x if, and only if, x satisfies the condition

$$x = \frac{\pi}{2} + 2n\pi$$

for some $n \in \mathbb{Z}$. We can express this with an existential quantifier as follows:

$$(\exists n \in \mathbb{Z} : x = \frac{\pi}{2} + 2n\pi)$$

In other words, we have that

$$\sin x = 1 \equiv (\exists n \in \mathbb{Z} : x = \frac{\pi}{2} + 2n\pi)$$

Using an existential quantifier and a universal quantifier in the same expression is also useful. Consider a function $f : A \rightarrow B$. The proposition

$$(\forall y \in B : (\exists x \in A : f(x) = y))$$

states that the function f is *surjective*, i.e., its range is B. This is an example of a logical proposition with two alternating quantifiers, here universal quantification followed by existential quantification. Alternating quantifiers have a central role in higher mathematics, and are discussed at length later on, in Chapter 17.

We usually avoid explicit use of quantifiers in upper secondary level mathematics, either because of mathematical tradition or because explaining quantifiers is considered to be too difficult. Instead, we use different notations in different situations to explain the same idea of quantification. Since quantifiers are so central to all mathematical reasoning and concepts, students will encounter them all the time, but disguised in different ways. One may ask whether sweeping quantifiers under the rug in upper secondary level mathematics is an effective strategy, rather than explaining this key concept once and for all, both how to understand quantified expressions, and how to argue with them.

We describe the syntax for quantifiers and the intuition behind quantifiers in this chapter. The next chapter shows how to use these new constructs in mathematical argumentation, and describes the inference rules for quantified expressions and how they are used in structured derivations. The following chapter then looks at alternating quantifiers, i.e. the use of both universal and existential quantification in the same expression. Alternating quantifiers are, e.g. used in the so called *epsilon-delta method*, which often makes life miserable for many beginning students of mathematics. Understanding this method is, however, straightforward when one is familiar with the basic principles of how to argue with quantified expressions.

15.1 Bound and Free Variables

Before defining quantifiers more precisely, we need to me more precise about variables in mathematical expressions and how they are used. We have until now talked about simple expressions, like

$$x^2 + 2x + 1, \quad \frac{e^{2x}}{\sin x}, \quad \log(2x) + z$$

and so on. We can freely *substitute* new expressions for variables in these expressions (i.e., replace a variables with some expression). For instance, we could substitute $(x + 2y)$ for x in the first expression and get

$$(x + 2y)^2 + 2(x + 2y) + 1$$

The same substitution in the second expression gives us

$$\frac{e^{2(x+2y)}}{\sin(x + 2y)}$$

and so on.

However, when going to more advanced mathematical concepts, we encounter expressions of a different kind. Consider, e.g., the expression

$$\sum_{n=0}^{k} x^n$$

This is a shorthand notation for the sum $x^0 + x^1 + x^2 + \ldots + x^k$, i.e.,

$$\sum_{n=0}^{k} x^n = x^0 + x^1 + x^2 + \ldots + x^k$$

Note that the variable n does not occur on the right hand side. The variable n is only used to indicate how the elements of the sum are created. We refer to n as a *bound variable*, in contrast to the *free variables* x and k in the sum expression. We may substitute expressions for the free variables, but not for the bound variables in an expression.

The sum operator $\sum$ binds the variable n in the expression. We can indicate the distinction between free and bound variables in an expression by coloring the bound variables red in the expression, leaving the free variables black. Our example expression would then be

$$\sum_{\color{red}n=0}^{k} x^{\color{red}n}$$

The operator $\sum$ colors all occurrences of variable n in its scope red. The scope is here the expression x^n.

Note. We only use coloring here as a pedagogical device, in practice one learns quickly to see the difference between free and bound variables, so coloring is not needed.

The same variable can occur both free and bound in an expression, as in

$$\sum_{\color{red}n=0}^{k} x^{\color{red}n} + \frac{1}{n} \sum_{\color{red}m=0}^{k} (x + 3)^{\color{red}m}$$

The sum operator only binds the term that comes immediately after it, in this case x^n for the first sum operator and $(x+3)^m$ for the second sum operator. The term $\frac{1}{n}$ is not in the scope of the first or second sum operator, so it is free, i.e., not colored red.

Free and bound variables are all over the place, once you start to look for these. Here is a short list of other expressions that are commonly used in high school mathematics, and which make use of bound variables:

$$\sum_{n=0}^{k} x^n$$

$$\prod_{n=0}^{k} x^n$$

$$\lim_{n\to\infty} \frac{a^n}{n}$$

$$\frac{d}{dx}(x^2 + a \cdot \sin x + y)$$

$$\int_{x=0}^{2} (ax^2 + 2bx)dx$$

$$\{(x,y) \in \mathbb{R} \times \mathbb{R} \,|\, x^2 + y^2 = r\}$$

Here the binding operators are $\sum$, $\prod$, lim, $\frac{d}{dx}$, $\int$, and the set forming operation.

We treat bound and free variables differently in an expression. We can *rename* (change the name of) a bound variable in an expression without changing the value of the expression. We can *substitute* new expressions for the free variables in an expression to calculate the value of the expression for these new values. However, in both cases, there are certain restrictions that have to be respected so that the meaning of an expression is not changed. Basically, we require that the coloring of variables does not change when we rename bound variables or substitute expressions for free variables. In other words, free variables should not become bound as the result of a renaming or a substitution.

Renaming bound variables Consider firsts *renaming* bound variables. We can rename n to m in our series expression, without changing the meaning of the expression, i.e.,

$$\sum_{m=0}^{k} x^m = \sum_{n=0}^{k} x^n$$

We see that the coloring does not change, only the variable names. We may not, however, rename n to a variable that already appears as a free variable in the scope of the binding operator, because this will change the coloring. If we, e.g., rename n

to x, we get

$$\sum_{x=0}^{k} x^{x} \neq \sum_{n=0}^{k} x^{n}$$

Here the variable x, which was free (colored black) in the original expression, is in the scope of the sum operator and is therefore colored red. In other words, a free variable has become bound in the new expression, which drastically changes the meaning of the expression.

In the expression

$$\lim_{x \to 0} \left(\sum_{n=0}^{k} x^{n} \right)$$

we cannot rename x or n to k, since this would change the value of the expression. But we can rename x and n to other variables that do not appear as free variables in the expression. We have, e.g., that

$$\lim_{x \to 0} \left(\sum_{n=0}^{k} x^{n} \right) = \lim_{y \to 0} \left(\sum_{m=0}^{k} y^{m} \right)$$

Here the coloring is not changed, so renaming is allowed.

The general rule is that we can rename bound variables in an expression freely, as long as no free variable in the expression becomes bound as the result of renaming.

Substituting for free variables Consider a proposition about sums, such as

$$\sum_{n=0}^{k} an = a\frac{k(k+1)}{2}$$

This proposition is true for every $k \in \mathbb{N}$ and $a \in \mathbb{R}$. By picking, e.g., $k = 3$ we get

$$\sum_{n=0}^{3} an = a\frac{3(3+1)}{2} = 6a$$

Since the proposition is true for every natural number, we can substitute any natural number expression for k and still have a true proposition. If we, e.g., substitute $k+i$ for k, we get the proposition

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

If we substitute $b+1$ for a we get the expression

$$\sum_{n=0}^{k} (b+1)n = (b+1)\frac{k(k+1)}{2}$$

which is also true, since the original proposition is true for every value $k \in \mathbb{N}$ and $a \in \mathbb{R}$.

It is, however, not permissible to substitute $b+n$ for a in this proposition. Then we get the proposition

$$\sum_{n=0}^{k}(b+n)n = (b+n)\frac{k(k+1)}{2}$$

The problem is here that the coloring has changed: the free variable n in the expression $b+n$ becomes bound (is colored red) in the expression we get after the substitution. This changes the meaning of the sum, and the proposition is not true anymore.

The general rule is that we are free to substitute an expression for a free variable in another expression, as long as no free variable in the substituting expression becomes bound in the resulting expression. We say that the expression e *is free for* x in expression E, if this condition holds.

How to avoid conflicts in substitutions The question now is what we should do if we, e.g., want to calculate the value of the expression

$$\sum_{n=0}^{k} an$$

when a is $b+n$. We may not substitute $b+n$ for a directly, since this would bind the free variable n in the expression $b+n$. We can, however, achieve the desired effect by a combination of renaming and substitution.

We first change the names of the bound variables, so that the substitution becomes permissible. In this case, we first rename n to another variable, say m. This gives us the expression

$$\sum_{m=0}^{k} am$$

which has the same meaning as the original expression. Now , $b+n$ is free for a in this new expression, so we are permitted to substitute $b+n$ for a. This gives us the expression

$$\sum_{m=0}^{k}(b+n)m$$

In other words, we can always substitute an expression e for a free variable x in an expression E, but sometimes we first have to rename bound variables in the expression E, to prevent a free variable in e from becoming a bound variable in E after the substitution.

15.2 The Universal Quantifier

We are now ready to introduce the two key operators in predicate calculus, the universal quantifier and the existential quantifier. We have previously treated propositional calculus, which does not involve quantifiers but only logical connectives. In predicate calculus, we study the logical properties of expressions that involve both quantifiers and logical connectives. Predicate calculus is an extension of propositional calculus, so all laws of propositional calculus remain true in predicate calculus.

Consider the expression $x + 1 > x$. Regardless of what value we substitute for x, the truth value will be T. This gives us the logical proposition

> $x + 1 > x$ is true for every value of x

We write this proposition using a *universal quantifier,* as

$$(\forall x \in \mathbb{R} : x + 1 > x)$$

We read this as "for all real numbers x: $x + 1 > x$". The parenthesis is part of the syntax, as is the colon that separate the quantified variable from the quantified expression. The operator $\forall x$ binds every variable x in the expression, i.e., x is bound in the expression within the parenthesis. Coloring the binding quantifier and the bound variables in the expression, as we did in the previous section, would look as follows:

$$(\forall x \in \mathbb{R} : x + 1 > x)$$

Since the proposition $x + 1 > x$ is true for every real number, we have

$$(\forall x \in \mathbb{R} : x + 1 > x) \equiv T$$

The proposition $(\forall x \in \mathbb{R} : x^2 \geq x)$, however, is not true, since there are counterexamples: when x takes the value $\frac{1}{2}$, the proposition $x^2 \geq x$ is false. Thus

$$(\forall x \in \mathbb{R} : x^2 \geq x) \equiv F$$

since it is not true that $x^2 \geq x$ for *every* real value of x.

The general form of a universal quantifier is

$$(\forall x \in A : p)$$

where p is a logical proposition. The variable x may be free in p, as in the example above (p is here $x + 1 > x$). There can also be other free variables in p. It is also possible that x does not appear as a free variable in p.

Successive universal quantifications of different variables can be written as a quantification of a sequence of variables. Thus, the expression

$$(\forall x \in \mathbb{R} : (\forall y \in \mathbb{R} : x + y + 1 > x + y - 1))$$

can also be written as

$$(\forall x \in \mathbb{R}, y \in \mathbb{R} : x + y + 1 > x + y - 1)$$

In this case, we can make the expression even shorter, as

$$(\forall x, y \in \mathbb{R} : x + y + 1 > x + y - 1)$$

since both variables x and y have the same domain. All the expressions say the same thing: that for every combination of the real numbers x and y, $x+y+1 > x+y-1$.

We can understand a universally quantified expression as a generalized conjunction. The proposition

$$(\forall x \in \mathbb{N} : x < x + 1)$$

would then stand for the infinite expression

$$0 < 0 + 1 \ \wedge \ 1 < 1 + 1 \ \wedge \ 2 < 2 + 1 \ \wedge \ 3 < 3 + 1 \ \wedge \ \ldots$$

Many properties of the universal quantifier are in fact generalizations of the corresponding properties of conjunction.

We have to introduce a special notation for universal (and existential quantifiers) to describe situations like this, since infinite expressions cannot be written down in practice. The rewriting above also only works for domains like $\mathbb{N}$, where we can enumerate the individual elements, but not anymore for domains that are not enumerable, like $\mathbb{R}$.

15.3 The Existential Quantifier

We can use an *existential quantifier* to express the fact that a proposition is true for some value. We can, e.g., propose that

> $x^2 > x$ is true for some value of x.

The proposition is true, since e.g. $2^2 = 4 > 2$. We write this proposition using the existential quantifier as

$$(\exists x \in \mathbb{R} : x^2 > x)$$

We read this as "there exists a real number x, such that $x^2 > x$". The operator $\exists x$ binds every variable x in the expression, i.e., x is bound in the expression within the parenthesis. Coloring the binding quantifier and the bound variables in the expression would look as follows:

$$(\exists x \in \mathbb{R} : x^2 > x)$$

The proposition is true, i.e.,

$$(\exists x \in \mathbb{R} : x^2 > x) \equiv T$$

We see this by choosing, e.g., $x = 2$.

The general form of an existential quantifier is

$$(\exists x \in A : p)$$

where p is a logical proposition that may contain free variables like x and possibly other variables.

In the same way as we could consider the universal quantifier as a generalized conjunction, we may consider an existentially quantified expression as a generalized disjunction. The proposition

$$(\exists x \in \mathbb{N} : x + 1 = 2x)$$

can be seen as the infinite proposition

$$0 + 1 = 2 \cdot 0 \ \vee \ 1 + 1 = 2 \cdot 1 \ \vee \ 2 + 1 = 2 \cdot 2 \ \vee \ 3 + 1 = 2 \cdot 3 \ \vee \ \ldots$$

We can omit the domain from a quantified expression if the context makes it obvious. We can write our example formulas as $(\forall x : x + 1 > x)$ and $(\exists x : x^2 > x)$, if we assume that x always takes real values. We will usually include the domain in our presentation, since we often work with different domains simultaneously in upper secondary level mathematics.

15.4 Manipulating Quantified Expressions

The discussion we had earlier on bound and free variables also applies to quantified expressions. A variable x is *bound* in a logical expression, if it appears in a context $(\forall x \in A : \ldots x \ldots)$ or $(\exists x \in A : \ldots x \ldots)$. We then say that x is bound by the quantifier ($\forall x$ or $\exists x$) in the expression. The same variable x can appear both as a bound variable and as a free variable in a logical expression. An example is the following expression, where we have colored the bound variables:

$$x \leq y \wedge (\forall x \in \mathbb{R} : x \geq 0 \Rightarrow x + y \geq 0) \qquad (*)$$

Here y is free in the entire expression, as is also the first occurrence of x. However, the second and subsequent occurrences of x are bound by the quantification $\forall x \in \mathbb{R}$.

The constraints that we have described in previous sections apply to quantified expressions as well:

- We can freely change the names of the bound variables in a quantified expression, provided that no free variable in the expression becomes bound after the name change. We can e.g. rename the bound variables x to z in the expression above,

  $$x \leq y \wedge (\forall z \in \mathbb{R} : z \geq 0 \Rightarrow z + y \geq 0)$$

  without the meaning of the expression changing, but we cannot rename x to y, since y appears as a free variable in the quantified expression before renaming, but would be bound after the renaming.

- We can freely substitute a new expression for a free variable, provided that no free variable in the substituted expression becomes bound after the substitution. We can e.g. substitute $y + z$ for y in the expression $(*)$ above, to get

  $$x \leq y + z \wedge (\forall x \in \mathbb{R} : x \geq 0 \Rightarrow x + y + z \geq 0)$$

  but we cannot substitute the expression $y + x$ for y, since x would the become bound in the resulting expression.

- However, we can do any substitution we want, if we first change the names of the bound variables in the expression so that the desired substitution becomes permissible. If we first change the bound variable x in the expression $(*)$ to z, we can then substitute $y+x$ for x without problem,

  $$x \leq y+x \wedge (\forall z \in \mathbb{R} : z \geq 0 \Rightarrow z+y+x \geq 0)$$

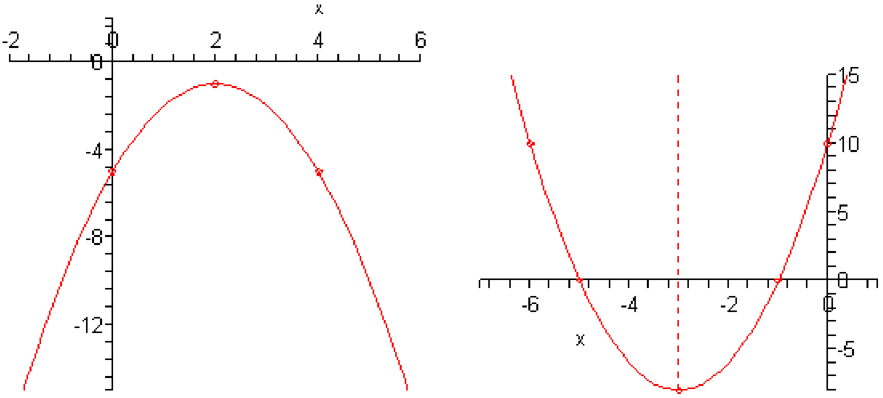
  since neither y nor x now become bound after the substitution.

15.5 Reasoning with Quantified Expressions

We will here give a first example of how to reason about quantified expressions in high school mathematics. We will give more examples later on, after we have presented the general rules for universal and existential quantification.

Example 75. We want to examine for what values of a the function $f(x) = -x^2 + ax + a - 3$ is always negative. We use the universal quantifier to formulate the problem. The curves in figure 15.1 reminds us of the two kinds of parabolas, those that open down and those that open up.

Figure 15.1: Parabolas opening down and up



•   Calculate the values of a for which the function f is always negative, when

- $f(x) = -x^2 + ax + a - 3$ for $x \in \mathbb{R}$.

⊩ $(\forall x \in \mathbb{R} : f(x) < 0)$

≡   {the function f is a parabola that opens down, because the coefficient of x^2 is negative; such a function is always negative if it has no zeros (the figure to the left)}

$(\forall x \in \mathbb{R} : f(x) \neq 0)$

$\equiv$ $\{$the condition is satisfied if the discriminant D_f of the function is less than zero$\}$

$D_f < 0$

$\equiv$ $\{$insert the value of $D_f\}$

> • D_f
>
> $=$ $\{$the discriminant of the function $f(x) = Ax^2 + Bx + C$ is $B^2 - 4AC\}$
>
> $a^2 - 4(-1)(a - 3)$
>
> $=$ $\{$simplify$\}$
>
> $a^2 + 4a - 12$
>
> □

$\ldots$ $a^2 + 4a - 12 < 0$

$\equiv$ $\{$the function $a^2 + 4a - 12$ opens up, since the coefficient of a^2 is positive; such a function is negative between the zeros (Figure 15.1)$\}$

> • Find the zeros of the function $a^2 + 4a - 12$
>
> $\Vdash$ $a^2 + 4a - 12 = 0$
>
> $\equiv$ $\{$the quadratic formula$\}$
>
> $$a = \frac{-4 \pm \sqrt{4^2 - 4 \cdot 1 \cdot (-12)}}{2 \cdot 1}$$
>
> $\equiv$ $\{$simplify the expression$\}$
>
> $a = 2 \vee a = -6$
>
> □

$\ldots$ $-6 < a < 2$

□

We have thus proven that the function f is always negative if, and only if, $-6 < a < 2$. ■

Note that we have to prove equivalence between the propositions in the example. If we only prove implication to the right, the condition for a might be to weak, i.e. we might get additional values of a, which do not satisfy the original condition. If we only prove implication to the left, we might miss some values of a that satisfy the original condition. With equivalence we show that the derived values of a exactly satisfy the condition.

15.6 Assignments

1. Determine which variables are bound and which are free in the statements a) $\lim\limits_{x\to\infty}\frac{a^x}{x}$, b) $\sum\limits_{a=0}^{b} ax^{k+a} + c$, and c) $\lim\limits_{x\to 2}\left(\sum\limits_{y=0}^{h} x^{y+z}\right)$

2. Express formally the following statement: There is a number n such that when it is multiplied by any number the result is n and when it is added to any number the result is that number. Does this statement hold when the universe of discourse (a.k.a. domain (of discourse), the set of entities over which the variables may range) is $\mathbb{R}$? How about $\mathbb{R}_+$?

3. Write down formally the following statement and then negate it: "For any integer x and any integer y , there exists an integer$z$ such that $y - z = x$ " Which one of the formulas is true, the original statement or its negation?

4. Let the formula φ (written \varphi in Latex) be

$$(\exists x : (P(x, y) \wedge (\forall y : (\neg Q(y, x) \wedge P(y, z)))))$$

Perform the following substitutions: (a) $\varphi [y := \omega]$, (b) $\varphi [y := f(x)]$, (c) $\varphi [z := g(y, z)]$

CHAPTER 16

Proof Strategies for Quantifiers

We described the basic natural deduction rules for logical connectives in previous chapters. We now turn to the natural deduction rules for quantifiers, and show how these are used when proving theorems involving quantifiers. We have four basic rules: *specialization* and *generalization* for universally quantified propositions, and the *witness rule* and *existential assumptions* for existentially quantified propositions. All other inference rules for quantifiers can be derived from these four basic rules. The natural deduction rules for quantifiers are shown in Table 16.1. We will consider each one in more detail below.

Introduction rules		
∀- intro (a not free in Φ or Q)	$\dfrac{\Phi,\, a \in A \vdash Q(a)}{\Phi \vdash (\forall x \in A : Q(x))}$	{*Generalization*}
∃-intro (t is free for x in Q)	$\dfrac{\Phi \vdash t \in A \quad \Phi \vdash Q(t)}{\Phi \vdash (\exists x \in A : Q(x))}$	{*Witness rule*}
Elimination rules		
∀- elim (t is free for x in Q)	$\dfrac{\Phi \vdash (\forall x \in A : Q(x)) \quad \Phi \vdash t \in A}{\Phi \vdash Q(t)}$	{*Specialization*}
∃ - elim (a not free in Φ or Q)	$\dfrac{\Phi \vdash (\exists x \in A : Q(x)) \quad \Phi,\, a \in A,\, Q(a) \vdash P}{\Phi \vdash P}$	{*Existential assumption*}

Table 16.1: Natural deduction rules for quantifiers

16.1 Generalization

We use the *generalization rule* ($\forall$*-intro*) to prove that a universally quantified formula is true. The template below shows how this rule is used in structured derivations. We can choose the name of the variable a freely, as long as we choose a name that does not already occur free in Φ or in $Q(x)$.

• $(\forall x \in A : Q(x))$

- Φ

⊩ $\{$*Generalization*, a not free in $Q(x)$ or $\Phi\}$

> • $Q(a)$
>
> \- $a \in A$

Table 16.2: Generalization rule

In other words, we show that the proposition $(\forall x \in A : Q(x))$ is true under the assumption Φ by showing that $Q(a)$ is true for an arbitrary value $a \in A$, under the same assumption Φ. The value a is arbitrary if it is not free in $Q(x)$ or in Φ (because then we know nothing about this value).

We will frequently use a a slightly weaker requirement for the generalization rule: we only require that a does not occur free in Φ. We can then choose a to be x when we want to prove that $\Phi \vdash (\forall x \in A : Q(x))$. In other words, we prove that $Q(x)$ is true under the assumption $x \in A$. (It is easy to see that this is permitted, because we may first rename the bound variable x in $(\forall x \in A : Q(x))$ to some other variable y that does not occur in Φ or Q. Then we can use the original inference rule as stated in Table 16.2, choosing a to be x.

Example 76. We show that $x^2 > x$ when $x > 1$. The proof uses the generalization rule for universal quantifiers, as well as the rule for proving an implication (direct proof).

• Show that $\big(\forall x \in \mathbb{R} : x > 1 \Rightarrow x^2 > x\big)$

⊩ {Generalization, chose an arbitrary a}

> • Show that $a > 1 \Rightarrow a^2 > a$ when
>
> \- $a \in \mathbb{R}$
>
> ⊩ {Direct proof}
>
> > • Show that $a^2 > a$ when
> >
> > \- $a > 1$
> >
> > ⊩ $a^2 > a$

$\equiv$ {adding $-a$ to both sides}

$a^2 - a > 0$

$\equiv$ {factoring}

$a(a-1) > 0$

$\equiv$ {both factors are positive according to the assumption}

T

□

□

□

The proof illustrates how we can combine reduction proofs (on the outer levels) with a calculation (on the innermost level). ■

16.2 Specialization

The second basic rule for the universal quantifier is the *specialization rule* ($\forall$-*elim*):

- $Q(t)$

- Φ

$\Vdash$ {*Specialization*, t is free for x in $Q(x)$}

- $(\forall x \in A : Q(x))$
- $t \in A$

Table 16.3: Specialization

The rule can be used when t can be substituted for x in $Q(x)$. The rule says that if we have proved that both $(\forall x \in A : Q(x))$ and $t \in A$ follow from Φ, then we know that $Q(t)$ follows from Φ.

Example 77. The distributive law $a \cdot (b+c) = a \cdot b + a \cdot c$ stands for the proposition

$$(\forall a, b, c \in \mathbb{R} : a \cdot (b+c) = a \cdot b + a \cdot c)$$

Using this rule, we can e.g. show that $x \cdot (x^2+1) = x \cdot x^2 + x \cdot 1$. We do this with the following reduction derivation:

- $x \cdot (x^2+1) = x \cdot x^2 + x \cdot 1$

$\Vdash$ {*Specialization* with $a := x$, $b := x^2$, $c := 1$}

- $(\forall a, b, c \in \mathbb{R} :\ a \cdot (b + c) = a \cdot b + a \cdot c)$
- $x, x^2, 1 \in \mathbb{R}$

The justification explains explicitly which substitution we have made: a is assigned x, b is assigned x^2, and c is assigned 1.

The specialization rule would more commonly be used in a forward derivation, of the form:

$\vdots$

(a) $(\forall a, b, c \in \mathbb{R} :\ a \cdot (b + c) = a \cdot b + a \cdot c)$

$\vdots$

$+$ $\{$*Specialization* on (a), $x, x^2, 1 \in \mathbb{R}\}$

$x \cdot (x^2 + 1) = x \cdot x^2 + x \cdot 1$

$\vdots$

Or, we could use the rule in a calculation, and just refer to the distribution rule for multiplication and addition:

$\vdots$

$x \cdot (x^2 + 1)$

$=$ $\{$*distribution rule*$\}$

$x \cdot x^2 + x \cdot 1$

$\vdots$

Here we have also omitted to state that $x, x^2, 1 \in \mathbb{R}$, as this is considered obvious. As in other applications of structured derivations, we are free to choose the level of detail in how we justify a derivation step. ■

16.3 The Witness Rule

The inference rule for proving an existential proposition is called the *witness rule* ($\exists$-*intro*).

• $(\exists x \in A : Q(x))$

- Φ

$\Vdash$ $\{$*Witness rule*, t is free for x in $Q(x)\}$

  • $Q(t)$

  • $t \in A$

Table 16.4: The witness rule

If a proposition $Q(t)$ is true for some value $t \in A$ of the existentially quantified variable x, then $(\exists x \in A : Q(x))$ is true. We call t a *witness* to the truth of the existential quantification. The rule can be used provided t can be substituted for x in $Q(x)$.

Example 78. Assume that the equation $3(\bar{a} - \bar{b}) = -4\bar{a} + 3\bar{b}$ holds for the vectors $\bar{a}$ and $\bar{b}$ ($\bar{a} \neq 0$, $\bar{b} \neq 0$). Show that the vectors are collinear, i.e. that $\bar{a} \upuparrows \bar{b}$.

• Show that the vectors $\bar{a}$ and $\bar{b}$ are collinear, when

(1) $\bar{a} \neq 0$ and $\bar{b} \neq 0$

(2) $3(\bar{a} - \bar{b}) = -4\bar{a} + 3\bar{b}$

$\Vdash$ $\bar{a} \upuparrows \bar{b}$

$\equiv$ {assumption (1) and the definition of collinear vectors}

$\left(\exists r \in \mathbb{R} : r > 0 \wedge a = r\bar{b}\right)$

$\equiv$ {the witness rule}

  • Find a value r that satisfies the conditions $\bar{a} = r\bar{b}$ and $r > 0$

  $\Vdash$ {calculating based on assumption (2)}

  $3(\bar{a} - \bar{b}) = -4\bar{a} + 3\bar{b}$

  $\equiv$ {expanding the parenthesis}

  $3\bar{a} - 3\bar{b} = -4\bar{a} + 3\bar{b}$

  $\equiv$ {adding $4\bar{a} + 3\bar{b}$ to both sides}

  $7\bar{a} = 6\bar{b}$

  $\equiv$ {dividing by 7}

  $\bar{a} = \frac{6}{7}\bar{b}$

  $\equiv$ $\{\frac{6}{7} > 0\}$

  $\bar{a} = \frac{6}{7}\bar{b} \wedge \frac{6}{7} > 0$

> □ $r = \frac{6}{7}$
>
> • Show that $\frac{6}{7} \in \mathbb{R}$
>
> $\Vdash$ {obvious}
>
> □

. . . T

□ ■

16.4 Existential Assumptions

The second rule for existential quantifiers, *existential assumptions* ($\exists$-*elim*), is shown in Table 16.5.

• P

- Φ

$\Vdash$ {*Existential assumption*, a not free in Φ, P or $Q(x)$}

> • $(\exists x \in A : Q(x))$
>
> • P
>
> \- $Q(a)$
>
> \- $a \in A$

Table 16.5: Existential assumption rule

We assume that the variable a is not free in P or $Q(x)$ or Φ. This rule shows how we can use an existential assumption when we prove a proposition P. We show that P is true under the assumption Φ, by showing that $(\exists x \in A : Q(x))$ and that P is true if we assume $Q(a)$ for an arbitrary $a \in A$, under the assumptions Φ.

Example 79. Show that $(m + n)$ is divisible by k if m and n are divisible by k.

• Show that $p|(m + n)$, when

- $p|n$, and $p|m$

$\Vdash$ {Existential assumption}

> • Show that $(\exists k \in \mathbb{N} : n = k \cdot p)$
>
> $\Vdash$ {Follows from assumption $p|n$}
>
> □

```
    •    Show that $p|(m+n)$, when
    -    $n = k \cdot p$
    -    $k \in \mathbb{N}$
    ⊩    {Existential assumption}
         •    Show that $(\exists l \in \mathbb{N} : m = l \cdot p)$
         ⊩    {Follows from assumption $p|m$}
         □
         •    Show that $p|(m+n)$, when
         -    $m = l \cdot p$
         -    $l \in \mathbb{N}$
         ⊩    $T$
         ≡    {assumptions}
              $n = k \cdot p \wedge m = l \cdot p$
         ⇒    {add $m$ and $n$}
              $n + m = k \cdot p + l \cdot p$
         ≡    {factor out a common term}
              $n + m = (k + l) \cdot p$
         ⇒    {the rule for existential quantification, $k + l$ is a witness}
              $(\exists r \in \mathbb{N} : n + m = r \cdot p)$
         ≡    {definition of divisibility}
              $p|(n+m)$
         □
    □
□
```

The next to last derivation step uses the witness rule as a logical rule rather than as an inference rule. We give this rule, and a collection of other rules for calculating with quantifiers, in the next section. ■

The use of the existential assumption rule is somewhat complicated, because it introduces new levels of nesting in the proof. The next section shows how this can be avoided using definitions.

16.5 Definitions

Consider the definition

```
    +    $c_1 \in A_1, \ldots, c_k \in A_k$

         {justification}

         $D(c_1, \ldots, c_k)$
```

The condition for the constants $c_1, \ldots, c_k$ to be well-defined in the definition is that

$$(\exists x_1 \in A_1 \ldots, \exists x_k \in A_k : D(x_1, \ldots, x_k))$$

it true in the context of the definition. A definition thus requires us to prove an existentially quantified formula.

We can avoid the complications of the existential assumption rule by using definitions. We show this by giving an alternative solution to Example 79.

Example 80. (Proof using definitions) Show that $(m+n)$ is divisible by p if both m and n are divisible by p.

• Show that $p|(m+n)$, when

- $p|n$, and

- $p|m$

+ $k \in N$

{k is well-defined, since $p|n$ means that $(\exists k \in \mathbb{N} : n = k \cdot p)$}

$n = k \cdot p$

+ $l \in N$

{l is well-defined, since $p|m$ means that $(\exists l \in \mathbb{N} : m = l \cdot p)$}

$m = l \cdot p$

$\Vdash$ $n = k \cdot p \wedge m = l \cdot p$

$\Rightarrow$ {add m and n}

$n + m = k \cdot p + l \cdot p$

$\equiv$ {factor out a common factor}

$n + m = (k + l) \cdot p$

$\Rightarrow$ {the rule for existential quantification, $k + l$ is a witness}

$(\exists r \in \mathbb{N} :\ n + m = r \cdot p)$

$\equiv$ {definition of divisibility}

$p|(n + m)$

□ ■

The next example illustrates the power of using definitions in proofs.

Example 81. A Diophantine equation is an equation for which you seek integer solutions. For example, the so-called Pythagorean triples (x, y, z) are positive integer solutions to the equation $x^2 + y^2 = z^2$. Another example is provided by the following theorem: There are no positive integer solutions to the Diophantine equation $x^2 - y^2 = 1$.

Let us start by formulating the theorem to be proved as as a task.

• Prove that $\neg(\exists x, y \in \mathbb{N}^+ : x^2 - y^2 = 1)$

We prove this by reductio ad absurdum.

• Prove that $\neg(\exists x, y \in \mathbb{N}^+ : x^2 - y^2 = 1)$

⊩ {Reductio ad absurdum}

  • Prove F

  - $(\exists x, y \in \mathbb{N}^+ : x^2 - y^2 = 1)$

□

Next, we introduce names for the two positive integers that we claim exist:

• Prove that $\neg(\exists x, y \in \mathbb{N}^+ : x^2 - y^2 = 1)$

⊩ {reductio ad absurdum}

  • Prove F

  - $(\exists x, y \in \mathbb{N}^+ : x^2 - y^2 = 1)$

  + $x, y \in \mathbb{N}^+$

  {x, y well-defined by assumption}

  $x^2 - y^2 = 1$

□

We can now prove the contradiction:

• Prove that $\neg(\exists x, y \in \mathbb{N}^+ : x^2 - y^2 = 1)$

⊩ {Reductio ad absurdum}

  • Prove F

  - $(\exists x, y \in \mathbb{N}^+ : x^2 - y^2 = 1)$

  + $x, y \in \mathbb{N}^+$

  {x, y well-defined by assumption}

  $x^2 - y^2 = 1$

$\Vdash$ $x^2 - y^2 = 1$

$\equiv$ {factoring}

$(x - y)(x + y) = 1$

$\equiv$ $\{x, y$ are both positive integers, by assumption$\}$

$(x - y = 1 \wedge x + y = 1) \vee (x - y = -1 \wedge x + y = -1)$

$\equiv$ {adding equations in both disjuncts}

$(2x = 2 \wedge x + y = 1) \vee (2x = -2 \wedge x + y = -1)$

$\equiv$ {solving equations}

$(x = 1 \wedge y = 0) \vee (x = -1 \wedge y = 0)$

$\equiv$ $\{x$ and y are both positive integers$\}$

$F \wedge F$

$\equiv$ {definition of conjunction}

F

□

□

A second examples of using definitions in proofs is given by the classical argument for the existence of infinitely many prime numbers.

Example 82. Show that there exists an infinite number of prime numbers. We prove this by showing that the counter assumption, that there only exists a finite number of prime numbers, leads to a contradiction. We denote the set of all prime numbers by $\mathbb{P}$.

- (*Euclid*) There exists an infinite number of prime numbers.

$\Vdash$ {indirect proof}

- Prove F, when

- there exists only a finite number of prime numbers in $\mathbb{P}$

[1] $k, p_1, \ldots, , p_k \in \mathbb{N}$

{We enumerate the k prime numbers $p_1, \ldots, p_k$, the nested derivation shows that these constants are well-defined}

- there exists only a finite number of prime numbers in $\mathbb{P}$

$\equiv$ {by the definition of a finite set}

$(\exists k \in \mathbb{N}, p : \{1, \ldots, k\} \to \mathbb{P} : p$ is a bijection $)$

$\equiv$ $\{$we write $p_i = p(i)\}$

$(\exists k \in \mathbb{N},\ p : \{1, \ldots, k\} \to \mathbb{P} : \{p_1, \ldots, p_k\} = \mathbb{P}\}$

$\ldots$ $\mathbb{P} = \{p_1, \ldots, p_k\}$

[2] $n \in \mathbb{N}$

{The product of a finite number of natural numbers is always well-defined}

$n = p_1 \cdot p_1 \cdot \ldots \cdot p_k$

[3] r, $p'_1, \ldots, p'_r \in \mathbb{N}$

{The Fundamental Theorem of Algebra states that we can express $n+1$ as the product of a finite number of prime numbers, i.e., $(\exists r \in \mathbb{N}, \exists p'_1, \ldots, p'_r \in \mathbb{P} : n+1 = p'_1 \cdot p'_2 \cdot \ldots \cdot p'_r)$}

$n+1 = p'_1 \cdot p'_2 \cdot \ldots \cdot p'_r \wedge p'_1, \ldots, p'_r \in \mathbb{P}$

$\Vdash$ {Derive a contradiction}

$n+1 = p'_1 \cdot p'_2 \cdot \ldots \cdot p'_r \wedge p'_1, \ldots, p'_r \in \mathbb{P}$

$\Rightarrow$ {choose a factor of $n+1$}

$p'_1 | (n+1) \wedge p'_1 \in \mathbb{P}$

$\Rightarrow$ {definition of n , $p'_1 \in \mathbb{P}$ }

$p'_1 | (n+1) \wedge p'_1 | n$

$\Rightarrow$ {p'_1 divides the difference between $n+1$ and n, according to an earlier lemma}

$p'_1 | (n+1-n)$

$\Rightarrow$ {calculation}

$p'_1 | 1$

$\Rightarrow$ {only the number 1 divides 1, but every prime number is different from 1}

F

□

□ ■

16.6 Calculating with Quantifiers

We have now presented the four main inference rules for quantified expressions. Quantified expressions occur quite frequently in calculations, so we also need calculation rules for quantifiers, similar to the calculation rules for connectives that we have described in Chapter 5. Table 16.6 lists a collection of useful rules for quantification. Many of these rules are generalizations of similar rules for conjunction and disjunction.

The table shows that the *specialization rule* and the *witness rule* can also be presented as logical formulas:

$$\begin{aligned} (\forall x \in A : Q(x)) \wedge t \in A &\quad\Rightarrow\quad Q(t) \\ t \in A \wedge Q(t) &\quad\Rightarrow\quad (\exists x \in A : Q(x)) \end{aligned}$$

$\vdash$	$(\forall x \in A : Q(x)) \Rightarrow Q(t)$	$\{$*specialization*$\}$
$\vdash$	$Q(t) \Rightarrow (\exists x \in A : Q(x))$	$\{$*witness rule*$\}$
$\vdash$	$(\forall x \in A : x = t \Rightarrow Q(x)) \equiv Q(t)$	$\{$*one-point rule for* $\forall\}$
$\vdash$	$(\exists x \in A : x = t \wedge Q(x)) \equiv Q(t)$	$\{$*one-point rule for* $\exists\}$
$\vdash$	$(\forall x \in A : Q) \equiv Q$	$\{\ \forall$ *trivial quantification*$\}$
$\vdash$	$(\exists x \in A : Q) \equiv Q$	$\{\ \exists$ *trivial quantification*$\}$
$\vdash$	$\neg(\forall x \in A : P(x)) \equiv (\exists x \in A : \neg P(x))$	$\{$*de Morgan for* $\forall\}$
$\vdash$	$\neg(\exists x \in A : P(x)) \equiv (\forall x \in A : \neg P(x))$	$\{$*de Morgan for* $\exists\}$
$\vdash$	$(\forall x \in A : P(x) \wedge Q(x))$ $\equiv (\forall x \in A : P(x)) \wedge (\forall x \in A : Q(x))$	$\{$*conjunction rule for* $\forall\}$
$\vdash$	$(\exists x \in A : P(x) \vee Q(x))$ $\equiv (\exists x \in A : P(x)) \vee (\exists x \in A : Q(x))$	$\{$*disjunction rule for* $\exists\}$
$\vdash$	$(\forall x \in A : P(x) \vee Q) \equiv (\forall x \in A : P(x)) \vee Q$	$\{$*disjunction rule for* $\forall\}$
$\vdash$	$(\exists x \in A : P(x) \wedge Q) \equiv (\exists x \in A : P(x)) \wedge Q$	$\{$*conjunction rule for* $\exists\}$
$\vdash$	$(\forall x \in A : P(x) \Rightarrow Q) \equiv ((\exists x \in A : P(x)) \Rightarrow Q)$	$\{$*implication rule 1 for* $\forall\}$
$\vdash$	$(\forall x \in A : Q \Rightarrow P(x)) \equiv (Q \Rightarrow (\forall x \in A : P(x)))$	$\{$*implication rule 2 for* $\forall\}$
$\vdash$	$(\exists x \in A : P(x) \Rightarrow Q) \equiv ((\forall x \in A : P(x)) \Rightarrow Q)$	$\{$*implication rule 1 for* $\exists\}$
$\vdash$	$(\exists x \in A : Q \Rightarrow P(x)) \equiv (Q \Rightarrow (\exists x \in A : P(x)))$	$\{$*implication rule 2 for* $\exists\}$

Table 16.6: Rules for quantifiers

The *one-point rules* are surprisingly useful in many contexts. These rules require that t can be substituted for x in $Q(x)$:

$$\begin{aligned}(\forall x \in A : x = t \Rightarrow Q(x)) &\equiv Q(t)\\ (\exists x \in A : x = t \wedge Q(x)) &\equiv Q(t)\end{aligned}$$

Quantification over a variable that does not occur in the formula can be ignored, as shown by the *trivial quantification* rules. Here Q is a formula where x does not occur free:

$$\begin{aligned}(\forall x \in A : Q) &\equiv Q\\ (\exists x \in A : Q) &\equiv Q\end{aligned}$$

The d*e Morgan rules* for conjunction and disjunction also carry over to universal and existential quantification:

$$\begin{array}{lcl}\neg(\forall x \in A :\, P(x)) & \equiv & (\exists x \in A :\, \neg P(x)) \\ \neg(\exists x \in A :\, P(x)) & \equiv & (\forall x \in A :\, \neg P(x))\end{array}$$

The *distribution rules* say that universal quantification distributes over conjunction and existential quantification distributes over disjunction. These rules show that we can consider universal quantification as a form of conjunction and existential quantification as a form of disjunction.

$$\begin{array}{lcl}(\forall x \in A :\, P(x) \wedge Q(x)) & \equiv & (\forall x \in A :\, P(x)) \wedge (\forall x \in A :\, Q(x)) \\ (\exists x \in A :\, P(x) \vee Q(x)) & \equiv & (\exists x \in A :\, P(x)) \vee (\exists x \in A :\, Q(x))\end{array}$$

We do not have similar convenient rules for distributing universal quantification over disjunction or for distributing existential quantification over conjunction. We do, however, have two very useful rules for handling universal quantification of a disjunction, and existential quantification of a conjunction. Here we assume that the variable x is not free in Q. These rules allow us to move part of a formula outside a quantification, when they do not refer to the quantified variable.

$$\begin{array}{lcl}(\forall x \in A :\, P(x) \vee Q) & \equiv & (\forall x \in A :\, P(x)) \vee Q \\ (\exists x \in A :\, P(x) \wedge Q) & \equiv & (\exists x \in A :\, P(x)) \wedge Q\end{array}$$

The rules for implication can be directly derived from these rules, when we characterize implication in terms of disjunction, i.e., that $(A \Rightarrow B) \equiv (\neg A \vee B)$. These rules assume that x does not occur free in Q.

$$\begin{array}{lcl}(\forall x \in A :\, P(x) \Rightarrow Q) & \equiv & ((\exists x \in A :\, P(x)) \Rightarrow Q) \\ (\forall x \in A :\, Q \Rightarrow P(x)) & \equiv & (Q \Rightarrow (\forall x \in A :\, P(x))) \\ (\exists x \in A :\, P(x) \Rightarrow Q) & \equiv & ((\forall x \in A :\, P(x)) \Rightarrow Q) \\ (\exists x \in A :\, Q \Rightarrow P(x)) & \equiv & (Q \Rightarrow (\exists x \in A :\, P(x)))\end{array}$$

Not all of these rules are independent of each other. We can, e.g., prove the last two rules using the conjunction rule for $\exists$.

Example 83. We prove the first implication rule for universal quantification. This proof illustrates how to derive new logical rules from earlier ones with structured derivations.

- •	Prove implication rule 1 for universal quantification, when
- -	x is not free in Q

$\Vdash$ $(\forall x \in T : P(x) \Rightarrow Q)$

$\equiv$ {write implication as disjunction}

$(\forall x \in T : \neg P(x) \vee Q)$

$\equiv$ {disjunction rule for $\forall$, x is not free in Q}

$(\forall x \in T : \neg P(x)) \vee Q$

$\equiv$ {de Morgans rule}

$\neg(\exists x \in T : P(x)) \vee Q$

$\equiv$ {rewrite as implication}

$(\exists x \in T : P(x)) \Rightarrow Q$

□ ■

16.7 Assignments

1. Negate the following formulas (latex:$\forall$ is written with \forall, $\exists$ is writen with \exists): (a) $(\exists x : P(x)) \Rightarrow (\forall y : Q(y))$ (b) $\neg(\exists x : \neg P(x)) \Rightarrow (\forall x : P(x))$
2. Negate the following formulae: a) $(\exists x : P(x) \Rightarrow (\forall y : Q(y)))$ b) $(\forall x : (\exists y : P(x) \vee Q(y)))$.
3. Explain why $p(t) \Rightarrow (\forall x \in \mathbb{R} : p(x))$ and $(\exists x \in \mathbb{R} : q(x)) \Rightarrow q(t)$ are not general rules in predicate logic.
4. It is possible to choose a in such a manner that $a(x+1) < (x+2)^2 - a$ holds for every value of x?
5. Prove $(\forall x : p(x)) \wedge \neg(\exists x : q(x)) \equiv (\forall x : \neg(p(x) \Rightarrow q(x)))$
6. Prove$(\exists x : p(x) \wedge q(x)) \equiv \neg(\forall x : p(x) \Rightarrow \neg q(x))$
7. Determine whether $(\forall x : p(x)) \vee (\forall x : \neg p(x))$ is valid (i.e. always true regardless of the statement $p(x)$) or not.
8. Assume that $(\exists x : p(x))$ and $(\forall x : p(x) \Rightarrow q(x))$. Prove that $(\exists x : q(x))$.
9. Given $(\forall x : p(x) \Rightarrow q(x))$ and $\neg q(m)$, prove $\neg(\forall x : p(x))$.
10. Prove that implication is not asymmetric. (Hint in order to be asymmetric, the relation must be antisymmetric and irreflexive)
11. Given the formula
$$(\exists x : (\forall y \cdot P(x,y) \wedge Q(y,x)))$$
prove that
$$(\forall y : (\exists x \cdot P(x,y)) \wedge (\exists z : Q(y,z)))$$
.

12. Derive

$$(\forall y : (\forall x \cdot \neg R\,(x, y)))$$

from

$$(\forall x : (\forall y \cdot R\,(x, y) \Rightarrow Q\,(x, y)))$$

and

$$(\forall y : (\forall x : \neg Q\,(x, y)))$$

13. Given that

$$(\forall x : (\forall y \cdot P\,(x, y) \Rightarrow Q\,(y, x)))$$

holds, show that

$$(\forall y : (\exists x : P\,(x, y)) \Rightarrow (\exists x : Q\,(y, x)))$$

Chapter 17

Derivations with Quantifiers

We present here some more advanced applications of reasoning with quantified formulas. We describe first how to prove universal properties about natural numbers using mathematical induction, and how to prove properties about recursively defined functions with complete induction. We then consider how to prove existentially quantified formulas using non-linear proofs, in order to apply the witness rules. Finally, we consider how to use non-linear proofs for proving formulas with alternating quantifiers. We will look at the *epsilon-delta method* as a prime example of reasoning about alternating quantifiers, and where non-linear proofs can greatly simplify the argumentation.

17.1 Mathematical Induction

Mathematical induction is a way to prove universally quantified propositions about natural numbers. The general inference rule is shown below in Table 17.1, on the left. On the right, we have a variant of this rule, called *complete induction.*

We illustrate induction proofs with two examples, a simpler proof that is more or less classical (calculating the sum of the first n natural numbers), and a more advanced one (characterization of the Fibonacci numbers using the golden ratio).

Example 84. (Proof by induction) We prove that the classical formula

$$0 + 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$$

holds for every natural number n.

•	$(\forall n \in \mathbb{N} : P(n))$	•	$(\forall n \in \mathbb{N} : P(n))$
-	Φ	-	Φ
⊩	{*Mathematical induction*}	⊩	{*Complete induction*}
•	$P(n)$	•	$P(n)$
-	$n = 0$	-	$(\forall m \in \mathbb{N} : m < n \Rightarrow P(m))$
•	$P(n')$		
-	$P(n)$		
-	$n' = n + 1$		

Table 17.1: Mathematical induction

The inference rule for mathematical induction shows that we can split the proof of the general statement into two steps, the *base case* and the *induction step*. This gives us the following initial attempt at a proof:

• Show that $0 + 1 + \ldots + n = \dfrac{n(n+1)}{2}$ for every integer n.

⊩ {Proof by mathematical induction}

  • *Base case*: Show that $0 + 1 + \ldots + n = \dfrac{n(n+1)}{2}$, when

  - $n = 0$

  • *Induction step*: Show that $0 + 1 + \ldots + n' = \dfrac{n'(n'+1)}{2}$, when

  - $0 + 1 + \ldots + n = \dfrac{n(n+1)}{2}$, and

  - $n' = n + 1$

□

We then prove the base case and the induction step separately (these proofs are shown in red):

• Show that $0 + 1 + \ldots + n = \dfrac{n(n+1)}{2}$ for every integer n.

⊩ {Proof by induction}

  • *Base case*: Show that $0 + 1 + \ldots + n = \dfrac{n(n+1)}{2}$, when

  - $n = 0$

$\Vdash$ $0 + 1 + \ldots + n$

$=$ $\{\text{assumption } n = 0\}$

0

$=$ $\{\text{arithmetics}\}$

$\dfrac{0(0+1)}{2}$

$=$ $\{\text{assumption}\}$

$\dfrac{n(n+1)}{2}$

□

- *Induction step*: Show that $0 + 1 + \ldots + n' = \dfrac{n'(n'+1)}{2}$, when

- $0 + 1 + \ldots + n = \dfrac{n(n+1)}{2}$, and

- $n' = n + 1$

$\Vdash$ $0 + 1 + \ldots + n'$

$=$ $\{\text{assumption}\}$

$0 + 1 + \ldots + n + (n + 1)$

$=$ $\{\text{induction hypothesis}\}$

$\dfrac{n(n+1)}{2} + (n + 1)$

$=$ $\{\text{find a common denominator}\}$

$\dfrac{n^2 + n + 2n + 2)}{2}$

$=$ $\{\text{simplify}\}$

$\dfrac{n^2 + 3n + 2}{2}$

$=$ $\{\text{factor}\}$

$\dfrac{(n+1)(n+2)}{2}$

$=$ $\{\text{assumption } n' = n + 1\}$

$\dfrac{n'(n'+1)}{2}$

□

□ ■

17.2 Recursive Definitions and Induction Proofs

Let us next consider *recursive definitions* of functions, and how to prove properties about recursively defined functions. A classical example is the *factorial function* $n!$.

We define the factorial function as follows. The *base case* is

$$0! = 1$$

and the *induction step* is

$$n! = n \cdot (n-1)!$$

for $n \geq 1$.

This gives us that $0! = 1$, by the *base case*. For $n = 1$, we have

$$1! = 1 \cdot 0! = 1 \cdot 1 = 1$$

For $n = 2$, we have

$$2! = 2 \cdot 1! = 2 \cdot 1 = 2$$

for $n = 3$, we have

$$3! = 3 \cdot 2! = 3 \cdot 2 = 6$$

and so on. Continuing, we will eventually define the value of $n!$ for every natural number n.

We can summarize the definition of the factorial function as follows:

$+$ {recursive definition, n is decreasing}

$0! = 1 \wedge$

$(\forall n \in \mathbb{N} : n \geq 1 \Rightarrow n! = n \cdot (n-1)!)$

Example 85. Compute the value of 5!. We have the following computation:

$\bullet$ $5!$

$=$ {induction step}

$5 \cdot 4!$

$=$ {induction step}

$5 \cdot 4 \cdot 3!$

$=$ {induction step}

$5 \cdot 4 \cdot 3 \cdot 2!$

$=$ {induction step}

$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1!$

$=$ {induction step}

$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 0!$

$=$ {base step}

$5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 1$

$=$ {calculation}

120

□

This is an example of how we compute the value of a recursively defined function: we expand the definition for a specific argument using the inductive step in the definition, until we reach a base case where we can determine the value directly, without further expansions. The recursion thus stops at the base case. ■

In a recursive function definition, the recursion always ends in a base case. This is because the value of the function for a specific argument is defined in terms of the same function, but for arguments that are in some sense simpler than the original argument. For natural numbers, we think of $n-1$ as being simpler than n. Ultimately, reducing the argument by one in each step gets us to the argument 0, for which we have a non-recursive definition of the function value.

A recursive definition and a *circular function definition* may look similar. However, a circular definition does not define the value of the function for a specific argument in terms of simpler values. This means that expanding the definition will go on forever, and we will never know the value of the function for the initial argument.

Proving a property of a recursively defined function is done by induction over the definition of the function. The following example illustrates this. We first give a recursive definition of a function (the Fibonacci numbers) and then prove a general property about all Fibonacci numbers.

Example 86. (Number theory) We want to characterize the Fibonacci numbers using the golden ratio. The solution gives a rather surprising and elegant combination of two very different concepts, which at first glance do not seem to have anything to do with each other. *The Fibonacci numbers* $F(0), F(1), F(2), F(3), \ldots$ form the sequence

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, \ldots$$

The *golden ratio* is defined as $\varphi = \dfrac{1+\sqrt{5}}{2}$

We want to prove that

$$F(n) = \frac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$$

for $n = 0, 1, 2, \ldots$.

We prove this proposition with a structured derivation. We start by defining the Fibonacci numbers and the golden ratio as assumptions:

$+$ $F : \mathbb{N} \to \mathbb{N}$, — we denote the Fibonacci numbers by F

{The function F is defined recursively}

$F(0) = 0$,

$F(1) = 1$, and

$F(n) = F(n-1) + F(n-2)$, when $n \geq 2$

$+$ $\varphi : \mathbb{R}$ — we denote the golden ratio by φ

{Fraction is well-defined}

$$\varphi = \frac{1+\sqrt{5}}{2}$$

Note that the definition of Fibonacci numbers has two base cases, $n = 0$ and $n = 1$. This is because the induction step defines $F(n)$ in terms of two smaller values, $n-1$ and $n-2$, so we need two base cases to get going. The defining condition for Fibonacci numbers could as well have been written as a single logical proposition,

$$F(0) = 0 \wedge F(1) = 1 \wedge (\forall n \in \mathbb{N} : n \geq 2 \Rightarrow F(n) = F(n-1) + F(n-2))$$

but the less formal writing is somewhat more easy to read, and means the same

We will use a more verbose presentation of structured derivations here, to show the overall structure of the argumentation more clearly: we write "Lemma" for the first two tasks, and "Theorem" for the last task. The first lemma shows that the golden ratio is a solution to a certain second-degree equation.

Lemma 1: $\big(x^2 - x - 1 = 0\big) \equiv (x = \varphi \vee x = (1-\varphi))$

$\Vdash$ $x^2 - x - 1 = 0$

$\equiv$ {solve the second-degree equation}

$$x = \frac{1+\sqrt{5}}{2} \vee x = \frac{1-\sqrt{5}}{2}$$

$\equiv$ {definition of the golden ratio, $\dfrac{1-\sqrt{5}}{2} = 1 - \dfrac{1+\sqrt{5}}{2}$}

$x = \varphi \vee x = (1-\varphi)$

□

The second lemma shows that the golden ratio satisfies a certain recursive relation:

Lemma 2: $\varphi^{n+1} = \varphi^n + \varphi^{n-1}$ and $(1-\varphi)^{n+1} = (1-\varphi)^n + (1-\varphi)^{n-1}$

- $n \geq 1$.

$\Vdash$ T

$\Rightarrow$ {Lemma 1}

$\varphi^2 - \varphi - 1 = 0 \wedge (1-\varphi)^2 - (1-\varphi) - 1 = 0$

$\equiv$ {regroup the terms}

$\varphi^2 = \varphi + 1 \wedge (1-\varphi)^2 = (1-\varphi) + 1$

$\Rightarrow$ {multiply both sides by φ^{n-1}, where $n \geq 1$}

$\varphi^{n+1} = \varphi^n + \varphi^{n-1} \wedge (1-\varphi)^{n+1} = (1-\varphi)^n + (1-\varphi)^{n-1}$

□

We can now prove our theorem using these two lemmas. We prove the theorem using complete induction. In the proof, we consider the two first cases, $n = 0$ and $n = 1$ separately, and then the general case, $n \geq 2$.

Theorem: $F(n) = \dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$, for $n = 0, 1, 2, \ldots$

$\Vdash$ {Proof by complete induction}

- *Base case 0:* Show that $F(n) = \dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$, when

- $n = 0$

$\Vdash$ $\dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$

$=$ {calculate the value of the expression}

0

$=$ {definition of the Fibonacci numbers}

$F(0)$

□

- *Base case 1:* Show that $F(n) = \dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$, when

- $n = 1$

$\Vdash$ $\dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$

$=$ {assumption $n = 1$}

$\dfrac{\varphi - (1-\varphi)}{\sqrt{5}}$

$=$ {calculate}

$\dfrac{2\varphi - 1}{\sqrt{5}}$

$=$ {insert the value of φ}

$\dfrac{2 \cdot \dfrac{1+\sqrt{5}}{2} - 1}{\sqrt{5}}$

$=$ {calculate}

1

$=$ {definition of the Fibonacci numbers}

$F(1)$

□

- *Induction step:* Show that $F(n) = \dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$, when

- $n \geq 2$, and

- $F(m) = \dfrac{\varphi^m - (1-\varphi)^m}{\sqrt{5}}$ for every $m < n$

$\Vdash$ $F(n)$

$=$ {definition of the Fibonacci numbers, $n \geq 2$}

$F(n-1) + F(n-2)$

$=$ {induction hypothesis}

$\dfrac{\varphi^{n-1} - (1-\varphi)^{n-1}}{\sqrt{5}} + \dfrac{\varphi^{n-2} - (1-\varphi)^{n-2}}{\sqrt{5}}$

$=$ {regroup the terms}

$\dfrac{\varphi^{n-1} + \varphi^{n-2} - ((1-\varphi)^{n-1} + (1-\varphi)^{n-2})}{\sqrt{5}}$

$=$ {Lemma 2}

$\dfrac{\varphi^n - (1-\varphi)^n}{\sqrt{5}}$

□

□ ■

17.3 Non-linear Proofs

The problem in proving existentially quantified formulas is to come up with an appropriate value for the witness. We want to prove the proposition $(\exists x \in \mathbb{R} :\ Q(x))$, but how do we find a value of t such that we can prove that $Q(t)$ is true? It is often difficult to directly see a suitable value for x in the proposition. The best strategy seems to be to postpone the decision while calculating ahead in the proof. If we are on the right track, then we will later see how to choose a value t for x that makes $Q(x)$ true. Since we are dealing with the existential quantifier, any value t that makes $Q(t)$ true is OK.

A proof that uses the witness rule is usually carried out as follows. On the left below, we show how we postpone the decision about a suitable value for x, writing a question mark for the value. As we proceed with the proof, we eventually find a value t that makes $Q(t)$ true. This part of the proof is marked in blue on the left. We then replace assumption $x =?$ with assumption $x = t$, and refer to this assumption in the step where we discovered that t would be a suitable witness. The proof then looks like the one to the right, and we have proved the existential proposition.

● Show that $(\exists x \in \mathbb{R} : Q(x))$
- Φ
$\Vdash$ {The witness rule}

  ● Show that $Q(x) \wedge x \in \mathbb{R}$
  - $x =?$
  $\Vdash$ $Q(x)$
  $\equiv$ $\{\dots\}$
  $\vdots$
  $\equiv$ $\{Q(t) \wedge t \in \mathbb{R}$ is true if we choose $x = t\}$
  $\vdots$
  T
  □

□

● Show that $(\exists x \in \mathbb{R} : Q(x))$
- Φ
$\Vdash$ {The witness rule}

  ● Show that $Q(x) \wedge x \in \mathbb{R}$
  - $x = t$
  $\Vdash$ $Q(x)$
  $\equiv$ $\{\dots\}$
  $\vdots$
  $\equiv$ {assumption $x = t$}
  $\vdots$
  T
  □

□

We have here an example of a *nonlinear* proof. In other words, we did not construct the proof in the order that it is written down. We started with one proof, and changed the proof when we got more information. For the finished proof, it is irrelevant in which order we wrote the steps, as long as we can justify each proof step. But it is important that we know the order in which the proof was constructed when we want to understand it. In this case, the order in which the proof was created shows that we did not invent the witness out of thin air, using some divine inspiration, but rather calculated it in a systematic and rational fashion.

We need a notation for nonlinear proofs that shows that we are correcting earlier steps. We use the notation

...old text...//...new text...

to mark that we have replaced the old text by new text. In handwritten proofs we can also use the notation

~~...old text...~~ ...new text...

but this is a more difficult notation in print.

We will in subsequent examples mark in red the point where we see which value is a suitable witness, and the assumption that we make based on this value. When encountering this notation in a nonlinear proof, we should first read the old version of the text, until we reach the point where we see what assumption we should make. Then we make the correction in the text, writing the new text that replaces the old, and continue with the proof. We construct the proof in the same way, i.e. we leave the assumption open until we can see for which value the proof checks out, and then we correct the assumption with an appropriate value.

Example 87. Show that the function $f(x) = \dfrac{1}{x^2+1}$ is bounded. We will describe the proof in a step-by-step manner. As before,we indicate how the different versions change by coloring new text red and old text black. Each step will prove a subproblem, possibly by adding new subproblems to solve.

Step 0 Let us start by formulating the problem.

- Show that f is bounded, when

- $f(x) = \dfrac{1}{x^2+1}$ for $x \in \mathbb{R}$

Step 1 We now try to reduce this problem to simpler problems to prove. The definition of boundedness involves two alternating quantifiers, so we will need nonlinear reasoning.

- Show that f is bounded, when

- $f(x) = \dfrac{1}{x^2+1}$ for $x \in \mathbb{R}$

$\Vdash$ f is bounded

$\equiv$ {definition of a bounded function}

$(\exists M > 0 : \forall x \in \mathbb{R} : f(x) \leq M)$

$\equiv$ {the witness rule}

- Show that $(\forall x \in \mathbb{R} : |f(x)| \leq M)$
  $M = ?$ — we look below for a suitable value for M

... T

□

This reduces the proof of boundedness to a simpler proof using the witness rule: showing that $(\forall x \in \mathbb{R} : |f(x)| \leq M)$ for a suitably chosen value (witness) for M. The specific value of M is left open for the moment ($M =?$). We will fill in the correct value when our calculations suggest a suitable candidate. (If we apply the witness rule strictly, we should also show that $M \in \mathbb{R}$. However, for simplicity, we omit this condition here and in the sequel in situations where any real number is acceptable as a witness.)

Step 2. Our next step is to apply the generalization rule to prove the formula $(\forall x \in \mathbb{R} : |f(x)| \leq M)$:

• Show that f is bounded, when

- $f(x) = \dfrac{1}{x^2+1}$ for $x \in \mathbb{R}$

⊩ f is bounded

≡ {definition of a bounded function}

$(\exists M > 0 : \forall x \in \mathbb{R} : f(x) \leq M)$

≡ {the witness rule}

> • Show that $(\forall x \in \mathbb{R} : |f(x)| \leq M)$
>
> $M = ?$ — we look below for a suitable value for M
>
> ⊩ {Generalization, choose an arbitrary x}
>
> > • Show that $|f(x)| \leq M$, when
> >
> > \- $x \in \mathbb{R}$
>
> □

... T

□

Step 3. We now proceed to find a suitable value for M.

• Show that f is bounded, when

- $f(x) = \dfrac{1}{x^2+1}$ for $x \in \mathbb{R}$

⊩ f is bounded

≡ {definition of a bounded function}

$(\exists M > 0 : \forall x \in \mathbb{R} : f(x) \leq M)$

≡ {the witness rule}

• Show that $(\forall x \in \mathbb{R} : |f(x)| \leq M)$

$M =?$ — we look for a suitable value for M

$\Vdash$ {Generalization, choose an arbitrary x}

• Show that $|f(x)| \leq M$, when

- $x \in \mathbb{R}$

$\Vdash$ $|f(x)| \leq M$

$\equiv$ {assumption (1)}

$$\left|\frac{1}{x^2+1}\right| \leq M$$

$\equiv$ {the expression in the absolute value is positive}

$$\frac{1}{x^2+1} \leq M$$

$\equiv$ {choose $M = 1$} — we have found a suitable value for M

$$\frac{1}{x^2+1} \leq 1$$

$\equiv$ {multiplying by $x^2 + 1$, the expression is always positive and therefore cannot switch the direction of the inequality sign}

$1 \leq x^2 + 1$

$\equiv$ {add -1 to both sides}

$x^2 \geq 0$

$\equiv$ {true when $x \in \mathbb{R}$}

T

□

□

$\ldots$ T

□

Step 4. Correcting the assumption now gives us the final proof.

• Show that f is bounded, when

- $f(x) = \dfrac{1}{x^2+1}$ for $x \in \mathbb{R}$

$\Vdash$ f is bounded

$\equiv$ {definition of a bounded function}

$(\exists M > 0 : \forall x \in \mathbb{R} : f(x) \leq M)$

$\equiv$ {the witness rule}

• Show that $(\forall x \in \mathbb{R} : |f(x)| \leq M)$

$M =?$ // $M = 1$

$\Vdash$ {Generalization, choose an arbitrary x}

• Show that $|f(x)| \leq M$, when

- $x \in \mathbb{R}$

$\Vdash$ $|f(x)| \leq M$

$\equiv$ {assumption (1)}

$\left|\dfrac{1}{x^2+1}\right| \leq M$

$\equiv$ {the expression in the absolute value is positive}

$\dfrac{1}{x^2+1} \leq M$

$\equiv$ {assumption $M = 1$}

$\dfrac{1}{x^2+1} \leq 1$

$\equiv$ {multiplying by x^2+1, the expression is always positive and therefore cannot switch the direction of the inequality sign}

$1 \leq x^2 + 1$

$\equiv$ {add -1 to both sides}

$x^2 \geq 0$

$\equiv$ {true when $x \in \mathbb{R}$}

T

□

□

$\ldots$ T

□ ■

17.4 The Epsilon-delta Method

We can prove a proposition of the form $(\forall x \exists y : \ldots)$ using the so called *epsilon - delta* method (i.e. the proposition is of the form $(\forall \epsilon \exists \delta : \ldots)$). The reason that students perceive this method as difficult is that this is the first time they are seriously confronted with nonlinear proofs. It is easier to understand the method when we explicitly describe the proof in the way that we explained above. We show with two examples how to prove propositions with the epsilon-delta method: a proof of continuity and a proof of uniform continuity. We write $(\forall x > 0 : \ldots)$ for $(\forall x \in \mathbb{R}^+ : \ldots)$, and similarly for existential quantification.

The first example will be presented in a stepwise manner.

Example 88. Show that the function $f(x) = \sqrt{x}$ is continuous for $x_0 > 0$.

Step 0. We start by formulating the problem:

- Show that f is continuous at point x_0, when

(1) $f(x) = \sqrt{x}$ for $x \in \mathbb{R}$, $x \geq 0$

(2) $x_0 > 0$

Step 1 We will start from the definition of continuity. This involves three alternating quantifiers. Our first task is to get rid of the outermost universal quantifier $\forall \epsilon > 0$. This reduces the overall proof to proving an existential statement, of the form $(\exists \delta > 0 : \forall x \geq 0 : \ldots)$.

- Show that f is continuous at point x_0, when

(1) $f(x) = \sqrt{x}$ for $x \in \mathbb{R}$, $x \geq 0$

(2) $x_0 > 0$

$\Vdash$ f is continuous at x_0

$\equiv$ {definition of continuity}

$(\forall \epsilon > 0 : \exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$

$\equiv$ {generalization, choose ϵ arbitrarily}

- Show that $(\exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when

- $\epsilon > 0$

... T

□

Step 2 Our next step is to prove the existentially quantified formula, using nonlinear reasoning.

- Show that f is continuous at the point x_0, when

(1) $f(x) = \sqrt{x}$ for $x \in \mathbb{R}$, $x \geq 0$

(2) $x_0 > 0$

$\Vdash$ f is continuous at x_0

$\equiv$ {definition of continuity}

$(\forall \epsilon > 0 : \exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$

$\equiv$ {generalization, choose ϵ arbitrarily}

- Show that $(\exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when

- $\epsilon > 0$

$\Vdash$ {The witness rule}

- • Show that $(\forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when
- - $\delta =?$ – find a suitable value for δ

□

... T

□

Step 3 Next, we start proving the universally quantified formula $(\forall x \geq 0 : \ldots)$

• Show that f is continuous at the point x_0, when

(1) $f(x) = \sqrt{x}$ for $x \in \mathbb{R}$, $x \geq 0$

(2) $x_0 > 0$

$\Vdash$ f is continuous at x_0

$\equiv$ {definition of continuity}

$(\forall \epsilon > 0 : \exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$

$\equiv$ {generalization, choose ϵ arbitrarily}

- • Show that $(\exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when
- - $\epsilon > 0$

$\Vdash$ {The witness rule}

- • Show that $(\forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when
- - $\delta =?$ – find a suitable value for δ

$\Vdash$ {Generalization, choose $x \geq 0$ arbitrarily}

- • Show that $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$, when
- - $x \geq 0$

□

□

... T

□

Step 4 We can now use a direct proof to prove the remaining implication $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$.

• Show that f is continuous at the point x_0, when

(1) $f(x) = \sqrt{x}$ for $x \in \mathbb{R}$, $x \geq 0$

(2) $x_0 > 0$

$\Vdash$ f is continuous at x_0

$\equiv$ {definition of continuity}

$(\forall \epsilon > 0 : \exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$

$\equiv$ {generalization, choose ϵ arbitrarily}

• Show that $(\exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when

- $\epsilon > 0$

$\Vdash$ {The witness rule}

• Show that $(\forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when

- $\delta =?$ – find a suitable value for δ

$\Vdash$ {Generalization, choose $x \geq 0$ arbitrarily}

• Show that $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$, when

- $x \geq 0$

$\Vdash$ {Proof of implication}

• Show that $|f(x) - f(x_0)| < \epsilon$, when

- $|x - x_0| < \delta$

□

□

□

. . . T

□

Step 5 Finally, we can start from the proposition to prove, $|f(x) - f(x_0)| < \epsilon$, and derive a value for δ that allows the proof to go through.

• Show that f is continuous at the point x_0, when

(1) $f(x) = \sqrt{x}$ for $x \in \mathbb{R}$, $x \geq 0$

(2) $x_0 > 0$

$\Vdash$ f is continuous at x_0

$\equiv$ {definition of continuity}

$(\forall \epsilon > 0 : \exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$

$\equiv$ {generalization, choose ϵ arbitrarily}

> • Show that $(\exists \delta > 0 : \forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when
>
> \- $\epsilon > 0$
>
> ⊩ {The witness rule}
>
>> • Show that $(\forall x \geq 0 : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon)$, when
>>
>> \- $\delta =?$ // $\delta = \epsilon \cdot \sqrt{x_0}$
>>
>> ⊩ {Generalization, choose $x \geq 0$ arbitrarily}
>>
>>> • Show that $|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \epsilon$, when
>>>
>>> \- $x \geq 0$
>>>
>>> ⊩ {Proof of implication}
>>>
>>>> • Show that $|f(x) - f(x_0)| < \epsilon$, when
>>>>
>>>> \- $|x - x_0| < \delta$
>>>>
>>>> ⊩ $|f(x) - f(x_0)|$
>>>>
>>>> $=$ {assumption (1)}
>>>>
>>>> $\left|\sqrt{x} - \sqrt{x_0}\right|$
>>>>
>>>> $=$ {expand by the conjugate of $\sqrt{x} - \sqrt{x_0}$, $\sqrt{x} + \sqrt{x_0}$, which is positive}
>>>>
>>>> $\dfrac{|x - x_0|}{\sqrt{x} + \sqrt{x_0}}$
>>>>
>>>> $\leq$ {make the denominator smaller, assumption (2)}
>>>>
>>>> $\dfrac{|x - x_0|}{\sqrt{x_0}}$
>>>>
>>>> $<$ {assumption $|x - x_0| < \delta$}
>>>>
>>>> $\dfrac{\delta}{\sqrt{x_0}}$
>>>>
>>>> $=$ {assumption $\delta = \epsilon \cdot \sqrt{x_0}$ } — we have found a suitable value for δ.
>>>>
>>>> ϵ
>>>>
>>>> □
>>>
>>> □
>>
>> □
>
> □

... T

□

The proof is nonlinear. Only when we reach the end of the innermost derivation do we see that we can prove the required proposition by choosing $\delta = \epsilon \cdot \sqrt{x_0}$. Since any value of δ that satisfies the condition $\delta > 0$ will do, we can return to the beginning of the proof and choose this value of δ as a witness in the assumption. The proof is then correct. ■

This proof has a total of four levels of indentations, one for each quantifier and another to prove implication. With more experience in using of quantifier rules, we can combine several inference steps into one bigger step, reducing the level of nesting and in that way simplify the proof. We show this in the next example.

The proof was constructed step by step, for pedagogical reasons. In practice, one would construct the proof in one go, and let the structured derivation format show how each step was constructed. This is made even simpler by the fact that at each step, there is essentially just one proof strategy that one can use, and this is determined by the outermost quantifier. For universal quantification, we use the generalization rule, and for existential quantifiers, we use the witness rule. We illustrate this with the next example.

Example 89. Prove that the function $f(x) = 2x$ is uniformly continuous. The definition of uniform continuity has three alternating quantifiers, $(\forall \epsilon > 0 : \exists \delta > 0 : \forall x, y : \ldots)$. We will here combine proof strategies, to reduce the number of required nested proofs.

• Prove that $f(x)$ is uniformly continuous

- when $f(x) = 2x$ for each $x \in \mathbb{R}$

⊩ $f(x)$ is uniformly continuous

≡ {definition of uniform continuity}

$(\forall \epsilon > 0 : \exists \delta \in \mathbb{R} : \forall x, y \in \mathbb{R} : y < x \wedge x - y < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$

≡ {generalization, choose an arbitrary ϵ}

> • Show that $(\exists \delta \in \mathbb{R} : \forall x, y \in \mathbb{R} : y < x \wedge x - y < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$, when
>
> \- $\epsilon > 0$
>
> ⊩ {The witness rule, find a suitable value for δ}
>
> > • Show that $(\forall x, y \in \mathbb{R} : y < x \wedge x - y < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$, when
> >
> > \- $\delta = ?$ $//\delta = \epsilon/2$ – find a suitable value for δ
> >
> > ⊩ {Generalization, choose arbitrary x and y, proof of implication}
> >
> > > • Show that $|f(x) - f(y)| < \epsilon)$, when
> > >
> > > \- $x, y \in \mathbb{R}$
> > >
> > > \- $y < x$ and $x - y < \delta$
> > >
> > > ⊩ $|f(x) - f(y)| < \epsilon)$
> > >
> > > ≡ {definition of $f(x)$}

$|2x - 2y| < \epsilon$

$\equiv$ $\{$assumptions$\}$

$2x - 2y < \epsilon$

$\equiv$ $\{$arithmetics$\}$

$x - y < \frac{\epsilon}{2}$

$\equiv$ $\{$assumption $\delta = \frac{\epsilon}{2}\}$ – suitable value for δ found

$x - y < \delta$

□

□

□

$\ldots \quad T$

□

This proof is a straightforward application of the inference rules for quantifiers, the witness rule for existential quantification and the generalization rule for universal quantification. It is easy to see that all the action happens in the innermost derivation, where we are looking for a suitable value for δ. Most epsilon-delta proofs will follow this pattern. Once one has understood the general structure of this kind of proofs, it should be rather straightforward to prove similar results for other functions. ■

17.5 Assignments

1. Given the formulae $(\forall x : s(x) \Rightarrow p(x))$, $(\forall x : s(x) \wedge p(x) \Rightarrow q(x))$ and $s(m)$, prove $q(m)$.
2. Prove by induction that $1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$.
3. Prove that sum of even numbers from 2 to $2n$ is given by $n(n + 1)$.
4. Prove using induction that the sum of the squares of the integers 1 to n is given by $\frac{n(n+1)(n+2)}{6}$.
5. Prove using induction that $1^3 + 2^3 + 3^3 + \cdots + n^3 = \frac{n^2(n+1)^2}{4}$
6. Prove $\left(\forall n \in \mathbb{N} : \frac{3^{2n+1}+1}{4} \in \mathbb{Z}\right) \equiv T$.
7. Using the Epsilon-Delta-method, prove that the function f, where $f(x) = 2x + 3$, is continuous at every point x_0 .
8. Using the Epsilon-Delta-method, prove that every linear function f, where $f(x) = ax + b$ and $a, b \in \mathbb{R} \wedge a \neq 0$, is continuous at every point x_0 .
9. Using the Epsilon-Delta-method, prove that the function f, given by $f(x) = 2x$, is uniformly continuous on $\mathbb{R}$.

10. Using the Epsilon-Delta-method, prove that the function f, given by $f(x) = \sin(x)$, is uniformly continuous on $\mathbb{R}$.

Chapter 18

Structuring the context

We have introduced structured derivations as a general method for describing a mathematical problem, the context in which the problem is solved, and the solution to the problem. Our main purpose has been to show how to structure mathematical arguments to make them more readable and easier to understand. We have in previous chapters shown how to structure tasks and general derivations. We will now look at how to structure the context itself, i.e., the environment in which the mathematical argument is carried out.

The mathematical context is essentially the specific model that we are working in, together with the general background of established mathematical theories. Rather than building the context as a single long derivation of declarations, facts, definitions, assumptions, and tasks, essentially putting all the stuff into one big background theory, we can chop up the derivation into a number of smaller theories, each covering some specific aspect of mathematics, and then combining these to create larger and more complex theories. This is the way mathematics itself has evolved, by identifying sub-branches of mathematics, defining these as theories of their own, analyzing the consequences that can be drawn in these theories, and then allowing the results to be reused in other mathematical studies.

Our approach here is practical, we want to show how to structure the mathematical context we use. A more detailed understanding of these issues would take us into the area of metamathematics and mathematical logic proper, for which there is ample literature available elsewhere. It is a fascinating area of research, and we encourage everybody who is interested to probe further into this area to get a deeper understanding of mathematics and logic. General mechanisms for combining theories is particularly needed when building automatic and interactive theorem provers. Here it is important to be very clear about what theories are available to the proof engine when it searches for a proof. Our presentation here is mainly based on the way mechanized theories are handled in PVS [28] and Isabelle [29].

18.1 Theories

A *mathematical theory* is essentially a structured derivation with a name. In other words, the theory consists of a sequence of assumptions, observations (declarations, facts and definitions), and tasks:

theory

§ *name*

derivation

■ *name*

The theory starts with the symbol "§", and ends with the symbol "■". The name follows immediately after the "§" symbol (and optionally also after the "■" symbol). This is then followed by the structured derivation that makes up the theory.

A simple example of a mathematical theory is a *group*. The theory of groups is described in Table 18.1 We explain this theory in more detail below, going through its different components one by one.

We start by giving a name to the theory of groups, here simply “Group”.

§ Group

The paragraph symbol “§” indicates that we are defining a theory. We could replace this with more intuitive words, like “Theory”, or something similar. However, this would then be language dependent, while “§” is the same in all languages.

The group declares G to be a set *(the group elements).* Here *Set* is some unspecified collection of sets[1].

$+ \quad G \in Set$

For any two sets A and B, we write $A \to B$ for the set of all (total) functions from A to B. Hence, we can write $f : A \to B$ also as $f \in A \to B$.

A group has an operation $*$ that allows us to form a new element $a * b$ in the group from two group elements a and b. We declare this as a function,

$+ \quad * \in G \times G \to G$

The declaration of $*$ means that the G is closed under the $*$- operation, i.e., that for any two elements a and b in a group G, the element $a * b$ also belongs to the group

We assume that G is non-empty and that the group operation is *associative* (i.e, the way in which the group operations are combined does not matter):

(a) $G \neq \emptyset$

(b) $(\forall x, y, z \in G : \; x * (y * z) = (x * y) * z)$

This means that we can leave out the parenthesis in group terms: $a*b*c = (a*b)*c = a * (b * c)$.

We will assume that there is an identity element e in a group. This means that multiplying any group element with the identity element does not change the value of the original group element.

(c) $(\exists y \in G : (\forall x \in G : \; y * x = x \wedge x * y = x))$

This assumption allows us to give a name to the identity element:

[1] We could have different interpretations of what *Sets* is, depending on whether we are working in *axiomatic set theory* or in *higher order logic.* In the former interpretation, *Sets* is the collection of all sets that satisfy the axioms of set theory. In the second interpretation, we have a predefined hierarchy of *types.* We start from a small collection of basic types, like *truth values* and *natural numbers,* and create new types from simpler types A and B as functional types $A \to B$, product types $A \times B$, and sum type $A + B$. A set is then a subset of some type. The naive interpretation of sets as a collection of elements will suffice in the sequel.

[1] Define $e \in G$

> {assumption (c) shows that such an element exists}
>
> $(\forall x \in G:\ e * x = x \wedge x * e = x)$

We show that there is only one identity element in a group (this justifies us to choose a name e for this element).

[2] {The group identity element is unique}

> • Prove that $(\forall d \in G : (\forall x \in G : d * x = x \wedge x * d = x) \Rightarrow d = e)$
>
> $\Vdash$ {generalization}
>
> > • $(\forall x \in G : d * x = x \wedge x * d = x) \Rightarrow d = e$, where
> >
> > \- $d \in G$
> >
> > $\Vdash$ {prove implication}
> >
> > > • $d = e$
> > >
> > > (i) $(\forall x \in G : d * x = x \wedge x * d = x)$
> > >
> > > $\Vdash$ e
> > >
> > > $=$ {assumption (i), choose $x = e$}
> > >
> > > $e * d$
> > >
> > > $=$ {definition [1], e is the identity element}
> > >
> > > d
> > >
> > > □
> >
> > □
>
> □

. . . $(\forall d \in G : (\forall x \in G : d * x = x \wedge x * d = x) \Rightarrow d = e)$

Each element in a group has an *inverse element.* Multiplying the group element with its inverse gives us the identity element. We will denote the inverse of an element x with $\mathrm{inv}(x)$.

(d) $(\forall x \in G : (\exists y \in G :\ x * y = e \wedge y * x = e))$

[3] Define $inv \in G \to G$

> {the element $inv(x)$ exists, for any $x \in G$, by assumption (d)}
>
> $(\forall x \in G:\ x * inv(x) = e \wedge inv(x) * x = e)$

We can show that the inverse of any group element is unique.

[4] {Inverse element is unique}

> • $(\forall x, b, c \in G:\ x * b = e \wedge b * x = e \wedge x * c = e \wedge c * x = e \Rightarrow b = c)$

```
⊩   {generalization}
    •   $x * b = e \wedge b * x = e \wedge x * c = e \wedge c * x = e \Rightarrow b = c$, where
    -   $x, b, c \in G$
    ⊩   {prove implication}
        •   $b = c$
        -   $x * b = e \wedge b * x = e \wedge x * c = e \wedge c * x = e$
        ⊩   $b$
        =   {$e$ is identity element}
            $b * e$
        =   {$c$ is inverse of $x$}
            $b * (x * c)$
        =   {associativity}
            $(b * x) * c$
        =   {$b$ is inverse of $x$}
            $e * c$
        =   {$e$ is identity element}
            $c$
    □
□
... $(\forall x, b, c \in G : \; x * b = e \wedge b * x = e \wedge x * c = e \wedge c * x = e \Rightarrow b = c)$
■   **Group**
```

This finishes the definition of the theory of groups.

18.2 Extending theories

Rather than presenting groups as one big theory, we may split up the definition of a group into smaller theories that build on each other, until we finally have a theory of groups. We can achieve this by allowing a theory to assume another theory, with the construct

$*\qquad N$

The symbol "$*$" means that we *import* the theory called N into the present theory, or alternatively, that we *extend* the theory N with the constructs in the present theory. The effect is the same as if we would just replace the line above with the derivation named N. We then exclude the first §-line and the last ■- line, and renumber the derivation steps so that we do not use the same identification for a derivation step in the importing theory and a derivation step in the imported theory.

We start with the theory of *semigroups.*

§	**Semigroup**
+	$G \in \mathit{set}$ – declare a set of group elements
+	$* \in G \times G \to G$ – declare an operation "*" on group elements (*multiplication*)
(a)	$G \neq \emptyset$ – assume that the set is non-empty
(b)	$(\forall x, y, z \in G :\ x*(y*z) = (x*y)*z)$ -- assume that multiplication is associative
■	**Semigroup**

Table 18.2: Definition of semigroups

A *monoid* is defined as an *extension* of a semigroup: it is a semigroup with an identity element. In other words, a monoid is a special kind of semigroup which satisfies some additional properties.

§	**Monoid**
*	Semigroup
(a)	$(\exists y \in G : (\forall x \in G :\ y * x = x \wedge x * y = x))$ – assume that there is an *identity element* in the group
[1]	Define $e \in G$ – define e to be the identity element
	{assumption (a) shows that such an element exists}
	$(\forall x \in G :\ e * x = x \wedge x * e = x)$
[2]	{identity element is unique} – show that the identity element is unique
...	$(\forall d \in G : (\forall x \in G : d * x = x \wedge x * d = x) \Rightarrow d = e)$
■	**Monoid**

Table 18.3: Definition of monoid

Finally, we will now define a *group* as a monoid which has an inverse for each group element.

§	**Group**
*	Monoid
(a)	$(\forall x \in G : (\exists y \in G : x * y = e \wedge y * x = e))$ – assume that there is an *inverse* for each group element
[1]	{inverse element is unique} – show that the inverse of each group element is unique
...	$(\forall x, y, z \in G : x * y = e \wedge y * x = e \wedge x * z = e \wedge z * x = e \Rightarrow y = z)$
[2]	$inv : G \to G$ – define inv to be the inverse operation
	{the element inv exists, for any $x \in G$, by assumption (a)}
	$(\forall x \in G : x * inv(x) = e \wedge inv(x) * x = e)$
■	**Group**

Table 18.4: Definition of groups

If we now expand the theory Monoid in the theory Group, and then expand the theory Semigroup that was imported into Monoid, we will get the same definition of Group as we had originally. Note, however, that the identification of entities in the derivation changes when we have theories. In the theory Group, we could refer to the observation that an identity element is unique as [Monoid.2], because the numbering of derivation steps starts anew in each theory. Or we could simply renumber the steps in either importing or imported theory, to avoid conflicts.

The advantage of splitting up the definition of a group into smaller theories is that we now can study what are the consequences of associativity in a semigroup, without assuming that there exists an identity element. Similarly for monoids and groups. This is useful, because we can often find structures which are associative but do not have an identity element, and we would like to know what properties follow from associativity in these structures.

We can extend the above theories with a commutativity property. We could, e.g., define commutative monoids and commutative groups, as follows.

§	**Commutative monoid**
*	Monoid
(a)	$(\forall x, y \in G : x * y = y * x)$ – assume that multiplication is commutative
■	**Commutative monoid**

Table 18.5: Definition of commutative monoids

§	**Commutative group**
*	Group
(a)	$(\forall x, y \in G : x * y = y * x)$ – assume that multiplication is commutative
■	**Commutative group**

Table 18.6: Definition of commutative groups

18.3 Reusing theories

In some situations, it is not sufficient to just extend a theory with new derivation steps. We may want to change the naming of the different entities that we are reusing in our theory, or use a theory in two different ways. An example of this is a *ring.* A ring has two operations, *addition* and *multiplication,* and defines unit elements for both operations. It also postulates a distributivity property for these two operations. In fact, we will later need multiplication to also be commutative, so we will define directly a *commutative ring.*

We can define a theory of commutative rings as follows:

§	**Commutative ring**
+	$R \in set$ – element of the ring
+	$+ \in R \times R \to R$ – addition in ring
+	$\cdot \in R \times R \to R$ – multiplication in ring
+	$0 \in R$ – identity element of addition
+	$1 \in R$ – identity element of multiplication
*	Commutative monoid $[G := R, * := \cdot, e := 1]$ – multiplication forms a commutative monoid
*	Commutative group $[G := R, * := +, e := 0, \mathrm{inv} := -]$ – addition forms a commutative group
(a)	$(\forall x, y, z \in R : x \cdot (y + z) = x \cdot y + x \cdot z)$ – multiplication distributes from the right
(b)	$(\forall x, y, z \in R : (x + y) \cdot z = x \cdot z + y \cdot z)$ – multiplication distributes from the left
■	**Commutative ring**

Table 18.7: Definition of commutative rings

The ring thus has two different binary operations on ring elements, addition and multiplication. Multiplication forms a commutative monoid on ring elements, which we indicate by writing

$*$ Commutative monoid $[G := R, * := \cdot, e := 1]$

The import statement copies the theory of Commutative monoid to the definition of a Commutative ring, but with the following adaptions:

- The declarations of $G, *$ are omitted (we use the declarations of R and "$\cdot$" instead)
- We substitute the name R for G, the name "$\cdot$" for "$*$", and 1 for e in the rest of the theory Commutative monoid, and then copy this to the theory of Commutative ring.
- We change the numbering of the different items in the resulting theory, so that references are still the same as in the original theory.

In other words, we say that the set R forms a commutative monoid with unit element 1 and monoid operation "$\cdot$". In a similar way, addition forms a commutative group in a ring, with R as the group elements, "$*$" as the group operation, 0 as the group identity element, and "$-$" as the group inverse.

$*$ Commutative group $[G := R, * := +, e := 0]$

The two additional assumptions describe how multiplication and addition behave when both occur in the same expression. In this case, we assume that multiplication distributes over addition, both from the left and from the right.

This is the way mathematics is usually presented, as a large collection of little theories. The beauty of this method is that the connections between the different theories are dynamic. Any new results that are derived for commutative monoids or for commutative groups are directly available also for commutative rings. In this way, mathematicians can work on different areas of mathematics (read: different theories), and still know that all result that they derive in their own theory will be immediately available for mathematicians working on theories that build on these theories.

The way these theories are connected to each other is illustrated in the diagram in Table 18.8. Here each contour is a theory. A contour nested in another theory is an extension, inheriting all the derivation steps of the surrounding theory. The hollow arrow shows that the theory reuses the theory pointed to, with the assignment indicated on the arrow.

18.4 Theory interpretations

A t*heory interpretation* is essentially a fact that states that some specific operations in the theory A form another theory B. As an example, the real numbers form a commutative group, when G is interpreted as $\mathbb{R} - \{0\}$, group multiplication is interpreted as multiplication on real numbers (“$\cdot$”), the group unit element e is 1, and the group inverse “*inv*“ is interpreted as the inverse value on reals. We can then show that all the assumptions of a commmutative group are satisfied with these changes:

$+$ Define $\mathbb{R}^+ \in 2^{\mathbb{R}}$

$\{\mathbb{R}^+$ is a subset of the real numbers $\mathbb{R}\}$

$\mathbb{R}^+ = \mathbb{R} - \{0\}$

$+$ {Check that the commutative group assumptions are all satisfied for real numbers different from 0, when group multiplication is multiplication on reals}

• $\mathbb{R}^+$ is not empty

$\Vdash$ {by definition of reals, $1 \in \mathbb{R}$}

□

• $\mathbb{R}^+$ is closed under multiplication

$\Vdash$ $\{a, b \neq 0 \Rightarrow a \cdot b \neq 0\}$

□

• Multiplication is associative in $\mathbb{R}^+$

$\Vdash$ $\{a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for reals $a, b, c\}$

□

• There is an identity element for multiplication

$\Vdash$ {1 is an identity element in $\mathbb{R}^+$: $1 \cdot a = a \cdot 1 = a$}

□

• There is an inverse element in $\mathbb{R}^+$ for each non-zero real number

$\Vdash$ {choose x^{-1} as the identity element of x; when $x \neq 0$,we have that $x \cdot x^{-1} = x^{-1} \cdot x = 1$}

□

• Multiplication is commutative for real numbers in $\mathbb{R}^+$

$\Vdash$ $\{x \cdot y = y \cdot x$ holds for any two real numbers x and $y\}$

□

... Commutative group $[G := \mathbb{R}^+, e := 1, * := \cdot \mathit{inv}{:=}^{-1}]$

The nested derivations show that each of the assumptions that we make for a commutative group is satisfied, with the interpretation of the group operations and constants that we have given. We need to also show that multiplication on reals is closed under this interpretation for group multiplication (the second nested task). Hence, with this interpretation, the non-zero real numbers form a commutative group. This means that we are now free to use any theorems that have been proved for commutative groups when we reason about non-zero real numbers, with the indicated interpretation.

18.5 Ordering

We can extend the theory of rings to a complete theory of real numbers. For this, we first define the theory of *partial orders*. A partial order is a set with an ordering relation $\leq$ that is *reflexive*, *transitive* and *antisymmetric*.

§	**Partial order**
+	$P \in Set$ – declare a set of elements
+	$\leq \in P \times P \to \mathbb{B}$ – declare an ordering relation on elements in the set
(a)	$P \neq \emptyset$ – assume that the set is non-empty
(b)	$(\forall x \in P : x \leq x)$ - the ordering is reflexive
(b)	$(\forall x, y, z \in P : x \leq y \wedge y \leq z \Rightarrow x \leq z)$ – the ordering is transitive
(c)	$(\forall x, y \in P : x \leq y \wedge y \leq x \Rightarrow x = y)$ – the ordering is antisymmetric
■	**Partial order**

Table 18.9: Definition of partial order

Note that we describe the relation $\leq$ as a function from pairs of elements to truth values. Then $\leq (a, b) = T$ means that $a \leq b$.

A *total order* is a special kind of partial order, where for any two elements a and b, either $a \leq b$ or $b \leq a$. We define the total orders as follows:

§	**Total order**
*	Partial order
(a)	$(\forall x, y \in P : x \leq y \vee y \leq x)$ – the ordering is total
■	**Total order**

Table 18.10: Definition of total order

A total order is sometimes referred to as a *chain* or *linear order,* because all elements are ordered one after the other. The natural numbers are a good example of a total order: for any two numbers x and y, either $x \leq y$ or $y \leq x$. Rational numbers are also totally ordered, as are real numbers.

We will need a specific total ordering, a *complete total order,* in order to define the real numbers. This is the essential feature needed to capture the notion of reals.

§	**Complete total order**
*	Total order
[1]	$ub \in P \times 2^P \to \mathbb{B}$ – upper bound relation
	$\{ub$ is a total function$\}$
	$ub(x, S) \equiv (\forall s \in S : s \leq x)$ – x is an upper bound of the set S, iff $s \leq x$ for any s in S.
[2]	$lub \in P \times 2^P \to \mathbb{B}$ – least upper bound relation
	$\{lub$ is a total function$\}$
	$lub(x, S) \equiv up(x, S) \wedge (\forall y \in P : ub(y, S) \Rightarrow x \leq y)$
(a)	$(\forall S \in 2^P : (\exists x \in P : ub(x, S)) \Rightarrow (\exists x \in P : lub(x, S))$
■	**Complete total order**

Table 18.11: Definition of complete total orders

Consider an arbitrary subset S of P (we write 2^P for the set of all subsets of P, so $S \in 2^P$). An element x in P is an upper bound of the set S, if it is greater or equal to all elements in S. A set S is said to be *bounded,* if it has an upper bound. The element x is a least upper bound of S, if it is the smallest upper bound of S. We can show that the least upper bound is unique, if it exists. The assumption in the theory of complete total orders is that each bounded subset S of P has a least upper bound.

18.6 Theory of real numbers

We are now ready to define the real numbers. First, we extend rings to *fields.* A field is a ring with an inverse operation for non-zero numbers.

§	**Field**
*	Commutative ring
+	$^{-1} \in R \to R$ – inverse number
(a)	$(\forall x \in R : x \neq 0 \Rightarrow x \cdot x^{-1} = 1 \wedge x^{-1} \cdot x = 1)$ – property of inverse numbers
■	**Field**

Table 18.12: Definition of fields

Finally, we define the theory of *real numbers* as an extension of the theory of fields.

§	**Real numbers**
+	$\mathbb{R} \in Set$
+	$+ \in \mathbb{R} \times \mathbb{R} \to \mathbb{R}$
+	$\cdot \in \mathbb{R} \times \mathbb{R} \to \mathbb{R}$
+	$^{-1} \in \mathbb{R} \to \mathbb{R}$
+	$\leq \in \mathbb{R} \times \mathbb{R} \to \mathbb{B}$
*	Field $[R := \mathbb{R}, + := +, \cdot := \cdot, 0 := 0, 1 := 1, ^{-1} := ^{-1}]$
+	$\leq \in \mathbb{R} \times \mathbb{R} \to \mathbb{B}$ – ordering of reals
*	Complete total order $[P := \mathbb{R}, \leq := \leq]$
(a)	$(\forall x, y, z \in R : x \leq y \Rightarrow x + z \leq y + z)$ – addition is monotonic for ordering
(b)	$(\forall x, y \in R : 0 \leq x \wedge 0 \leq y \Rightarrow 0 \leq x \cdot y)$ – product of two non-negative numbers is non-negative
■	**Real numbers**

Table 18.13: Definition of real numbers

Thus, the theory of real numbers extends the theory of fields. In addition, it assumes that the ordering defined on real numbers is a complete total order. Two additional axioms are needed, that addition is monotonic for the ordering of real numbers, and that the product of two non-negative real numbers is non-negative. That is all that is needed to get us going with creating the whole complex theory of real numbers, needed for algebra, analysis, geometry, etc.

We can describe these different theories more intuitively as nested diagrams, as shown in Table 18.14. Here each box stands for a theory. A box inherits all properties of the outer boxes. The diagram describes how rings are related to the previous theories. The hollow arrows that are annotated with assignments correspond to the two import assumptions in the definition of a ring.

We can collapse all the theories above into a single theory for reals, which collects all axioms for real numbers into one place. This is shown in Table 18.15. We have here renumbered the derivation steps, to get a uniform numbering.

18.7 Consistency

We can choose the assumptions in a theory freely. However, if we are careless, our assumptions can make the theory useless. This happens when the assumptions are *inconsistent* with each other. The assumptions $A_1, \ldots, A_m$ are *inconsistent*, if their combination is equivalent to false, i.e., $A_1 \wedge \ldots \wedge A_m \equiv F$. Because F implies every proposition, this means that $A_1, \ldots, A_m$ is inconsistent when $A_1 \wedge \ldots \wedge A_m \Rightarrow F$. In a inconsistent theory, every proposition P follows from the assumptions.

The set of assumptions $A_1, \ldots, A_m$ is *consistent,* if they are not inconsistent, i.e., if $A_1 \wedge \ldots \wedge A_m \not\Rightarrow F$. A theory is consistent, if its assumptions are consistent. A consistent theory has a *mathematical model.* This means that the constants in the theory can be interpreted in such a way that each assumption is true in the interpretation. An inconsistent theory does not have any mathematical model.

Example 90. Assume that $\leq$ is a partial order, and that $a < b$, $b < c$, $c < a$. We can then derive a contradiction (i.e., F) from these assumptions, as follows:

• F

* Partial order

- $a < b \wedge b < c \wedge c < a$

$\Vdash$ $a < b \wedge b < c \wedge c < a$

$\Rightarrow$ {transitivity of $<$}

$a < a$

$\equiv$ {definition of $<$ }

$a \leq a \wedge a \neq a$

$\equiv$ {reflexivity of $=$}

$a \leq a \wedge a \neq a \wedge a = a$

$\equiv$ {associativity of conjunction}

$a \leq a \wedge (a \neq a \wedge a = a)$

$\equiv$ {contradiction}

$a \leq a \wedge F$

$\equiv$ {conjunction with false}

F

□

This shows that F follows from the assumptions that $\leq$ is a partial order, and that $a < b \wedge b < c \wedge c < a$. Hence, adding the latter in an extension of the partial order theory is not a good idea, it will make the resulting theory inconsistent. ■

We can show mathematically that a theory is consistent by giving a model for the theory. For example, we can show that partial order theory is consistent by interpreting the theory in natural numbers: choose P to be $\mathbb{N}$, and let the ordering of P be the $\leq$ ordering of natural numbers. Then we see that all the axioms of the partial order theory are true for natural numbers: the ordering is reflexive ($n \leq n$ for all natural numbers), it is transitive ($m \leq n \leq k$ implies that $m \leq k$), and it is antisymmetric ($m \leq n$ and $n \leq m$ implies that $m = n$).

Example 91. We write this same argument as a structured derivation:

* Natural numbers

+ {prove that natural numbers form a partial order}

- • $(\forall n \in \mathbb{N} : n \leq n)$
- ⊩ {property of natural numbers}
- □
- • $(\forall m, n, k \in \mathbb{N} : m \leq n \wedge n \leq k \Rightarrow m \leq k)$
- ⊩ {property of natural numbers}
- □
- • $(\forall m, n \in \mathbb{N} : m \leq n \wedge n \leq m \Rightarrow m = n)$
- ⊩ {property of natural numbers}
- □

... Partial order $[P := \mathbb{N}, \leq := \leq]$

We have shown that partial orders form a consistent theory by interpreting this theory in a stronger theory, that of natural numbers. This proves that the partial order theory is consistent, provided that we believe that the theory of natural numbers is consistent. Essentially, we just shift the burden of proof to the consistency of another theory. However, the consistency of the theory of natural numbers has a strong basis in our intuition, so we can feel reasonably sure that the theory of partial orders is consistent. ■

We can show that a theory is inconsistent by deriving a contradiction from the axioms, i.e., proving

$$A_1, \ldots, A_m \vdash F$$

which we again can derive by finding some property P such that

$$A_1, \ldots, A_m \vdash P \text{ and } A_1, \ldots, A_m \vdash \neg P$$

We can also show inconsistency by showing that one of the assumptions, say A_m, contradicts the other assumptions $A_1, \ldots, A_{m-1}$. We show that

$$A_1, \ldots, A_{m-1} \vdash \neg A_m$$

Then we can choose A_m as the offending assumption P, because we now know that $A_1, \ldots, A_m \vdash \neg A_m$ (because we can always add extra assumptions to a sequent) and $A_1, \ldots, A_m \vdash A_m$ (because A_m is an assumptions).

We are usually concerned with consistency when we make an extension of a theory. Assume that we extend theory T with the new axioms $A_1, \ldots, A_m$. Let Φ be the assumptions in T. Then the extension is inconsistent, if we can prove that

$$\Phi, A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_m \vdash \neg A_i$$

for some axiom A_i.

In conclusion, consistency is something that we need to be concerned with. There is no simple way of checking the assumptions that we make for consistency. Proving that a set of assumptions is inconsistent is straightforward, we just need to derive a contradiction from the assumptions. But proving that the set of assumptions is consistent is harder. The most natural way is to use our intuition and check that there is a real-world or mathematical model that satisfies all the assumptions.

§	**Group**
+	$G \in Set$ – declare a set of group elements
+	$* \in G \times G \to G$ – declare an operation "*" on group elements (*multiplication*)
(a)	$G \neq \emptyset$ – assume that the set is non-empty
(b)	$(\forall x, y, z \in G :\ x * (y * z) = (x * y) * z)$ -- assume that "*" is associative
(c)	$(\exists y \in G : (\forall x \in G :\ y * x = x \wedge x * y = x))$ – assume that there is an *identity element* in the group
[1]	Define $e \in G$ – define e to be the identity element
	{assumption (c) shows that such an element exists}
	$(\forall x \in G :\ e * x = x \wedge x * e = x)$
[2]	{identity element is unique} – show that the identity element is unique
...	$(\forall d \in G : (\forall x \in G : d * x = x \wedge x * d = x) \Rightarrow d = e)$
(d)	$(\forall x \in G : (\exists y \in G :\ x * y = e \wedge y * x = e))$ – assume that there is an *inverse* for each group element
[3]	{inverse element is unique} – show that the inverse of each group element is unique
...	$(\forall x, y, z \in G :\ x * y = e \wedge y * x = e \wedge x * z = e \wedge z * x = e \Rightarrow y = z)$
[4]	Define $inv \in G \to G$ – define inv to be the inverse operation
	{the element $inv(x)$ exists, for any $x \in G$, by assumption (d)}
	$(\forall x \in G :\ x * inv(x) = e \wedge inv(x) * x = e)$
■	**Group**

Table 18.1: Definition of groups

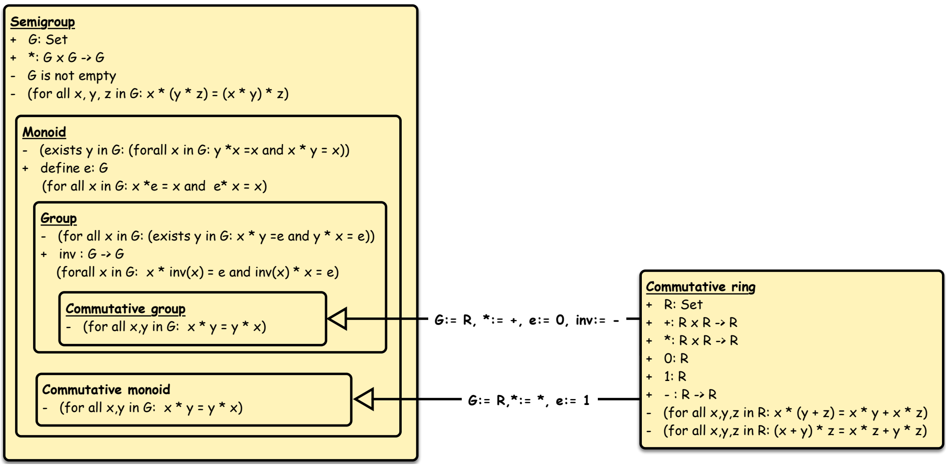

Semigroup
+ G: Set
+ *: G x G -> G
- G is not empty
- (for all x, y, z in G: x * (y * z) = (x * y) * z)
Monoid
- (exists y in G: (forall x in G: y *x =x and x * y = x))
+ define e: G
(for all x in G: x *e = x and e* x = x)
Group
- (for all x in G: (exists y in G: x * y =e and y * x = e))
+ inv : G -> G
(forall x in G: x * inv(x) = e and inv(x) * x = e)
Commutative group
- (for all x,y in G: x * y = y * x)
Commutative monoid
- (for all x,y in G: x * y = y * x)
G:= R, *:= +, e:= 0, inv:= -
G:= R,*:= *, e:= 1
Commutative ring
+ R: Set
+ +: R x R -> R
+ *: R x R -> R
+ 0: R
+ 1: R
+ - : R -> R
- (for all x,y,z in R: x * (y + z) = x * y + x * z)
- (for all x,y,z in R: (x + y) * z = x * z + y * z)

Table 18.8: Diagram of rings

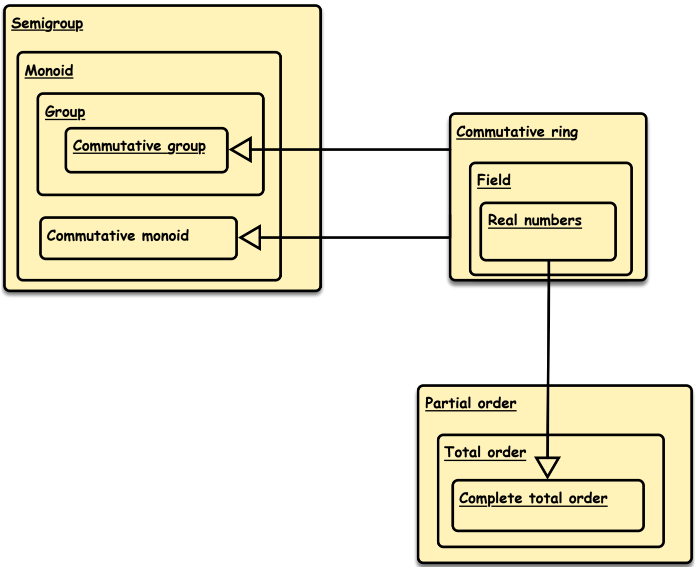


Table 18.14: Real numbers

§	**Real numbers**
+	$\mathbb{R} \in Set$ – Semigroup
(a)	$\mathbb{R} \neq \emptyset$ – Semigroup
+	$+ \in \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ – Semigroup
(b)	$(\forall x, y, z \in \mathbb{R} :\ x + (y + z) = (x + y) + z)$ – Semigroup
(c)	$(\exists y \in \mathbb{R} : \forall x \in \mathbb{R} :\ y + x = x \wedge x + y = x)$ – Monoid
[1]	Define $0 \in \mathbb{R}$ – define 0 to be the identity element {Assumption (c)} – Monoid
	$(\forall x \in \mathbb{R} :\ 0 + x = x \wedge x + 0 = x)$
(d)	$(\forall x \in \mathbb{R} : \exists y \in \mathbb{R} :\ x + y = 0 \wedge y + x = 0)$ – Group
[2]	Define $- \in \mathbb{R} \to \mathbb{R}$ {the element $-x$ exists, for any $x \in R$, by assumption (d)} – Group
	$(\forall x \in \mathbb{R} :\ x + (-x) = 0 \wedge (-x) + x = 0)$
(e)	$(\forall x, y \in \mathbb{R} :\ x + y = y + x)$ – Commutative group
+	$\cdot \in \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ – Semigroup
(f)	$(\forall x, y, z \in \mathbb{R} :\ x \cdot (y \cdot z) = (x \cdot y) \cdot z)$ – Semigroup
(g)	$(\exists y \in \mathbb{R} : \forall x \in \mathbb{R} :\ y \cdot x = x \wedge x \cdot y = x)$ – Monoid
[3]	Define $1 \in \mathbb{R}$ {Assumption (c) shows that such an element exists} – Monoid
	$(\forall x \in \mathbb{R} :\ 1 \cdot x = x \wedge x \cdot 1 = x)$
(h)	$(\forall x, y \in \mathbb{R} :\ x \cdot y = y \cdot x)$ – Commutative monoid
(i)	$(\forall x, y, z \in \mathbb{R} :\ x \cdot (y + z) = x \cdot y + x \cdot z)$ – Ring
(j)	$(\forall x, y, z \in \mathbb{R} :\ (x + y) \cdot z = x \cdot z + y \cdot z)$ – Ring
+	$^{-1} \in \mathbb{R} \to \mathbb{R}$ – Field
(k)	$(\forall x \in \mathbb{R} :\ x \neq 0 \Rightarrow x \cdot x^{-1} = 1 \wedge x^{-1} \cdot x = 1)$ – Field
+	$\leq \in \mathbb{R} \times \mathbb{R} \to \mathbb{B}$ – Partial order
(l)	$(\forall x \in R :\ x \leq x)$ - Partial order
(m)	$(\forall x, y, z \in R :\ x \leq y \wedge y \leq z \Rightarrow x \leq z)$ – Partial order
(n)	$(\forall x, y \in R :\ x \leq y \wedge y \leq x \Rightarrow x = y)$ – Partial order
(o)	$(\forall x, y \in R :\ x \leq y \vee y \leq x)$ – Total order
[4]	Define $ub \in R \times 2^{\mathbb{R}} \to \mathbb{B}$ {ub is a total function} – Complete total order
	$ub(x, S) \equiv (\forall s \in S : s \leq x)$.
[5]	Define $lub \in \mathbb{R} \times 2^{\mathbb{R}} \to \mathbb{B}$ {lub is a total function} – Complete total order
	$lub(x, S) \equiv up(x, S) \wedge (\forall y \in \mathbb{R} : ub(y, S) \Rightarrow x \leq y)$
(p)	$(\forall S \in 2^{\mathbb{R}} : (\exists x \in \mathbb{R} : ub(x, S)) \Rightarrow (\exists x \in \mathbb{R} : lub(x, S))$ – Complete total order
(q)	$(\forall x, y, z \in \mathbb{R} :\ x \leq y \Rightarrow x + z \leq y + z)$ –Real numbers
(r)	$(\forall x, y \in \mathbb{R} :\ 0 \leq x \wedge 0 \leq y \Rightarrow 0 \leq x \cdot y)$ – Real numbers
■	**Real numbers**

Table 18.15: Expanded definition of real numbers

Chapter 19

Syntax of Structured Derivation

The previous chapters have tried to explain the basic ideas of structured derivations and show how the method can be used in mathematic education in practice. The last three chapters in this book will now be devoted to explaining in more detail, and with a more precise terminology, the mathematical and logical underpinnings for structured derivations, and how this can be exploited to build computer support for structured derivations. The syntax of structured derivations is presented more precisely in this chapter. The next chapter will explain how to check that a structured derivation is mathematically correct. Chapter 21 will discuss computer support for structured derivations. Finally, Appendix A shows that structured derivations is both sound and complete, in the same way that natural deduction is sound and complete.

A central goal of structured derivations is to support different levels of rigor in mathematical argumentation, from rather informal arguments to very formal derivations and proofs. We can achieve this by making a distinction in the syntax between the *syntax in the large* and the *syntax in the small.* The former describes the overall structure of the mathematical argument: how it is organized into tasks, assumptions, facts, definitions, calculation steps, answers, and justifications (with possible nested tasks): The latter describes the specific syntax that is needed when carrying out argumentations on a more precise, logical level, making use of the laws and inference rules that we have presented in this book. We will not fix a specific logic for the detailed syntax, but assume that we use some standard logical system with the traditional syntactic categories of propositions, terms, axioms, and inference rules.

19.1 General Syntax for Structured Derivations

The syntax of structured derivations is designed so that it allows different kinds of notational conventions and traditions to be used in a derivation, as well as making it possible to choose the right level of rigor in a derivation. The general syntax of a mathematical argument is described by a *(structured) derivation.* This is a sequence of *derivation steps*. The box next here defines the syntax of a structured derivation. The box should

derivation
*derivation step**

be interpreted as follows: a (structured) derivation is written as a list of successive derivation steps, written one under the other, each step starting on a new line. The start indicates that there may be zero, one or more derivation steps in this list. We color concepts that are defined later blue, while concepts that are taken as primitive are marked with other colors.

A *derivation step* is either an *assumption*, an *observation* or a *task*. The vertical bar "|" is used to separate the alternatives from each other. We explain below how to write assumptions, observations and tasks.

derivation step
assumption | *observation* | *task*

assumption
aid   *proposition*

An assumption is a *proposition*. We mark an assumption with an *assumption identifier (aid),* written in the first column. The assumption identifier can be either "-", or a small letter in parenthesis, like (a), (b), (c), ... We write the *proposition* in the second column. An assumption is a logical statement that we may assume is true without further justification.

An *observation* is either a *declaration*, *fact* or a *definition*. We identify an *observation* with an *observation identifier (oid)* in the first column. This is either a "+" sign, or a number in square brackets, like in [1], [2], [3], ... The observation identifier is followed by a *fact* or a *definition*, written in the second column.

observation
oid   *declaration* | *fact* | *definition*

fact
justification
proposition

definition
declaration
justification
proposition

A *fact* is a *proposition* which follows from earlier assumptions and observations. The justification is an argument that is intended to convinces ourselves (and others) that the fact follows from the preceding steps of the derivation. A *definition* introduces one or more new *names* with a *declaration*, together with a *justification* that shows that these names are well-defined by the *proposition* on the next line.

There are two kinds of *tasks*: *calculation tasks* and *general tasks.*

task
calculation_task | *general_task*

A *general task* starts with a *task identifier (tid)* in the first column. This can be either a bullet "•", or a capital letter like A., B., C.,. . . . The task ends with a square "□" in the first column. The *justification* following the *proof sign* "⊩" explains why the answer is correct. The star indicate that we can have zero or more assumptions, and similarly for *observation*. A *calculation task* is simpler, it only has a task identifier, a calculation and an ending square.

general task

tid *question*

*assumption**

*observation**

⊩ *justification*

calculation

□ *answer*

calculation task

tid *calculation*

□

justification

{ *explanation* }

*task**

A *justification* can be a simple *explanation*, enclosed in curly brackets, or it may include nested tasks. The nested *tasks* are written one step to the right, i.e., they start in the same column as the explanation in curly brackets. The derivation returns to the previous level after the nested tasks. The nested tasks thus begins in the second column of the original task, since they are indented one step. There may be zero or more nested *tasks*.

calculation

expression

rel *justification*

expression

⋮

rel *justification*

expression

The three dots "⋮" in a *calculation* show that we can add zero or more steps to the first step. Every calculation step has two lines, one with a *relation* and a *justification*, and one with an *expression*.

Note that a *task* (both a general task and a calculation task) is explained in terms of *justifications*, and *justifications* are in turn explained in terms of *tasks*. We thus have a *recursive* definition of tasks: a task contains justifications, which in turn can contain nested tasks. The nested tasks can then again contain justifications, which again can contain nested tasks, etc. We can thus have any number of tasks nested inside each other. The recursion ends when we justify a step without introducing new nested tasks.

19.2 Detailed Syntax of Structured Derivations

The syntax definition for general structured derivations leaves a number of basic constructs undefined. These are *propositions, explanations, expressions, relations, declarations, questions,* and *answers.* The syntax for these constructs can be freely chosen in the general syntax, and depends on the specific mathematical domain that we are working in. However, we need to determine the syntax of these categories also, at least in somewhat more detail, if we want to make use of the logical laws and inference rules that we have described earlier. .

This more precise level of syntax will support support mathematical proofs and reasoning "*in the small*". We assume that we work in some logical system, like first order logic or higher order logic. Such a system comes with its own definitions of logical propositions, terms, and rules for how to explain that an inference step is permitted. We then make the following more precise definitions.

proposition : a logical proposition in the chosen system

explanation : application of an inference rule in the chosen system

expression : a term in the chosen system

rel : a binary relation in the chosen system

answer : a logical proposition in the chosen system

This leaves two syntactic constructs not accounted for, declarations and questions. We give the syntax for these below.

A *declaration* is a list of a names together with a domain of acceptable values that can be assigned to each name. The general form for a declaration is

$$x_1 \in U_1, \ldots, x_m \in U_m$$

Note that the declaration is also a logical proposition, it stands for $x_1 \in U_1 \wedge \ldots \wedge x \in U_m$. Here $x_1, \ldots, x_m$ are constant or variable names, and $U_1, \ldots, U_m$ are the domains that these constants take their values from. A domain can, e.g., be the set of real numbers $\mathbb{R}$, or the natural numbers $\mathbb{N}$, or the positive natural numbers $\mathbb{N}^+$, or it can be the set of functions from real numbers to real numbers, $\mathbb{R} \to \mathbb{R}$. We then write $f \in \mathbb{R} \to \mathbb{R}$ to indicate that f is a function in this domain. We may also use the more traditional notation for this, $f : \mathbb{R} \to \mathbb{R}$.

A *question* is either an *all-question,* of the form

$$!\, x_1 \in U_1, \ldots, x_m \in U_m : P(x_1, \ldots, x_m)$$

which says that we are looking for **all values** $x_1, \ldots, x_m$ that satisfy the condition $P(x_1, \ldots, x_m)$, or a *some-question,* of the form

$$?\, x_1 \in U_1, \ldots, x_m \in U_m : P(x_1, \ldots, x_m)$$

when we are looking for **some values** $x_1, \ldots, x_m$ that satisfy the condition $P(x_1, \ldots, x_m)$.

19.3 Derivations with Theories

We consider theories as a syntactic layer on top of structured derivations. All occurrences of theories can be removed from a derivation by systematically replacing the import statement with the derivation defined in the theory, renaming constants and renumbering steps as required. Theory definitions can then all be removed, as they are not imported anywhere. This leaves us with a standard structured derivation that has the same logical meaning as the derivation with theories that we started from.

An *extended derivation* is a structured derivation with theories. Syntactically, we define an *extended derivation* as a sequence of steps, where each step is either a derivation step or a theory. We define a *theory* as a named sequence of steps, where each step is either a derivation step or an import.

extended derivation

(*derivation step* | *theory*)*

theory

§ *name*

(*derivation step* | *import*)*

■ *name*

An *import* statement imports a theory, and has the syntax

import

* *name* [*assignment*]

An *assignment* is of the form

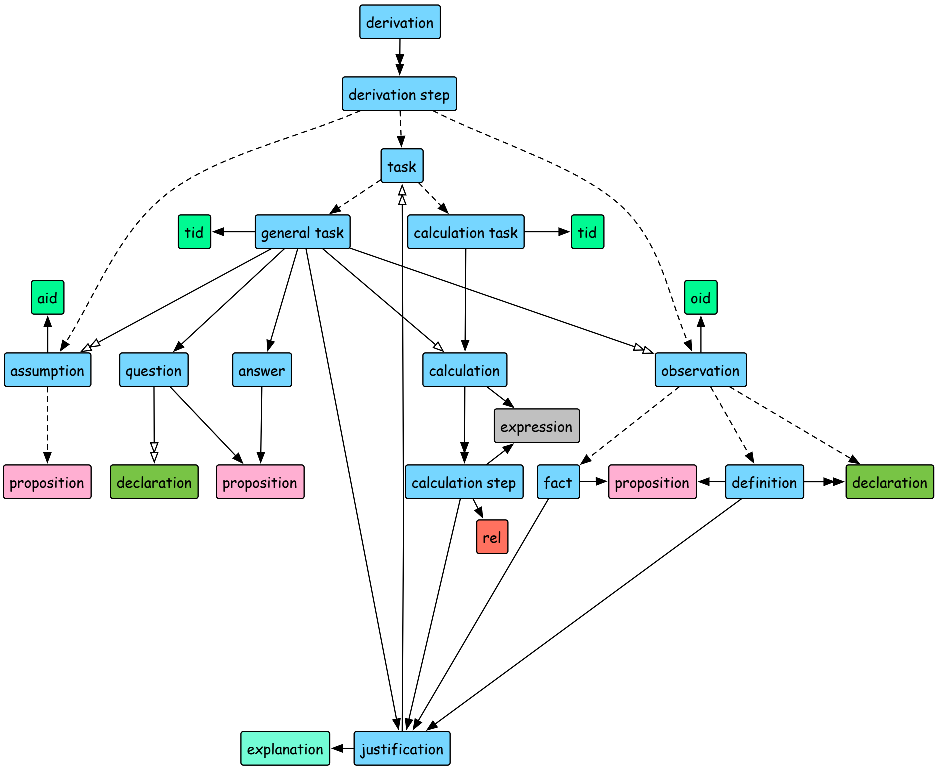
$$c_1 := t_1, \ldots, c_m := t_m$$

This assigns expression t_i to the constant c_i, for $i = 1, \ldots, m$.

The syntax guarantees that we can only have theories at the outermost derivation level, i.e., there are no nested theories. We assume that each import statement refers to a theory that has been defined before in the extended derivation. Recursive imports are thus not allowed, i.e., a theory cannot import itself, neither directly nor indirectly. This means that the substitution of theory definitions for imports will eventually end.

19.4 Abstract Syntax of Structured Derivations

The following diagram shows the different syntactic categories of structured derivations graphically, and how they are related to each other (a so called *abstract syntax*). We include here both the general and detailed syntax.



Legend:
- blue boxes aare nonterminals,
- other colors are terminals
- dashed arrows show alternatives
- solid arrows show components
- single black arrowhead shows one component is required
- single open arrowhead shows optional component
- multiple black arrowhead shows one or more components
- multiple open arrowheads show zero or more components

Chapter 20

Correctness of Structured Derivations

We will here define in more precise terms what it means for a structured derivation to be correct. Because of the recursive nature of structured derivations, the definition of correctness must also be recursive. We explained the idea of recursive definitions of functions in Section 17.2, and gave an example of how to prove properties about recursively defined functions (the relationship between Fibonacci numbers and the golden ratio, Example 86). Here we will apply these same principles to defining the correctness of a structured derivation.

The essential steps are as follows:

1. The correctness of a basic inference step is defined by the mathematical and logical framework used in the derivation.
2. The correctness of a fact in a given context is defined in terms of (i) the correctness of a basic inference step and (ii) the correctness of the possible nested tasks in its justification.
3. The correctness of a definitions or a calculations in a given context is defined in terms of correctness of facts.
4. The correctness of a structured task in a given context is defined in terms of correctness of the facts, definitions, and calculations in the task, as well as the correctness of the answer given in the task.
5. Finally, the correctness of a structured derivation is defined in terms of the correctness of the facts, definitions and tasks that make up the derivation.

We illustrate the correctness notions in Figure 20.1. The figure follows the overall structure of structured derivations: the top-level notion is that of a structured derivation. Correctness of a structured derivation is reduced to the correctness of the facts, definitions and tasks that (together with assumptions) make up the derivation. Correctness of a task is reduced to the correctness of the facts, definitions,

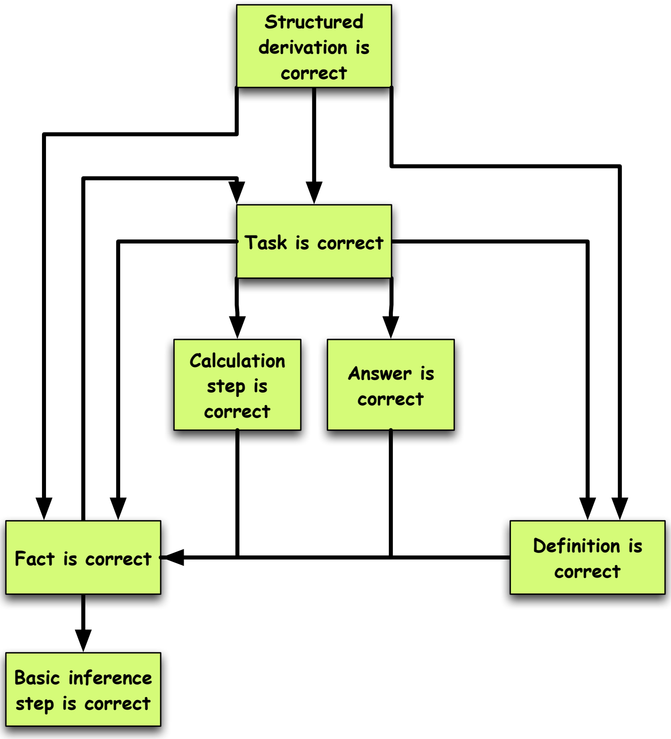

Structured derivation is correct
Task is correct
Calculation step is correct
Answer is correct
Fact is correct
Definition is correct
Basic inference step is correct

Figure 20.1: Correctness of structured derivations

and the calculation steps in the task, together with the correctness of the answer. Correctness of definitions and calculation steps are again reduced to the correctness of facts. Finally, correctness of a fact is reduced to the correctness of some basic inference step together with the correctness of the nested tasks in the justification for the fact. These reduction relations are shown as arrows in the figure.

Correctness of a specific construct in a structured derivation is thus reduced to the correctness of other constructs, which are syntactically simpler that the original construct. We are essentially defining correctness of a structured derivation recursively over the syntax of structured derivations.

20.1 Properties of Derivation Steps

Consider a task T, of the form

• Q
- A
$\vdots$
$\Vdash$ J
$\vdots$
□ R

where $A = A_1, \ldots, A_m$ are the assumptions in the task. The question Q and the answer R in the task are used to focus the attention on what we are supposed to do. Once we have found the answer to the question, the conclusion J gives the explanation for why the answer is correct. That is, proving that the answer is correct is the same as proving that some proposition is true. For a some-task, the situation is shown below: on the left, we have the some-task, and on the right the corresponding proof task. A task is considered to be correct if the corresponding proof task is correct.

• $?\, x \in U : P(x)$
- $A(x)$
$\vdots$
$\Vdash$ J
$\vdots$
□ $R(x)$

• $(\exists x \in U : A(x) \wedge R(x) \wedge P(x))$
$\vdots$
$\Vdash$ J
$\vdots$
□

If $R(x)$ is of the form $x = t$, then we can simplify the property to be proved using the the one- point rule for existential quantification to

$$t \in U \wedge A(t) \wedge P(t)$$

In other words, the answer $x = t$ is correct, if t belongs to the set of permitted values U, and t satisfies both the assumption and the query condition.

Similarly, we can rewrite an all-task as an equivalent proof task. Below, the all-task is on the left and the corresponding proof task is on the right. Again, the task is considered to be correct if the corresponding proof task is correct.

$$\begin{array}{ll} \bullet & !\,x \in U :\ P(x) \\ - & A(x) \\ \vdots & \\ \Vdash & J \\ \vdots & \\ \Box & R(x) \end{array}$$

$$\begin{array}{ll} \bullet & (\forall x \in U : A(x) \Rightarrow R(x) \equiv P(x)) \\ \vdots & \\ \Vdash & J \\ \vdots & \\ \Box & \end{array}$$

The answer $R(x)$ to an all-question is thus correct, if $R(x)$ is equivalent to $P(x)$ for any value of $x \in U$ that satisfies the assumption $A(x)$.

We will introduce the notation $Q\#R$ for the proposition that we need to prove in order to show that answer R to the question Q is correct. This will allow us to treat some- and all-questions in a uniform manner.

Definition 1. Let Q be a question, $A(x)$ the assumption, and $R(x)$ and answer of a task We then define $A\#Q\#R$ as follows:

$$\begin{array}{rcl} A\#Q\# & = & A \Rightarrow Q \\ A\#(?x \in U : P(x))\#R(x) & = & (\exists x \in U : A(x) \wedge R(x) \wedge P(x)) \\ A\#(!x \in U : P(x))\#R(x) & = & (\forall x \in U : A(x) \Rightarrow R(x) \equiv P(x)) \end{array}$$

(the first case covers proof tasks, for which there is no answer).

A structured derivation is a sequence of assumptions, facts, definitions and tasks. A structured task can in addition have calculation steps. These derivation steps are listed in a certain order, $D_1, \ldots, D_m$. Each fact, definition, calculation step, or task D_i *proves* that some specific property P'_i holds in a certain context Φ and *establishes* some property P_i that may be assumed in subsequent steps $D_{i+1}, D_{i+2}, \ldots$. In other words, the established property P_i is added to the context Φ of the subsequent steps. For facts, calculation steps and tasks, the property proved is also the property that may be assumed in the subsequent steps, but for a definition, these are different. An assumption does not need any proof, but it does establish a property.

Definition 2. We define the *property established* by a derivation step D in context Φ as follows:

(a) if D is an assumption $-A$, then D establishes the property A in context Φ

(b) if D is a fact $+J;\ P$, then D establishes the property P in context Φ,

(c) if D is a definition $+y \in U;\ J;\ P(y)$, then D establishes the property

$$y \in U \wedge P(y)$$

in context Φ,

(d) if D is a calculation step $t;\ \sim J; t'$, then D establishes the property $t \sim t'$ in context Φ, and

(e) if D is a task $\bullet Q;\ -A;\ \ldots;\ \Box R$, then D establishes the property

$$A \# Q \# R$$

in context Φ. ■

Here $A = A_1;\ \ldots;\ A_m$ and $\wedge A = A_1 \wedge \ldots \wedge A_m$.

Definition 3. For facts, calculation steps and tasks, the *property proved* by a derivation step D in a context Φ is the same as the property established in this context. For a definition $+y \in U;\ J;\ P(y)$, the property proved in context Φ is $(\exists y \in U : P(y)$.

For a definition, we need to prove that it is well-defined, i.e., that $(\exists y \in U : P(y))$. If this is the case, then we may assume that $y \in U \wedge P(y)$ is true in subsequent derivation steps. Here y must be a new variable, i.e., a variable that does not occur free in the context of the definition.

20.2 Correctness of Derivation Steps

Let us now define more precisely what it means for a derivation step to be correct. Consider first a fact $+J;\ P$ where

$$J = \{E\};\ [T_1;\ \ldots;\ T_m],$$

and

$$T_i = \bullet Q_i;\ -A_i;\ \ldots;\ \Box R_i$$

for $i = 1, \ldots, m$. We use square brackets to indicate that the tasks are indented one step to the right). Here A_i stands for the sequence of assumptions in task T_i. The fact is thus of the following form:

$$
\begin{array}{lll}
+ & \{E\} & \\
 & \bullet & Q_1 \\
 & \text{-} & A_1 \\
 & \vdots & \\
 & \square & R_1 \\
 & \vdots & \\
 & \bullet & Q_m \\
 & \text{-} & A_m \\
 & \vdots & \\
 & \square & R_m \\
\ldots & P &
\end{array}
$$

Definition 4. The *assumption* $-A$ *is correct in the context* Φ, denoted $\Phi \vdash -A$.

An assumption needs no proofs.

Definition 5. The *fact* $+J;\ P$ *is correct in the context* Φ, denoted $\Phi \vdash +J;\ P$, if

- $\Phi \vdash T_i$, for $i = 1, \ldots, m$, and
- $\dfrac{\Phi \vdash A_1 \# Q_1 \# R_1 \qquad \ldots \Phi \vdash A_m \# Q_m \# R_m}{\Phi \vdash P} \{E\}$ ■

The second condition here is a *basic inference step.* We will discuss this second condition in more detail below.

A fact is thus correct in a given context Φ, if each nested task in the justification is correct in the same context, and we can infer $\Phi \vdash P$ with E from the properties established by the tasks. .

The correctness of a justification, a calculation step and a task can be reduced to the correctness of a fact.

Definition 6. The *definition* $+y \in U;\ J;\ P(y)$ *is correct in the context* Φ, denoted $\Phi \vdash +y \in U;\ J;\ P(y)$, if

$$\Phi \vdash +J;\ (\exists y \in U : P(y))$$

■

Definition 7. The *calculation step* $t;\ \sim J;\ t'$ *is correct in the context* Φ, denoted $\Phi \vdash t;\ \sim J;\ t'$, if

$$\Phi \vdash +J;\ (t \sim t')$$

■

Definition 8. Let A denote the assumptions and $D_1, \ldots, D_r$ the sequence of facts, definitions and calculation steps in task T. Let $P_1, \ldots, P_r$ be the properties established by $D_1, \ldots, D_r$. Let Q be the question and R the answer in the task, and let J be the justification for why the answer is correct. Then T *is correct in the context* Φ, denoted $\Phi \vdash T$, if

- $\Phi, A, P_1 \ldots, P_{i-1} \vdash D_i$ for $i = 1, \ldots, r$, and
- $\Phi, P_1 \ldots, P_r \vdash +J;\ A\#Q\#R$ ■

We thus reduce the correctness of a structured task to the correctness of the facts, definitions and calculation steps in the task, together with the correctness of the answer of the task.

Correctness of tasks and facts are thus defined in terms of each other. The recursion stops at facts which are justified without nested tasks.

Let us finally defined the correctness of a general structured derivation.

Definition 9. Consider a structured derivation $D = D_1;\ \ldots;\ D_r$. Let $P_1, \ldots, P_r$ be the properties established by the derivation steps $D_1, \ldots, D_r$. The derivation D is then correct in the context Φ, denoted $\Phi \vdash D$, if

- $\Phi, P_1, \ldots, P_{i-1} \vdash D_i$, for $i = 1, \ldots, r$. ■

A structured derivation is thus correct in a given context, if each derivation step is correct in the original context extended with the properties established by the previous derivation steps.

The above definitions shows how the correctness of a structured derivation is ultimately reduced to the correctness of a collection of basic inference steps.

20.3 Basic Inference Steps

The correctness of a fact is reduced to the correctness of the nested tasks T_i and to a *basic inference step* of the form

$$\frac{\Phi \vdash A_1\#Q_1\#R_1 \qquad \ldots \Phi \vdash A_m\#Q_m\#R_m}{\Phi \vdash P}\{E\}$$

This says that $\Phi \vdash P$ can be inferred from the properties established by the nested tasks $T_1, \ldots, T_m$. Here E is the justification for this.

We assume in the derivation that all basic inference steps are mathematically correct. Then the structured derivation itself will also be correct. If even one basic inference step is incorrect, then the whole structured derivation is incorrect. Basic inference steps are the device that we use to build a structured derivation on some established collection of mathematical theories. We do not need to give new proofs for the theorems in these theories, we use them straight on as basic inference steps in our derivation.

An example of a basic inference step could be, e.g.,

$$\frac{}{\Phi \vdash 2 \cdot (x^2 + y) = 2 \cdot x^2 + 2 \cdot y}\ \{\mathit{distribution\ rule}\}$$

The context Φ would here include the definition of real numbers, the definition of addition and multiplication of reals, and the distribution rule for reals. We may assume that the distribution rule has already been proved in the theory of reals, so we need not repeat the proof here. There are no premises in this inference step.

Another example is

$$\frac{\Phi \vdash x + y \geq 0}{\Phi \vdash \sqrt{(x+y)^2} = x + y}\ \{\mathit{square\ root\ of\ square}\}$$

Here the premise is that the expression we square is non-negative.

A basic inference step can also be the application of a standard inference rule of natural deduction, like the proof strategies described in Chapter 11. We can apply these safely in structured derivations, because they are known to be correct. An example of this would be the following basic fact

$$\frac{\Phi, k \in \mathbb{N}, \mathit{even}(k) \vdash \mathit{even}(k^2 + k) \qquad \Phi, k \in \mathbb{N}, \neg \mathit{even}(k) \vdash \mathit{even}(k^2 + k)}{\Phi, k \in \mathbb{N} \vdash \mathit{even}(k^2 + k)}\ \{\mathit{case\ analysis}\}$$

We know that this inference step is correct, because it is a straightforward application of the case rule of natural deduction.

This means that if we only use mathematical theorems and inference rules that have been proved correct in some established theory, then the structured derivation itself will also be correct, with the same level of rigor. If our basic facts are based more on intuition than on exact rules, then the structured derivation is on the same level of rigor. If even one of the basic inference steps that we have used in the derivation is false, then we know nothing about the correctness of the result of the derivation.

20.4 Checklists for Structured Derivations

We will summarize the rules for checking a structured derivation in three checklists: checking the correctness of a fact (Table 20.2), checking the correctness of a task (Table 20.3), and checking the correctness of derivation (Table 20.4). These checklists just repeat the definitions that we have given earlier, but may provide a more comprehensive overview of what we have to check for each derivation step. Ultimately, we need to check that each basic inference step in the derivation is correct.

We write the structured derivation on the left in the table, and the facts that we have to check on the right hand side of the table. The context is written in blue and the property to prove in red. We assume for simplicity that facts come before the definitions in the task. The property that we need to check for each derivation step is either that a fact is correct in a certain context, i.e., $\Phi \vdash +J; P$, or that a task is correct in a given context, i.e., $\Phi \vdash T$. In the first case, we use Table 20.2, in the second case we use Table 20.3, to see how to continue. .

External context is Φ:			Check that
$+$	$\{E\}$		$\dfrac{\Phi \vdash A_1\#Q_1\#R_1 \qquad \ldots \Phi \vdash A_m\#Q_m\#R_m}{\Phi \vdash P}\{E\}$
	$\bullet$	Q_1	check that this task is correct in context Φ
	-	A_1	
	$\vdots$		
	$\square$	R_1	
	$\vdots$		
	$\bullet$	Q_m	check that this task is correct in context Φ
	-	A_m	
	$\vdots$		
	$\square$	R_m	
$\ldots$	P		

Figure 20.2: Checking correctness of a fact

External context is Φ:		Check that
•	Q	
-	A	
$+$	J_1	Φ, A
	P_1	$\vdash + J_1;\ P_1$
$\vdots$		
$+$	J_n	$\Phi, A, P_1, \ldots, P_{n-1}$
	P_n	$\vdash + J_n;\ P_n$
$+$	$y_1 \in U_1 \quad J_{n+1}$	$\Phi, A, P_1, \ldots, P_n$
	P'_1	$\vdash + J_{n+1};\ (\exists y_1 \in U_1 : P'_1)$
$\vdots$		
$+$	$y_h \in U_h \quad J_{n+h}$	$\Phi, A, P_1, \ldots, P_n, P'_1, \ldots, P'_{h-1}$
	P'_h	$\vdash + J_{n+h};\ (\exists y_h \in U_h : P'_h)$
$\Vdash$	J	$\Phi, P_1, \ldots, P_n, P'_1, \ldots, P'_h, t_0 \sim_1 t_1, \ldots, t_{k-1} \sim_k t_k$
	t_0	$\vdash + J;\ A\#Q\#R$
$\sim_1$	J_{n+h+1}	$\Phi, A, P_1, \ldots, P_n, P'_1, \ldots, P'_h$
	t_1	$\vdash + J_{n+h+1};\ t_0 \sim_1 t_1$
$\vdots$		
$\sim_k$	J_{n+h+k}	$\Phi_0, A, P_1, \ldots, P_n, P'_1, \ldots, P'_h, t_0 \sim_1 t_1, \ldots, t_{k-2} \sim_{k-1} t_{k-1}$
	t_k	$\vdash + J_{n+h+k};\ t_{k-1} \sim_k t_k$
□	R	

Figure 20.3: Checklist for task

External context is Φ		Check that
[1]	P_1	
$\vdots$		
[n]	J_n	$\Phi, P_1, \ldots, P_{n-1}$
	P_n	$\vdash + J_n;\ P_n$
$\vdots$		
[k]	$y_k \in U_k$	
	J_k	$\Phi, P_1, \ldots, P_{k-1}$
	P_k	$\vdash +J_k;\ (\exists y_k \in U_k : P_k)$
$\vdots$		
[r]	T_r	$\Phi, P_1, \ldots, P_{r-1} \vdash T_r$
$\vdots$		

We assume that we have already checked the previous steps for fact [n] , definition [k], and task [r]. Note that when [r] is a task $T = \bullet Q;\ A;\ \ldots;\ \Box R$, then P_r stands for $A \# Q \# R$. When [k] is a definition, then P_k stands for $y_k \in U_k \wedge P_k(y_k)$.

Figure 20.4: Checklist for derivation

Chapter 21

Computer Support for Structured Derivations

Structured derivations were originally designed with computer support in mind, partly inspired by the design of *programming languages* and *programming editors* (also known as *source code editors*). A programming editor is a text editor with a number of useful features that support writing program code: color syntax highlighting, automatic indentation and de-indentation, auto completion of keywords, bracket matching, on-line syntax checking with error messages, and support for using different kinds of plugins. These features help the programmer to build the program in an incremental fashion, check that the program is syntactically correct, edit the program code, debug the code directly from the editor, and test the program on different data sets. We would like to have the same facilities for creating and editing mathematical derivations: a computer based *editor for structured derivations.*

Writing a structured derivation with an editor means that we can use the computer to analyze the derivation, as it is now in machine readable form. The editor would be similar to a compiler for a programming language: it checks the syntax of the derivation (to guarantee that the derivation is meaningful), and then translates the derivation into a form that the computer can understand. We would then use the computer to check whether the derivation is mathematically and logically correct. Alternatively, the computer can help us create a correct derivation in the first place. We will consider both these approaches below, and show how they can be used in checking the correctness of a structured derivation.

The description of editors and checkers for structured derivations below is based on ongoing work on building an integrated learning environment for mathematics based on structured derivations. You can find out more about the work done in our projects by visiting our web sites (www.emathstudio.fi, www.fourferries.fi).

21.1 An Editor for Structured Derivations

Dijkstra's and his colleagues original work on calculational proofs were mostly presented as handwritten or printed proofs. Structured derivations add a number of new features to calculational proofs, in particular assumptions, observations and nested tasks, so there is more syntax to remember, and more special symbols used in derivations (“•”, “⊩”, “□”,”§" , and "■"). There is a certain learning threshold for students to start using structured derivations, to remember and understand the meaning of these different symbols and to organize the derivation so that it is syntactically correct. Computer support for writing structured derivations would therefore be useful, both for creating the derivation an for checking its syntax.

We have built a special editor for structured derivations, the *SD editor*, in our research project. The SD editor knows the syntax of structured derivations, and constrains the user to only create syntactically correct structured derivations. It is thus a *syntax driven editor,* the derivation is constructed with special editor instructions for the different kinds of derivation steps. The editor also keeps track of the indentation level of any derivation step, which seems to be the main issue when writing structured derivations by hand. The syntax driven part is, however, restricted to the overall structure of the derivation. The specific fields in structured derivations, like assumptions, facts, calculation terms, and justifications, are written with a text editor that understands and displays standard mathematics notation for mathematical formulas (a LaTeX based *wysiwyg math text editor*).

The advantages of the SD editor are the same as for any text editor. The user can add new derivation steps as needed, delete a specific derivation step, and edit a derivation step (e.g., modifying the proposition, expressions or justification in a step). The user can also insert new facts and definitions at appropriate places in the derivation, as well as create nested derivations in justifications. The derivation can be constructed in the order that is most natural for the problem at hand, which often is not from the beginning to the end. The editor thus supports the incremental approach to constructing a derivation described in Section 8.4.

Another advantage of the SD editor is that one can show and hide nested derivations at will. The user can inspect the derivation at different levels of detail, by selectively showing some nested derivations and hiding others, as well as work on different levels of detail in different parts of the derivation. Hiding most nested derivations shows the overall structure of the proof, while showing a nested derivation allows working on a more detailed level. This means that it is possible to manage very large derivations in the SD editor without loosing the overall view of what we are doing.

An important advantage of the SD editor is that derivations can be analyzed by computers. For instance, the editor can tell the user beforehand what type of information is expected in a specific place (a mathematical expression, a logical proposition, a justification, a declaration, and so on), and it can check that the entered text is of the correct syntactic form. As an example, assume that we have started a derivation as follows:

• $2x + 3 = y - x \wedge 5y + 4 = x$

We then ask the syntax directed editor to generate a calculation step. The editor completes the derivation as follows (new text written in color):

• $2x + 3 = y - x \wedge 5y + 4 = x$

rel {justification}

$2x + 3 = y - x \wedge 5y + 4 = x$

The editor shows where the relation symbol (in red) and the justification field (in green) are written, and creates a copy of the first expression as the next expression (the blue color indicates that this is a copy). The editor copies the original expression to the next line, so that the new expression can be quickly created by editing the old expression.

The user then fills in the missing details, and edits the new expression. The result could, e.g., be:

• $2x + 3 = y - x \wedge 5y + 4 = x$

$\equiv$ {substitute the value of x given in the second equation for the value of x in the first equation}

$2 \cdot (5y + 4) + 3 = y - (5y + 4) \wedge 5y + 4 = x$

The construction of the derivation proceeds in this way step by step. The user can freely add, delete and edit assumptions, facts, definitions, nested derivations and tasks in the derivation, as long as the syntax of structured derivations is respected.

An important feature of the editor is that is uses standard mathematical notation. The editor uses LaTeX as the underlying representation of a mathematical formula, but displays the formula using typeset LaTeX syntax. The formula can also be edited in the standard display notation. This means that the user never needs to see the underlying LaTeX representation. The computer has access to the LaTeX notation and can process it mechanically in order to analyze it in different ways.

21.2 Checking the Correctness of Structured Derivation

All mathematical proofs, derivation, calculation, and constructions can contain errors. An error in a proof makes the theorem useless, so it is important to guard against errors. There are different ways to do this. Going over the proof once more, but checking each step very carefully, is a good way of detecting errors, and should always be the first step. Explaining the proof to somebody else and discussing each step carefully is also a very efficient, as is asking somebody else (a fellow student, a teacher, etc) to read through it.

Computers have long been used to prove mathematical theorems, and to check that a given proof is correct. There are basically two different approaches here: *automatic* and *interactive theorem proving*. Automatic theorem proving is what the name suggest: we feed a theorem to a computer, and ask the computer to generate a mathematical proof for this theorem. The computer will either come up with a proof, or then give up, after a certain time limit has been passed. Automatic theorem proving is essentially a branch of artificial intelligence, where the computer searches for a proof in a very large search tree. If the computer finds a proof, then this proof is correct by construction.

Interactive theorem proving takes another approach. Here the proof is constructed by a user with the help of a computer. The user starts with the overall goal: to prove the given theorem. The next step is to choose which inference rule to apply to this goal. The computer applies this inference rule, and creates zero or more new subgoals to be proved. Thus, interactive theorem proving is essentially computer supported backward proofs of theorems. The theorem is proved when all subgoals generated during the proof process have been proved. The system guarantees that if a proof is found, then it is correct. Interactive theorem provers are nowadays quite powerful. The user can formulate and apply *proof strategies* when generating the next proof obligations. A proof strategy is essentially a search program that tells the computer how to look for a good way of either directly proving a goal, or reduce the goal to other goals that are simpler to prove.

Automatic theorem proving is obviously much to be preferred, whenever possible. However, many mathematical problems are too difficult for automatic theorem provers. Automatic theorem provers excel in certain areas of mathematics, like real numbers and propositional logic, whereas they can be quite weak in other areas, like proofs with quantifiers. Interactive theorem provers are much more powerful, allowing us to prove very hard and demanding theorems. However, they requires hard work and very specialized skills. In practice, interactive theorem proving is usually combined with automatic theorem proving: the interactive theorem prover is used to split up the proof into a collection of lemmas, which we then try to prove automatically. Those lemmas that are not proved automatically are then manipulated with the interactive theorem prover, applying inference rules to create new lemmas, and so on. This process continues until the whole theorem is proved, or the user gives up, or the user finds out that the theorem is in fact not true.

The problem with these computer based tools is that they require that theorems are expressed in a formal logical system (usually predicate calculus, set theory, or higher order logic). In interactive theorem proving, the user also has to create the proof using only formal inference rules (usually some version of the natural deduction rules that we described above). Both the theorem and its proof are thus formulated in a completely rigorous logical system. Such a formulation is very far from the language used by ordinary mathematicians, and the proofs are very difficult to follow and in most cases also not very intuitive.

Structured derivations is an attempt to bridge the gap between standard, informal mathematical proofs and completely formal computer checkable proofs. The notation used in structured derivations is quite close to the standard way of writing mathematical proofs, except that the overall structure of the proof is shown explicitly (rather than being implicit in the narrative of the proof). This makes it much easier for the computer to understand how the proof is structured, and how to split up the proof into smaller, more manageable proof steps. The computer can then try to prove these smaller proof steps automatically. If the automatic proof fails, then the user can use interactive theorem proving to chop up the original proof step into smaller proof steps, and try to prove these automatically.

There are a number of other difficulties that still need to be overcome when proving correctness of a structured derivation. First, standard mathematical notation is quite far from the traditional, ascii based notation used by automatic and interactive theorem provers. The computer needs to translate the standard mathematical notation to the language that the computer based theorem provers understand. This can sometimes be quite challenging, because standard mathematical notation has a lot of ambiguities that needs to be resolved before translation.

The second problem is that both automatic and interactive theorem provers need to know the area of mathematics in which they are expected to carry out the proof. This means that we need to have *mechanized theories,* which formalize a specific mathematical area in a way that is understood by the theorem prover. Computer based theorem proving has mainly been developed for proving that computer software and hardware is correct. This means that there are good mechanized theories available for fields of mathematics needed in software and hardware development, like propositional and predicate calculus, natural numbers, real numbers, and some calculus. For other areas, there are few or no mathematical theories available (vectors, probability theory among others).

21.3 Automatic Theorem Provers

Chapter 20 showed how to reduce the correctness of a structured derivation to a collection of theorems for the correctness of the individual proof steps in the derivation. If all proof steps are correct, then the structured derivation is also correct. This is a consequence of the soundness of structured derivation that we prove in the appendix. Soundness means that if we have constructed a proof for a mathematical theorem as a structured derivation, where all basic steps are correct, and we only have used natural deduction inference rules, then this theorem is necessarily true. We show this by reducing an arbitrary structured derivation to an equivalent proof in natural deduction. As natural deduction is known to be sound, it follows that the structured derivation method is also sound.

An automatic proof checker for structured derivations works more or less as shown in the next derivation. The checker adds marks for each derivation step, which shows the proof status of that step.

$\bullet$ $2x + 3 = y - x \wedge 5y + 4 = x$

$\equiv$ {substitute the value of x given in the second equation for the value of x in the first equation} ok

$2 \cdot (5y + 4) + 3 = y - (5y + 4) \wedge 5y + 4 = x$

$\equiv$ {simplify the first equation} ok

$10y + 11 = -4y - 4 \wedge 5y + 4 = x$

$\equiv$ {rearrange terms in the first equation} not proved

$14y = 7 \wedge 5y + 4 = x$

$\equiv$ {solve the first equation} ok

$$y = \frac{7}{14} \wedge 5y + 4 = x$$

$\equiv$ {insert the value of y into the second equation} ok

$$y = \frac{7}{14} \wedge 5 \cdot (\frac{7}{14}) + 4 = x$$

$\equiv$ {solve the second equation} ok

$$y = \frac{7}{14} \wedge x = \frac{35}{14} + 4$$

$\equiv$ {simplify the equations} ok

$$y = \frac{1}{2} \wedge x = 6\frac{1}{2}$$

□

The checker has found a proof for all steps except the third derivation step. The step is in fact wrong, so it is not even possible to prove it. However, the checker cannot say that the step is incorrect (for this it would need to generate a counter example), it can only say that it was not able to prove the step in the allocated amount of time. The message that this step could not be proved is usually sufficient information to detect the error in the derivation. Once the error has been corrected, the checker is able to prove each step correct.

There is a number of reasons why the checker may not be able to prove a basic step in the derivation:

1. the step may in fact be incorrect, so there is no proof of it,
2. the step is correct, but the checker is not able to find a proof in the given amount of time,
3. the step is correct, but the checker does not have sufficient information to be able to prove it (missing mechanized theories), or
4. the step may be correct, but the checker does not understand it, and hence is not able to translate it into a formal mathematical theorem.

All of these reasons are possible and occur rather frequently. There are different ways of handling the unproved message:

- We can check whether the step is in fact wrong. There can be a typing error, or a rule that has been applied in the wrong way. Often, the error may not be in the step itself, but some assumptions may be wrong or are missing in the derivation. Adding/correcting these assumptions makes the step correct.

- We may decide that a manual proof or careful justification is sufficient here. We think that the step is in fact correct, based on earlier experience and understanding of the underlying mathematical theory, and accept the step without proof.
- We may think that the step is correct, but want the theorem prover to accept it. We can try to reorganize the derivation, so that it becomes easier to check automatically, e. g., by adding a nested derivation to prove the offending step with smaller proof steps, or by inserting extra derivation steps in a calculation.
- We could suspect that the checker does not know some mathematical fact that is needed to prove the step. We can add this mathematical fact to the context, and try again. Of course, we need to be careful not to add incorrect facts to the context.
- If the checker does not understand the step (a syntax error), then we should correct the syntax and check again..

The automatic theorem prover does not check whether the justification given for a derivation step is correct, it only checks whether the step itself is correct, as a mathematical theorem. Consider for instance the fifth step in the derivation above:

$$y = \frac{7}{14} \wedge 5 \cdot (\frac{7}{14}) + 4 = x$$

$\equiv$ {solve the second equation} ok

$$y = \frac{7}{14} \wedge x = \frac{35}{14} + 4$$

The theorem prover has to prove the theorem

$$\vdash (y = \frac{7}{14} \wedge 5 \cdot (\frac{7}{14}) + 4 = x) \equiv (y = \frac{7}{14} \wedge x = \frac{35}{14} + 4)$$

When looking for a proof, it does not look at all at the justification for the the derivation step. There could be any justification inside the curly brackets, or the justification may be missing altogether, the theorem prover does not care.

This is of course a shortcoming as far as teaching mathematics goes. One would like to have the checker to also say whether the justification for a derivation step is correct. One could envision an automatic checker for structured derivations that also checks the justifications. However, this would require that we also formalize the justification language, so that the checker can understand the justifications. This means that the derivation as a whole is quite formal, essentially amounting to reasoning in a formal logical system. We have avoided going down this path, as it could easily stifle the mathematical intuition and would make derivations more difficult to create (on the same level as using interactive theorem provers). This level is too difficult for ordinary school mathematics, and is also not what we want to teach in schools. We want to encourage the use of mathematical intuition in creating derivations, and may accept that some of the justifications are more or less loose, as long as they convey the correct intuition about why the proof step is correct.

Checking the correctness of justifications is thus left to the teacher. The justifications provide an excellent window into seeing how the student has understood the mathematics behind the derivation. If the derivation is checked with a computer, then the teacher is left with checking those steps that were not proved, as well as the justification for the derivation steps. Our experience from using structured derivations in class indicates that checking a structured derivation is much faster than checking a standard mathematical argument, and that it is much easier to detect errors in structured derivations. Using automatic checking takes away also the tedium of checking that the computation steps are in fact correct, thus reducing the teachers work load even further.

21.4 Interactive Theorem Provers

Besides using an automatic proof checker for checking that a structured derivation is correct, we can also use an interactive theorem prover. Such a theorem prover should be integrated into the editor for structured derivations. The editor would provide a command for applying a specific inference rule that generates the next step in the derivation.

Consider the derivation above, as an example. Our task is to solve the equation pair $2x+3 = y-x \wedge 5y+4 = x$. We start by writing down the start of the calculation:

$$\begin{array}{ll} \bullet & 2x+3 = y-x \wedge 5y+4 = x \end{array}$$

We then ask the editor to substitute the value of x in the second equation for the value of x in the first equation. We do this by filling in the justification for the next step (shown in green). The editor then generates the following step (in red):

$$\begin{array}{ll} \bullet & 2x+3 = y-x \wedge 5y+4 = x \\ \equiv & \{x := 5y+4 \text{ in first equation}\} \\ & 2\cdot(5y+4)+3 = y-(5y+4) \wedge 5y+4 = x \end{array}$$

This means that the editor knows that the relation between the terms is equivalence, and calculates the new expression (it also needs to check that the assignment actually results in an equivalence between the two logical propositions).

The next step would be to ask the editor to simplify the expressions in the first equation:

$$\begin{array}{ll} \bullet & 2x+3 = y-x \wedge 5y+4 = x \\ \equiv & \{x := 5y+4 \text{ in first equation}\} \\ & 2\cdot(5y+4)+3 = y-(5y+4) \wedge 5y+4 = x \end{array}$$

$\equiv$ {simplify first equation}

$10y + 11 = -4y - 4 \wedge 5y + 4 = x$

In this way, we could construct the derivation step by step, using commands that the editor (using a background interactive theorem prover) understands. The interactive theorem prover guarantees that each constructed step is correct.

As another example, consider the following proof of the theorem that $k^2 + k$ is even for all natural numbers k. We start by formulating the problem as a task:

• Show that $k^2 + k$ is even, when

- k is a natural number

We can then ask the editor to apply the case analysis rule to this step. We need to tell the editor that the case we consider is whether k is even or not. The user request is shown as a justification (in green). The editor would then fill in the following step (in red):

• Show that $k^2 + k$ is even, when

- k is a natural number

$\Vdash$ {Case analysis: k is even}

 • Show that $k^2 + k$ is even, when

 - k is even

 • Show that $k^2 + k$ is even, when

 - $\neg$(k is even)

□

We can then continue by constructing the proofs for the two nested tasks. Once we are satisfied, we can use the automatic checker to check that all proof steps are correct. The automatic checker does not need to check that the steps generated by the interactive theorem prover are correct, they are correct by construction. The editor for structured derivations thus works as a calculator that allows us to construct the derivation step by step, all the time being assured that all steps created thus far are correct.

We see that there is a use for both automatic theorem proving and interactive theorem proving for structured derivations. These theorem provers should be integrated into the editor for structured derivations. The user can then choose between checking a given derivation step with the automatic theorem prover, or using the interactive theorem prover to generate the next proof step. This kind of integration of automatic and interactive theorem prover is already done in most industrial scale interactive theorem provers, like PVS [28] (uses the automatic theorem prover GRIND and Z3), Isabelle [29] (uses the automatic theorem prover Sledgehammer), and Coq.

Chapter 22

Background on Structured Derivations

Edsger W. Dijkstra, one of the great pioneers of computer science, worked together with his colleagues Wim Feijen and Nettie van Gasteren on methods for making proofs about program correctness as easy and intuitive as possible. They developed a notation that is known as *calculational proofs* [16, 36, 17]. They wanted to carry out mathematical proofs and derivations in the same way as in traditional calculations, like when solving equations, simplifying expressions or calculating values of functions. They used logical rules to calculate the truth of mathematical statements in the same way as we use algebraic rules to simplify expressions. The calculational style introduced the idea of explicit justifications on separate lines. This proof style has been adopted quite widely in articles and text books on programming methods, in particular in the context of formal (or logical, mathematical) methods for constructing correct programs. The approach is used in, e.g., the university level textbook on discrete mathematics by David Gries and Fred Schneider [22]. Jan van de Snepscheuts [35] and Ann Kaldewejs [24] both wrote textbooks on programming methods based on the calculational proof style. Gries and Schneider have also proposed using calculational proofs in high school teaching [**?**, 21, 23] and have argued for the advantages of this method in practical mathematics education.

Dijkstra's calculational proof style corresponds to what we in this book call *structured calculations.* Dijkstra's and his colleagues' work has been the main inspiration and the starting point for our own work. Joakim von Wright and I developed *structured derivations* as an extension of Dijkstra's calculational style. We originally presented the method in our book on *refinement calculus* [12], as well as in a conference paper, a technical report and a journal article [14, 13, **?**]. We used structured derivations throughout the book to prove a large number of theorems and lemmas of varying complexity, mainly in lattice theory and programming logic. While Dijkstra's original calculational proofs were based on a version of first-order predicate calculus and a Hilbert-style proof system, we have adopted Gentzen's natural deduction and Church's higher order logic as the foundations for structured derivations. This has allowed us to add nested derivations with a simple logical interpretation. Higher order logic was invented by Alonzo Church [15] in the 1940s. We based our approach on a variant of this logic described by Michael Gordon and Tom Melham [20] that was developed for the interactive theorem prover HOL.

There are not that many alternative approaches to building more precise but still human readable mathematical proofs. Leslie Lamport proposed a Gentzen-like proof style where indentation was used as a structuring device [25]. The Hilbert-like proof style for geometry has been tried in schools using a two-column proof format (see e.g., en.wikibooks.org/wiki/Geometry/Chapter_2). Interactive theorem provers like Isabelle [29] (e.g., the Isar front end [37]), Mizar [33, 34] and PVS [28] have also been equipped with more user friendly front ends for reading and writing proofs. However, these front ends usually target advanced users, and are not suitable as such for teaching mathematical proofs at the secondary education levels. Structured derivations is our attempt at creating a language for proofs that are somewhere between fully formal mechanized proofs and the intuitive proofs that are now used in education, a language that both humans and computers can understand.

The experiences that we had of using structured derivations in our book were very positive, we felt that they made the proofs easy to understand and also made it easier to construct these proofs in the first place. This prompted us to look at whether structured derivations also could be used in ordinary mathematics teaching [13]. Starting in the year 2000, we have conducted a large number of pilot studies on the use of structured derivations in class teaching, focusing on high school mathematics and introductory mathematics courses in universities [8, 10, 27, 11]. The results have been very encouraging. The students see the method as different but not particularly difficult. They say that the teacher's proofs and derivations are easier to understand when they are presented as structured derivations. They also gain a better understanding of their own proofs when written in this way, and find it is easier to detect errors in the proofs. The teachers appreciate the method because it makes it is easier to check students' solutions, to see where they made mistakes and how they had misunderstood things. The biggest drawback of structured derivations mentioned by students is that the derivations tend to become longer. This is because each step has to be explicitly justified. We actually see this is as an important advantage, since it means that the students are carefully thinking about and justifying each step in their solution. The teachers feel that requiring explicit justifications leads to a deeper understanding of mathematics and to a better competence in applying mathematics to practical problem solving.

We continued to develop structured derivations, based on feedback from these pilot studies. A later version of structured derivations [4] added observations as new features for derivations, and showed how structured derivations now could be seen as a unification of the three main proof paradigms in use today: forward derivation, backward derivation and calculation. The structured derivations method presented in this book is a further extension of the method presented in [4], adding definitions and a more precise treatment of questions and answers in tasks, as well as a more general notion of structured derivations to support more advanced mathematical modeling.

The structured derivation method has been developed in a sequence of research projects at the *Learning and Reasoning Laboratory* of *TUCS* (Abo Akademi University and University of Turku) in 2000 - 2016. The laboratory was jointly chaired by

Tapio Salakoski and me. The research has been funded by the Academy of Finland, the Technology Development Center of Finland (TEKES), the Technology Industry in Finland, the European Union and the Swedish Cultural Foundation in Finland. The method has been tested on a larger scale in an EU-project, the E-math project in 2011-13 (see http://emath.eu/ for more information about this project). This project piloted structured derivations in 15 high schools in Finland, Sweden and Estonia. Approximately 1 000 students participated in these pilots, which covered first year high school courses in mathematics. We have recently created a comprehensive series of interactive text books in e-book format for high school mathematics based on structured derivations. The work on this textbook series was funded by EU, the Swedish Cultural Foundation of Finland, and the Technology Industry Foundation of Finland. This series covers the full advanced mathematics curriculum for Finnish high schools and is available in Swedish, Finnish and English.

Appendix **A**

Soundness and Completeness of Structured Derivations

We have above showed how to prove that a structured derivation is correct. But the question still remains whether the structured derivation method itself is correct. This is known as the *soundness problem*: does the fact that we have proved a mathematical theorem with structured derivations guarantee that this theorem is true. And we have the converse question, the *completeness problem:* if a mathematical theorem is true, does this mean that there is a structured derivation proof of the theorem.

These questions have been answered for natural deduction. Natural deduction for first order logic has been proved to be both sound and complete. Soundness means here that if we have proved some logical theorem with natural deduction, then this theorem will be true in any possible interpretation (i.e., model) of the constant, function and predicate symbols that occur in the theorem. Completeness means that if some logical theorem is true in every possible interpretation of the constant, function and predicate symbols occurring in the theorem, then there is a natural deduction proof of this theorem. Gödel's incompleteness theorem shows that completeness only holds for the case that the theorem is true in every interpretation. A theorem that is true in only some interpretations, for instance in the standard model for arithmetic, need not have a natural deduction proof.

What is the relationship between structured derivations and natural deduction. Structured derivations have a richer syntax that allows combining calculations, forward derivations and backward derivations in a single format for mathematical arguments. It is aimed at general problem solving, with questions and answers. Natural deduction has only backward derivations, and only supports proofs of mathematical theorems. However, we can use the soundness and completeness of natural deduction to prove that the structured derivation method is sound and complete. We do this by showing that any result that we prove using structured derivations also has a natural deduction proof. Soundness of natural deduction then implies that any theorem that we prove using structured derivation is true in every interpretation. In particular, this means that any theorem we prove with structured derivations about some specific mathematical domain, like arithmetic or algebra, is also true in this domain. For completeness, we have an analogous argument. We show that any natural deduction proof can be expressed as a structured derivation proof. This then means that if a theorem is true in all models, then there is a structured derivation proof of it.

A.1 Soundness of Derivation Steps

Any task can be transformed into an equivalent proof task, as we have shown earlier. Hence, without loss of generality, we will in the sequel restrict ourselves to structured derivations where all tasks are proof tasks. A proof task in context Φ, with theorem Q and assumptions A, establishes the property $\Phi \vdash A \Rightarrow Q$. This is equivalent to $\Phi, A \vdash Q$. In the sequel, we will mostly be using the latter characterization.

We will first show the soundness of structured derivation steps, before looking at the soundness of a structured derivation itself. We prove the following theorem.

Theorem 1. *Let D be a fact, definition, calculation step or a (proof) task. Assume that D is correct in context Φ, i.e., $\Phi \vdash D$, and that there is a natural deduction proof for each basic inference step of D. Then there is a natural deduction proof of $\Phi \vdash P'$, where P' is the property proved by D in context Φ.*

Proof. We will prove this theorem by induction over the definition of correctness of structured derivations. Assume that the correctness of D in context Φ is reduced to the correctness of derivation steps $D_1, \ldots, D_m$, in respective contexts $\Phi_1, \ldots, \Phi_m$, $m \geq 0$. The induction hypothesis is that there is a natural deduction proof H_i of $\Phi_i \vdash P'_i$, for each $i = 1, \ldots, m$, where P'_i is the property proved by D_i. Our task is then to prove that there is a natural deduction proof of $\Phi \vdash P'$.

We prove this by cases.

a) Assume first that D is a fact, $D = +J; P$, and that $J = \{E\}; [T_1; \ldots; T_m]$. Assume that the task T_i has assumptions A_i and proves Q_i, for $i = 1, \ldots, m$. The property proved by D in context Φ is P. By definition, $\Phi \vdash D$ holds when

- $\Phi \vdash T_i$, for $i = 1, \ldots, m$, and
- $\dfrac{\Phi, A_1 \vdash Q_1 \quad \ldots \quad \Phi, A_m \vdash Q_m}{\Phi \vdash P}\{E\}$

By the induction hypothesis, there is a natural deduction proof H_i that proves $\Phi, A_i \vdash Q_i$, for $i = 1, \ldots, m$. In addition, there is a natural deduction proof which shows that the conclusion $\Phi \vdash P$ follows from these premises, because $\{E\}$ is a natural deduction inference rule. Combining these two, we see that there is a natural deduction proof for $\Phi \vdash P$.

b) Assume that D is a calculation step $t; \sim J; t'$. The property proved by D is $t \sim t'$. By definition, $\Phi \vdash D$ holds when

$$\Phi \vdash +J;\ (t \sim t')$$

i.e., we reduce the correctness of a calculation step to the correctness of a fact. By the induction hypothesis, there is then a natural deduction proof of $\Phi \vdash t \sim t'$.

c) Assume that D is a definition $+y \in U;\ J;\ P(y)$. The property proved by D in context Φ is $(\exists y \in U : P(y))$. By definition, $\Phi \vdash D$ holds when

$$\Phi \vdash +J;\ (\exists y \in U : P(y))$$

By the induction hypothesis, there is then a natural deduction proof of $\Phi \vdash (\exists y \in U : P(y))$.

d) Assume that D is a proof task $T = \bullet Q;\ -A;\ \ldots; \Vdash J_0;\ \ldots; \Box$, with r derivation steps $D_1, \ldots, D_r$. Each derivation step is either a fact, a definition, or a calculation step. Because T is correct in context Φ, we know that

- $\Phi, A, P_1, \ldots, P_{i-1} \vdash +J_i;\ P_i'$, for $i = 1, \ldots, r$, and
- $\Phi, A, P_1, \ldots, P_r \vdash +J_0;\ Q$

Here P_i is the property established by D_i and P_i' is the property proved by D_i. By the induction hypothesis, we have a natural deduction proof H_i of

$$\Phi, A, P_1, \ldots, P_{i-1} \vdash P_i'$$

for $i = 1, \ldots, m$, and we have a natural deduction proof H_0 of

$$\Phi, A, P_1, \ldots, P_r \vdash Q$$

Our task is now to construct a natural deduction proof for $\Phi, A \vdash Q$.

Consider the last step D_r. We know that there is a natural deduction proof H_r for $\Phi, A, P_1 \ldots, P_{r-1} \vdash P_r'$. We now consider three possible cases, depending on whether the step D_r is a fact, a calculation step or a definition.

d1) Assume that D_r is the fact $+J_r;\ P_r$. The property established by the derivation step and the property proved for the derivation step are the same, i.e., $P_r' = P_r$. Then we can construct the following natural deduction proof:

$$\dfrac{\dfrac{H_r}{\Phi, A, P_1, \ldots, P_{r-1} \vdash\ P_r} \quad \dfrac{H_0}{\Phi, A, P_1, \ldots, P_r \vdash Q}}{\Phi, A, P_1, \ldots, P_{r-1} \vdash Q}\ \{lemma\,rule\}$$

In other words, we have a natural deduction proof for $\Phi, A, P_1 \ldots, P_{r-1} \vdash Q$.

d2) D_r is a calculation step t_{r-1}; $\sim_r$ J_r; t_r. In that case, the property assumed and the property proved for D_r are again the same, i.e., $P'_r = P_r = (t_{r-1} \sim_r t_r)$. We construct a natural deduction proof for $\Phi, A, P_1 \ldots, P_{r-1} \vdash Q$, in the same way as in case (d1).

d3) D_r is a definition, of the form $+y_r \in U_r$; J_r; $P_r(y_r)$. Then $P'_r = (\exists y_r \in U_r : P_r(y_r))$ and $P_r = y_r \in U_r \wedge P_r(y_r)$. We observe that P' can be written as $(\exists y_r : y_r \in U_r \wedge P_r(y_r))$. We then use the natural deduction proof for existential assumptions:

$$\frac{\dfrac{H_r}{\Phi, A, P_1, \ldots, P_{r-1} \vdash (\exists y_r : y_r \in U_r \wedge P_r(y_r))} \quad \dfrac{H_0}{\Phi, A, P_1, \ldots, P_{r-1}, (y_r \in U_r \wedge P_r(y_r)) \vdash Q}}{\Phi, A, P_1, \ldots, P_{r-1} \vdash Q}$$

In all three cases, we have proved that there exists a natural deduction proof for

$$\Phi, A, P_1, \ldots, P_{r-1} \vdash Q$$

We can now proceed in the same way, and show that there exists a natural deduction proof of

$$\Phi, A, P_1, \ldots, P_{r-2} \vdash Q$$

and so on. Eventually, this will show that there is a natural deduction proof H of

$$\Phi, A \vdash Q$$

We have now shown for all possible cases that $\Phi \vdash P'$ ■

We have thus proved our theorem, i.e. that each property proved by a correct structured derivation in a given context can also be proved using natural deduction in the same context.

A.2 Soundness of Structured Derivations

We now turn to the soundness of structured derivations. We prove the following theorem.

Theorem 2. *Let $D = D_1; \ldots; D_m$ be a structured derivation. Let P_i be the property established by a fact D_i, and P'_i the property proved by D_i. Let $A^{(i)}$ be the sequence of assumptions in $D_1, \ldots, D_i$. Assume that D is correct in the context Φ, i.e, $\Phi \vdash D$, and that there is a natural deduction proof for each basic inference step in D. Then*

$$\Phi, A^{(i-1)} \vdash P'_i$$

for each step D_i in the derivation D.

Proof. Consider a definition $D_r = +y_r \in U_r$; J_r; $P_r(y_r)$ in the derivation D. We can replace this definition, shown on the left below, with a fact followed by an assumption, as shown on the right below:

```
⋮
+   $+y_r \in U_r$
    $J_r$
    $P_r(y_r)$
⋮
```

```
⋮
+   $J_r$
    $(\exists y_r \in U_r : P_r(y_r))$
-   $y_r \in U_r \wedge P_r(y_r)$
⋮
```

The property to prove is the same in both derivations, and so is also the property that may be assumed in subsequent derivation steps. Hence, if we have proved that the left derivation is correct in a certain context, then we also have a proof that the right derivation is correct, in the same context, and vice versa.

Similarly, consider a task in the derivation, shown on the left, and the fact shown on the right below:

```
⋮
•   $Q_r$
-   $A_r$
⋮
□
⋮
```

```
⋮
+   {prove implication}
    •   $Q_r$
    -   $A_r$
    ⋮
    □
... $A_r \Rightarrow Q_r$
⋮
```

In both cases, we have to prove that the task is correct in its context. On the right hand side, we have an additional inference step (prove implication), but this is a standard inference rule of natural deduction that does not need to be proved.

This means that we can replace each definition and task in a structured derivation with the corresponding facts and assumptions, as shown above. Therefore, without loss of generality, we can prove the original theorem by showing that it holds for any structured derivation that only consists of assumptions and facts.

Let us therefore assume that D is a structured derivation that only contains assumptions and facts, and that $\Phi \vdash D$. By the definition, this means that

- $\Phi, P_1, \ldots, P_{i-1} \vdash D_i$, for $i = 1, \ldots, m$

By the previous theorem, we know that there is a natural deduction proof of

$$\Phi, P_1, \ldots, P_{i-1} \vdash P_i$$

for every fact $D_i = +J;\ P_i$. Let $P^{(i-1)}$ be the sequence of properties proved in the derivation D by facts prior to D_i. The assumptions in a sequent denote a set, so they can be reordered freely. This means that the previous sequent can also be written as

$$\Phi, A^{(i-1)}, P^{(i-1)} \vdash P_i$$

Here all assumptions before D_i are listed before the properties established by facts before D_i.

Consider now a specific fact D_r. We know that for each fact D_i, $i < r$, there is a natural deduction proof of the sequent above. We can now use the widening rule of natural deduction:

$$\frac{\Phi, A^{(i-1)}, P^{(i-1)} \vdash P_i}{\Phi, A^{(r-1)}, P^{(i-1)} \vdash P_i}\ \{\mathit{widening}\}$$

The assumptions in the conclusion now contain $A^{(r-1)}$ rather than $A^{(i-1)}$, i.e., we have added all assumptions before D_r for the conclusion. We may now conclude that we have a natural deduction proof H_i for each sequent

$$\Phi, A^{(r-1)}, P^{(i-1)} \vdash P_i$$

when $i \leq r$.

We can now use the lemma rule:

$$\frac{\Phi, A^{(r-1)} \vdash P_{j_1}, \ldots \quad , \Phi, A^{(r-1)}, P_{j_1}, \ldots, P_{j_{k-1}} \vdash P_{j_k} \quad \Phi, A^{(r-1)}, P_{j_1}, \ldots, P_{j_k} \vdash P_r}{\Phi, A^{(r-1)} \vdash P_r}$$

Here $D_{j_1}, \ldots, D_{j_k}$ are the facts in derivation D that precede derivation step D_r. This shows that we have a natural deduction proof of $\Phi, A^{(r-1)} \vdash P_r$ for each fact D_r in the derivation D. This proves the required theorem. ■

A.3 Completeness of Structured Derivations

The completeness of structured derivations is easier to prove. We have the following theorem.

Theorem 3. *Assume that there is a natural deduction proof of* $\Phi \vdash P$. *Then there is also a structured task that proves this fact.*

Proof. Assume that there is a natural deduction proof of $\Phi \vdash P$. As we explained in Section 11.1, the weakening rule is adequate, so there is then also a natural deduction proof H of $\Phi \vdash P$ that does not use the weakening rule. We prove that there is a structured derivation of $\Phi \vdash P$, by complete induction on the structure of the natural deduction proof H.

Assume that H is of the form

$$\cfrac{\cfrac{H_1}{\Phi, \Phi_1 \vdash P_1}\{R_1\}, \quad \ldots, \cfrac{H_m}{\Phi, \Phi_m \vdash P_m}\{R_m\}}{\Phi \vdash P}\{R\}$$

where $m \geq 0$. The induction hypothesis is that for each $i \leq m$, there is a structured proof task $T_i = \bullet P_i;\ -\Phi_i; \ldots ; \square$ that proves $\Phi, \Phi_i \vdash P_i$ (i.e., $\Phi, \Phi_i \vdash T_i$). Then the following is a structured task T that proves $\Phi \vdash P$(i.e., $\Phi \vdash T$):

• P

- Φ

⊩ $\{R\}$

T_1

$\vdots$

T_m

□

This follows directly from the definition of correctness of a structured task, which requires that each nested task is correct, and that the basic inference rule R is valid. The former follows from the induction hypothesis, and the latter from the fact that R is an inference rule in natural deduction. ■

The proof shows that we can turn any natural deduction proof of $\Phi \vdash P$ into a structured task that proves the same fact. The transformation is done step by step, starting from the last inference in the natural deduction proof.

Appendix B

Answers to exercises

Chapter 2

1. a^{-1}

5. 2

6. $2x\cos(2x) - 2x^2\sin(2x)$

Chapter 4

1. a)

2. c)

3. False

4. a)

5. d)

6. c)

7. True

9. T

10. $x > y$

11. T

12. Yes

13. T

14. T

15. a) T, c) Yes

16. $(2 < x < 3 \vee 4 < x < 5) \wedge x = 4, 5$

Chapter 5

1. $x = 5$ and $y = 7$

2. $x = \sqrt{\frac{3}{2}} \wedge y = 0$

3. $-2 < x < 3$

8. Yes, $a = 8$ (despite that the two equations are actually the same equation this only means that it holds for the points that satisfy the equation, not that it would be universally true)

9. $x = 1 \wedge y = 2 \wedge z = 4$

10. 4 liters of the 15% solution and 8 liters of the 30% solution.

14. $x = 0$

15. $x = 3 \vee x = -8$

16. $x = -\frac{1}{2} \vee x = 0 \vee x = 7$

17. $x = 5 \wedge y = 7$

18. $x = 12 \vee x = -4$

Chapter 6

1. $x = -\frac{9}{5}$

2. $x = 1$

3. $x = \sqrt[3]{2}$

4. $x = e$

5. Yes

7. 3

9. $2^n n!$

Chapter 9

1. $x = 5 \wedge y = 17 \ \wedge \ z = 9$

2. $x = 3 \ \wedge \ y = \frac{1}{3} \ \wedge \ z = 6$

3. $x = \frac{13}{5}$

4. $x < -8 \ \vee \ x > -2$

5. $x = -3 \ \vee \ x = 5$

6. Yes

7. It intersects the xy-plane in the point $(5.2,\ 3.8,\ 0)$

8. The number is 13

Chapter 10

3. $2 < x \leq \sqrt{10}$

5. $(4,\ 0,\ 1)$

Chapter 13

1. 2.5 pizzas

2. It is $\frac{1}{10}$

3. Anna mixed the juice at a ratio of $3 : 11$

4. The person could be on their death bed arranging cards and they would still have barely started the task (the age of the universe would also be woefully insufficient as would the squared and cubed ages of the universe)

5. The ratio of the mixtures should be 3 parts of the stronger and 4 parts of the weaker sauce

6. The increase in price did not pay off!

7. Approximately $2.5 \cdot 10^{22}$ joules of energy was released

Chapter 14

1. $x = 0 \ \vee \ x = 1$, when n is even, $x = -1 \ \vee \ x = 0 \ \vee \ x = 1$, when n is odd

2. $x < -1 \ \vee \ -1 < x < 0 \ \vee \ x > 1$

3. a) Approximately 1600 km b) The volume of Cube-Sedna is approximately $1.5 \cdot 10^9 \text{km}^3$. This is approximately 63% of the original volume

4. a) $102\,\text{dm}^3$ b) Approximately $91\,\text{dm}^3$ of gunpowder

5. a) $\frac{2}{11}$, b) $\frac{10}{33}$, c) $\frac{4}{165}$.

6. a) $4x^3 + 5$ b) 37 c) $f(x) = \left(\frac{5}{4}\right)^{\frac{4}{3}} - 5 \cdot \sqrt[3]{\frac{5}{4}} + 2 \approx -2.03956\ldots$ at $x = -\sqrt[3]{\frac{5}{4}}$

7. a) 12 000 volume units, b) $1.54\,\text{m}^3$c) 1 080 kg of wood d) 2.76 liters of varnish

Chapter 15

1. a) bound: x, free: a, b) bound: a, free: $b,\ x,\ k,\ c$, c) bound: $x,\ y$, free: $h,\ z$.

2. $(\exists n \in A : (\forall m \in A : nm = n \wedge n + m = m))$. This works for $\mathbb{R}$, but not for $\mathbb{R}_+$.

3. $(\forall x \in \mathbb{Z} :\ (\forall y \in \mathbb{Z} :\ (\exists z \in \mathbb{Z} :\ y - z = x)))$. The original statement is true.

4. a) $(\exists x : (P\,(x,\ \omega) \wedge (\forall y : (\neg Q\,(y,\ x) \wedge P\,(y,\ z)))))$,
b) $(\exists x' : (P\,(x',\ f\,(x)) \wedge (\forall y : (\neg Q\,(y,\ x') \wedge P\,(y,\ z)))))$,
c) $(\exists x : (P\,(x,\ y) \wedge (\forall y' : (\neg Q\,(y',\ x) \wedge P\,(y',\ g\,(y,\ z)))))$

Chapter 16

1. a)$(\exists x \cdot P\,(x)) \wedge (\exists y \cdot \neg Q\,(y))$, b) F

2. a) $(\forall x \cdot P\,(x) \wedge (\exists y \cdot \neg Q\,(y)))$, b) $(\exists x \cdot (\forall y \cdot \neg P\,(x) \wedge \neg Q\,(y)))$

4. No, $x = -2$ will always makes the inequality false.

Bibliography

[1] R. J. Back, M. Sjöberg, and J. von Wright. Field tests of the structured derivations method. Tech. Rpt. 491, Turku Centre for Computer Science, November 2002.

[2] Ralph-Johan Back. *Correctness Preserving Program Refinements: Proof Theory and Applications*, volume 131 of *Mathematical Center Tracts*. Mathematical Centre, Amsterdam, The Netherlands, 1980.

[3] Ralph-Johan Back. A calculus of refinements for program derivations. *Acta Informatica*, 25:593–624, 1988.

[4] Ralph-Johan Back. Structured Derivations: a Unified Proof Style for Teaching Mathematics. *Formal Aspects of Computing*, 22(5):629–661, 2010.

[5] Ralph-Johan Back. *Structured Derivations: Teaching Mathematical Reasoning in High School*. Four Ferries Publishing, 2015.

[6] Ralph-Johan Back, Stefan Asikainen, Matti Hutri, Joonatan Jalonen, Antti Lempinen, Marie Linden-Slotte, Saara Mäkinen, Petri Sallasmaa, and Petri Salmela. *eMath: Textbooks for High School Mathematics*. Four Ferries Publishing, 2016.

[7] Ralph-Johan Back, Jim Grundy, and Joakim von Wright. Structured calculational proof. *Formal Aspects of Computing*, 9:469–483, 1998.

[8] Ralph-Johan Back, Linda Mannila, Mia Peltomäki, and Patrick Sibelius. Structured derivations: A logic based approach to teaching mathematics. In *FORMED 2008: Formal Methods in Computer Science Education, Budapest*, 2008.

[9] Ralph-Johan Back, Linda Mannila, and Solveig Wallin. "it takes me longer, but i understand better" - student feedback on structured derivations. Technical Report 943, 2009.

[10] Ralph-Johan Back, Linda Mannila, and Solveig Wallin. Student justifications in high school mathematics. In *CERME 6*, January 2009.

[11] Ralph-Johan Back, Linda Mannila, and Solveig Wallin. "It Takes Me Longer, but I Understand Better" - Student Feedback on Structured Derivations. In *International Journal of Mathematical Education in Science and Technology*, volume 41, pages 575–593, 2010.

[12] Ralph-Johan Back and Joakim von Wright. *Refinement Calculus: A Systematic Introduction.* Springer-Verlag, 1998. Graduate Texts in Computer Science.

[13] Ralph-Johan Back and Joakim von Wright. A method for teaching rigorous mathematical reasoning. In *Proceedings of Int. Conference on Technology of Mathematics*, University of Plymouth, UK, Aug 1999.

[14] Ralph-Johan Back and Joakim von Wright. Structured derivations: a method for doing high-school mathematics carefully. TUCS Technical Report 246, TUCS - Turku Centre for Computer Science, Turku, Finland, Mar 1999.

[15] A. Church. A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5:56–68, 1940.

[16] Edsger W. Dijkstra and C. S. Scholten. *Predicate Calculus and Program Semantics.* Springer-Verlag, 1990.

[17] E.W. Dijkstra. The notational conventions I adopted, and why. *Formal Aspects of Computing*, 14:99 – 107, 2002.

[18] Maria Joai Frade. Classical first-order logic. Lecture notes, www4.di.uminho.pt/ mjf/pub/SFV-FOL-2up.pdf.

[19] Gerhard Gentzen. Untersuchungen uber das logische schließen. *Mathematische Zeitschrift*, 39(2):176–210, 1934.

[20] M.J.C. Gordon and T.F. Melham. *Introduction to HOL.* Cambridge University Press, New York, 1993.

[21] David Gries. Teaching calculation and discrimination: A more effective curriculum. *Communications of the ACM*, (34):45 – 54, 1991.

[22] David Gries and Fred Schneider. *A Logical Introduction to Discrete Mathematics.* Springer-Verlag, 1993.

[23] David Gries and Fred Schneider. Teaching math more effectively through calculational proofs. *Am. Math. Monthly*, pages 691–697, October 1995.

[24] Anne Kaldewaij. *Programming: The Derivation Of Algorithms.* Prentice Hall, 1990.

[25] Leslie Lamport. How to write a proof. *American Math. Monthly*, 102(7):600–608, 1995.

[26] Linda Mannila and Solveig Wallin. Promoting students justification skills using structured derivations. In *ICMI 19 studies*, 2009.

[27] Linda Mannila and Solveig Wallin. Promoting Students' Justification Skills Using Structured Derivations. In *Proceedings of the ICMI Study 19 Conference: Proof and Proving in Mathematics Education*, pages 64–69, Taipei, Taiwan, 2009. National Taiwan Normal University.

[28] Sam Owre, Natarajan Shankar, and John Rushby. PVS: A prototype verification system. In *CADE 11*, Saratoga Springs, NY, June 1992.

[29] L. C. Paulson. Isabelle: the next 700 theorem provers. In P. Odifreddi, editor, *Logic and Computer Science*, pages 361–386. Academic Press, 1990.

[30] Mia Peltomäki and Ralph-Johan Back. An empirical evaluation of structured derivations in high school mathematics. In *ICMI-19 studies*, Taipei, Taiwan, 2009.

[31] Pisa. *Draft Mathematics Framework*, 2015. https://www.oecd.org/pisa/pisaproducts/Draft

[32] M. E. Szabo, editor. *The Collected Papers of Gerhard Gentzen.* Studies in Logic and the Foundations of Math. North Holland, Amsterdam, 1969.

[33] A. Trybulec. The Mizar logic information language. In *Studies in Logic, Grammar and Rhetoric*, volume 1. Bialystok, 1980.

[34] A. Trybulec and P. Rudnicki. On equivalents of well-foundedness: An experiment in Mizar. *Journal of Automated Reasoning*, 23:197–234, 1999.

[35] Jan L. A. van de Snepscheut. *What computing is all about.* Springer Verlag, 1993.

[36] A. J. M. van Gasteren. *On the Shape of Mathematical Arguments.* Lecture Notes in Computer Science. Springer-Verlag, Berlin, 1990.

[37] Markus Wenzel. Isar - a generic interpretative approach to readable formal proof documents. In Y. Bertot, G. Dowek, A. Hirschowitz, C. Paulin, and L. Thery, editors, *Theorem Proving in Higher Order Logics, 12th International Conference, TPHOLs'99*, volume 1690 of *LNCS*. Springer Verlag, 1999.

Teaching Mathematics in the Digital Age
with Structured Derivations

This book describes an alternative way of presenting mathematical arguments, **structured derivations**, that aims at making the reasoning more transparent and easier to understand. A structured derivation shows clearly the overall structure of the argumentation, while at the same time requiring that each step in the derivation is carefully justified. Structured derivations can be used in any area of mathematics, and at any level of education.

The format has been specially designed for teaching mathematics in a digital environment, with editors for creating solutions to mathematics problems, the web for rapid feedback from teachers, and computers for analyzing the correctness of students solutions, just to name a few examples. The method has been developed in close cooperation with mathematics teachers. It has been tested in a large number of pilot courses on high school level, with very good feedback from both teachers and students.

4Ferries
www.fourferries.com

ISBN 978-952-7147-01-6