Automated Reasoning 2016/2017
Coursework
Theorem Proving in Isabelle

Daniel Raggi Jacques Fleuriot

October 16, 2016

Introduction

The coursework for Automated Reasoning is designed to test and help to develop your understanding and practical skills using the interactive theorem prover Isabelle/HOL.

Part 1 consists of a set of exercises involving the use of basic inference rules in Isabelle/HOL. You will have to prove simple theorems in propositional and first order logic.

In Part 2 you will have to prove more complex theorems about miscellaneous mathematical concepts. Thankfully, you will also be granted access to more powerful reasoning tools!

In Part 3 you will delve into some simple formalisations of geometry in Isabelle/HOL. You will be asked to formalise axiomatic systems for geometry and prove theorems resulting from these axioms. Moreover, you will explore some surprising models for these axiomatic systems, which give rise to a fascinating area of mathematics: finite geometries.

To get started, download the file practical.thy from:

`http://www.inf.ed.ac.uk/teaching/courses/ar/`

Essential Reading

As you will be using Isabelle interactively, you will need to be familiar with the system before you start. Formal mathematics is not trivial! You will find this assignment much easier if you attend the lectures, attempt the various Isabelle exercises given on the course webpages, and ask questions about using Isabelle before you start. You can find a lot of useful information on using Isabelle in the Isabelle/HOL tutorial, located at:

`http://www.cl.cam.ac.uk/research/hvg/Isabelle/dist/Isabelle2016/doc/tutorial.pdf`,

In particular, Chapter 5 has a lot of information on all the basic reasoning tools and rules. Moreover, Chapter 6 contains some information about the representation of Sets and Functions in Isabelle/HOL.

You may also find it very useful to take a look at Chapter 5 of Concrete Semantics, where a clear overview of declarative (structured) proofs is presented:

`http://www.concrete-semantics.org/concrete-semantics.pdf`.

Part 1: Some propositional and first-order proofs [15%]

For the first part of this assignment, you should attempt to prove a number of simple propositional and first-order statements in Isabelle.

For this part of the assignment use only the following proof methods: rule, rule_tac, drule, drule_tac, erule, erule_tac, frule, frule_tac and assumption. You are also restricted to using only the introduction and elimination rules taught in lectures (e.g. conjI, conjE, exI, allE, ccontr) in the usual procedural style (a sequence of rule applications).

You should also be aware of the tactic cut_tac, which inserts a known rule or fact as an assumption in your proof. For example, the known fact:

excluded_middle: "¬ P ∨ P"

can be inserted as an assumption in your proof by using the command

```
apply (cut_tac excluded_middle)
```

If you wish to rename the variable P, to A say, then you can simply give the command

```
apply (cut_tac P=A in excluded_middle)
```

If you are struggling to mechanize a lemma or theorem in Isabelle, then the command sorry can be used. This allows the lemma or theorem to be asserted as true without completing the proof. It will enable you to make progress in the practical, however no marks will be allocated for the missing part of the proof. You should not use other people's proofs or formalisations.

Prove the following statements:

1. $(P \longrightarrow Q) \longrightarrow (P \longrightarrow \neg Q) \longrightarrow (P \longrightarrow R)$ **(2 marks)**
2. $(P \longrightarrow Q) \vdash (R \longrightarrow S) \wedge P \wedge R \longrightarrow Q \wedge S$ **(2 marks)**
3. $(\neg P \wedge \neg Q) \longleftrightarrow \neg(P \vee Q)$ **(3 marks)**
4. $\neg(\exists x.\ \neg P\, x) \vdash \forall x.\ P\, x$ **(3 marks)**

For the next problem, you are also allowed to use any rules you proved in the previous exercises (in addition all the rules taught in lectures):

5. $\exists x.\,(\texttt{drunk}\ x \longrightarrow (\forall y.\ \texttt{drunk}\ y))$ **(5 marks)**

Hint: To use a rule, you first have to give it a name. For example, if you have proved a statement myrule: $P \vdash Q$, and later you find $A \vdash Q$ as a goal in the proof of another theorem, writing apply (rule myrule) will yield $A \vdash P$ as a new goal.

Part 2: Structured proofs & powerful reasoning tools [25%]

In this part of the assignment you are asked to prove some more complex theorems of mathematical miscellanea. For this purpose, you are given access to some of Isabelle's more powerful reasoning tools. Specifically, you can use the tactics: `auto`, `simp`, `blast`, `fast`, `force`, `fastforce`, `presburger`, `algebra`, `unfold`, `induction`, as well as all the methods from Part 1. Moreover, now you can write proofs in Isabelle's declarative (structured proof) style:

```
theorem x:  ···
proof ···
  assume ···
    ⋮
  show ···
qed
```

For this part you are **not** allowed to use the tactics `metis`, `meson` or `smt`, unless explicitly stated. You can use the tools `sledgehammer`, `try` and `try0` to suggest methods, but take into account that if they suggest you to use `metis`, `meson` or `smt` you should then find an alternative proof.

Problem 6 (3 marks). Recall that $\exists x.\, \forall y.\, P\, x\, y$ implies $\forall y.\, \exists x.\, P\, x\, y$. However, it is not necessarily the case that $\forall y.\, \exists x.\, P\, x\, y$ implies $\exists x.\, \forall y.\, P\, x\, y$. This fact is captured by the sentence $\neg \forall P.\, (\forall y.\, \exists x.\, P\, x\, y \longrightarrow \exists x.\, \forall y.\, P\, x\, y)$, which you have to prove, assuming that the universe is the domain of integer numbers. The first step of the proof (an application of the tactic `simp`) is given, which you will see yields the goal $\exists P.\, (\forall y.\, \exists x.\, P\, x\, y) \wedge (\forall x.\, \exists y.\, \neg P\, x\, y)$. You have to write the rest of the proof in declarative style.

Problems 7, 8, 9 (2, 3, 3 marks respectively). For these problems a relation R_{12} is defined (given). This relation is such that two integers are related by R_{12} if and only if 12 divides their difference. You have to prove that R_{12} is reflexive (Problem 7), symmetric (Problem 8) and transitive (Problem 9). Some steps in these proofs are given. You have to write the rest of the proofs in declarative style.

For the next few problems you will need to define *primitive recursive functions* using Isabelle's command primrec. You should have in mind that if a function f is defined using primrec then Isabelle creates *simplification rules* f.simps. These rules are equalities which are incorporated automatically into the workings of tactics `simp` and `auto`.

Problem 10 (2 marks). You have to define a function gauss_sum where

$$\text{gauss\_sum}\ n = 1 + 2 + \cdots + n.$$

You need to give this definition recursively using the command primrec. Thus, you need to define it for two cases: when the argument is 0 and when the argument is the successor of something. You are given the structure of this definition and you only need to fill the missing parts. After you have defined the function, you need to state and prove the theorem 2(gauss_sum n) $= n(n+1)$. Here, you can use the tactic `induction`. You can complete this proof either in declarative or procedural (sequence of tactic applications) style.

Problem 11 (2 marks). Like above, you have to define a function sum_of_odds using the command primrec. Define it so that sum_of_odds n is the sum of the first n odd numbers. After you have defined it, state and prove the theorem sum_of_odds $n = n^2$. You can use `induction`.

Problem 12 (5 marks). For this problem the primitive recursive definition of the function power_sum is given. Specifically, it is such that

$$\textsf{power\_sum}\; n\, m = (n-1) + (n-1)n^1 + (n-1)n^2 + \cdots + (n-1)n^m.$$

You have to prove that if $n > 0$ then power_sum $n\,m = n^{m+1} - 1$. You can build this proof either declaratively or procedurally. This problem is hard because the natural number definition of subtraction ("minus") in Isabelle is such that if $a < b$ then $a - b = 0$. A consequence of this is that $(a-b)+c$ will not necessarily equal $(a+c)-b$. Thus, you might find it hard to translate your intuitions into a working proof.

In traditional mathematical writing it is common to show $a = d$ with a 'chain' of equalities, as follows:

$$\begin{aligned} a &= b \\ &= c \\ &= d. \end{aligned}$$

Behind the scenes of such a proof, one is proving the chained equalities $a = b$, $b = c$ and $c = d$ and implicitly using the transitivity of equality to show $a = d$. The equivalent way of writing this pattern inside an Isabelle declarative proof is:

```
have            ‘‘a = b’’ by ⟨some method⟩
moreover have ‘‘... = c’’ by ⟨some method⟩
moreover have ‘‘... = d’’ by ⟨some method⟩
ultimately show ‘‘a = d’’ by auto
```

The dots '...' allow you to avoid writing terms redundantly. The keyword moreover collects a bunch of facts (in this case the facts are the equalities in the chain), and the keyword ultimately uses all of the collected facts to be used in the proof of the following statement. In this case, they will be chained equalities, which will allow auto to prove $a = d$ by transitivity.

For Problem 12 you may want to write (on paper) a proof in the traditional chained style, and then try to formalise it in Isabelle using the hint above. If you cannot link two consecutive steps in the chain, try adding a helpful intermediate step.

Problem 13 (5 marks). If all the elements of a set are equal, then the set's cardinality cannot be greater than 1. Here you are asked to prove this statement. You can construct this proof either declaratively or procedurally. For this problem you **can** use any tactic except smt. **Hint:** given an *existence* theorem $\exists x.\, P\, x$, you can use the command `obtain` in a declarative proof to introduce an element that satisfies P.

Part 3: Reasoning about Geometries [60%]

Geometry has a long history of being presented and represented in terms of *axiomatic systems*. Here you will work with one axiomatisation which conceives the geometric *plane* as a set (of *points*), and *lines* as sets of points (of the plane). For these concepts to match our geometric intuitions, some axioms must be satisfied. We represent this in terms of Isabelle's *locales*.

For the proofs in this section you can use any tactic except smt. You can choose whether you want to construct proofs declaratively or procedurally (or mixed, if you need it). Note that structured proofs (declarative) are more congenial to formalised mathematics, and more similar to how mathematicians tend to present their arguments. You can also use the help of sledgehammer, try and try0.

A locale called Simple_Geometry is declared, with a pair of constants 'plane' and 'lines'. The first three axioms (A1, A2 and A3) are given. They state that the plane is not empty (A1), that every line is a non-empty subset of the plane (A2), and that for every pair of points in the plane there is a line that contains both (A3).

Problem 14 (2 marks). Using the same syntax used for the first three axioms, formalise Axioms 4 and 5:[1]

A4: *Two different lines intersect in no more than one point.*[2]

A5: *For every line L there is a point in the plane outside of L.*

As you may notice, these 5 axioms are not everything we will be using. In fact, by defining the plane as a set, and lines as subsets of it, we inherit results from the theory `Set` of Isabelle/HOL. In fact, we inherit everything from Isabelle's `Main` theory. This background allows us to use various lemmas (e.g., those that `sledgehammer` may suggest), and makes tactics like `simp` and `auto` quite powerful.

Given these axioms we can start proving some basic theorems. For some theorems the statement is given and you only have to find the proof. For some theorems you have to both formalise the statement and find the proof.

As a general hint, consider that every theorem that you prove can be used as a lemma for constructing the proof of another theorem. Moreover, if you get stuck in one proof you can skip it by writing the command `sorry` (in place of proof methods). Naturally, you cannot get full marks for a proof that contains `sorry`. However, you can get partial marks for it, and it allows you to move on with the rest of the assignment. Moreover, `sorry` can be used so that you can *pretend* that you have completed the proof. In other words, `sorry` is like an (invalid) inference rule that allows you to prove **anything**. Therefore, if you use `sorry` in the proof of one lemma you can use that lemma afterwards in the proof of something else, as if it was a proven fact. Thus, you must be very careful not to use sorry in a **false** statement, as this will result in the introduction of an inconsistency (and recall that, in classical logic, any inconsistency allows you to prove everything!). You should know that if you use `sorry` in the solution of a problem P we will penalise your marks for P,

[1] Be very, very careful with the way that you formalise these axioms. The wrong axiomatisation can make the locale too weak (it does not entail certain necessary properties), too strong (it entails more than we are asking), inefficient (it is difficult to reason with it) or outright inconsistent (it entails a contradiction!). If any of these things happen, marks will be deducted proportionally to the magnitude of the problem. Furthermore, it may make it very hard for you to move forward in the assignment.

[2] We recommend that you **do not** use the function `card` of Isabelle for this axiom. In Isabelle, the cardinality of infinite sets is defined as 0, so using `card` may have unintended disastrous consequences (e.g., allowing different lines to intersect in an infinite number of points).

but only once. Thus, if you use P as a lemma for the proof of a problem Q, we will not take marks off Q; unless P introduces an inconsistency. If you introduce an inconsistency then we will take off marks every time that inconsistency is used, even if unknowingly (e.g., by one of the automatic reasoning tools).

Problem 15 (1 mark). State the fact that there exists at least one line and prove it.

Problem 16 (2 marks). Prove that there exist at least two different points in the plane (the statement is given).

Problem 17 (3 marks). Prove that there exist at least three different points in the plane (the statement is given).[3]

Problem 18 (3 marks). Prove that the cardinality of the plane is greater than 3 (if the plane is finite, because recall that `card` of infinite sets is 0!).

One of the interesting results of our simple axiomatisation of geometry is that there exist *models* of simple geometry which only contain a finite number of points.

Problem 19 (2 marks). We have proved that there are *at least* 3 points. Now we will show that there is a model of Simple Geometry with no more than three points! Using the command interpretation, give a model of Simple Geometry where the plane has only 3 elements. **Hint:** the *points* can be anything, e.g., integers[4]. This command will ask you to prove that the 5 axioms are satisfied by your given model. It should be easy to show that it is the case.

Problem 20 (5 marks). Suppose you have a line l and two different points a and b on it. Suppose that you have a point p outside of l. Let n be the line which contains both a and p, and let m be the line that contains both

[3] Notice the use of predicate distinct, which applies to lists. It simply generalises the notion given by $\neq$ to various items. Specifically, it means that for **every pair** a, b of items in the list $a \neq b$.

[4] If you want to use numerals (integers or naturals) as the 'points' of the plane, be sure to specify their type, e.g., by writing, '`2::int`'. Otherwise, Isabelle will not know facts such as $2 \neq 1$ (in Isabelle, numerals can be such that $2 = 1$, because these may represent objects other than integers or naturals).

b and p. Prove that the lines n and m are different.[5] **Hint:** Prove it by contradiction. Assume that m and n are the same. Then notice that m must intersect l in two points (a and b). From this conclude that $m = l$ and use this to show that p is in l. Show that this contradicts your assumptions. See figure 1 if it helps you to visualise it.

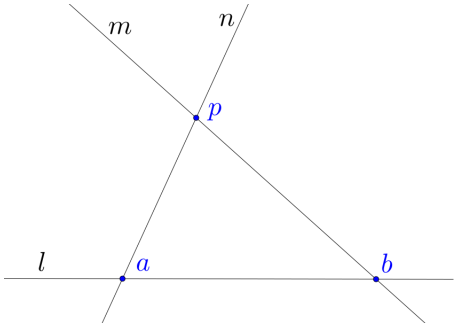

Figure 1

Problem 21 (4 marks). Suppose you have a line l and two points a and b on it. Suppose that you have a point p outside of l. Let n be the line which contains both a and p, and let m be the line that contains both b and p. Let c be a point on n different from p, and let d be any point on m. Prove that c and d are different. See figure 2.

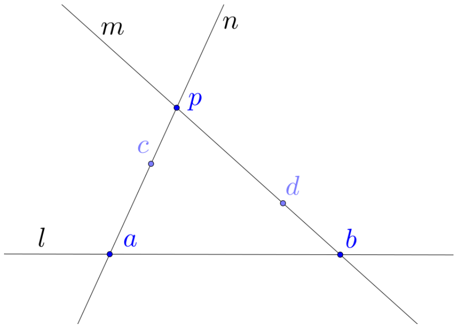

Figure 2

[5] For this problem the statement is given in terms of the keywords **assumes** and **shows**. Recall that within the proof you can invoke the whole collection of assumptions with the name assms. To invoke individual assumptions (e.g., the i^{th}) write assms(i).

Notice that the 5 axioms of Simple Geometry do not specify anything about parallel lines. Interestingly, there are models of this Geometry where there are no parallel lines (every two lines intersect). These are called *projective geometries.* Conversely, there are other models of Simple Geometry where there are some parallel lines or even *too many* of them! We will explore briefly what happens if we add the requirement that parallels must exist. We call this Non-Projective Geometry, and we define it as a new locale. Notice that we define it as an extension of Simple Geometry. Doing this allows us to use every theorem of Simple Geometry in Non-Projective Geometry.

Problem 22 (1 mark). Formalise an axiom which captures the following: *if a point p lies outside of a line l then there must exist at least one line m that passes through p, which does not intersect l.*

Problem 23 (2 mark). Give a model of Non-Projective Geometry where the plane has 4 points. Use the command `interpretation` to show that it is indeed a model.

Problem 24 (3 marks). Formalise the statement *it is not true that every pair of lines intersect in a point.* Prove it.

Enough of Non-Projective Geometries! Now let us define a new locale Projective Geometry, which includes two extra axioms. The first one states that every two lines intersect in a point (A6), and the second one states that there are at least three points (A7). The formalisation of these statements is given. Do not change them. Now we will prove some theorems in Projective Geometry.

Problem 25 (3 marks). You are asked to prove an alternative version of axiom A7. The statement is given. Naturally, you should use A7 in the proof.

Problem 26 (3 marks). You are asked to prove yet another alternative to axiom A7. The statement is given.

Problem 27 (5 marks). For every point in the plane there are at least two lines that pass through it. The statement is given and you have to give a proof.

Problem 28 (4 marks). For every point in the plane, there is one external line. The statement is given and you have to give a proof.

Notice that this fact is analogous to axiom A5, which states that for every line there is a point outside of it.

Problem 29 (6 marks). For every point p in the plane, there are at least three lines that pass through it. See figure 3 for an idea on how to prove this.

Notice that this fact is analogous to axiom A7, which states that every line contains at least three points.

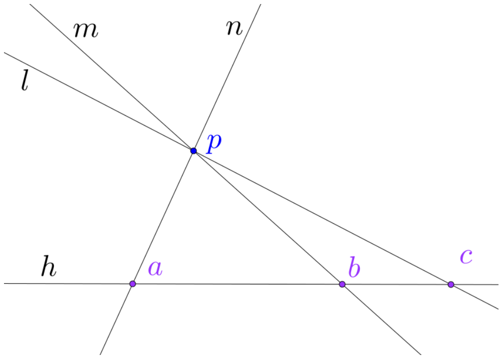

Figure 3: *If you find a line h and three different points a, b and c that lie on it you may be able to prove that their respective lines l, m and n are different. Can you? Maybe you need the objects in the figure to satisfy some extra properties!*

Fun fact: the theorems proved for problems 28 and 29 show a *duality* between some facts of projective geometry, which results from swapping points for lines and lines for points. The extent of this duality is even stronger. For example, notice that axiom A3 states that *every two points have one line in common.* Interestingly, its dual statement is axiom A6, which states that *every two lines have one point in common.* Moreover, axiom A4 is its own dual!

Problem 30 (8 marks). Prove that there are at least 7 points on the plane.

This implies, in particular, that there are no models of Projective Geometry with fewer than 7 points. But, is there a model with 7?

Problem 31 (3 marks). Give a model of Projective Geometry with exactly 7 points.

Demonstrator Hours

The demonstrator, Daniel Raggi (D.Raggi@sms.ed.ac.uk), will be available to give advice on Mondays from 9am-11am in 1 FH (Forrest Hill) Room 3.D02, or any time by email.

Submission

By 4pm on 21th November 2016 you must submit your solution in electronic form. This should consist of your theory file practical.thy and can be submitted using the command:

```
submit ar 1 practical.thy
```

Late coursework will be penalised in accordance with the Informatics standard policy. Please consult your course guide for specific information about this. Also note that, while we encourage students to discuss the practical among themselves, we take plagarism **seriously** and any suspected case will be treated appropriately. Please consult your student guide for more information about this matter.