Logic II
Lecture Notes

Josef Schicho
RISC Linz
Johannes Kepler University, A-4040 Linz
e-mail: Schicho@risc.uni-linz.ac.at
www: http://www.risc.uni-linz.ac.at/people/jschicho

March 30, 1998

1 Introduction

There are two motivations for the desire to formalize reasoning. On the one hand, one wants to have some sort of "quality control", in order to check and improve reliability and efficiency of proofs. This was the impetus behind Euclid's efforts to formalize geometry. It was also the motivation of Frege, Russell, Hilbert, Dedekind, ..., to construct a universe of mathematics based on formal systems. Historically, the efforts of axiomatization led to the discovery of antinomies, which in turn increased the necessity to improve on reliability. A more modern example are the efforts of Buchberger to improve the "quality of proofs" of his students in his famous Thinking-Speaking-Writing course.

A completely different goal is the desire to automatically generate or verify theorems. This idea can be traced back to Lullus (around 1300, "ars magna") and Leibniz (1646–1716), who distinguishes decision methods "ars iudicandi" and generation methods "ars inveniendi". Modern computer science adopted this vision as the core of the "Artificial Intelligence" program. Even though limitations of this approach became clear, we have now many useful algorithms. Most notably are Robinson's resolution algorithm and Collins' quantifier elimination algorithm.

In order to take the first aspect into account, we will restrict metatheory to finitary methods (as in the textbook of Kleene). In order to take the second aspect into account, we will in particular focus on algorithmic theorems.

Predicate calculus (also called first order logic). This part culminates in the completeness proof for Robinson's resolution procedure for semi-decision of predicate formulas. We also give proofs for Church's result of undecidability of predicate calculus, Gödel's completeness theorem, and Herbrand's theorem.

Elementary theory of reals . We give an axiomatization and discuss the Collins-Tarski decision method. We also give an axiomatization (due to Tarski) of geometry, which is closely related to the elementary theory of reals.

Formal number theory . Two famous negative results - Gödel's incompleteness theorem and the unsolvability of Hilbert's 10th problem - are obtained as immediate consequences of two positive results (which are closely related): any recursively enumerable set is arithmetic (Gödel) and any recursively enumerable set is diophantine (Matijasevic).

For the preparation of these notes, I leaned on the books [12, 2, 6], but not following them completely. The algorithm for quantifier elimination over the reals was taken from [1]. The discussion of Hilbert's 10th problem was taken from [11].

For further study, I recommend [21, 7, 17, 9, 18, 13, 5, 14] for studying predicate logic and [8, 4, 15, 3, 16, 19, 20] for studying the theory of reals and geometry. For automated theorem proving, I refer to the lectures [10].

2 Formal Theories

A *formal theory* consists of

- linguistical part:
  - symbols (predicate symbols, function symbols, logical connectives, variables, ...
  - terms (described by a syntax)
  - formulas, sentences (described by a syntax)
- logical part:
  - logical axioms
  - inference rules
- material part:
  - material axioms

A *formal theorem* is either an axiom or an immediate consequence of other formal theorems by some inference rule. A **proof** is a sequence of theorems such that each theorem is either an axiom or an immediate consequence of theorems before it.

The purpose of a formal theory is to formally describe reasoning/proving in some informal “real world” theory (number theory, analysis, set theory ...).

Semantics investigates the relation of the formal system and the informal theory which is going to be formalized. Typical problems:

- Construction of a formal theory for some given informal theory.
- Correctness (relative consistency).
- Completeness w.r.t. the interpretation (relative completeness).
- Definability of informal statements within the formal theory.
- Categoricity.

From the methodological point of view, semantics belongs to the given informal theory (number theory, set theory, ...). It is allowed to use the informal methods of this informal theory. (Consequently, its results belong to the informal theory. For instance, Gödel's completeness theorem is a result of set theory and not of logic.)

Metatheory is reasoning about the formal theory (which is also called *object theory* when the distinction to metatheory is emphasized). Metatheory is again an informal theory.

In contrast to semantics, a metatheory investigates a formal system as formal system, abstracting from the informal theory which is going to be formalized. In finding problems and solutions, it may be guided by the informal system, but in the final statement and in the proof, the use of the interpretation is prohibited. The allowed methods are the *finitary* methods:

- Any object is finite. The number of objects may be infinite (i.e. countable), but the set of all these objects is not an object of metatheory.
- Universal quantification is allowed if one can prove some assertion for each object in a finite number of steps.
- Existence proofs must contain a construction method (at least impliclitly).

Typical problems:

- Consistency.
- Completeness.
- Decidability.

An example of a finitary method is induction over natural numbers.

The viewpoint of *formalism* (Hilbert) is the following: once we have a formal theory, the informal object theory can be disregarded completely, because all statements can be proven formally. In order to construct a formal theory, the only care one has to take is consistency. The informal reasoning of the object theory is justified by a metatheoretical proof of consistency of the formal theory. 'Proving existency is nothing more than proving that sth may by described consistently' (Hilbert). This thesis lead to "Hilbert's program", which is to give metatheoretical consistency proofs for formalizations of all important theories in mathematics. We know that Hilbert's program broke down on Gödel's second incompleteness theorem: there is no metatheoretical proof for the consistency of formal number theory.

On the other hand, there is the viewpoint of *intuitionism* (Brouwer): a formalization of any mathematical theory is just an approximate description of the reasoning/proving within that theory. The quality of a proof has to be judged on the conviction it carries. It is not allowed to overtake "logical principles" from another theory without justification. Consequently, Brouwer attacks the universal use of the principle of excluded middle, because it is not convincing when applied to undecided statements. Consistency alone is not sufficient: 'An incorrect theory which is not stopped by a contradiction is none the less incorrect, just as a criminal policy unchecked by a reprimanding court is none the less criminal.' But still, formalizing informal theories and investigating the formalism by metatheory has its merits. We have now an intuitionistic propositional logic, an intuitionistic predicate logic, an intuitionistic formal number theory, an intuitionistic formal set theory, an intuitionistic formal analysis, and (since recently) also an intuitionistic proof of the Gröbner bases algorithm.

A property is called *semi-decidable* iff there exists an *accepting procedure*, i.e. the procedure accepts the input after finitely many steps iff the input enjoys the property. In metatheory, it does not make much sense to speak about properties which are not semi-decidable.

A procedure is called *decidable* iff there exists a *decision algorithm*, i.e. an algorithm that always terminates and tells if the input has the property or not.

3 Propositional Logic

The object theory is the theory of statements (propositions). Several theorems have been formulated by Aristotle:

"If A implies B and B implies C, then A implies C".

"Either A or not A" (principle of excluded middle).

Linguistical part:

- atoms P, Q, R, ... (infinitely many)
- connectives: $\wedge$, $\vee$, $\rightarrow$, $\neg$, $\perp$

- brackets, syntax, formulas

Binding assumptions: $\neg$ binds stronger than $\wedge$ and $\vee$, which bind stronger than $\rightarrow$. Moreover $\rightarrow$ binds from right to left, i.e. $A \rightarrow B \rightarrow C$ means $A \rightarrow (B \rightarrow C)$.

Logical part:

- Axioms:

$$A \rightarrow B \rightarrow A$$

$$(A \rightarrow B) \rightarrow (A \rightarrow B \rightarrow C) \rightarrow (A \rightarrow C)$$

$$A \rightarrow B \rightarrow A \wedge B$$

$$A \wedge B \rightarrow A$$

$$A \wedge B \rightarrow B$$

$$A \rightarrow A \vee B$$

$$B \rightarrow A \vee B$$

$$(A \rightarrow C) \rightarrow (B \rightarrow C) \rightarrow (A \vee B \rightarrow C)$$

$$(A \rightarrow B) \rightarrow (A \rightarrow \neg B) \rightarrow \neg A$$

$$\neg\neg A \rightarrow A$$

$$\neg\bot$$

  where A, B, C are formulas.

- Inference rules:

$$A, A \rightarrow B \vdash B$$

The consequence relation $\vdash$:

$$A_1, \ldots, A_r \vdash C$$

iff there exists a sequence $B_1, \ldots, B_n = C$ (called derivation), such that each B_i is either an axiom, or one of the A_i, or an immediate consequence of two of $B_1, \ldots, B_{i-1}$ by the inference rule (modus ponens).

As an example for a formal proof, we prove $P \vee Q \rightarrow Q \vee P$.

$$(P \rightarrow Q \vee P) \rightarrow (Q \rightarrow Q \rightarrow P) \rightarrow P \vee Q \rightarrow Q \vee P \text{ axiom 8}$$

$$P \rightarrow Q \vee P \text{ axiom 7}$$

$$(Q \rightarrow Q \rightarrow P) \rightarrow P \vee Q \rightarrow Q \vee P \text{ modus ponens from 1 and 2}$$

$$Q \rightarrow Q \vee P \text{ axiom 6}$$

$$P \vee Q \rightarrow Q \vee P \text{ modus ponens from 3 and 4}$$

Exercise 1 Prove that

$$A, B \vdash A \wedge B$$

$$A \wedge B \to C, A, B \vdash C$$

$$\vdash A \to A$$

Theorem 1 (Deduction theorem) *If* $\Gamma, A \vdash B$, *then* $\Gamma \vdash A \to B$.

Proof: By induction on the length k of the derivation of B.

If $k = 1$, then we distinguish two cases.

Case 1: B is an axiom or a formula in Γ. Then one can derive $A \to B$ by B and the axiom $B \to A \to B$.

Case 2: B is A. See above exercise.

Inductive step. Suppose that there exists a derivation of B of length $k+1$. If B is an axiom, a formula in Γ, or A, then the theorem follows as in the base case $k = 1$. Suppose that B is an immediate consequence of two formulas C and $C \to B$, which occur earlier in the proof. By induction hypothesis, the formulas $A \to C$ and $A \to C \to B$ can be derived from Γ. Then, we apply modus ponens to times to the axiom

$$(A \to C) \to (A \to C \to B) \to (A \to B)$$

and derive $A \to B$, as desired.

As an application of the deduction theorem, we prove the theorem $(A \to B) \to \neg B \to \neg A$. First, we prove that $A \to B, \neg B \vdash \neg A$:

$$\neg B \to A \to \neg B \text{ axiom 1}$$

$$\neg B \text{ assumption 2}$$

$$A \to \neg B \text{ modus ponens from 1 and 2}$$

$$(A \to B) \to (A \to \neg B) \to \neg A \text{ axiom 9}$$

$$A \to B \text{ assumption 1}$$

$$(A \to \neg B) \to \neg A \text{ modus ponens from 4 and 5}$$

$$\neg A \text{ modus ponens from 6 and 3}$$

Now, the theorem follows from the deduction theorem applied twice.

We write $A \sim B$ as an abbreviation for $\vdash (A \to B) \wedge (B \to A)$.

Exercise 2 Prove that $\sim$ is an equivalence relation.

Theorem 2 (Replacement theorem) *Let* C_A *be a formula with a subformula* A, *and let* C_B *be the formula obtained when* A *is replaced by* B *in* C_A.

If $A \sim B$, *then* $C_A \sim C_B$.

Proof: By induction over connectives in the context C_{-}, it suffices to prove the following lemma.

Lemma 1 *If $A{\sim}B$, then*
a) $A{\wedge}C{\sim}B{\wedge}C$.
b) $C{\wedge}A{\sim}C{\wedge}B$.
c) $A{\vee}C{\sim}B{\vee}C$.
d) $C{\vee}A{\sim}C{\vee}B$.
e) $A{\rightarrow}C{\sim}B{\rightarrow}C$.
f) $C{\rightarrow}A{\sim}C{\rightarrow}B$.
g) $\neg A{\sim}\neg B$.

Exercise 3 Prove three of the six points (b),...,(g).

Theorem 3 (Consistency) *$\bot$ is not a theorem.*

Proof: The natural way to prove consistency is to give a model. From the metatheoretic point of view, an interpretation is just a way of assigning truth to formulas in an algorithmic way. For several formal systems, it is possible to construct an interpretation with finitary means (semantics entering by the backdoor).

An *interpretation* $\mathcal{I}$ of a formula A is a function from the set of atoms occuring in A to {`True`,`False`}. We define the relation 'A is true under the interpretation $\mathcal{A}$' or '$\mathcal{I}$ is a model of A'

$$\mathcal{I} \models A$$

by induction over the subformulas, in the obvious way:

- An atom is true iff it is mapped to `True`.
- $A{\wedge}B$ is true iff A is true and B is true.
- $A{\vee}B$ is true iff A is true or B is true.
- $A{\rightarrow}B$ is true iff A is not true or B is true.
- $\neg A$ is true iff A is not true.
- $\bot$ is not true.

It is easy to see that all axioms are true under all interpretations, and for any two formulas which are true under all interpretations, their immediate consequence is true under all interpretations. Therefore, all theorems are true under all interpretations. On the other hand, $\bot$ is false under all interpretations, so $\bot$ is not a theorem.

Corollary 1 (Correctness theorem) *If A is a theorem, then A is true under all interpretations.*

Theorem 4 (Completeness) *If A is true under all interpretations, then A is a theorem.*

Proof: By the equivalences

$$A \rightarrow B \sim \neg A \vee B, \bot \sim A \wedge \neg A$$

(which must be shown, this is tedious but straightforward), we eliminate the symbols $\rightarrow$ and $\bot$. By the the Morgan rule

$$\neg(A \wedge B) \sim \neg A \vee \neg B, \neg(A \vee B) \sim \neg A \wedge \neg B,$$

we move negation signs at the innermost places. By the distributive laws

$$(A \wedge B) \vee C \sim (A \vee C) \wedge (B \vee C),$$

$$C \vee (A \wedge B) \sim (C \vee A) \wedge (C \vee B),$$

we move $\wedge$ outside. By the equivalence $\neg\neg A \sim A$, we cancel double negations. We have also commutativity and associativity of $\vee$, therefore we can disregard the brackets and permute factors in the disjunctions. Doing all this, we obtain a formula which is equivalent to the original formula and is in conjunctive normal form (I.e. a conjunction of disjunctions of atoms or negation of atoms). Therefore, it suffices to prove the theorem for conjunctive normal forms.

Suppose that $A = A_1 \wedge \ldots A_n$ is in conjunctive normal form and is true under all interpretations. Then the disjunctions A_i are true under all interpretations. This implies that any A_i is $P_i \vee \neg P_i \vee B_i$ for some atom P_i. By the theorem $A \vee \neg A$ and axiom 6, A_i is a theorem. By repeated application of modus ponens with axiom 3, we obtain that A is a theorem.

Theorem 5 (Decidability) *There is an algorithm that decides whether a given formula is a theorem.*

Proof: By theorem 1 and theorem 4, it suffices to check truth under all possible interpretations. There are only finitely many.

Remark 1 The above algorithm is obviously exponential. If the conjecture "$P \neq NP$" is true, then the problem is not polynomial (the problem is known to be co-NP-hard).

4 Predicate Logic

The object theory is the theory of functions and predicates (or subsets) in some set universe. Aristotle formulated several theorems (syllogisms), for instance:

"If all elements with property A have property B, and x has property A, then x has property B."

Linguistical part:

- predicate symbols, function symbols (infinitely many for each arity)
- connectives (same as in propositional logic)
- variables
- quantifiers
- terms, formulas

An occurence of a variable x in a formula A is said to be *bound* iff the occurence is in a quantifier $\forall x$ or $\exists x$ or in the scope of a quantifier $\forall x$ or $\exists x$. Otherwise, the occurence is called *free*.

A term t is said to be *free for x in A* iff no free occurence of x in A is in the scope of a quantifier where the quantified variable occurs in t.

If t is free for x in A, then we write $A(x \leftarrow t)$ for the formula obtained by substituting x by t in all free occurences of x in A.

Exercise 4 Give examples for the following informal facts:
a) substituting in a bound occurence is nonsense,
b) substituting a term which is not free for x is nonsense.

A formula is called *closed* (or: a sentence) iff it has no free variables. Sometimes, we write $A(x)$ instead of A to indicate that x may occur free in A (but not necessarily).

Logical part:

- Axioms: all axioms as in propositional logic (but with different types of formulas),
$$A(x \leftarrow t) \rightarrow \exists x A(x)$$
$$\forall x A(x) \rightarrow A(x \leftarrow t)$$
where t is free for x in $A(x)$.
- Inference rules:
$$A, A \rightarrow B \vdash B$$
$$C \rightarrow A(x) \vdash C \rightarrow \forall x A(x)$$
$$A(x) \rightarrow C \vdash \exists x A(x) \rightarrow C$$
where x is not free in C.

Theorem 6 (Deduction theorem) *If $\Gamma, A \vdash B$ and A is closed, then $\Gamma \vdash A \to B$.*

Proof: Analogous as in the propositional case.

Exercise 5 To see what happens when the assumption that A is closed is dropped, consider the following. Prove that

$$P(x) \vdash \forall x P(x).$$

Give a semantic argument that

$$P(x) \to \forall x P(x)$$

is not a theorem.

Theorem 7 (Replacement theorem) *Let C_A be a formula with a subformula A, and let C_B be the formula obtained when A is replaced by B in C_A.*
If $A \sim B$, then $C_A \sim C_B$.

Proof: Analogous as in the propositional case.

Theorem 8 (Consistency) *$\bot$ is not a theorem.*

Proof: In order to give a finitary proof, we consider finite semantics. Let k be a positive integer. A k-interpretation $\mathcal{I}$ of a formula A consists of

- a set S with k-elements,
- for each n-ary function symbol in A, a function from S^n to S;
- for each n-ary predicate symbol in A, a relation (i.e. a subset) of S^n;
- for each variable in A, an element.

We define a value function from the set of all subterms of A into S in the obvious way (by induction on subterms): each variable is mapped to its corresponding element, and each function symbol is mapped to the function value of the corresponding function term, with the values of the argument subterms as arguments.

As in the propositional case, one defines the relation 'A is true under the interpretation $\mathcal{I}$' or '$\mathcal{I}$ is a model for A' or $\mathcal{I} \models A$ in the obvious way (by induction on subformulas of A):

- A predicate symbol is true iff the corresponding predicate holds for the values of the argument terms.

- $A \wedge B$ is true iff A is true and B is true.
- $A \vee B$ is true iff A is true or B is true.
- $A \rightarrow B$ is true iff A is not true or B is true.
- $\neg A$ is true iff A is not true.
- $\perp$ is not true.
- $\forall x A(x)$ is true iff $A(x \leftarrow y)$ is true for all interpretations $\mathcal{I}'$ extending $\mathcal{I}$, where y is a new variable y (which may be assigned any value of S).
- $\exists x A(x)$ is true iff $A(x \leftarrow y)$ is true for one interpretation $\mathcal{I}'$ extending $\mathcal{I}$, where y is a new variable y.

It is easy to see that all axioms are true under all interpretations, and for any two formulas which are true under all interpretations, their immediate consequence is true under all interpretations. Therefore, all theorems are true under all interpretations. On the other hand, $\perp$ is false under all interpretations, so $\perp$ is not a theorem.

Remark 2 The correct metatheoretical interpretation of theorem 8 is not '$\perp$ has the property non-theorem' - because this property is not semi-decidable and therefore it cannot be verified with finitary means that $\perp$ has the property - but 'any theorem is not $\perp$', or, more explicitly, 'at the end of a proof, one always finds a formula different from $\perp$'. In fact, that is exactly what we have proved.

Corollary 2 *A necessary condition for being a theorem is that a formula is true under all of its k-interpretations.*

Exercise 6 Prove that the formula $P(x) \rightarrow \forall x P(x)$ is not a theorem.

Exercise 7 Try to define 0-interpretations exactly as we have defined k-interpretations for $k > 0$. Show that there exists a theorem which is not true in any 0-interpretation.

Exercise 8 Show that the formula

$$\neg(\neg P[x,x] \wedge ((P[x,y] \wedge P[y,z]) \rightarrow P[x,z]) \wedge \exists y P[x,y]),$$

is valid in all k-interpretations, for all $k > 0$.

Give a semantic argument showing that the formula above is not a theorem.

5 The Completeness Theorem

This section uses informal set theory including AC.

Let Γ be a set of formulas. The consequence relation $\Gamma \vdash C$ is defined as in the finite case: $\Gamma \vdash C$ iff there exists a sequence $B_1, \ldots, B_n = C$, such that each B_i is either an axiom, or an element of Γ, or an immediate consequence of two of $B_1, \ldots, B_{i-1}$. The set Γ is called *consistent* iff not $\Gamma \vdash \perp$.

An interpretation $\mathcal{I}$ of a formula Γ consists of

- a set S,
- for each n-ary function symbol in Γ, a function from S^n to S;
- for each n-ary predicate symbol in Γ, a relation (i.e. a subset) of S^n;
- for each free variable in Γ, an element.

Again, one defines the relation 'Γ is true under the interpretation $\mathcal{I}$' or '$\mathcal{I}$ is a model for A' or $\mathcal{I} \models A$ in the obvious way (as with k-interpretations). The set Γ is called satisfiable iff it has a model.

Theorem 9 (Correctness theorem) *If A is a theorem, then A is true under all interpretations.*

Proof: Similar to the proof of the consistency theorem (easy).

Theorem 10 (Completeness theorem) *If A is a formula which is true under all interpretations, then A is a theorem.*

This follows from the more general statement:

Theorem 11 (Completeness theorem for infinitely many formulas) *Let Γ be a set of formulas. If Γ is consistent, then Γ is satisfiable.*

Let A be a formula which is true under all interpretations. Then $\forall(x_1, \ldots, x_n)A$ is also true under all interpretations, where $x_1, \ldots, x_n$ are the free variables of A. Therefore the set

$$\Gamma = \{\neg\forall(x_1, \ldots, x_n)A\}$$

is not satisfiable. If the second theorem is true, then Γ is inconsistent, hence $\Gamma \vdash \perp$. By the deduction lemma,

$$\neg\forall(x_1, \ldots, x_n)A \rightarrow \perp$$

is a theorem. This is equivalent to $\forall(x_1, \ldots, x_n)A$, and from this we can deduce A. Hence A is a theorem. We have shown that the second theorem implies the first.

Theorem 12 (Prenex normal form) *Let A be a formula (predicate logic). There exists a formula $Pr(A)$, called* prenex normal form *of A, of the form $Q_1x_1\ldots Q_nx_nB(x_1,\ldots,x_n)$, where B is a formula without quantifiers, such that $A{\sim}Pr(A)$.*

Proof: By repeated application of the replacement theorem 7 and the following equivalences.

$$A \to B{\sim}\neg A{\vee}B,$$
$$\neg\forall xA(x){\sim}\exists x\neg A(x),$$
$$\neg\exists xA(x){\sim}\forall x\neg A(x),$$
$$\forall xA(x){\sim}\forall yA(x{\leftarrow}y)(*),$$
$$\exists xA(x){\sim}\exists yA(x{\leftarrow}y)(*),$$
$$\forall x(A(x){\wedge}B(x)) \sim \forall xA(x){\wedge}\forall xB(x),$$
$$\exists x(A(x){\vee}B(x)) \sim \exists xA(x){\vee}\forall xB(x),$$
$$B{\wedge}/{\vee}\forall/\exists xA(x) \sim \forall/\exists x(B{\wedge}/{\vee}A(x)),$$
$$\forall/\exists xA(x){\wedge}/{\vee}B \sim \forall/\exists x(A(x){\wedge}/{\vee}B),$$

where in $(*)$, we assume that y is neither free or nor bound in $A(x)$, and in $(**)$, we assume that x is not free in B.

Let A be a theorem in prenex normal form. We define a formula $Sk(A)$, the skolemization of A, as follows. If $A = Q_1x_1\ldots Q_n(x_n)B(x_1,\ldots,x_n)$, and Q_i is an existential quantifier, then we substitute x_i in B by the term $x_i(x_{i_1},\ldots,x_{i_k})$, where $x_{i_1},\ldots,x_{i_k}$ are the $\forall$-quantified variables with index less than i, and x_i is a new function symbol (which incidentally has the same name as the now superfluos variable). For instance, the skolemization of

$$\forall x_1\exists y_1\forall(x_2,x_3,x_4)\exists y_2\forall x_4B(z_1,z_2,x_1,y_1,x_2,x_3,y_2,x_4)$$

is

$$\forall(x_1,x_2,x_3,x_4)B(z_1,z_2,x_1,y_1(x_1),x_2,x_3,y_2(x_1,x_2,x_3),x_4).$$

Note that we do not have $A{\sim}Sk(A)$ in general (the statement itself makes only sense in the extended system).

Exercise 9 Compute the skolemization of the prenex normal form for the formula

$$\neg(\forall xP(x) \to \exists y\forall zQ(y,z)).$$

Let Γ be a set of formulas. The Herbrand universe $H(\Gamma)$ is the set of all terms formed by all function symbols and all free variables occuring in Γ. An interpretation of Γ with set $H(\Gamma)$ and obvious functions and obvious assignment is called a Herbrand interpretation. A Herbrand interpretation is specified by specifying the predicates.

Remark 3 If Γ has no free variables and no 0-ary function symbols, then the Herbrand universe as defined above is empty. In this case, one introduces a constant symbol a and defines the Herbrand universe as the set of all terms formed by all function symbols and by a.

Let Γ be a set of formulas in Skolem form. A ground instance of Γ is a formula of the form $B(t_1, \ldots, t_n)$, where $\forall(x_1, \ldots, x_n)B(x_1, \ldots, x_n)\in\Gamma$ and $t_1, \ldots, t_n\in H(\Gamma)$. The set of all ground instances is denoted by $Gr(\Gamma)$. We regard ground instances as formulas in the propositional calculus, where the atoms are $P(t_1, \ldots, t_n)$, where P is a predicate symbol and $t_1, \ldots, t_n\in H(\Gamma)$.

Lemma 2 *Let Γ be a consistent set of prenex normal forms. Then $Gr(Sk(\Gamma))$ is consistent in the propositional sense.*

Proof: Suppose that $Gr(Sk(\Gamma))$ is inconsistent, i.e. we have a derivation $Gr(Sk(\Gamma))\vdash\perp$. Only finitely many propositions are used in the derivation. Let $A_1, \ldots, A_n$ be the formulas in Γ of which a ground instances are used. We show that we can construct a derivation $A_1, \ldots, A_n\vdash\perp$.

Suppose that A_1 is $\forall x\exists y\forall zB(x, y, z)$. Suppose that the derivation uses the ground instances $B(x_1, y(x_1), z_1)$ and $B(x_2, y(x_2), z_2)$ (where $x_1, \ldots, z_2$ are elements from the Herbrand universe). The ground instances are consequences of the formulas

$$\forall zB(x_1, y(x_1), z), \forall zB(x_2, y(x_2), z),$$

therefore we may exchange the two ground instances by these two assumptions. If $x_1\neq x_2$, then $y(x_1)\neq y(x_2)$. Moreover, both terms do not occur anywhere an $Gr(Sk(\Gamma))$. Then we may replace $y(x_1)$ by the variable y_1 and $y(x_2)$ by the variable y_2 (since the only thing which matters in the propositional calculus is whether propositional atoms are equal or not). If $x_1 = x_2$, then we replace $y(x_1)$ by y_1 throughout the proof.

Up to now, the proof uses only specialization with respect to z and propositional rules, i.e. we can apply the deduction lemma and get a derivation of the formulas

$$\forall zB(x_1, y_1, z)\to\perp, \forall zB(x_2, y_2, z)\to\perp$$

from the remaining assumptions. By the inference rule for $\exists$, we can construct a derivation of the formula

$$\exists y_1\forall zB(x_1, y_1, z)\to\perp.$$

By modus ponens, we can construct a derivation of $\perp$ from the formula

$$\exists y_1\forall zB(x_1, y_1, z)$$

But this formula can be derived from A_1. Thus, we have exchanged in the assumptions the ground instances by A_1 itself.

Similarily, one can treat $A_2, \ldots, A_n$.

Lemma 3 *Let Σ be a set of formulas in the proposition logic, which is maximally consistent. Let A be a formula. Then $\Sigma \vdash A$ or $\Sigma \vdash \neg A$.*

Proof: Assume not $\Sigma \vdash A$. Because Σ is maximally consistent, $\Sigma \cup A$ is inconsistent, i,e, $\Sigma, A \vdash \bot$. By the deduction theorem 1, $\Sigma \vdash A \rightarrow \bot$. But $\neg A$ is a consequence of $A \rightarrow \bot$, hence $\Sigma \vdash \neg A$.

Lemma 4 (Herbrand's theorem) *Let Γ be a set of formulas in Skolem form, such that $Gr(\Gamma)$ is consistent. Then Γ has a Herbrand model.*

Proof: Let Σ be a maximally consistent set containing $Gr(\Gamma)$, which exists by virtue of Zorn's lemma.

For any atom P, either $\Sigma \vdash P$ or $\Sigma \vdash \neg P$, by lemma 3. We define a Herbrand interpretation $\mathcal{I}$ as follows. If P is an n-ary predicate symbol and $t_1, \ldots, t_n \in H(\Gamma)$, then $P(t_1, \ldots, t_n)$ holds iff the corresponding propositional atom is in Σ.

It is easy to check that $\mathcal{I}$ is a Herbrand model for Γ.

Proof of theorem 11: Let Γ be a consistent set of formulas. By theorem 12, we may assume that the formulas in Γ are in prenex normal form. By lemma 2, the set $Gr(Sk(\Gamma))$ is consistent. By lemma 4, $Sk(\Gamma)$ has a Herbrand model and is therefore satisfiable. But any interpretation which satisfies $Sk(\Gamma)$ also satisfies Γ.

Remark 4 Because the Herbrand universe is countable, we actually proved that any consistent set has a countable model (Löwenheim's theorem).

6 Mechanical Theorem Proving

As any formal theory, predicate logic is semi-decidable: one can give a procedure that finds a proof for any theorem, but this algorithm does not necessarily terminate for any formula. However, the obvious procedure (enumerate all sequences and check is we have a proof) is totally impracticable.

Theorem provers do a more directed search of the proof using Herbrand semantics. The idea is the following.

A formula A.

A proof of A if $\vdash A$. Otherwise, the procedure may not terminate.

1. Reduce to consistency: it suffices to show that $\neg A$ is inconsistent.
2. Reduce to Skolem form: it suffices to show that $B := Sk(Pr(\neg A))$ is inconsistent.
3. Find an inconsistent subset of $Gr(B)$.

Due to technical reasons, we only consider the case that A has no free variables. The general case can be reduced by $\exists$-quantification of all free variables.

Some technical definitions:

A *literal* is an atom or the negation of an atom.

A *clause* is a Skolem formula without free variables, such that the matrix is a disjunction of literals. Because $\vee$ is commutative and associative up to equivalence, we usually omit brackets and permute literals in clauses when convenient. We also regard $\perp$ as a clause (disjunction of zero literals). Note that a clause is uniquely determined by its matrix: all variables are universally quantified, and the order of quantifiers is irrelevant (exercise).

Lemma 5 *Let B be a Skolem formula without free variables. Then there exists a finite set of clauses $C = \{C_1, \ldots, C_m\}$, such that*

$$B \sim C_1 \wedge \ldots \wedge C_m.$$

Proof: The matrix of B can be brought to conjunctive normal form (i.e. a conjunction of clauses). Then we can use the equivalence

$$\forall x(A(x) \wedge B(x)) \sim \forall x A(x) \wedge \forall x B(x).$$

Now, we have reduced semi-decision of $\vdash A$ to semi-decision of inconsistency a finite set of clauses, by theorem 12, lemma **??**, and lemma 5 (the reduction is algorithmic because these theorems have been given a constructive proof). This is done by a procedure which can detect inconsistency of a given finite set of clauses, but which does not terminate if the given set is consistent: Robinson's resolution method.

In the following, we usually omit the quantifiers in clauses.

An *expression* is a term or a literal or a clause.

A *substitution* is a finite set of pairs $(x_i \leftarrow t_i)$, where the x_i is a variable and t_i is a term, and the variables are mutually distinct. If θ is a substitution and E is an expression, then $E\theta$ is the expression obtained by substituting t_i for x_i in each term. Composition of substitutions is defined in the following natural way: Let

$$\theta = \{x_1 \leftarrow t_1, \ldots, x_m \leftarrow t_m\},$$

$$\lambda = \{y_1 \leftarrow s_1, \ldots, y_n \leftarrow s_m\},$$

and suppose that $y_1, \ldots, y_r$ are the variables in λ that do not occur as variables in θ. Then

$$\theta\lambda = \{x_1 \leftarrow t_1\lambda, \ldots, x_m \leftarrow t_m\lambda, y_1 \leftarrow s_1, \ldots, y_r \leftarrow s_r\}.$$

Exercise 10 a) Show that the composition is associative.

b) Show that $E(\theta\lambda) = (E\theta)\lambda$ (i.e. substitutions act on expressions).

Let E_1, E_2 be expression. A *unifier* for E_1 and E_2 is a substitution θ with the property $E_1\theta = E_2\theta$. A *most general unifier* is a unifier σ such that for any unifier θ there exists a substitution λ with $\theta = \sigma\lambda$.

Theorem 13 (Unification) *If E_1 and E_2 are unifiable, then there exists a most general unifier.*

Proof: By induction over the number of variables in E_1 and E_2.

If there are no variable, the theorem is trivially true.

The *disagreement set* is the set of the two subexpressions beginning with the first symbol where E_1 and E_2 disagree. Any unifier must unify the disagreement set; this is possible only if one of the expressions is a variable x and the other is a term t which does not contain x. Let $\alpha := \{x \leftarrow t\}$, $E_1' := E_1\alpha$, $E_2' := E_2\alpha$. Then any unifier θ can be written as $\theta = \alpha\theta'$, where θ' is a unifier of E_1' and E_2'. If E_1' and E_2' are not unifiable, then E_1 and E_2 are not unifiable and the theorem is true. Otherwise, there exists a most general unifier σ' for E_1' and E_2' by the induction hypothesis, because E_1' and E_2' do not contain the variable x. Then, $\sigma := \alpha\sigma'$ is a most general unifier for E_1, E_2.

Exercise 11 Determine whether E_1 and E_2 are unifiable. If yes, obtain a most general unifier.

$$E_1 := P(x, y, z), E_2 := P(f(y, y), f(z, z), f(x, x)).$$

Let C_1 and C_2 be two clauses with no variables in common. Suppose that $L_1 = P_1$ and $L_2 = \neg P_2$ are two literals of opposite signs of C_1, C_2, respectively, such that P_1, P_2 are unifiable. Let θ be a most general unifier. Then the clause

$$(C_1 - L_1)\theta \vee (C_2 - L_2)\theta.$$

is called a *resolvent* of C_1 and C_2.

If C_1 and C_2 are arbitrary clauses, then a resolvent is obtained by first renaming variables in C_2 (note that this gives an equivalent clause) so that C_1 and C_2 have no common variables, and then apply resolution as above.

Lemma 6 *If C_3 is a resolvent of C_1 and C_2, then $C_1, C_2 \vdash C_3$.*

Let C be a clause, and let L_1 and L_2 be two literals in C which are unifiable. Let θ be a most general unifier. Then the clause $(C - L_1)\theta$ $(= (C - L_2)\theta$ up to a permutation of literals) is called a *factor* of C.

Lemma 7 *If C' is a factor of C, then $C \vdash C'$.*

Exercise 12 a) Compute all factors of

$$P(x) \vee P(f(y)) \vee Q(x, y) \vee Q(z, f(z)).$$

b) Compute all resolvents of the clauses

$$P(x) \vee Q(y) \vee R(f(z)), \neg P(z) \vee \neg Q(y) \vee \neg R(z).$$

Let A be a finite set of clauses. A *resolution refutation* of A is a finite sequence of clauses $C_1, \ldots C_n$, such that C_i is in A or a resolvent of C_j, C_k, where $j, k < i$, or a factor of C_j, where $j < i$, and $C_n = \bot$.

A *ground clause* is a clause without variables.

Lemma 8 *Let A be a set of ground clauses. Then A is inconsistent iff there exists a resolution refutation for A.*

Proof: Note that this is essentially a theorem on propositional logic, because we may replace any atom by a proposition letter.

Only if: obvious from lemma 6, lemma 7, and the definition of a refutation.

If: Any derivation of $\bot$ from A uses only finitely many formulas in A. Hence, it suffices to prove the statement for finite sets A. We proceed by induction on the number n of atoms in A.

If $n = 0$, then any clause in A is $\bot$, and A is inconsistent iff it is non-empty. But then, there exists a refutation with one step.

Let P be an atom in A. We define the following sets of clauses.

- A_1 is the set of clauses in A which contain P but not $\neg P$.
- A_1' is obtained from A_1 by deleting all literals P.
- A_2 is the set of clauses in A which contain $\neg P$ but not P.
- A_2' is obtained from A_2 by deleting all literals $\neg P$.
- $A_{12} := \{C_1 \vee C_2 \mid C_1 \in A_1', C_2 \in A_2'\}$.
- A_3 is the set of clauses in A which contain neither P nor $\neg P$.
- A_4 is the set of clauses in A which contain both P and $\neg P$.

Let $A' := A_{12} \cup A_3$. We will show that inconsistency of A implies inconsistency of A'. Assume, indirectly, that we have a model $\mathcal{I}$ for A'. Then $\mathcal{I} \models A_{12}$ and $\mathcal{I} \models A_3$. Because the conjunction of all clauses in A_{12} is equivalent to the disjunction of the conjunction of all formulas in A_1' and the conjunction of all formulas in A_2', we have either $\mathcal{I} \models A_1'$ or $\mathcal{I} \models A_2'$. Assume $\mathcal{I} \models A_1'$. Then we can extend $\mathcal{I}$ by assigning `False` to P. The result is a model for A_1, A_2, A_3, and A_4, hence a model for A. This contradicts the assumption that A is inconsistent.

Now, suppose that A is inconsistent. Any clause in A' can be derived from clauses in A by resolution or factor rule. By the induction hypothesis, there is a resolution refutation for A'. Hence there is a resolution refutation for A.

Lemma 9 (Lifting lemma) *Let C_1, C_2 be two clauses with ground instances C_1', C_2', respectively. Let C_3' be a resolvent of C_1', C_2'. Then there exists a resolvent C_3 of C_1, C_2, such that C_3' is a ground instance of C_3.*

Proof: W.l.o.g., we may assume the following.

- C_1 and C_2 have no variables in common.
- The resolved literals in C_1' and C_2' are the last ones. It is a positive in C_1' and negative in C_2'.
- The instantiating substitutions θ_1, θ_2 change only the variables occuring in C_1/C_2. (Especially, the variables in θ_1 and θ_2 are disjoint.)

We may write $C_1 = D_1 \vee P_1$ and $C_2 = D_2 \vee \neg P_2$ for suitable clauses D_1, D_2, and literals P_1, P_2. Let $\theta := \theta_1\theta_2 = \theta_2\theta_1$. Then $P_1\theta = P_2\theta =: P$ is the resolved atom in the resolution of C_1', C_2'. Hence, P_1 and P_2 are unifiable. Let C_3 be the resolvent obtained by resolving this pair of literals, using the most general unifier σ. Let λ be such that $\theta = \sigma\lambda$. Then $C_3\lambda = C_3'$.

Lemma 10 (Lifting lemma for the factor rule) *Let C be a clause with ground instances C'. Let C_1' be a factor of C'. Then there exists a factor C_1 of C, such that C_1' is a ground instance of C_1.*

Proof: Analogous as for the lifting lemma for the resolution rule.

Theorem 14 (Resolution theorem) *Let A be a finite set of clauses. Then A is inconsistent iff there exists a resolution refutation for A.*

Proof: Only if: obvious.

If: Suppose that A is inconsistent. Then A does not have a Herbrand model. By Herbrand's theorem (lemma 4), $Gr(A)$ is inconsistent. By lemma 8, there is a resolution refutation for $Gr(A)$. Using lemma 9 and lemma 10, this refutation can be lifted to a resolution refutation of A.

Note that this is a semantic (i.e. set-theoretic) proof for a metatheoretical statement, because we used Herbrand's theorem.

Starting from a given set of clauses, the resolution procedure starts producing systematically all possible resolutions and factors. If the set is inconsistent, the algorithm will stop and obtain a resolution refutation. Otherwise, it does not terminate.

A disadvantage of the resolution procedure is that it produces a lot of irrelevant clauses. Therefore, one uses various strategies that forbid applications of the resolution rule in particular situations (hyperresolution, lock resolution, linear resolution, set of support strategy). Any of these strategies is complete, but one cannot arbitrarily combine the strategies without losing completeness.

Remark 5 Writing $A \to B$ is $\neg A \vee B$, the modus ponens inference rule

$$A, \neg A \vee B \vdash B$$

appears as a special case of resolution. But one has to keep in mind an important difference between these two inference rules: while resolution can just resolve literals, modus ponens allows to "resolve" arbitrarily complex formulas (lemmata). Automatic theorem provers have no chance to use lemmata, because the lemma cannot be obtained from the given theorem alone, i.e. it "disappears" during the proof procedure. The computer does not know what to put as a lemma. This is an essential difference between human proofs and computer-generated proofs.

7 Predicate theories with axioms

Predicate logic is expressive enough to formalize various theories within it. When doing so, the function symbols symbolize functions from the informal theory, and the predicate symbols symbolize predicates from the informal theory. The number of function symbols and predicate symbols may be finite or infinite. Usually, one has additional axioms. The set of axioms may be infinite, but it must be decidable (this is needed to be able to check proofs).

One particular predicate is rather common: equality, symbolized by "=" and written in infix notation. It has the axioms

$$x = x,$$

$$x_1 = y_1 \to \ldots \to x_n = y_n \to f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n)$$

for any n-ary function symbol,

$$x_1 = y_1 \to \ldots \to x_n = y_n \to P(x_1, \ldots, x_n) \to P(y_1, \ldots, y_n)$$

for any n-ary predicate symbol.

Exercise 13 Prove symmetry and transitivity.

Remark 6 It is usual to regard "=" as belonging to the logics (rather than to the material part). Obviously, predicate logic with "=" is consistent.

Theorem 15 (Replacement theorem for Equality) *a) Let u_s be a term with a subterm s, and let u_t be the term obtained when s is replaced by t in u_s. Then*

$$\vdash s = t \to u_s = u_t.$$

b) Let A_s be a formula with a term s (not as a quantified variable), and let A_t be the formula obtained when s is replaced by t in A_s. Then

$$\vdash s = t \to (A_s \leftrightarrow A_t).$$

Proof: a) Easy by induction on function symbols in the context u_{-}.

b) If A_s is an atom, then the statement follows immediately from (a) and the equality axiom for the corresponding predicate symbol. The general case follows from the replacement theorem 7.

A sufficient criterion for consistency is the existence of a finitary model (then, consistency can be proved as in the proof of theorem 8). We have the following stronger set-theoretic result.

Theorem 16 *A predicate theory is consistent iff it has a model.*

Proof: Consistency means just that the set Γ of axioms is consistent.

If: Suppose that Γ is consistent. By Gödels completeness theorem for infinitely many formulas (theorem 11), Γ has a model.

Only if: There is no model for $\perp$.

Here is a rather simple example of a predicate theory. Its intended model is a universe with a single element.

- predicate symbols: "$=$".
- no function symbols.
- axioms:
$$x = y$$

Exercise 14 Give a formalization of a universe with two elements.

We say that a formal theory is *complete* iff it is consistent and there exists an algorithm that computes for any given formula A either a proof of A or a derivation $A \rightarrow \perp$. (If we adjoin a formula which is not already a theorem, then the theory is already inconsistent.) Any complete theory is decidable.

The theory above is complete. Its theorems are exactly those formulas which are true in the intended model.

Exercise 15 Prove the following theorem:

A theorem is complete iff, for any formula without free variables, we have $\vdash A$ or $\vdash \neg A$ (exclusive or).

There are several interesting predicate theories. Some will be discussed in the following: theory of groups, theory of real numbers, formal number theory.

An interesting theory which will not be discussed here is formal set theory. Its intended interpretation is a universe (not a set!) of sets. It allows to express the predicate "countable", and to prove that there exists a non-countable set. On the other hand, we know by Löwenheim's theorem (remark 4) that the theory has a countable model, with universe a countable set. From the semantic point of view, this is paradox.

7.1 The Theory of Groups

The intended interpretation is a group.

- predicate symbols: "$=$".
- function symbols: e (0-ary), "$\cdot$" (binary infix), "$\_^{-1}$" (unary postfix).
- axioms:

$$x\cdot(y\cdot z) = (x\cdot y)\cdot z,$$

$$x\cdot e = e\cdot x = e,$$

$$x^{-1}\cdot x = x\cdot x^{-1} = e.$$

As usually, we allow ourselves the following metatheoretical abbreviations: we omit brackets and we write integer exponents.

Exercise 16 Give a metatheoretical proof for the consistency of the theory of groups.

Exercise 17 Give a metatheoretical proof for the incompleteness of the theory of groups: there exists a non-formula which can be added as an axiom without losing consistency.

Group theory can be shown to be undecidable. We give a semantic proof for this fact, reducing the decision problem to the word problem.

The *free group* generated by a, b is the group of all *words* $x_1^{e_1} \ldots x_n^{e_n}$, where $e_i \in \mathbf{Z} - \{0\}$, $x_i \in \{a, b\}$, $x_i \neq x_{i+1}$.

The *word problem* is the following.

Given: A finite set of words $S = \{A_1, \ldots, A_r\}$ (called g-axioms);
A word B (called g-formula).
Decide: Is there a g-derivation of B from S?

For deriving g-formulas, the following inference rules are allowed:

- Composition of g-axioms or previously derived g-formulas.
- Inversion of a g-axiom or a previously derived g-formula.
- Conjugation of a g-axiom or of a previously derived g-formula by an arbitrary word.

For example, one can show that the g-formulas which can be derived from the axiom a are the words for which the sums of all exponents of b is zero.

Exercise 18 Can you give an easy description of the g-formulas derivable from the axiom $aba^{-1}b^{-1}$?

Theorem 17 *The word problem is undecidable.*

Proof: Well-known (see A. Turing: The word problem in semi-groups with cancellation. Ann. math. (52), 1950, pp 491–505). The proof idea is to reduce to the halting problem for Turing machines.

Lemma 11 *Let $S = \{A_1, \ldots, A_r\}$ be a set of g-axioms, and let B be a g-formula. Then the formula*

$$A_1 = e \to \ldots \to A_r = e \to B = e,$$

where a and b are considered as variables, is a theorem in the theorem of groups iff B is g-derivable from S.

Proof: If: Let G be the free group generated by a, b. Let H be the subgroup generated by S, Let N be its normalization (i.e. N is the set of all g-theorems). Let $Q := G/N$ be the quotient group. If the formula above is a theorem, then it is true under the interpretation Q, where the free variables a and b are assigned to the classes of the words a and b. Obviously, the subformulas $A_i = e$, $i = 1, \ldots, r$ are true under this interpretation. Hence, the subformula $B = e$ must also be true, hence $B \in N$.

Only if: Using the following easy provable inference rules

$$x = e, y = e \vdash x \cdot y = e,$$

$$x = e \vdash x^{-1} = e,$$

$$x = e \vdash y^{-1} \cdot x \cdot y = e,$$

one can lift any g-derivation of B to a derivation

$$A_1 = e, \ldots, A_r = e \vdash B = e.$$

Then the fact that the above formula is a theorem follows easily from the deduction theorem.

Theorem 18 *The theory of groups is undecidable.*

Proof: Suppose indirectly that we had an algorithm for deciding the theory of groups. By the lemma, we could use it to decide the word problem. But the word problem is undecidable.

Theorem 19 (Church) *Predicate logic is undecidable.*

Proof: Suppose indirectly that we had an algorithm $\mathcal{M}$ for deciding predicate logic. Then, we could use it to decide the theory of groups, as follows. Let A be a given formula in the theory of groups. The theory of groups has only finitely many axioms, which we now abbreviate by $A_1, \ldots, A_r$. Then A is a theorem of group theory iff the formula

$$A_1 \to \ldots \to A_r \to A$$

is a theorem of predicate logic. But this can be decided using the algorithm $\mathcal{M}$.

But the theory of groups is undecidable, a contradiction.

Exercise 19 Which of the following properties of formulas in first order theory is semi-decidable?

- provable
- not provable
- consistent
- inconsistent
- satisfiable
- not satisfiable
- satisfiable by a k-interpretation
- not satisfiable by a k-interpretation

7.2 Definability

Let $\mathcal{T}$ and $\mathcal{T}'$ be predicate theories. We say that $\mathcal{T}'$ is a *subtheory* of $\mathcal{T}$ iff

- any function symbol of $\mathcal{T}'$ is a function symbol of $\mathcal{T}$,
- any predicate symbol of $\mathcal{T}'$ is a predicate symbol of $\mathcal{T}$,
- any axiom of $\mathcal{T}'$ is a theorem of $\mathcal{T}$.

The subtheory $\mathcal{T}'$ is called a *full subtheory* iff any formula of $\mathcal{T}'$ which is a theorem in $\mathcal{T}$ is also a theorem in $\mathcal{T}'$.

Exercise 20 The theory of monoids is the subtheory of the theory of groups with predicate symbol "=" and function symbols $\cdot$ and e, and with all the axioms that do not contain $\_^{-1}$. Prove that the theory of monoids is not a full subtheory.

The subtheory $\mathcal{T}'$ is called an *equivalent subtheory* iff there is an algorithm that computes for any formula A of $\mathcal{T}$ a formula $F(A)$ in $\mathcal{T}'$, such that

1. $F(A)$ is A for any formula A in $\mathcal{T}'$,
2. $A{\sim}F(A)$ in $\mathcal{T}'$,
3. if A is an axiom of $\mathcal{T}$, then $F(A)$ is a theorem of $\mathcal{T}'$,
4. if B is an immediate consequence of $A_1, \ldots, A_n$ in $\mathcal{T}$, then $F(B)$ is a consequence of $F(A_1), \ldots, F(A_n)$ in $\mathcal{T}'$.

Theorem 20 *a) A subtheory of a consistent theory is consistent.*
b) A subtheory of a complete theory is complete.
c) A subtheory of a decidable theory is decidable.
d) A theory with a full consistent subtheory is consistent.
e) An equivalent subtheory is full.
f) A theory with a complete equivalent subtheory is complete.
g) A theory with a decidable equivalent subtheory is decidable.

Let $\mathcal{T}'$ be an equivalent subtheory of $\mathcal{T}$. Then the function/predicate symbols in $\mathcal{T} - \mathcal{T}'$ are said to be *eliminable* with respect to $\mathcal{T}$, or *definable* with respect to $\mathcal{T}'$.

Theorem 21 *An n-predicate symbol P is eliminable iff*

$$P(x_1, \ldots, x_n) \sim A(x_1, \ldots, x_n)$$

for some suitable formula A not containing P without free variables other than $x_1, \ldots, x_n$.

Vice versa, let $A(x_1, \ldots, x_n)$ be a formula without free variables other than $x_1, \ldots, x_n$. Then we can define a new n-ary predicate symbol P by adjoining

$$P(x_1, \ldots, x_n) \leftrightarrow A(x_1, \ldots, x_n)$$

as a new axiom N.

Proof: If P is eliminable, then we have a formula A equivalent to $P(x_1, \ldots, x_n)$ in some equivalent subtheory. By condition 2 in the definition of equivalent subtheories, we have $P(x_1, \ldots, x_n){\sim}A$. If A contains a free variable y distinct from $x_1, \ldots, x_n$, then one can also show $P(x_1, \ldots, x_n){\sim}\forall y A(x_1, \ldots, x_n, y)$. This shows that such an equivalence for P is necessary for eliminability.

Assume that we have such an equivalence for P. To compute F, we rewrite any $P(t_1, \ldots, t_n)$ by $A(x_1{\leftarrow}t_1, \ldots, x_n{\leftarrow}t_n)$. As the axioms of the theory $\mathcal{T}'$, we take the translations $F(A)$ for all material axioms A of $\mathcal{T}$.

We now have to show that F makes $\mathcal{T}'$ to an equivalent subtheory. Condition 1 is trivial, condition 2 follows from the replacement theorem and the substitution axiom $\forall x A(x) \rightarrow A(x \leftarrow t)$. For any axiom $F(B)$ of $\mathcal{T}'$, we have $F(B) \sim B$ and B is an axiom, therefore $\mathcal{T}'$ is indeed a subtheory. Condition 3 holds by definition for all material axioms. Moreover, it is easy to show that $F(B)$ is a logical axiom if B is a logical axiom. Condition 4 holds also, we can even say stronger that the function F preserves immediate consequences.

Now, start with a theory $\mathcal{T}'$ without P and a formula $A(x_1, \ldots, x_n)$, and define the theory $\mathcal{T}$ with P by adjoining the formula N as a new axiom. Then $P(x_1, \ldots, x_n) \sim A$ in $\mathcal{T}$, and hence we can construct an equivalent subtheory $\mathcal{T}''$ by the above construction. The axioms of $\mathcal{T}''$ are exactly the axioms of $\mathcal{T}'$ plus $F(N)$, which is $A \leftrightarrow A$. But this is obviously a theorem, hence it can be cancelled from the list of axioms. Thus, $\mathcal{T}'$ is an equivalent subtheory, hence P is definable.

As a shorthand for unique existence, we write $\exists! x A(x)$ for

$$\exists x A(x) \wedge (A(x) \rightarrow A(y) \rightarrow x = y).$$

(One can easily make the shorthand formal by introducing a new logical axiom.)

Theorem 22 *An n-ary function symbol is eliminable iff*

$$y = f(x_1, \ldots, x_n) \sim A(x_1, \ldots, x_n, y)$$

for some suitable formula A not containing f without free variables other than $x_1, \ldots, x_n, y$.

Vice versa, let $A(x_1, \ldots, x_n, y)$ be a formula without free variables other than $x_1, \ldots, x_n, y$. Assume that the formula

$$\exists! y A(x_1, \ldots, x_n, y)$$

is a theorem. Then we can define a new n-ary function symbol P by adjoining

$$A(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$$

as a new axiom.

Proof: The condition is clearly necessary, because we must have some way to rewrite the formula $y = f(x_1, \ldots, x_n)$, and a similar argument as above shows that we may assume that the translated formula $A(x_1, \ldots, x_n, y)$ contains no free variables other than $x_1, \ldots, x_n, y$.

Assuming that

$$y = f(x_1, \ldots, x_n) \sim A(x_1, \ldots, x_n, y),$$

we give an algorithm to translate a formula containing f into one not containing f. Namely, we rewrite any $P(s_1, \ldots, s_{k-1}, f(t_1, \ldots, t_n), s_{k+1}, \ldots, s_m)$ by

$$\exists y (P(s_1, \ldots, s_{k-1}, y, s_{k+1}, \ldots, s_m) \wedge A(t_1, \ldots, t_n, y)),$$

where y is a new variable. As the axioms of our subtheory $\mathcal{T}'$, we take the transforms of all material axioms plus the formula

$$\exists! y A(x_1, \ldots, x_n, y),$$

which we call N.

We now have to show that F makes the subtheory to an equivalent one. Condition 1 is trivial, condition 2 follows from the replacement theorem and the equivalence

$$B(f(t_1, \ldots, t_n)) \sim \exists y(B(y) \wedge A(t_1, \ldots, t_n, y)),$$

which is a consequence of the assumption $y = f(x_1, \ldots, x_n) \sim A(x_1, \ldots, x_n, y)$. If $F(B)$ is an axiom of $\mathcal{T}'$, then $B \sim F(B)$ and B is an axiom, hence $F(B)$ is a theorem of $\mathcal{T}$. Moreover, N is also an obvious theorem of $\mathcal{T}$, hence $\mathcal{T}'$ is indeed a subtheory. Condition 3 is fulfilled by definition for all material axioms.

Among the logical axioms, only the two substitution problems make problems. These two axioms, namely

$$B(h(f(t_1, \ldots, t_n))) \to \exists x B(x)$$

and

$$\forall x B(x) \to B(h(f(t_1, \ldots, t_n)))$$

rewrite to

$$\exists y(B(h(y)) \wedge A(t_1, \ldots, t_n, y)) \to \exists x B(x)$$

and

$$\forall x B(x) \to \exists y(B(h(y)) \wedge A(t_1, \ldots, t_n, y)),$$

respectively. The first formula is a theorem, and the second formula is a consequence from $\exists y A(x_1, \ldots, x_n, y)$, which is again a consequence of the assumption $y = f(x_1, \ldots, x_n) \sim A(x_1, \ldots, x_n, y)$.

All equality axioms in $\mathcal{T}$ are transformed to axioms except for one, namely

$$x_1 = y_1 \to \ldots \to x_n = y_n \to f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n).$$

It rewrites to

$$x_1 = y_1 \to \ldots \to x_n = y_n \to \exists u \exists v(A(x_1, \ldots, x_n, u) \wedge A(y_1, \ldots, y_n, u) \wedge u = v).$$

This can be derived from $\exists y A(x_1, \ldots, x_n, y)$, by the replacement theorem for equality (theorem 15).

Condition 4 is also clear: as in the above proof, one can even show that the function F preserves immediate consequence.

Now, start with a theory $\mathcal{T}'$ without f and a theorem N as above. Define the theory $\mathcal{T}$ with f by adjoining the formula

$$A(x_1, \ldots, x_n, f(x_1, \ldots, x_n))$$

as a new axiom, called N'. Then

$$y = f(x_1, \ldots, x_n) \sim A(x_1, \ldots, x_n, y)$$

in $\mathcal{T}$, and hence we can construct an equivalent subtheory $\mathcal{T}''$ by the above construction. The axioms of $\mathcal{T}''$ are exactly the axioms of $\mathcal{T}'$ plus $F(N')$ plus N. Both formulas are theorems of $\mathcal{T}'$, hence they can be cancelled from the list of axioms. Thus, $\mathcal{T}'$ is an equivalent subtheory, hence f is definable.

For instance, consider the constant e in the theory of monoids. We can prove $x = eleftrightarrow\forall y\ xy = yx = e$: The direction left to right is an axiom, and the direction from right to left is obvious from the same axiom when we specify $y = e$. The equivalent subtheory without e has the axiom of associativity and the axiom

$$\exists! x \forall y\ xy = yx = e.$$

Actually, one can prove uniqueness from existence, so it suffices to take

$$\exists x \forall y\ xy = yx = e$$

as axiom instead.

Exercise 21 Show that $\_^{-1}$ is eliminable in the theory of groups. Write down axioms for the subtheory.

8 Theory of Real Numbers

Let $\mathcal{F}$ be the theory of fields, with predicate symbol "=", function symbols $+, -, 0, \cdot, 1$, ring axioms, and the axiom

$$\neg x = 0 \to \exists y\ xy = 1.$$

The theory $\mathcal{OF}$ (ordered fields) is an extension with additional predicate symbol $<$ and additional axioms

$$\neg x < x, x < y \to y < z \to x < z,$$

$$x < y \vee y < x \vee x = y,$$

$$x < y \to x + z < y + z,$$

$$0 < z \to x < y \to xz < yz.$$

Obviously, one can define the predicates $>, \leq, \geq, \neq$.

By composition of the available function symbols, we get all polynomials (with integral coefficients). We "identify" polynomials which provably coincide for all arguments, e.g. $(x + x)\cdot x = x\cdot x + x\cdot x + 0\cdot x$. If a polynomial p can be written as $t_0 + \ldots + t_n x^n$, where $t_0, \ldots, t_n$ are terms not containing x, then we say that p has formal degree n (with respect to x). Formal degree is not unique, any polynomial having degree n has also degree $n+1$. (By definition, we say that the zero polynomial has any formal degree, including negative ones.) If $a_0, \ldots, a_n$ are variables, then the polynomial

$$\Gamma_n := a_0 + \ldots + a_n x^n$$

is called a *generic polynomial* of formal degree n.

The theory $\mathcal{R}$ of real numbers is an extension of $\mathcal{OF}$ by an infinite number of axioms, one for each natural number n:

$$\Gamma_n(x) < 0 \rightarrow \Gamma_n(y) > 0 \rightarrow \exists z(x < z < y \wedge \Gamma_n(z) = 0)$$

(called intermediate value axioms).

Exercise 22 Prove

$$x > 0 \rightarrow \exists y(y^2 = x).$$

Theorem 23 *Let* $\Gamma_n := a_0 + \ldots + a_n x^n$*,* n *odd. Then*

$$\vdash a_n \neq 0 \rightarrow \exists x \Gamma_n(x) = 0.$$

Proof: Assume $a_n \neq 0$. For $i = 0, \ldots, n-1$, we know $\exists b_i \; a_n b_i = a_i$. For these b_i,

$$\exists b(b > -nb_0 \wedge \ldots \wedge b > -nb_{n-1} \wedge b > 1),$$

since $<$ is a total order. For this b, we have $b^n > -nb_i x^i$, $i = 0, \ldots, n-1$. Therefore

$$b_0 + \ldots + b_{n-1}b^{n-1} + b^n > 0.$$

Similarily, we construct an a such that

$$b_0 + \ldots + b_{n-1}a^{n-1} + a^n < 0.$$

By the intermediate value axiom, the polynomial $b_0 + \ldots + b_{n-1}x^{n-1} + x^n$ has a zero c. But then, we also have

$$a_0 + \ldots + a_n c^n = 0.$$

Exercise 23 Let $\mathcal{R}'$ be the extension of $\mathcal{OF}$ by adjoining the $\mathcal{R}$-theorems in the exercise above. Give a semantic proof of the intermediate value theorem, using the following algebraic fact:

> Let R be an ordered field, such that any positive element has a square root, and any polynomial of odd degree has a root. Then $R(\sqrt{-1})$ is algebraically closed.
>
> (Jacobson, Basic Algebra)

[Hint: if $R(\sqrt{-1})$ is algebraically closed, then any polynomial of degree ≥ 3 is reducible.]

We cannot define an inverse, because its value on 0 is not defined. However, we can define a new function which allows us to divide if the argument is not zero. The following function has the additional advantage that it also allows case distinction between positive and negative arguments.

$$i(x) := \begin{cases} x^{-1} & \text{if } x > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Exercise 24 *Prove that the function above is definable.*

By composition all polynomial functions and i, we get all *piecewise rational functions*. For instance,

$$sgn(x) = i(x)\cdot x - i(-x)\cdot x$$

is piecewise rational.

Let $g := t_0 + \ldots + t_n x^n$ be a polynomial with formal degree n. Then we define the *leading coefficient* as

$$\text{lcoeff}(g) := \begin{cases} t_n, & \text{if } t_1 \neq 0, \\ \text{else} t_{n-1}, & \text{if } t_{n-1} \neq 0, \\ \ldots & \\ \text{else} t_0, & \text{if } t_0 \neq 0, \\ \text{else} 1. & \end{cases}$$

This is a piecewise linear function obtained by case distinction from $t_0, \ldots, t_n$ (which is in general not identical to any of these terms). The *degree* is a piecewise constant function (obtained in a similar way) with values in the finite set $\{n, n-1, \ldots, 0, -\infty\}$. (Formally, the degree cannot be regarded as a function symbol; rather, we have to write the informal statement degree$(g) = m$ as a conjunction of equalities and inequations.)

Given two polynomials $g = t_0 + \ldots + t_n x^n$ and $h = s_0 + \ldots + s_m x^m$, we define the *quotient* and *remainder* as usual (these are polynomials with piecewise rational coefficients of formal degree n and $m-1$, respectively). In the cases where h vanishes identically, we define that quotient and remainder both be zero.

The *gcd* of g and h is a polynomial which divides g and h (with zero remainder) and with the property that any polynomial dividing f and g also divides the gcd. It can be computed by the Euclidean algorithm, i.e. by repeated application (at most m times) of taking remainder. It is a polynomial of formal degree $\max(m, n)$ with piecewise rational coefficients. (The maximum is necessary because there might be cases where g vanishes identically and the gcd has to coincide with h.)

A polynomial g is called *squarefree* iff $g{\equiv}0$ or if there exists no polynomial h of positive degree such that h^2 divides g. This is equivalent to saying that g and g' are relatively prime. (Two polynomials are relatively prime iff their gcd is a constant.) This property can be expressed as formula in the coefficients. For any given polynomial $g = t_0 + \ldots + t_n x^n$, we can compute its *squarefree factorization*

$$g = q_1 q_2^2 \ldots q_n^n,$$

$$q_i = \frac{\mathrm{gcd}(g, g', \ldots, g^{(i-1)}) {\cdot} \mathrm{gcd}(g, g', \ldots, g^{(i+1)})}{\mathrm{gcd}(g, g', \ldots, g^{(i)})^2},$$

by a finite number of case distinctions.

Exercise 25 Compute the squarefree factorization of Γ_2.

For any given finite set of squarefree polynomials $g_1, \ldots, g_m$, we can compute a *factorization into relatively prime*

$$g_i = q_1^{e_{i,1}} \ldots q_n^{e_{i,n}}, i = 1, \ldots, m,$$

using the gcd and quotient operations. By definition, the zero polynomial is relatively prime to any polynomial.

Now, there are a lot of well-known theorems for polynomials that can be stated and proved - for fixed formal degree - as theorems in $\mathcal{R}$. We will use them in the following without further notification.

Theorem 24 *Any nonzero polynomial of degree n has at most n roots.*

Theorem 25 *Any root of $a_0 + \ldots + a_n x^n + x^{n+1}$ lies in the interval $[-b, b]$, where $b := \max(1, |a_0| + \ldots + |a_n|)$.*

Theorem 26 *Let g be a squarefree polynomial. Let x be a root of g. Then there exist two intervals $I_1 = (x_1, x)$ and $I_2 = (x, x_2)$, such that ($g(I_1) < 0$ and $g(I_2) > 0$ and $g'(I_1 \cup I_2) > 0$) or ($g(I_1) > 0$ and $g(I_2) < 0$ and $g'(I_1 \cup I_2) < 0$).*

Theorem 27 (Rolle) *If $a < b$ and $g(a) = g(b) = 0$, then there exists $c {\in} (a, b)$ such that $g'(c) = 0$.*

Exercise 26 Write down formulas expressing the above statements.

If g and h are polynomials, then we define its *Sturm sequence* by $p_0 := g$, $p_1 := h$, $p_{i+2} := -\mathrm{remainder}(p_i, p_{i+1})$.

Let $\vec{t} := (t_0, \ldots, t_n)$ be a sequence of terms. The *number of sign variations* $V(\vec{t})$ is the number of pairs of indices $i < j$, such that $t_i {\cdot} t_j < 0$ and $t_{i+1} = \ldots t_{j-1} = 0$. The statement $V(\vec{t}) = k$ can be formally written as a disjunction of conjunctions of equalities and inequalities.

Theorem 28 (Sturm) *Let g be a squarefree polynomial. Let h be a polynomial relatively prime to g. Let $p_0, \ldots, p_m$ be the Sturm sequence of $g, g'h$. Let a, b be two terms not containing x. Let V_a be the number of sign variations of $(p_0(a), \ldots, p_m(a))$. Let V_b be the number of sign variations of $(p_0(b), \ldots, p_m(b))$. Then the following "formula" is a theorem:*

'If $g(a) \neq 0$ and $g(b) \neq 0$ and $a < b$, then the number of roots of g such that $h(x) > 0$ minus the number of roots of g such that $h(x) < 0$ is equal to $V(a) - V(b)$.'

Proof: Assume $g(a) \neq 0$ and $g(b) \neq 0$. Because g and $g'h$ are relatively prime, the Sturm sequence contains some constant polynomial p_r (the gcd), followed by zero polynomials. Also, we know that two consecutive polynomials do not have a common zero.

By definition of the Sturm sequence, $p_{i+2} + p_i = q \cdot p_{i+1}$. If $p_{i+1}(x) = 0$ for some x, then p_i and p_{i+2} have opposite signs at x. This implies that the number of sign variations of the sequence p_i, p_{i+1}, p_{i+2} is constant in some interval containing x as an inner point.

If x is a zero of g, then $g \cdot g'$ is negative in some interval $I_1 = (x_1, x)$ and positive in some interval $I_2 = (x, x_2)$. Therefore, the number of sign variations of the sequence p_0, p_1 in I_1 is one more or one less than the number of sign variations in I_2, depending on the sign of h at x.

The number of roots of any of the polynomials $p_0, \ldots, p_r$ is finite. We decompose the interval $[a, b]$ into subintervals by these roots $a = a_0 < a_1 < \ldots < a_n = b$ (taking as much new variables $a_1, \ldots, a_{n-1}$ as needed). By the intermediate value axioms, the number of sign variations is constant in any of the open subintervals.

Collecting all the sign changes in the subintervals, we get the theorem.

Remark 7 Sturm stated this theorem for $h = 1$, and Tarski made the above generalization. Sturm's theorem allows to calculate the number of real roots of a polynomial in some interval.

Exercise 27 Determine the number of roots of $x^3 - 7x - 7$ in the interval $(-2, -1)$.

Exercise 28 Determine the number of roots of $x^4 + 12x^2 + 5x - 9$.

Theorem 29 (quantifier elimination) *Let $A(x)$ be a formula in $\mathcal{R}$ without quantifiers. Then there exists a formula B without quantifiers, such that $\exists x A(x) \sim B$.*

Proof: Write A as a disjunction of conjunctions of atoms of the form $g(x) = 0$ or $g(x) > 0$, where g is a polynomial. By the equivalence

$$* : \exists x (A_1(x) \vee A_2(x)) \sim \exists x A_1(x) \vee \exists x A_2(x),$$

we may assume that A is a conjunction of atoms, i.e. we have to eliminate the quanitifier in

$$\exists x(g_1(x) = 0 \wedge \ldots \wedge g_n(x) = 0 \wedge g_{n+1}(x) > 0 \wedge \ldots \wedge g_{n+r}(x) > 0).$$

We compute squarefree factorization of each g_i and do the factorization into relatively prime polynomials:

$$g_i = q_1^{e_{i,1}} \ldots q_m^{e_{i,m}}, i = 1, \ldots, n+r$$

for some squarefree relatively prime polynomials $q_1, \ldots, q_m$ with piecewise rational coefficients. The next step is to reduce again to ordinary polynomials, by clearing denominators and extracting the case distinctions using the formula

$$\exists x(C \wedge A(x)) \sim C \wedge \exists x A(x).$$

The next step is removing all zero polynomials (maybe using more case distinctions) in $q_1, \ldots, q_n$. In the case $q_i \equiv 0$, we do the following: If q_i is a factor of some equation, then we can cancel the equation, and if q_i is a factor of some inequality, then we may return $\perp$.

The next step is to simplify the equational part. Obviously, it can be replaced by one equation which is the vanishing of the product of all common factors of $g_1, \ldots, g_n$. This one equation is rewritten as a disjunction of equations $q_i(x) = 0$, the formula is expanded by the distributive law, and then we use the rule $(*)$ to reduce to our standard form. Thus, we have to eliminate the quantifier in

$$\exists x(q_0(x) = 0 \wedge g_{n+1}(x) > 0 \wedge \ldots \wedge g_{n+r}(x) > 0).$$

If q_0 appears as a factor in the remaining g's, then we return $\perp$. Hence we may assume that g_i is relatively prime to q_0 for $i = n+1, \ldots, n+r$. We can also replace any odd exponent of some factor by one and discard all factors with even exponents.

Let S be a subset of $\{1, \ldots, n\}$. By Sturm's theorem, we may compute (making further case distinctions) u_S, which is defined as the number of roots of q_0 such that $g > 0$ minus the number of roots such that $g < 0$, where $g = \prod_{i \in S} q_i$. If v_S is the number of roots of q_0 such that $q_i < 0$ iff $i \in S$, then we have $\vec{u} = H\vec{v}$ for some matrix H with entries in $\{-1, +1\}$. One can show that this matrix is invertible (see exercise below), hence we can compute the numbers v_S. Then we can also compute the number of roots such that $g_i > 0$ for $i = n+1, \ldots, n+r$. If this number is positive, then we return $\neg\perp$, if it is zero, we return $\perp$.

It remains to treat the case where we do not have equalities. This case can be reduced to the other one by adding one equality $q = 0$, where q is the derivative of the product of the q_i times $x^2 - b^2$, where b is a bound for the roots of the q_i. Any interval bounded by zeroes of q_i's contains at least one zero of q by Rolle's theorem. Since the signs are constant in each such interval, it suffices to check the formula for all zeroes of q.

Exercise 29 Compute the inverse of the matrix H above.

Note that we can also eliminate $\forall$-quantifiers, because $\forall x A(x) \sim \neg\exists x \neg A(x)$. If A is an arbitrary formula, we can construct a formula A' without quantifiers, such that $A \sim A'$, by eliminating quantifiers inside out.

As an application, we give a useful "geometric" classification of formulas.

Theorem 30 *Let $A(x, y_1, \ldots, y_n)$ be a formula with a free variable $x, y_1, \ldots, y_n$. Then there exist piecewise rational functions $a_1, \ldots, a_m$ in $y_1, \ldots, y_n$, such that A is equivalent to a formula built from atoms $x < a_i$. In other words: any solution set of a formula-constraint is a finite union of intervals.*

Proof: We may assume that A is quantifier-free, i.e. built from atoms of the kind $P(x, y_1, \ldots, y_n) < Q(x, y_1, \ldots, y_n)$. By bringing Q to the left hand side, we may assume that Q is zero. By the intermediate value theorem and by finiteness of roots, the solution set of $P(x) < 0$ is a finite union of intervals.

Exercise 30 Prove that the formulas

$$\exists x \forall y (A(y) \to y \leq x) \to \exists! z (\forall y (A(y) \to y \leq z) \wedge \forall x (\forall y (A(y) \to y \leq x) \to x \leq z))$$

(existence of lowest upper bound) are theorems of $\mathcal{R}$.

Theorem 31 *$\mathcal{R}$ is decidable.*

Proof: To decide provability of closed formulas, it suffices to decide formulas without variables (neither free nor bound). But this is easy.

From the meta-theoretic point of view, the above proof has a gap: we do not know yet whether $\mathcal{R}$ is consistent, since we do not know a finitary model. However, the decision algorithm allows to give a finitary proof for consistency, too.

Theorem 32 *$\mathcal{R}$ is consistent.*

Proof: Fix an order of variables. For any formula A, compute the output $v(A)$ of the decision method applied to $\forall(x_1, \ldots, x_n)A$, where $x_1, \ldots, x_n$ are the free variables of A in ascending order. The consistency proof is accomplished by showing the following.

- Any logical axiom has value `True`.
- Any material axiom has value `True`.
- If $A \to B$ and A have value `True`, then B has value `True`.

- If B is a consequence of A by one of the other logical inference rules, and A has value `True`, then B has value `True`.

It follows that any theorem has value `True`. Since $\perp$ has value `False`, $\perp$ is not a theorem.

The given proof for decidability is algorithmic, but the algorithm which is implicitly contained in the proof is useless in practice: even for very small examples, the computation time for decision is too large to be awaited.

A big improvement has been achieved by the decision algorithm of G. Collins, which is double exponential in the number of variables but polynomial in the input when the number of variables is kept fixed. It is also known that the double exponential growth in the number of variables cannot be improved in general. In practice, Collins' method can be used to solve formulas with up to 6 variables.

A refinement of the algorithm given here (the Ben-Or-Kozen-Reif method) has quite good asymptotic complexity and has the advantage (in comparison with Collins' method) that it takes good advantage from parallelization.

For particular classes of formulas, it is possible to give faster algorithms (e.g. Grigorev's decision method for prenex formulas with $\exists$-quantifiers only, which is only single exponential).

The quantifier elimination problem can also be considered as "decision problem with parameters". Collins also gave a quantifier elimination algorithm with a similar complexity. Other quantifier elimination algorithms have been given by Hong, by Ben-Or and Kozen and Reif, and by Renegar. Since this is an important problem to which many other problems can be reduced, it is still a very active field of research.

9 Formal Geometry

Usually, euclidean geometry is formalized using two types, points and lines (or points, lines, and planes). In our formalization, we follow Tarski, who observed that the type "lines" is eliminable.

In addition to equality, we have two predicates. The ternary predicate $Bxyz$ has the following semantic interpretation: the three points x, y, z are collinear and y is between x and z (or coincides with x or z). The 4-ary predicate $xy{\cong}zw$ has the following semantic interpretation: the distance between x and y is equal to the distance of z and w. We define a new ternary predicate by

$$Cabc :\leftrightarrow Babc \vee Bbca \vee Bcab,$$

which has the interpretation that a, b, c are collinear.

We have 10 axioms and one axiom scheme (which gives infinitely many axioms).

$$ab \cong ba,$$

$$ab \cong pq \to ab \cong rs \to pq \cong rs,$$

$$ab \cong cc \to a = b,$$

$$\exists x(Bqax \wedge ax \cong bc),$$

$$a{\neq}b \to Babc \to Ba'b'c' \to ab{\cong}a'b' \to bc{\cong}b'c' \to ad{\cong}a'd' \to bd{\cong}b'd' \to cd{\cong}c'd',$$

$$Baba \to a = b,$$

$$Bapc \to Bbqc \to \exists x(Bbxp \wedge Bqxa)(\text{Pasch's axiom}),$$

$$\exists a\exists b\exists c\neg Cabc(\text{lower dimension axiom}),$$

$$p{\neq}q \to ap \cong aq \to bp \cong bq \to cp \cong cq \to Cabc(\text{upper dimension axiom}),$$

$$Badt \to Bbdc \to a{\neq}d \to \exists x\exists y(Babx \wedge Bacy \wedge Bxty)$$

(the last is equivalent to Euclid's axiom of parallels). The axiom scheme is

$$\exists a\forall x\forall y(A(x){\to}B(y){\to}Baxy) \to \exists b\forall x\forall y(A(x){\to}B(y){\to}Bxby).$$

This is a geometric formulation of Dedekind's axiom of cuts.

Here are some easy theorems.

Theorem 33 *The following formulas are theorems.*

$$ab \cong ab,$$

$$ab \cong cd \to cd \cong ab,$$

$$ab \cong cd \to cd \cong ef \to ab \cong ef,$$

$$aa \cong bb,$$

$$Babc \to Ba'b'c' \to ab{\cong}a'b' \to bc{\cong}b'c' \to ac{\cong}a'c',$$

$$q{\neq}a \to \exists!x(Bqax{\wedge}ax{\cong}bc),$$

$$Babb,$$

$$Babc \to Bcba,$$

$$Babc \to Bbac \to a = b,$$

$$\exists x\exists y\ x{\neq}y.$$

Proof: We prove the first, and the eigth assuming the seventh.

$$ba \cong ab \to ba \cong ab \to ab \cong ab \text{ axiom 2}$$

$$ba \cong ab \text{ axiom 1}$$

$$ba \cong ab \to ab \cong ab \text{ modus ponens from 1 and 2}$$

$$ab \cong ab \text{ modus ponens from 3 and 2}$$

To show the eigth, we assume $Babc$ and $Bbcc$ (by the seventh) and derive Bca.

$$Babc \to Bbcc \to \exists x(Bbxb \wedge Bcxa) \text{ axiom 7}$$

$$\exists x(Bbxb \wedge Bcxa) \text{ modus ponens two times}$$

By axiom 6, such an x is unique and it must be b, hence $Bcba$.

Exercise 31 Prove three of the above formal theorems (not the first and the eighth).

Here is a theorem which is geometrically more interesting.

Theorem 34 *The following formula is a theorem.*

$$Babc \to Bdbe \to ab{\cong}db \to bc{\cong}be \to ae{\cong}cd$$

Proof: We distinguish the cases $a{\neq}b$ and $a = b$.

Assume $a = b$. Then $d = b$ by the assumption $ab{\cong}db$ and axiom 3 (using the symmetry facts on $\cong$ which have been proved in the above exercise). Then the assumption $bc{\cong}be$ implies the conclusion $ae{\cong}cd$.

Now, assume $a{\neq}b$. By the specialization $a' = d$, $b' = b$, $c' = e$, $d' = a$, from axiom 5, we obtain

$$a{\neq}b \to Babc \to Bdbe \to ab{\cong}db \to bc{\cong}be \to ad{\cong}da \to bd{\cong}ba \to cd{\cong}ea.$$

All assumptions of the axioms are fulfilled in our case, hence also the conclusion $cd{\cong}ea$, which implies $ae{\cong}cd$.

We define a new binary function s (point reflection) by

$$Bbas(a,b) \wedge as(a,b){\cong}ab.$$

One can show that $s(a, \_)$ is an involutory map preserving distance and betweenness. Then, we define a new ternary predicate (right angle) by

$$Rabc :\leftrightarrow ac \cong s(b,a)c.$$

One can show that $\exists! x(Baxb \wedge ax{\cong}xb)$. We use this to define a binary function m (midpoint) by

$$Bam(a,b)b \wedge am(a,b){\cong}m(a,b)b.$$

We define a 4-ary predicate (parallel lines) by

$$ab||cd :\leftrightarrow \exists(x,y)(Rbax \wedge Raby \wedge Ccxd \wedge Ccyd \wedge ax{\cong}by).$$

Exercise 32 Prove

$$Babc \to Bade \to ab{\cong}ad \to bc{\cong}de \to be{\cong}cd$$

Rene Descartes had the idea to reduce formulas of geometry to formulas of $\mathcal{R}$, i.e. he gave an algorithm with the following specification.

Input: A formula A of geometry.

Output: A formula $B = F(A)$ of $\mathcal{R}$, which is a theorem iff A is a theorem.

To describe Descartes' algorithm, we describe how to deal with the formulas $a = b$, $Babc$, and $ab{\cong}cd$. These will be translated into certain formulas $E(a_1, a_2, b_1, b_2)$, $B(a_1, a_2, b_1, b_2, c_1, c_2)$, and $D(a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2)$. The logical connectives $\wedge, \vee, neg, \to, \perp$ are tranlated 1:1. The quantifiers are doubled, i.e. $\exists x$ translates to $\exists x_1 \exists x_2$, and $\forall x$ translates to $\forall x_1 \forall x_2$.

$$E \text{ is } a_1 = b_1 \wedge a_2 = b_2,$$

$$B \text{ is } a_1(b_2-c_2)+b_1(c_2-a_2)+c_1(a_2-b_2) = 0 \wedge (b_2-c_2)(a_2-b_2) \geq 0 \wedge (b_1-c_1)(a_1-b_1) \geq 0,$$

$$D \text{ is } (a_1 - b_1)^2 + (a_2 - b_2)^2 = (c_1 - d_1)^2 + (c_2 - d_2)^2.$$

Exercise 33 Give a translation of the predicates R and $||$ (without formal proof).

Theorem 35 *If A is a theorem of geometry, then $F(A)$ is a theorem of $\mathcal{R}$.*

Proof: It suffices to prove the statement for all material axioms, because we did not do any changes in the logical structur (except for the duplication of quantifiers, but here all axioms/inference rules carry over).

The first 10 axioms can be checked automatically, this is cumbersome but easy.

Now, let $A(x)$ and $B(y)$ be formulas. Let D be the axiom arising from A and B by the Dedekind scheme. Let $A(x_1, x_2)$ and $B(x_1, x_2)$ be the real translations of A and B. Using some case distinction, we can derive $F(D)$ from the formula

$$(A'(x) \to B'(y) \to x \leq y) \to \exists b (A'(x) \to B'(y) \to x \leq b \leq y),$$

which follows from theorem 30.

Theorem 36 *Formal geometry is consistent.*

Proof: Obvious from the consistency of $\mathcal{R}$.

To show that Descartes' algorithm is correct, we have to show that no non-theorem translates to a theorem. To do this we need to translate proofs in **R** into geometric proofs.

The theory $\mathcal{C}$ has all symbols and axioms from formal geometry, three new constant symbols o, u, v and the new axioms

$$o \neq u,$$

$$ou \cong ov,$$

$$Ruov.$$

Lemma 12 *Formal geometry is a full subtheory of* $\mathcal{C}$.

Proof: Let $C(o,u,v)$ be the conjunction of the three axioms above. Let A be a formula of geometry (i.e. not containing o,u,v) which is a theorem in $\mathcal{C}$. Then $C(o,u,v) \to A$ is a theorem of geometry, where o,u,v are considered as variables. By the third inference rule, $\exists(o,u,v)C(o,u,v) \to A$ is a theorem of geometry. We can also show that $\exists(o,u,v)C(o,u,v)$ is a theorem of geometry. By modus ponens, A is a theorem of geometry.

We extend Descartes' translation algorithm to the theory $\mathcal{C}$ by translating the constants o,u,v to the pairs of constants $(0,0),(1,0),(0,1)$ (note that we have to translate to the pairs because the formulas B and C have double "arity"). It is trivial to check that the three new axioms are theorems of $\mathcal{R}$.

We want to use the points o,u,v as base points for a coordinate system. As possible values of the coordinate functions, we do not have numbers available, just points. But we can use one of the coordinate axis as "the line of numbers".

We define unary functions c_1, h, c_2 by

$$Couc_1(a) \wedge Roc_1(a)a,$$

$$Covh(a) \wedge Roh(a)a,$$

$$Couc_2(a) \wedge c_2(a)h(a)||uv.$$

(Of course, one has to prove unique existence. The function h was just just auxiliary to define c_2.)

For doing some arithmetic on our "line of numbers", we define functions new functions. Their interpretation is that the sum/product of the first coordinates of the arguments is the first coordinate of the function value.

$$a +_g b := s(o, m(a,b)),$$

$$-_g a := s(a, o),$$

$$c = a \times_g b :\leftrightarrow \exists x(Couc \wedge Covx \wedge xc_1(a)||uv \wedge xc||vc_1(b)).$$

Moreover, we introduce a new order relation on points on the x-axis by

$$a <_g b :\leftrightarrow \neg Bc_1(a -_g b)ou.$$

Now, we are ready to describe a translation algorithm G from formulas of $\mathcal{R}$ into formulas of $\mathcal{C}$. First, we arrange the variables such that any variable either has an index 1 or 2. Then, we follow the following steps recursively.

1. Replace $+$, $-$, $\cdot$, 0, 1, $<$, by $+_g$, $-_g$, $\times_g$, o, u, $<_g$.
2. Replace each indexed variable x_1/x_2 by $c_1(x)/c_2(x)$.
3. If x_1 is an indexed variable not in the scope of a quantifier Qx_2, replace $\forall x_1 A(x_1)$ by
$$\forall(x, x_1)(c_1(x) = x_1 \rightarrow A(c_1(x)))$$
and $\exists x_1 A(x_1)$ by
$$\exists(x, x_1)(c_1(x) = x_1 \wedge A(c_1(x))).$$
Likewise for x_2.
4. If x_1 is an indexed variable in the scope of a quantifier Qx_2, replace $\forall x_1 A(x_1)$ by
$$\forall x_1(c_1(x) = x_1 \rightarrow A(c_1(x)))$$
and $\exists x_1 A(x_1)$ by
$$\exists x_1(c_1(x) = x_1 \wedge A(c_1(x))).$$
Likewise for x_2.

Lemma 13 *If A is a theorem of $\mathcal{R}$, then $G(A)$ is a theorem of $\mathcal{C}$.*

Proof: If A is a logical axiom coming from a propositional axiom scheme, then $G(A)$ is also a logical axiom coming from the same scheme.

Suppose that B is a logical axiom coming from the scheme $A(t) \rightarrow \exists x_1 A(x_1)$, where x_2 is not bound in A. Then $G(A)$ is

$$A_g(t_g) \rightarrow \exists(x, x_1)(c_1(x) = x_1 \wedge A_g(c_1(x))),$$

where A_g is an abbreviation for $G(A)$, and t_g is obtained from t by replacing the function symbols and variables as described. By recursion on subterms, one can show $\vdash Cous_g$ for all terms which come from translation (involving $+_g$ and so on). Moreover, the formula

$$Coux_1 \rightarrow \exists x \; c_1(x) = x_1$$

is a theorem. Using this, it follows that $G(B)$ is a theorem. If x_2 is bound in A, then one has to take care of the inner quantifier also, because the algorithm has a different effect on the inner quantifier.

Similarily, the remaining axiom scheme $\forall x A(x) \rightarrow A(t)$ can be treated.

If C is a modus ponens from A and B, then $G(C)$ is a modus ponens from $G(A)$ and $G(B)$.

Suppose that, in $\mathcal{R}$, we infer $C \rightarrow \forall x_1 A(x_1)$ from $C \rightarrow A(x_1)$. The transform of the assumption is $C_g \rightarrow A_g(c_1(x))$. From this and the theorem $\forall x_1 \exists x\ c_1(x) = x_1$, we can derive

$$C_g \rightarrow \forall (c_1(x) = x_1 \rightarrow A_g(c_1(x))).$$

Similarily, one can show that the third inference rule is produces only theorems (supposed that the assumption is a theorem).

Now, we have to translate the material axioms and show that they are theorems. This is somewhat cumbersome, but it works out. For the intermediate value axioms, it is more convenient to prove them from the remaining axiom and the Dedekind scheme

$$(A'(x) \rightarrow B'(y) \rightarrow x \leq y) \rightarrow \exists b (A'(x) \rightarrow B'(y) \rightarrow x \leq b \leq y),$$

and then prove the transform of the Dedekind scheme.

Lemma 14 *Let A be a formula of $\mathcal{C}$. Then $G(F(A)) \sim A$.*

Proof: Let A be the formula $a = b$. Then $F(A)$ is $a_1 = b_1 \wedge a_2 = b_2$, and $G(F(A))$ is

$$c_1(a) = c_1(b) \wedge c_2(a) = c_2(b).$$

One can show that this is equivalent to $a = b$.

If A is $Babc$ or $ab \cong cd$, then the proof is similar, although more complicated.

If A is a disjunction, conjunction, implication, or negation of subformulas, then the statement follows immediately (assuming that the statement holds for subformulas by induction).

Let $A = \forall x B(x)$, and assume that the statement is true for B. The real transform of B is $B_r(x_1, x_2)$, and the transform back is some formula $B_{rg}(c_1(x), c_2(x))$ which is equivalent to $B(x)$. Then, $G(F(A))$ is the formula

$$\forall (x, x_1, x_2)(c_1(x) = x_1 \rightarrow c_2(x) = x_2 \rightarrow B_{rg}(c_1(x), c_2(x))),$$

which is equivalent to $\forall x B_{rg}(c_1(x), c_2(x))$, which is equivalent to $\forall x B(x)$ by assumption. The case $A = \exists x B(x)$ is similar.

Theorem 37 *Formal geometry is complete (hence decidable).*

Proof: Let A be a closed geometric formula. Since $\mathcal{R}$ is complete, either $F(A)$ or $\neg F(A) = F(\neg A)$ is a theorem. In the first case, $G(F(A))$ is a theorem, hence A is a theorem. In the second case, $G(\neg F(A)) = \neg G(F(A))$ is a theorem, hence $\neg A$ is a theorem.

References

[1] Ben-Or, M., Kozen, D., and Reif, J. H. The complexity of elementary algebra and geometry. *J. Comput. System Sci. 32*, 2 (1986), 251–264.

[2] Chang, B., and Lee, C. *Symbolic Logic and Mechanical Theorem Proving.* Academic Press, 1973.

[3] Collins, G. E. Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. In *Lecture Notes In Computer Science* (1975), Springer, pp. 134–183. Vol. 33.

[4] Dedekind, R. *Stetigkeit und irrationale Zahlen.* Vieweg, 1872.

[5] Ebbinghaus, H., Flum, J., and Thomas, W. *Mathematical Logic.* Springer, 1983.

[6] H. Schwabhäuser and W. Szmielev and A. Tarski. *Metamathematische Methoden in der Geometrie.* Springer, 1983.

[7] Heyting, A. *Intuitionism.* North-Holland, 1966.

[8] Hilbert, D. *Grundlagen der Geometrie.* Teubner, 1977.

[9] Hilbert, D., and Bernays, P. *Grundlagen der Mathematik I und II.* Springer, 1970.

[10] Jebelean, T. Automatic theorem proving 1 and 2. Lectures at the university of Linz, winter 1998/99 and summer 1999.

[11] Jones, A., and Matijasevic, Y. Proof of the recursive unsolvability of hilbert's 10th problem. *Am. Math. Monthly 98* (1991), 689–709.

[12] Kleene, S. *Introduction to Metamathematics.* North-Holland, 1967.

[13] Manin, Y. *A course in mathematical logic.* Springer, 1977.

[14] Margaris, A. *First Order Mathematical Logic.* Xerox, 1967.

[15] Pfalzgraf, J., and Wang, D. *Automated practical reasoning.* Springer, 1995.

[16] Renegar, J. On the computational complexity and geometry of the first-order theory of the reals. *Journal of Symbolic Computation 13*, 3 (1992), 255–300.

[17] Shoenfield, J. *Mathematical Logic.* Addison-Wesley, 1973.

[18] Smullyan, R. *First order logic.* Springer, 1968.

[19] Tarski, A. *A Decision Method for Elementary Algebra and Geometry*, second ed. Univ. of California Press, Berkeley, 1951.

[20] Tarski, A. What is elementary geometry? In *Proceedings of an International Symposium on the Axiomatic Method, with Special Reference to Geometry and Physics* (1959), A. T. L. Henkin, P. Suppes, Ed., Studies in Logic and the Foundations of Mathematics, Amsterdam: North Holland, pp. 16–29.

[21] Troelstra, A., and van Dalen, D. *Constructivism in Mathematics.* North-Holland, 1988.