Logic and Proof Hilary 2026

Decidable Theories (I)

James Worrell

1 Logical Theories

In this lecture we work exclusively with first-order logic with equality.

Fix a signature σ. A *theory* $\boldsymbol{T}$ is a set of σ-sentences that is closed under semantic entailment, i.e., if $\boldsymbol{T} \models F$ then $F \in \boldsymbol{T}$. Given a σ-structure $\mathcal{A}$ it is clear that the set of sentences that hold in $\mathcal{A}$ is a theory. We denote this theory by $\mathrm{Th}(\mathcal{A})$ and call it the *theory of* $\mathcal{A}$. For example, below we will consider the theory of the ordered set $(\mathbb{Q}, <)$.

Another important source of theories is from sets of axioms. Given a set of sentences $\boldsymbol{S}$, the set $\boldsymbol{T} = \{F : \boldsymbol{S} \models F\}$ is a theory. We call $\boldsymbol{S}$ a set of *axioms* for the theory $\boldsymbol{T}$. For example, if $\boldsymbol{S}$ comprises the group axioms (over a suitable signature σ) then $\boldsymbol{T}$ is the theory of groups, i.e., the set of all σ-sentences that are true in every group.

We say that a theory $\boldsymbol{T}$ is *complete* if for any sentence F, either $F \in \boldsymbol{T}$ or $\neg F \in \boldsymbol{T}$. Clearly the theory of any particular structure is complete; however the theory of an axiomatically presented class of structures can easily fail to be so. For example, the theory of groups is not complete: if m denotes the binary multiplication operation then the theory of groups neither contains the sentence $\forall x \forall y\, (m(x,y) = m(y,x))$ nor its negation (some groups are abelian and other groups are non-abelian). More simply, the set of valid σ-formulas is an example of a theory that is not complete.

We say that a theory $\boldsymbol{T}$ admits *quantifier elimination* if for any formula $\exists x\, F$, with F quantifier-free, there exists a quantifier-free formula G with the same set of free variables as $\exists x\, F$ such that $\boldsymbol{T} \models \exists x\, F \leftrightarrow G$. Concretely, this means that for any model $\mathcal{A}$ of $\boldsymbol{T}$ and every valuation $\boldsymbol{a}$ of the free variables of $\exists x\, F$, we have $\mathcal{A} \models \exists x\, F(\boldsymbol{a})$ if and only if $\mathcal{A} \models G(\boldsymbol{a})$. We furthermore say that $\boldsymbol{T}$ has a *quantifier elimination procedure* if there is an algorithm to obtain G given F.

Example 1. Let $\boldsymbol{T}$ denote the theory of the structure $(\mathbb{R}, +, \cdot, 0, 1)$ and consider the formula $F := \exists x\, (ax^2 + bx + c = 0)$ in free variables a, b, c. This formula asserts that the quadratic equation $ax^2 + bx + c = 0$ has a real solution. By the quadratic formula we have $\boldsymbol{T} \models F \leftrightarrow b^2 \geq 4ac$. As another example, consider the formula

$$F := (x_1 a + x_2 c = 1) \wedge (x_1 b + x_2 d = 0) \wedge (x_3 a + x_4 c = 0) \wedge (x_3 b + x_4 d = 1)\,.$$

F can be written $\begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ in matrix notation. Thus $\exists x_1 \exists x_2 \exists x_3 \exists x_4 F$ asserts that the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has a multiplicative inverse. Thus $\boldsymbol{T} \models \exists x_1 \exists x_2 \exists x_3 \exists x_4 F \leftrightarrow ad - bc \neq 0$.

The definition of quantifier elimination refers only to the existential quantifier. The universal quantifier can be handled using duality. Consider a formula $\forall x\, F$ with F quantifier-free. If a theory $\boldsymbol{T}$ has quantifier elimination then we can find a quantifier-free formula G such that $\boldsymbol{T} \models \exists x\, \neg F \leftrightarrow G$. But then $\boldsymbol{T} \models \forall x\, F \leftrightarrow \neg G$.

A theory $\boldsymbol{T}$ is *decidable* if there is an algorithm that, given a sentence F, determines whether or not $F \in \boldsymbol{T}$. A theory $\boldsymbol{T}$ is decidable if it has a quantifier elimination-procedure and a procedure for determining whether or not $F \in \boldsymbol{T}$ for a variable-free atomic formula F. Given an arbitrary formula F, to determine whether $F \in \boldsymbol{T}$, first convert F to an equivalent formula in prenex normal form, and eliminate quantifiers from the inside out. In particular, if $\boldsymbol{T} \models \exists x\, F^* \leftrightarrow G$ then $\boldsymbol{T} \models Q_1 x_1 \ldots Q_n x_n Q x\, F^* \leftrightarrow Q_1 x_1 \ldots Q_n x_n\, G$, where $Q_i, Q \in \{\exists, \forall\}$. Eventually one obtains a sentence F' such that $\boldsymbol{T} \models F \leftrightarrow F'$. Thus $F \in \boldsymbol{T}$ if and only if $F' \in \boldsymbol{T}$. But by assumption we have a procedure to decide this last membership query.

2 Unbounded Dense Linear Orders

Consider a signature with a single binary relation $<$. The theory $\boldsymbol{T}_{UDLO}$ of *unbounded dense linear orders* is the set of sentences entailed by the following set of axioms:

$$
\begin{array}{ll}
F_1 & \forall x\, \neg(x < x) \\
F_2 & \forall x\, \forall y\, \forall z\, (x < y \wedge y < z \rightarrow x < z) \\
F_3 & \forall x\, \forall y\, (x < y \vee y < x \vee x = y) \\
F_4 & \forall x\, \forall y\, (x < y \rightarrow \exists z\, (x < z \wedge z < y)) \\
F_5 & \forall x\, \exists y\, \exists z\, (y < x < z)\,.
\end{array}
$$

Theorem 2. The theory $\boldsymbol{T}_{UDLO}$ of unbounded dense linear orders is complete, decidable, and has quantifier elimination.

Proof. The main step of the proof is to show that $\boldsymbol{T}_{UDLO}$ has an effective quantifier-elimination procedure. Consider a formula $\exists x\, F$, with F quantifier-free. We give a quantifier-free formula G, with the same free variables as $\exists x\, F$, such that for any structure $\mathcal{A}$ that satisfies all sentences in $\boldsymbol{T}_{UDLO}$ and any valuation $\boldsymbol{a}$ in $\mathcal{A}$ of the free variables, we have $\mathcal{A} \models \exists x\, F(\boldsymbol{a})$ if and only if $\mathcal{A} \models G(\boldsymbol{a})$. The quantifier-elimination procedure has two phases: first we simplify the formula F through logical manipulations and then we show how to eliminate quantifiers within formulas in simplified form.

As a first step, we can convert F into a logically equivalent formula in DNF. We can moreover eliminate negative literals by replacing the subformula $\neg(x_i < x_j)$ with $x_i = x_j \vee x_j < x_i$ and replacing the subformula $\neg(x_i = x_j)$ with $x_i < x_j \vee x_j < x_i$.

Henceforth we assume that F is in DNF and negation-free. Now using the equivalence $\exists x\, (F_1 \vee F_2) \equiv \exists x\, F_1 \vee \exists x\, F_2$ it suffices that we be able to eliminate the quantifier $\exists x$ in case F is a conjunction of atomic formulas. Finally, using the equivalence $\exists x\, (F_1 \wedge F_2) \equiv \exists x\, F_1 \wedge F_2$ in case x is not free in F_2, it suffices that we be able to eliminate the quantifier $\exists x$ in case F is a conjunction of atomic formulas all of which mention x. Such formulas have the form $x = y$, $x < y$ or $y < x$ for some variable y.

For the final case above, we proceed as follows. If F contains a conjunct $x < x$ then we have $\boldsymbol{T}_{UDLO} \models \exists x\, F \leftrightarrow \textbf{false}$. Otherwise, if F contains a conjunct $x = y$ for some other variable y then we have that $\boldsymbol{T}_{UDLO} \models \exists x\, F \leftrightarrow F[y/x]$.

If neither of the above applies then (after deleting conjuncts of the form $x = x$ if present) we can write F in the form

$$
F = \bigwedge_{i=1}^{m} l_i < x \wedge \bigwedge_{j=1}^{n} x < u_j\,,
$$

where the l_i and u_j are variables different from x. Now if $m = 0$, i.e., there are no lower bounds on x, then $\boldsymbol{T}_{UDLO} \models \exists x\, F \leftrightarrow \mathbf{true}$ (since we're considering the theory of unbounded orders). Likewise if $n = 0$, i.e., there are no upper bounds on x, then $\boldsymbol{T}_{UDLO} \models \exists x\, F \leftrightarrow \mathbf{true}$. Otherwise, by density of the order relation, we have

$$\boldsymbol{T}_{UDLO} \models \exists x\, F \leftrightarrow \bigwedge_{i=1}^{m} \bigwedge_{j=1}^{n} l_i < u_j\,.$$

Decidability of $\boldsymbol{T}_{UDLO}$ follows straightforwardly from the existence of a quantifier-elimination procedure. Starting from a sentence F, after eliminating all quantifiers from F we are left with a variable-free formula G such that $\boldsymbol{T} \models F \leftrightarrow G$. But G must be a propositional combination of **true** or **false**, and therefore logically equivalent to either **true** or **false**. The same reasoning shows *inter alia* that $\boldsymbol{T}_{UDLO}$ is complete: given a sentence F, either F holds on all unbounded dense linear orders, or its negation holds on all unbounded dense linear orders. □

Theorem 2 shows that $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$ satisfy the same first-order sentences. (This finally answers Exercise 7 from the lecture introducing first-order logic.) You may recall that $(\mathbb{R}, <)$ is *Dedekind complete*: any non-empty set of reals that is bounded above has a least upper bound. This property fails for the rationals since, e.g., $\{x \in \mathbb{Q} : x^2 < 2\}$ has no least upper bound in the rationals. Evidently Dedekind completeness cannot be expressed in first-order logic in the language of linear orders.

3 Ordered Divisible Abelian Groups

Consider a signature with a binary relation symbol $<$, binary function symbol $+$, and a constant symbol 0. Via an obvious notational shortcut, it will be convenient to admit $\mathbb{Z}$-linear expressions in variables as terms. For example, we write $3x + y$ for the term $x + (x + (x + y))$ and we write $x - 2y < z$ for the formula $x < (z + y) + y$.

The set of axioms $\{F_1, \ldots, F_9\} \cup \{G_n : n \in \mathbb{N}_+\}$, shown below, determines the theory $\boldsymbol{T}_{\mathrm{ODAG}}$ of (non-trivial) ordered divisible abelian groups.

$$\begin{array}{ll}
F_1 & \forall x\, \neg(x < x) \\
F_2 & \forall x\, \forall y\, \forall z\, (x < y \wedge y < z \to x < z) \\
F_3 & \forall x\, \forall y\, (x < y \vee y < x \vee x = y) \\
F_4 & \forall x\, \forall y\, (x + y = y + x) \\
F_5 & \forall x\, \forall y\, \forall z\, ((x + y) + z = x + (y + z))) \\
F_6 & \forall x\, \exists y\, (x + y = 0) \\
F_7 & \forall x\, (x + 0 = x) \\
F_8 & \forall x\, \forall y\, \forall z\, (x < y \to x + z < y + z) \\
F_9 & \exists x\, \neg(x = 0) \\
G_n & \forall x\, \exists y\, (ny = x)\,.
\end{array}$$

A model of $\boldsymbol{T}_{\mathrm{ODAG}}$ is a structure $(A, <, +, 0)$ such that $(A, <)$ is a linearly ordered set, $(A, +, 0)$ is a divisible Abelian group, addition $+$ is monotone in both variables, and A is non-trivial as a group (it has some non-zero element). Examples include $(\mathbb{R}, <, +, 0)$ and $(\mathbb{Q}, <, +, 0)$. We leave it as an exercise to show that for any model $(A, <, +, 0)$ of $\boldsymbol{T}_{\mathrm{ODAG}}$ the order $<$ is unbounded and dense.

Theorem 3. $\boldsymbol{T}_{\mathrm{ODAG}}$ has quantifier elimination.

Proof. Following the proof of Theorem 2, it suffices to show how to eliminate the quantifier $\exists x$ in $\exists x\, F$, where F is a conjunction of atomic formulas all of which mention x. Each such atomic formula has the form $t_1 < t_2$ or $t_1 = t_2$ for terms t_1 and t_2, where at least one of t_1 or t_2 mentions x. We can further simplify to assume that there exists a positive integer m such that each formula in F has the form $mx < t$, $t < mx$, or $mx = t$ for some term t (where some variables may occur in t with negative coefficients). Suppose that F has the form

$$\bigwedge_{i=1}^{n_1} t_i < mx \wedge \bigwedge_{j=1}^{n_2} mx < s_j \wedge \bigwedge_{k=1}^{n_3} mx = u_k\,.$$

Then, by the divisibility axioms, $\boldsymbol{T}_{\mathrm{ODAG}} \models \exists x\, F \leftrightarrow \exists y\, G$ where

$$G := \bigwedge_{i=1}^{n_1} t_i < y \wedge \bigwedge_{j=1}^{n_2} y < s_j \wedge \bigwedge_{k=1}^{n_3} y = u_k\,,$$

for some fresh variable y not occurring in F. If $n_3 > 0$ then $\exists y\, G \equiv G[u_1/y]$. If $n_3 = 0$ then $T_{\mathrm{ODAG}} \models \exists y\, G \leftrightarrow \bigwedge_{i=1}^{n_1} \bigwedge_{j=1}^{n_2} t_i < s_j$. □

The quantifier elimination procedure for $\boldsymbol{T}_{\mathrm{ODAG}}$ can be used to solve certain elementary problems in convex geometry. In this context quantifier elimination is sometimes called *Fourier-Motzkin elimination*. For example, given matrices A and C and vectors $\boldsymbol{b}$ and $\boldsymbol{d}$, all with rational entries, determining the whether the polygon $\{\boldsymbol{x} \in \mathbb{R}^n : A\boldsymbol{x} \leq \boldsymbol{b}\}$ is included in the polygon $\{\boldsymbol{x} \in \mathbb{R}^n : C\boldsymbol{x} \leq \boldsymbol{d}\}$ can straightforwardly be reduced to the decision problem for $\boldsymbol{T}_{\mathrm{ODAG}}$.