The following lemma is a consequence of Gödel's completeness theorem for FOL. It makes use of our default assumption that the (sets of predicate and function symbols of the) given vocabulary and variable-set are recursively enumerable.

**Lemma 1.4.9.** *Each axiomatizable FO-theory is recursively enumerable.*

*Proof.* Let $\mathfrak{D}$ be a sound and complete Hilbert proof system for FOL. Let $\mathfrak{T}$ be a FO-theory and $\mathfrak{F}$ a decidable set of formulas such that $\mathfrak{T} = \{\, \phi : \mathfrak{F} \Vdash \phi \,\}$. But then

$$\mathfrak{T} \;=\; \{\, \phi : \mathfrak{F} \vdash_{\mathfrak{D}} \phi \,\} \;=\; \mathit{Cl}(\mathfrak{F}).$$

On the basis of this observation, we obtain a procedure that generates all formulas in $\mathfrak{T}$ as follows.

Let us first suppose that the set of all formulas over the given vocabulary *Voc* and variable-set *Var* is decidable, which is the case if the sets $\mathit{Pred}_n$ and $\mathit{Func}_m$ in *Voc* and the variable-set *Var* are decidable. We consider a recursive enumeration of all nonempty finite sequences of formulas and check for each of these sequences $\psi_1, \ldots, \psi_m$ whether it constitutes a derivation in the given proof system $\mathfrak{D}$ using the axiom $\mathfrak{F}$. If so, then the algorithm outputs the last formula, otherwise $\psi_1, \ldots, \psi_m$ is ignored. Note that the question whether a given finite sequence of formulas is a derivation from $\mathfrak{F}$ is decidable since $\mathfrak{F}$ and all axioms and proof rules are assumed to be decidable.

**FOR ALL** nonempty finite sequences $\psi_1, \ldots, \psi_m$ of formulas **DO**
  **IF** $\psi_1, \ldots, \psi_m$ is $\mathfrak{D}$-proof from $\mathfrak{F}$ **THEN**
    return $\psi_m$
  **FI**
**OD**

If the vocabulary *Voc* or variable-set *Var* are recursively enumerable, but at least one of them is not decidable, then we make use of Remark 1.2.15 on page 29 and consider a sound and complete Hilbert proof system $\mathfrak{D}$ with abstract symbols for the formulas, variables, terms and predicate and function symbols, taken from infinite, decidable sets. The set of all meta-formulas, i.e., formulas with the abstract symbols rather than concrete ones from *Voc* and *Var*, is decidable, and so is the question whether a given finite sequence of meta-formulas is a $\mathfrak{D}$-proof from $\mathfrak{F}$. For simplicity, we suppose here that the symbols that appear in $\mathfrak{F}$ are identified with corresponding abstract symbols. This premits to consider $\mathfrak{F}$ as a decidable set of meta-formulas with fixed meaning over (*Var*, *Voc*).

A recursive enumeration of all $\mathfrak{T}$ is obtained as follows. We consider a recursive enumeration of all pairs $\langle (\Xi_1, \ldots, \Xi_m), \sigma \rangle$ consisting of a nonempty sequence $\Xi_1, \ldots, \Xi_m$ of meta-formulas and an assigment $\sigma$ that maps the abstract symbols appearing in $\Xi_1, \ldots, \Xi_m$ to concrete ones. Note that $\Xi_1, \ldots, \Xi_m$ can only contain finitely many abstract symbols and that the sets of all formulas, variables, terms and predicate and function symbols over (*Voc*, *Var*) are recursively enumerable. If $\Xi_1, \ldots, \Xi_m$ constitutes a $\mathfrak{D}$-proof from $\mathfrak{F}$ then the procedure returns $\Xi_m\sigma$, i.e., the formula over (*Voc*, *Var*) that is obtained from $\Xi_m$ by instantiating the abstract symbols uniformly with the concrete symbols as indicated by $\sigma$. Given the soundness and completeness of $\mathfrak{D}$, this procedure encounters precisely the formulas in $\mathfrak{T}$.

Alternatively, one could even relax the definition of axiomatizations as pairs consisting of a decidable set of meta-formulas (that might be used as axioms in $\mathfrak{D}$-proofs) and a recursive enumerable set of assignments mapping the abstract symbols to concrete ones (that are used in the instantiating phase). □

Since complete theories are either not recursively enumerable or decidable (see Theorem 1.4.7 on page 48), each complete and axiomatizable FO-theory is decidable. Vice versa, each decidable theory $\mathfrak{T}$ is axiomatizable since we may deal with $\mathfrak{T}$ as an axiomatization. (Recall that an axiomatization for $\mathfrak{T}$ is a decidable set whose logical closure agrees with $\mathfrak{T}$.) Hence, we get:

**Corollary 1.4.10 (Axiomatizability vs decidability for complete FO-theories).** *Let $\mathfrak{T}$ be a complete FO-theory. Then, $\mathfrak{T}$ is axiomatizable iff $\mathfrak{T}$ is decidable.*

A sound and complete deductive calculus for an FO-theory $\mathfrak{T}$ means a proof system $\mathfrak{D}$ (e.g., in the Hilbert-style) such that $\vdash_{\mathfrak{D}} \phi$ iff $\phi \in \mathfrak{T}$ for all sentences $\phi$.

**Theorem 1.4.11 (Axiomatizability and sound and complete deductive calculi).** *Let $\mathfrak{T}$ be a FO-theory. Then, $\mathfrak{T}$ is axiomatizable iff $\mathfrak{T}$ has a sound and complete deductive calculus.*

*Proof.* "$\Longrightarrow$": If $\mathfrak{T}$ is axiomatizable, say $\mathfrak{T} = \{\phi : \mathfrak{F} \Vdash \phi\}$ where $\mathfrak{F}$ is decidable, then any sound and complete Hilbert proof system $\mathfrak{D}$ for FOL together with the additional axiom $\mathfrak{F}$ yields a Hilbert proof system $\mathcal{D} + \mathfrak{F}$ such that for all FOL-sentences $\phi$:

$$\vdash_{\mathcal{D}+\mathfrak{F}} \phi \quad \text{iff} \quad \mathfrak{F} \vdash_{\mathfrak{D}} \phi \quad \text{iff} \quad \mathfrak{F} \Vdash \phi \quad \text{iff} \quad \phi \in \mathfrak{T}$$

(Recall that for Hilbert proof systems, axioms are just decidable sets of formulas. This permits to use $\mathfrak{F}$ as an axiom.) Thus, $\mathcal{D} + \mathfrak{F}$ is sound and complete for $\mathfrak{T}$.

"$\Longleftarrow$": The existence of a sound and complete proof system for $\mathfrak{T}$ implies that $\mathfrak{T}$ is recursively enumerable. This follows by an argument as in the proof of Lemma 1.2.16 on page 29. Let $\psi_1, \psi_2, \psi_3, \ldots$ be a recursive enumeration of the formulas in $\mathfrak{T}$. We define

$$\phi_n \stackrel{\text{def}}{=} \psi_n \wedge \underbrace{\mathit{true} \wedge \ldots \wedge \mathit{true}}_{n \text{ times}}$$

*Claim:* $\mathfrak{F} \stackrel{\text{def}}{=} \{\phi_n : n \geqslant 1\}$ is an axiomatization for $\mathfrak{T}$.

*Proof of the claim.* Clearly, any formula in $\mathfrak{T}$ is equivalent to some formula in $\mathfrak{F}$ (as $\psi_n \equiv \phi_n$). As $\mathfrak{T}$ is a theory, $\mathfrak{T}$ agrees with its closure and we get:

$$\mathfrak{T} \;=\; \mathrm{Cl}(\mathfrak{T}) \;=\; \mathrm{Cl}(\mathfrak{F})$$

It remains to show that $\mathfrak{F}$ is decidable. For this, we have to show that there is an algorithm that takes as input a FOL-sentence $\phi$ and decides whether $\phi \in \mathfrak{F}$. Given a formula $\phi$, the algorithm first calculates the length $|\phi| = m$. Since

$$|\phi_n| \;=\; |\psi_n| + n \;\geqslant\; n,$$

the only chance for $\phi \in \mathfrak{F}$ is $\phi = \phi_n$ for some $n \leqslant m$. Hence, we can apply the procedure that generates the first $m$ formulas $\phi_1, \ldots, \phi_m$ of $\mathfrak{F}$ and checks whether one of them agrees with $\phi$. If so, then $\phi \in \mathfrak{F}$. Otherwise, i.e., if $\phi \notin \{\phi_1, \ldots, \phi_m\}$ then $\phi \notin \mathfrak{F}$. □

Note that the above proof applies to any theory of a logic that has sound and complete proof systems. Furthermore, in the proof of the implication "$\Longleftarrow$" we just use the fact that the existence of sound and complete proof systems implies the existence of a recursive enumeration. Hence, if a FO-theory is recursively enumerable then it is axiomatizable. Together with Lemma 1.4.9 this yields:

**Corollary 1.4.12 (Equivalence of axiomatizability and semi-decidability).** *Let $\mathfrak{T}$ be a FO-theory. Then, $\mathfrak{T}$ has an axiomatization iff $\mathfrak{T}$ is recursively enumerable.*

We conclude this section by listing some results on popular FO-theories. The proofs for these results are rather involved and will not be presented here.

- The FO-theory of arithmetic $\mathit{Th}(\mathbb{N}, +, *, =)$ is not decidable. Hence, $\mathit{Th}(\mathbb{N}, +, *, =)$ is not recursively enumerable and neither axiomatizable nor does $\mathit{Th}(\mathbb{N}, +, *, =)$ have sound and complete deductive calculi. This result is the famous *(first) Gödel's incompleteness theorem*.
- If we drop the multiplication and just consider the FO-theory $\mathit{Th}(\mathbb{N}, +, =)$ then we obtain the so-called *Presburger arithmetic* which is known to be decidable, and hence, axiomatizable.
- If we switch from the natural numbers to the rationals then we obtain the FO-theory $\mathit{Th}(\mathbb{Q}, +, *, =)$, which is undecidable as well. However, when dealing with the reals rather than the rationals then we obtain the FO-theory $\mathit{Th}(\mathbb{R}, +, *, =)$ of *real closed fields*. This theory, also called *Tarski algebra*, is known to be axiomatizable (and therefore decidable) as it has been shown by Tarski.

Although the FO-theories $\mathit{Th}(\mathbb{N}, +, =)$ and $\mathit{Th}(\mathbb{R}, +, *, =)$ are decidable, the decision problems for both are computationally hard. Presburger arithmetic $\mathit{Th}(\mathbb{N}, +, =)$ is decidable in double exponential deterministic space, but not in double exponential nondeterministic time. The FO-theory of real closed fields $\mathit{Th}(\mathbb{R}, +, *, =)$ is known to be complete for the complexity class *EXPSPACE*, and hence solvable in double exponential time. However, special subclasses of $\mathit{Th}(\mathbb{R}, +, *, =)$, e.g., the class of formulas where the depth of alternating universal and existential quantifiers is fixed, can be decided in (single) exponential time.

We will not provide the proofs for these complexity-theoretic statements. The decidability results for the above FO-theories rely on the concept of *quantifier elimination*. This means an algorithmic transformation of a given FOL-formula $\exists x.\phi$ into a quantifier-free formula $\psi$ such that $\exists x.\phi$ and $\psi$ are "equivalent" for the given FO-theory $\mathfrak{T}$. We will illustrate this concept by a simpler example, namely the structure $(\mathbb{Q}, \leqslant)$ of the rationals with the natural order.

### 1.4.3 The FO-theory of the ordered rationals

We consider the FO-theory $Th(\mathbb{Q}, \leqslant)$, i.e., the theory of all FOL-formulas that evaluate to true when interpreted over $(\mathbb{Q}, \leqslant)$. The underlying vocabulary consists of a single binary predicate symbol, denoted by $\sqsubseteq$. Atomic formulas will be written in infix notation $x \sqsubseteq y$ rather than $\sqsubseteq (x, y)$. Occasionally, we also use the symbol $\leqslant$ in formulas that are interpreted over the ordered rationals (or other subsets of the reals with the natural order). The goal is to provide an algorithm that takes as input a FOL-sentence $\phi$ over this vocabulary and checks whether $(\mathbb{Q}, \leqslant) \models \phi$. Before presenting such a quantifier elimination algorithm for the rational numbers, we give some definitions on order relations that will also be needed later for other purposes.

**Definition 1.4.13 (Quasi order, partial order, linear order).** Let $A$ be a set and $\sqsubseteq$ a binary relation on $A$, i.e., a subset $A \times A$. The pair $(A, \sqsubseteq)$ is called a *quasi order* on $A$, also often called *preorder*, if $\sqsubseteq$ that is transitive and reflexive. If in addition $\sqsubseteq$ is *antisymmetric* then $(A, \sqsubseteq)$ is called a *partial order*. That is, partial orders are structures $(A, \sqsubseteq)$ where the following FOL-formulas hold:

$$\forall x \forall y \forall z. \big( (x \sqsubseteq y \wedge y \sqsubseteq z) \rightarrow x \sqsubseteq z \big) \qquad \text{(transitivity)}$$

$$\forall x. (x \sqsubseteq x) \qquad \text{(reflexivity)}$$

$$\forall x \forall y. \big( (x \sqsubseteq y \wedge y \sqsubseteq x) \rightarrow x = y \big) \big) \qquad \text{(antisymmetry)}$$

A partial order is called a *linear* order, or a *total* order, iff every two elements of $A$ are related via $\sqsubseteq$, i.e., if the following FOL-formula holds for $(A, \sqsubseteq)$:

$$\forall x \forall y. \big( x \sqsubseteq y \vee y \sqsubseteq x \big)$$

For $a, b \in A$ the notation $a \sqsubset b$ means $a \sqsubseteq b \wedge a \neq b$. Similarly, $a \sqsupseteq b$ stands for $b \sqsubseteq a$, while $a \sqsupset b$ means $b \sqsubset a$. Note that for any linear order $(A, \sqsubseteq)$ the derived relation $\sqsubset$ is transitive, antireflexive and satisfies the *trichotomy property* on $A$. The latter means that for all $a, b \in A$ exactly one of the three possibilities $a = b$ or $a \sqsubset b$ or $b \sqsubset a$ holds. ■

Clearly, $(\mathbb{Q}, \leqslant)$ and $(\mathbb{R}, \leqslant)$ but also $(\mathbb{N}, \leqslant)$ are linear orders. An example for a partial order which is not linear is the structure $(2^A, \subseteq)$ for some set $A$ with at least two elements, i.e., the powerset of $A$ ordered by inclusion. Note that here the singleton sets $\{a\}$ and $\{b\}$ for $a, b \in A$, $a \neq b$, are not related via $\subseteq$. What is typical for $(\mathbb{Q}, \leqslant)$, but also $(\mathbb{R}, \leqslant)$, is the density and that there are no endpoints:

**Definition 1.4.14 (Density, endpoints of linear orders).** A linear order $\mathcal{A} = (A, \sqsubseteq)$ is called *dense* if between every two elements $a, b \in A$ where $a \sqsubset b$ there is an element which is strictly less than $b$ and strictly greater than $a$, i.e., if

$$\mathcal{A} \models \forall x \forall y. \big( x \sqsubset y \rightarrow \exists z. (x \sqsubset z \wedge z \sqsubset y) \big)$$

As before, $x \sqsubset y$ is a short hand notation for $x \sqsubseteq y \wedge x \neq y$. An element $a \in A$ is said to be *minimal* if $a \sqsubseteq b$ for all elements $b \in A$. Similarly, a maximal element means an element $a \in A$ such that $b \sqsubseteq a$ for all $b \in A$. (By the antisymmetry, any partial order has at most one minimal and at most one maximal element.) By an *endpoint*, we mean a minimal or maximal element. ■

The natural order $\leqslant$ on rational numbers is linear and dense, and has no endpoints. The latter means that

$$(\mathbb{Q}, \leqslant) \models \forall x. \big( \exists y. (x \sqsubset y) \wedge \exists z. (z \sqsubset x) \big)$$

Note that this property also holds for $(\mathbb{R},\leqslant)$, but not for $(\mathbb{N},\leqslant)$ where 0 is an endpoint. The linear order $(\mathbb{Z},\leqslant)$ of the integers has no end points, but is not dense. In fact, the conditions "density" and "having no endpoints" is characteristic for $(\mathbb{Q},\leqslant)$ in the following sense:

**Lemma 1.4.15 (Isomorphism of countable dense partial orders without endpoints).** *Given two countable linear orders $(A,\sqsubseteq_A)$ and $(B,\sqsubseteq_B)$, that are dense and have no endpoints, and strictly increasing sequences $a_1 \sqsubset_A \ldots \sqsubset_A a_k$ and $b_1 \sqsubset_B \ldots \sqsubset_B b_k$ of the same length $k \geqslant 0$ consisting of elements $a_i \in A$ and $b_i \in B$, $1 \leqslant i \leqslant k$, then the extended structures*

$$(A,\sqsubseteq_A, a_1,\ldots,a_k) \text{ and } (B,\sqsubseteq_B, b_1,\ldots,b_k) \text{ are isomorphic,}$$

*i.e., there exists a bijection $h : A \to B$ such that:*

*(1) for all $a$, $a' \in A$ we have: $a \sqsubseteq_A a'$ iff $h(a) \sqsubseteq_B h(a')$*

*(2) $h(a_i) = b_i$ for $1 \leqslant i \leqslant k$*

In particular, each countable linear order $(A,\sqsubseteq_A)$ that is dense and has no endpoints, is isomorphic to $(\mathbb{Q},\leqslant)$. The assumption "countable" is necessary as, e.g., the structure $(\mathbb{R},\leqslant)$ of the ordered reals is also dense without endpoints.

*Proof.* To establish an isomorphism $h : (A,\sqsubseteq_A, a_1,\ldots,a_k) \to (B,\sqsubseteq_B, b_1,\ldots,b_k)$, we fix enumerations for $A$ and $B$, where each element of $A$ and $B$ appears exactly once. (We do not require monotonicity for these enumerations.) Such enumerations $\alpha = a'_{i_1}, a'_{i_2}, a'_{i_3},\ldots$ and $\beta = b'_{j_1}, b'_{j_2}, b'_{j_3},\ldots$ of $A$ and $B$, respectively, exist as both $A$ and $B$ are supposed to be countable. The assumption that there are no endpoints yields that $A$ and $B$ are infinite.

The isomorphism $h$ is now defined inductively, using a *back-and-forth argument*. We start with the partial isomorphism $h$ that maps $a_i$ to $b_i$ for $1 \leqslant i \leqslant k$ and successively extend $h$ to obtain an isomorphism from $\mathcal{A}$ to $\mathcal{B}$.

"forth": Let $a = a_{k+1} \in A$ be the first element in $\alpha$ such that $a \in A \setminus \{a_1,\ldots,a_k\}$. Since $\sqsubseteq_B$ is linear, dense and has no endpoints there exists an element $b_{k+1} \in B \setminus \{b_1,\ldots,b_k\}$ that relates to $b_1,\ldots,b_k$ according to $\sqsubseteq_B$ in the same way as $a = a_{k+1}$ is related to $a_1,\ldots,a_k$ according to $\sqsubseteq_A$. That is, $b_{k+1}$ is an element of $B \setminus \{b_1,\ldots,b_k\}$ such that:

$$a_j \sqsubseteq_A a_i \quad \text{iff} \quad b_j \sqsubseteq_B b_i, \quad i,j = 1,\ldots,k,k+1$$

We then define $h(a_{k+1}) \stackrel{\text{def}}{=} b_{k+1}$. This yields a local isomorphism

$$h : \{a_1,\ldots,a_k,a_{k+1}\} \to \{b_1,\ldots,b_k,b_{k+1}\}, \;\; h(a_i) = b_i \text{ for } 1 \leqslant i \leqslant k+1.$$

"back": In the next step we apply the same technique with exchanged roles of $A$ and $B$. We pick the first element $b' = b_{k+2}$ in $\beta$ such that $b' \in B \setminus \{b_1,\ldots,b_k,b_{k+1}\}$. As $\sqsubseteq_A$ is linear, dense and has no endpoints there exists an element $a_{k+2} \in A \setminus \{a_1,\ldots,a_k,a_{k+1}\}$ that relates to $a_1,\ldots,a_k,a_{k+1}$ according to $\sqsubseteq_A$ in the same way as $b' = b_{k+2}$ is related to $b_1,\ldots,b_k,b_{k+1}$ according to $\sqsubseteq_B$. We then define $h(a_{k+2}) \stackrel{\text{def}}{=} b_{k+2}$. Thus, we have constructed a bijection

$$h : \{a_1,\ldots,a_{k+2}\} \to \{b_1,\ldots,b_{k+2}\}, \qquad h(a_i) = b_i \;\text{ for } 1 \leqslant i \leqslant k+2$$

which enjoys the property

$$a_i \sqsubseteq_A a_j \quad \text{iff} \quad b_i = h(a_i) \sqsubseteq_B h(a_j) = b_j, \quad \text{for all } i,j \in \{1,\ldots,k{+}2\}.$$

We repeat this back-and-forth definition of $h$ ad infinity and obtain an isomorphism

$$h : (A, \sqsubseteq_A, a_1, \ldots, a_k) \to (B, \sqsubseteq_B, b_1, \ldots, b_k).$$

Since in each step we take the first element in the enumerations $\alpha$ and $\beta$, respectively, where $h$ is not yet defined, eventually all elements of $A$ and $B$ will be picked in either the back- or the forth-step. This ensures the bijectivity of the constructed function $h$. □

As a consequence of the above lemma we get: whenever we fix rational numbers $a_1,\ldots,a_k$ and $b_1,\ldots,b_k$ such that $a_1 < \ldots < a_k$ and $b_1 < \ldots < b_k$, then there is an isomorphism from $(\mathbb{Q}, \leqslant, a_1, \ldots, a_k)$ to $(\mathbb{Q}, \leqslant, b_1, \ldots, b_k)$. In particular, these structures – viewed as interpretations with structure $(\mathbb{Q}, \leqslant)$ and variable valuations $x_i \mapsto a_i$ or $x_i \mapsto b_i$ for $1 \leqslant i \leqslant k$ – are equivalent for all FOL-formulas $\phi(x_1,\ldots,x_k)$ over the vocabulary consisting of a single binary predicate symbol for $\leqslant$. This observation provides the key for the quantifier elimination algorithm for $Th(\mathbb{Q}, \leqslant)$. Let us first sketch the main ideas by means of an example. The goal is to decide whether the formula

$$\phi \;=\; \exists x \forall y \exists z. \big( x < y \wedge (z \leqslant x \vee y = z) \big)$$

holds for $(\mathbb{Q}, \leqslant)$. The idea is to remove the quantifiers $\exists x$, $\forall y$, $\exists z$ by choosing appropriate values (rational numbers) for $x, y, z$ that serve as representatives for all other interpretations for $x, y, z$ in $\mathbb{Q}$. By Lemma 1.4.15, the inner subformula $\forall y \exists z.(\ldots)$ either holds for all values for $x$ or for none of them. Hence, we may pick an arbitrary element $a \in \mathbb{Q}$ and check whether

$$\phi(a) \stackrel{\text{def}}{=} \forall y \exists z. \big( a < y \wedge (z \leqslant a \vee y = z) \big)$$

holds for $(\mathbb{Q}, \leqslant)$. Here, $a$ is viewed as a constant symbol with fixed semantics (namely the rational number $a$). To check whether $\phi(a)$ holds we have to verify whether

$$\phi(a,b) \stackrel{\text{def}}{=} \exists z. \big( a < b \wedge (z \leqslant a \vee b = z) \big)$$

holds for all rational numbers $b$. However, there are just three essential possibilities for $b$: either $a < b$ or $a = b$ or $a > b$. Let $b_0, b_1, b_2$ be representatives for these three possibilities, say $b_0 < a$, $b_1 = a$ and $b_2 > a$. All other possibilities are "isomorphic" to one of these by Lemma 1.4.15. It now remains to choose appropriate values $c \in \mathbb{Q}$ for $z$ to eliminate the quantifier $\exists z$ for each of the three formulas $\phi(a, b_0)$, $\phi(a, b_1)$ and $\phi(a, b_2)$. For $\phi(a, b_0)$ we have $b_0 < a$ which gives five alternatives for the value $c$ for $z$: either $c < b_0 < a$ or $c = b_0 < a$ or $b_0 < c < a$ or $b_0 < c = a$ or $p_0 < a < c$. Similarly, for $\phi(a, b_1)$ there are three alternatives, namely $c < b_1 = a$ or $c = b_1 = a$ or $b_1 = a < c$, while for $\phi(a, b_2)$ there are again five possibilities. For each of these representatives $a$, $b$, $c$ for the possible interpretations for $x, y, z$, we can simply check the truth of the inner quantifier-free formula

$$\phi(a,b,c) \stackrel{\text{def}}{=} a < b \wedge (c \leqslant a \vee b = c).$$

Then, $\phi$ holds if for all choices $b$ for $y$ there is some choice $c$ for $z$ where $\phi(a,b,c)$ holds. The value $a$ for $x$ is irrelevant as we mentioned above.

We now describe the decision algorithm in more detail. For this we assume that the given FOL-sentence $\phi$ has been transformed into prenex form, i.e.,

$$\phi \;=\; Q_1 x_1 \ldots Q_n x_n . \psi$$

where $Q_i \in \{\exists, \forall\}$ and $\psi$ is quantifier-free. Recall that any FOL-formula $\phi$ can be transformed into an equivalent formula $\phi'$ in prenex form of the same asymptotic length and word-length. See page 12. Furthermore, we assume $\phi$ is a sentence, i.e., $\mathit{Free}(\psi) \subseteq \{x_1, \ldots, x_n\}$. A *linear arrangement* for the first $k$ variables $x_1, \ldots, x_k$ is a sequence

$$\alpha \;=\; x_{i_1} \bowtie_1 x_{i_2} \bowtie_2 \ldots \bowtie_{k-1} x_{i_k}$$

where $x_{i_1}, \ldots, x_{i_k}$ is a permutation of $x_1, \ldots, x_k$ and $\bowtie_j \in \{<, =\}$ for $1 \leqslant j < k$. Furthermore, if $\bowtie_j$ is the equality symbol (i.e., $x_{i_j} = x_{i_{j+1}}$ is a constraint in $\alpha$) then we require $i_j < i_{j+1}$. That is, for $k = 3$, $x_3 < x_1 = x_2$, $x_1 = x_2 = x_3$ and $x_2 < x_3 < x_1$ are linear arrangements, while $x_3 < x_2 = x_1$ and $x_3 = x_2 = x_1$ are not since they violate the last condition. An *instance* of $\alpha$ means a $k$-tuple $(a_1, \ldots, a_k) \in \mathbb{Q}^k$ of rational numbers such that

$$a_{i_1} \bowtie_1 a_{i_2} \bowtie_2 \ldots \bowtie_{k-1} a_{i_k}$$

A variable evaluation $\mathcal{V} : \{x_1, \ldots, x_k\} \to \mathbb{Q}$ is said to be *consistent* with $\alpha$ iff $(\mathcal{V}(x_1), \ldots, \mathcal{V}(x_k))$ is an instance of $\alpha$. By Lemma 1.4.15, all variable valuations $\mathcal{V} : \{x_1, \ldots, x_k\} \to \mathbb{Q}$ that are consistent with $\alpha$ yield the same truth value for

$$\phi_k \;\stackrel{\text{def}}{=}\; Q_{k+1} x_{k+1} \ldots Q_n x_n . \psi$$

Hence, if $\alpha$ is a linear arrangement for the first $k$ variables $x_1 \ldots x_k$ then for all instances $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ of $\alpha$ we have:

$$(\mathbb{Q}, \leqslant, a_1, \ldots, a_k) \models \phi_k \quad \text{iff} \quad (\mathbb{Q}, \leqslant, b_1, \ldots, b_k) \models \phi_k$$

This observation allows use linear arrangements as *symbolic* representations for sets of variable valuations and to define a satisfaction relation $\models$ for linear arrangements for the first $k$ variables $x_1, \ldots, x_k$ and the subformula $\phi_k \;=\; Q_{k+1} x_{k+1} \ldots Q_n x_n . \psi$ as follows.

$$\alpha \models \phi_k \;\stackrel{\text{def}}{\Longleftrightarrow}\; (\mathbb{Q}, \leqslant, a_1, \ldots, a_k) \models \phi_k \text{ for all/some instance(s) } (a_1, \ldots, a_k) \text{ of } \alpha$$

This yields the following recursive characterization of the satisfaction relation for linear arrangements. If $k = n$ then $\alpha$ is a linear arrangement for all variables that appear free in $\psi$. Then, $\alpha \models \psi$ iff $\psi$ holds for all/some variable valuation(s) that are (is) consistent with $\alpha$. Let us now assume that $1 \leqslant k < n$ and let $\alpha$ be a linear arrangement for $x_1, \ldots, x_k$. A linear arrangement $\beta$ for $x_1, \ldots, x_k, x_{k+1}$ is called an *extension* of $\alpha$ if $\beta$ arises from $\alpha$ by adding constraints for $x_{k+1}$. E.g., if $k = 2$ and $\alpha$ is $x_1 < x_2$ then there are five extensions:

$$x_1 < x_2 < x_3 \qquad x_1 < x_2 = x_3 \qquad x_3 < x_1 < x_2$$
$$x_1 = x_3 < x_2 \qquad x_1 < x_3 < x_2$$

For the general case, the number of extensions for $\alpha$ is bounded above by $2k{+}1$ (e.g., if $\alpha$ equals $x_1 < \ldots < x_k$) and bounded below by 3 (if $\alpha$ equals $x_1 = \ldots = x_k$). We then have:

$$\alpha \models \exists x_{k+1} Q_{k+2} x_{k+2} \ldots Q_n x_n.\psi \quad \text{iff} \quad \begin{cases} \text{there exists an extension } \beta \text{ of } \alpha \text{ such that:} \\ \quad \beta \models Q_{k+2} x_{k+2} \ldots Q_n x_n.\psi \end{cases}$$

$$\alpha \models \forall x_{k+1} Q_{k+2} x_{k+2} \ldots Q_n x_n.\psi \quad \text{iff} \quad \begin{cases} \text{for all extensions } \beta \text{ of } \alpha: \\ \quad \beta \models Q_{k+2} x_{k+2} \ldots Q_n x_n.\psi \end{cases}$$

The decision algorithm for the FO-theory $Th(\mathbb{Q},\leqslant)$ relies on a recursive procedure that successively eliminates the quantifiers $Q_k x_k$, $k = 1,\ldots,n$, by considering the extensions of the current linear arrangement for $x_1,\ldots,x_{k-1}$ according to the above satisfaction relation for linear arrangements. This procedure can be implemented such that the memory requirement is polynomially bounded. For instance, we may use a recursive approach as sketched in Algorithm 1.4.3 on page 57, where the initial call is *check*(1,$x_1$). If *check*(1,$x_1$) returns *true* then the given FOL-sentence $\phi = Q_1 x_1 \ldots Q_n x_n.\psi$ belongs to $Th(\mathbb{Q},\leqslant)$. Otherwise, it does not.

**Algorithmus 1** Recursive algorithm *check*(k, $\alpha$)

```
(* Given: 1 ⩽ k ⩽ n and α is a linear arrangement for x_1,...,x_k                *)
(* Checks whether α ⊨ φ_k = Q_{k+1} x_{k+1} ... Q_n x_n.ψ.                       *)

IF k = n THEN
  Check whether α ⊨ ψ.
  If so then return true. Otherwise return false.
ELSE
  IF Q_{k+1} = ∃ THEN
    FOR ALL extensions β for α DO
      IF check(k+1,β) THEN
        return true                          (* α ⊨ φ_k = ∃x_{k+1}.φ_{k+1} as β ⊨ φ_{k+1} *)
      FI
    OD
    return false      (* α ⊭ ∃x_{k+1}.φ_{k+1} as there is no extension β of α s.t. β ⊨ φ_{k+1} *)
  FI
  IF Q_{k+1} = ∀ THEN
    FOR ALL extensions β for α DO
      IF ¬ check(k+1,β) THEN
        return false                         (* α ⊭ ∀x_{k+1}.φ_{k+1} as β ⊭ φ_{k+1} *)
      FI
    OD
    return true          (* α ⊨ ∀x_{k+1}.φ_{k+1} as β ⊨ φ_{k+1} for all extensions β of α *)
  FI
FI
```

The depth of recursion is $n-1$. Thus, the space requirements are polynomial in $n$ and $\|\psi\|$. The time complexity of Algorithm 1.4.3 is, however, exponential in $n$, as it relies on a traversal of the tree spanned by the linear arrangements. This tree has height $n$ and the inner nodes of depth $k$ (representing linear arrangements for the first $k$ variables) have between 3 and $2k+1$ sons, provided that the inner FOR-loops does not abort before having considered all extensions of the current linear arrangement. However, we cannot expect a much more efficient algorithm to

check the truth of a FOL-sentence in $(\mathbb{Q}, \leqslant)$ since this problem is complete for the complexity class *PSPACE*.