1.7 Simple extensions of FOL

The results of the last two sections show that FO-definability is rather limited. We now study several extensions of FOL that are more powerful than FOL and expressive enough to characterize, e.g., reachability or the class of finite structures. The price we have to pay for increasing the expressiveness is the loss of nice properties like the compactness property or the upward Löwenheim-Skolem theorems.

1.7.1 Countable conjunctions and disjunctions

We now study the extension of FOL by countable disjunctions and conjunctions. We use the symbol FOL($\bigwedge$) for the resulting logic. (For historical reasons, this logic is often denoted by $\mathfrak{L}_{\omega 1\omega}$.) Formally, FOL($\bigwedge$) extends FOL by formulas of the form

$$\bigwedge_{i \in I} \phi_i$$

where $(\phi_i)_{i \in I}$ is a family of FOL($\bigwedge$)-formulas and I an arbitrary (possibly infinite) countable set of indices. As binary conjunction $\phi_1 \wedge \phi_2$ arises as a special case (with $I = \{1, 2\}$), it can be dropped from the syntax. Countable disjunctions are derived in the standard way:

$$\bigvee_{i \in I} \phi_i \stackrel{\text{def}}{=} \neg \big(\bigwedge_{i \in I} \neg \phi_i \big)$$

The semantics of $\bigwedge$ is defined in the expected way. That is, for each structure $\mathcal{A}$ and each variable valuation $\mathcal{V}$:

$$(\mathcal{A}, \mathcal{V}) \models \bigwedge_{i \in I} \phi_i \quad \text{iff} \quad (\mathcal{A}, \mathcal{V}) \models \phi_i \text{ for all } i \in I$$

Thus, $(\mathcal{A}, \mathcal{V}) \models \bigvee_{i \in I} \phi_i$ iff $(\mathcal{A}, \mathcal{V}) \models \phi_i$ for some $i \in I$. In particular:

$$\bigwedge_{i \in \varnothing} \phi_i \equiv \mathit{true} \quad \text{and} \quad \bigvee_{i \in \varnothing} \phi_i \equiv \mathit{false}$$

At first glance the generalization of binary conjunctions to countable conjunctions does not seem to be a big deal. However, FOL($\bigwedge$) is strictly more expressive than FOL. We illustrate this by providing several examples for FOL($\bigwedge$)-formulas that characterize properties of structures that are not FO-definable. We start with the class of finite structures that can be characterized by FOL($\bigwedge$) with equality for each vocabulary:

Example 1.7.1 (The class of finite structures is FOL($\bigwedge$)-definable). In the proof of Theorem 1.5.2 (page 72) we used the FOL-formulas

$$\psi_n \;=\; \exists x_1 \ldots \exists x_n . \bigwedge_{1 \leqslant i < j \leqslant n} (x_i \neq x_j)$$

to encode structures with n or more elements. We now consider the FOL($\bigwedge$)-formula

$$\psi_{\text{fin}} \stackrel{\text{def}}{=} \bigvee_{n \geqslant 2} \neg \psi_n$$

which holds exactly for finite structures. Note that if $\mathcal{A} \models \psi_{\text{fin}}$ then there exists some integer $n \geqslant 2$ with $\mathcal{A} \not\models \psi_n$. But then the domain of $\mathcal{A}$ has at most $n-1$ elements. In particular, $\mathcal{A}$ is finite. Vice versa, if $\mathcal{A}$ is a finite structure with n elements then $\mathcal{A} \not\models \psi_m$ for all $m > n$, and therefore $\mathcal{A} \models \psi_{\text{fin}}$. ■

Example 1.7.2 (The class of finite structures of even size is FOL($\bigwedge$)-definable). We saw in the previous section that for certain vocabularies there is no FOL-sentence that distinguishes the finite structures with an even number of elements from finite structures with an odd number of elements. We proved this for the empty vocabulary and the vocabulary of linear orders (see page 85 and page 112). However, there is a FOL($\bigwedge$)-sentence over the empty vocabulary that holds exactly for the finite structures of even size. For $n \geqslant 2$, let

$$\phi_n \ \stackrel{\text{def}}{=}\ \exists x_1 \ldots \exists x_n. \Big(\bigwedge_{1 \leqslant i < j \leqslant n} x_i \neq x_j \ \wedge\ \forall y. \bigvee_{1 \leqslant i \leqslant n} y = x_i \Big)$$

Clearly, $\mathcal{A} \models \phi_n$ iff the domain of $\mathcal{A}$ consists of exactly n elements. We now define:

$$\phi \ \stackrel{\text{def}}{=}\ \bigvee_{n \geqslant 1} \phi_{2n}$$

We then have $\mathcal{A} \models \phi$ iff there exists $n \geqslant 1$ such that $\mathcal{A} \models \phi_{2n}$ iff the domain of $\mathcal{A}$ consists of an even number of elements. A FOL($\bigwedge$)-characterization of the finite linear orders of even length is obtained by considering the conjunction of ϕ with the FOL-formulas that specify the conditions of linear orders. ■

Example 1.7.3 (FOL($\bigwedge$)-characterization of the natural numbers). Recall from Theorem 1.6.6 (page 86) that there is no FOL-formula over $\mathit{Voc}_{\text{nat}}$ that characterizes $(\mathbb{N}, \mathit{succ}, 0)$ up to isomorphism, where $\mathit{Voc}_{\text{nat}}$ denotes the vocabulary consisting of a constant symbol 0 and an unary function symbol *succ*. However, the class of structures that are isomorphic to $(\mathbb{N}, \mathit{succ}, 0)$ is FOL($\bigwedge$)-definable. Let ϕ be the conjunction of the following FOL($\bigwedge$)-formulas:

$$\forall x. \mathit{succ}(x) \neq 0$$

$$\forall x \forall y. (\mathit{succ}(x) = \mathit{succ}(y) \rightarrow x = y)$$

$$\forall x. \bigvee_{n \geqslant 0} x = \mathit{succ}^n(0)$$

Obviously, the natural numbers with the successor function *succ* and the constant 0 constitute a model for all three formulas. That is, $(\mathbb{N}, \mathit{succ}, 0) \models \phi$. It remains to show that each model for ϕ is isomorphic to $(\mathbb{N}, 0, \mathit{succ})$. Let $\mathcal{A} = (A, \mathit{succ}^{\mathcal{A}}, 0^{\mathcal{A}})$ be a model for ϕ and let

$$a_0 \stackrel{\text{def}}{=} 0^{\mathcal{A}}, \quad a_{n+1} \stackrel{\text{def}}{=} \mathit{succ}^{\mathcal{A}}(a_n) \text{ for all } n \geqslant 0.$$

Thus, a_n is the interpretation of the term $\mathit{succ}^n(0)$ under $\mathcal{A}$, i.e., $a_n = \mathit{succ}^n(0)^{\mathcal{A}}$.

We now show that the function $h : \mathbb{N} \to A$, $h(n) \stackrel{\text{def}}{=} a_n$ is an isomorphism from $(\mathbb{N}, \mathit{succ}, 0)$ to $\mathcal{A}$. Obviously, we have $h(0) = 0^{\mathcal{A}}$ and

$$h(\mathit{succ}(n)) \ = \ h(n+1) \ = \ a_{n+1} \ = \ \mathit{succ}^{\mathcal{A}}(a_n) \ = \ \mathit{succ}^{\mathcal{A}}(h(n))$$

for all $n \in \mathbb{N}$. Thus, h is a homomorphism. It remains to show that h is bijective. The third formula yields that all elements in $\mathcal{A}$ arise by finite applications of $succ^{\mathcal{A}}$ to $0^{\mathcal{A}} = a_0$. Hence:

$$A = \{a_n : n \geqslant 0\}$$

this shows that h is surjective. It remains to show that $a_n \neq a_m$ if $n < m$. Assuming that $a_n = a_m$ and $1 \leqslant n < m$ yields $a_{n-1} = a_{m-1}$ by the second formula. Repeating this argument we get:

$$a_0 = a_{m-n} = succ^{\mathcal{A}}(a_{m-n-1})$$

But this contradicts the first formula. ■

Other examples that illustrate that FOL($\bigwedge$) is more powerful than FOL are the definability of reachability and the characterization of periodic groups.

Example 1.7.4 (Reachability is FOL($\bigwedge$)-definable). To see that reachability is FOL($\bigwedge$)-definable, let $\phi_0(x,y)$ be the formula "$x = y$" and

$$\phi_n(x,y) \stackrel{\text{def}}{=} \exists z_0 \ldots \exists z_n. \big(x = z_0 \wedge y = z_n \wedge \bigwedge_{0 \leqslant i < n} E(z_i, z_{i+1})\big)$$

for $n \geqslant 1$. Intuitively, $\phi_n(x,y)$ states that node y is reachable from the node x by a path of length n, and the variables $z_0, z_1, \ldots, z_n$ encode a path from x to y of length n. (Recall that the length of a path in a graph denotes the total number of edges that are taken in that path.) Thus, reachability can be expressed in FOL($\bigwedge$) by

$$\bigvee_{n \geqslant 0} \phi_n(x,y),$$

stating there is some n and a path $z_0 z_1 \ldots z_n$ of length n which leads from $z_0 = x$ to $z_n = y$. ■

Example 1.7.5 (The class of periodic groups is FOL($\bigwedge$)-definable). The class of periodic groups can be characterized in FOL($\bigwedge$) by the conjunction of the group axioms (FOL-formula ϕ_{group} of Example 1.6.2 on page 83) with the formula

$$\forall x. \bigvee_{n \geqslant 1} x^n = e$$

which states that all elements are of finite order. ■

The fact that finite structures are FOL($\bigwedge$)-definable implies that the compactness theorem does not hold for FOL($\bigwedge$):

Theorem 1.7.6 (Lack of compactness property for FOL($\bigwedge$)). *There exists an unsatisfiable countable set of* FOL($\bigwedge$)*-sentences that is finitely satisfiable.*

Proof. We consider the FOL($\bigwedge$)-formula set

$$\mathfrak{F} \stackrel{\text{def}}{=} \{\psi_{\text{fin}}\} \cup \{\psi_n : n \geqslant 2\}$$

which is finitely satisfiable, but not satisfiable. Formula ψ_{fin} and the ψ_n's are defined as in Example 1.7.1 on page 118. That is, ψ_n characterizes all structures with at least n elements and ψ_{fin} is the disjunction of the formulas $\neg\psi_n$ and holds for exactly the finite structures. Unsatisfiability of $\mathfrak{F}$ is obvious. To see that $\mathfrak{F}$ is finitely satisfiable we observe that for any finite subset $\mathfrak{F}'$ of $\mathfrak{F}$ there is some $m \in \mathbb{N}$ such that $\psi_n \notin \mathfrak{F}'$ for all $n > m$. Thus:

$$\mathfrak{F}' \subseteq \{\psi_{fin}\} \cup \{\psi_2, \psi_3, \ldots, \psi_m\}$$

But then any finite structure with m (or more) elements is a model for $\mathfrak{F}'$. □

As FOL($\bigwedge$) has the power to characterize the class of finite structures, the upward Löwenheim-Skolem theorem *"from finite to infinite models"* (cf. Theorem 1.5.2 on page 72) is lost. To see this, we consider the FOL($\bigwedge$)-sentence ψ_{fin} which holds exactly for the finite structures. Then, for each $n \in \mathbb{N}$, the sentence ψ_{fin} has a finite model of size $\geqslant n$, but there is no infinite model. The fact that the natural numbers with the successor function and constant 0 have a characterization in FOL($\bigwedge$) yields that the upward Löwenheim-Skolem theorem *"from infinite to larger models"* (cf. Theorem 1.5.4 on page 73) does not hold for FOL($\bigwedge$). Recall from Example 1.7.3 on page 119 that there is a FOL($\bigwedge$)-sentence ϕ which holds for the structure $(\mathbb{N}, succ, 0)$, but ϕ does not have uncountable models (as all models for ϕ are isomorphic to $(\mathbb{N}, succ, 0)$).

Since the compactness theorem does not hold for FOL($\bigwedge$) there is *no deductive calculus* for FOL($\bigwedge$) which is sound and complete.

> Let us recall the arguments. Since proofs in any deductive calculus $\mathfrak{D}$ are finite, any formula ϕ that is $\mathfrak{D}$-provable from some formula-set $\mathfrak{F}$ is also $\mathfrak{D}$-provable from a finite subset $\mathfrak{G}$ of $\mathfrak{F}$. The soundness and completeness of $\mathfrak{D}$ then yields that $\mathfrak{F} \models \phi$ implies $\mathfrak{G} \models \phi$ for some finite subset $\mathfrak{G}$ of $\mathfrak{F}$. But this implies that each finitely satisfiable formula-set $\mathfrak{F}$ is satisfiable. Otherwise, i.e., if $\mathfrak{F}$ would not be satisfiable, then we would have $\mathfrak{F} \models \mathit{false}$, and therefore $\mathfrak{G} \models \mathit{false}$ for some finite subset $\mathfrak{G}$ of $\mathfrak{F}$. But then $\mathfrak{F}$ would not be finitely satisfiable.

However, this is no surprise since FOL($\bigwedge$)-formulas might have infinite length and so we should not expect to have finite proofs for all valid FOL($\bigwedge$)-formulas. We state here (without proof) that sound and complete deductive calculi for FOL($\bigwedge$) exist, if one allows for *infinite deductions* for formulas involving infinite conjunctions (or disjunctions). E.g., for a Hilbert-style calculus one might add the axiom schemata

$$\bigwedge_{i \in I} \Phi_i \to \Phi_j \quad \text{and} \quad \neg\Phi_j \to \neg \bigwedge_{i \in I} \Phi_i$$

for all $j \in I$. Although the compactness theorem and the upward Löwenheim-Skolem theorems do not hold for FOL($\bigwedge$), some properties of FOL are preserved when adding infinite conjunctions (and disjunctions). We consider here the downward Löwenheim-Skolem theorem.

Remark 1.7.7 (Subformulas and number of free variables). Since the length of FOL($\bigwedge$)-formulas can be infinite, the set of FOL($\bigwedge$)-formulas is uncountable. However, the set $\mathit{subf}(\phi)$ of subformulas of a given FOL($\bigwedge$)-formula ϕ is countable. This can be shown by structural induction. In case ϕ is $\bigwedge_{i \in I} \phi_i$, the set of subformulas of ϕ arise by the union of the set of subformulas of the ϕ_i's. Since countable unions of countable sets are countable, the induction hypothesis yields the claim.

For each FOL($\bigwedge$)-formula ϕ where $Free(\phi)$ is finite the variable-sets $Free(\psi)$ for the subformulas ψ of ϕ are finite too. In particular, if ϕ is a FOL($\bigwedge$)-sentence then $Free(\psi)$ is finite for all $\psi \in subf(\phi)$. To prove this, we may use an inductive argument to show that $Free(\psi)$ is finite for all subformulas ψ of a formula ϕ where $Free(\phi)$ is finite. If ϕ is $\bigwedge\limits_{i \in I} \phi_i$, then

$$Free(\phi) \;=\; \bigcup_{i \in I} Free(\phi_i),$$

but then with $Free(\phi)$ each of the sets $Free(\phi_i)$ must be finite too. ■

Theorem 1.7.8 (Downward Löwenheim-Skolem Theorem for FOL($\bigwedge$)). *Each satisfiable* FOL($\bigwedge$)*-sentence has a countable (possibly finite) model.*

Note that Theorem 1.7.8 also holds for countable sets of FOL($\bigwedge$)-sentences since we then may consider the formula that results from the conjunction of all formulas in the given set. The observation that finite structures are FOL($\bigwedge$)-definable yields that we cannot guarantee the existence of infinite countable models.

Proof. Let ϕ be a satisfiable FOL($\bigwedge$)-sentence. It is no restriction to assume that the underlying vocabulary has at least one constant symbol. If not then we simply extend the vocabulary by a fresh constant symbol and regard ϕ as a formula over the extended vocabulary. We show that if $\mathcal{B}$ is an uncountable model for ϕ then ϕ is satisfiable over some countable substructure $\mathcal{A}$ of $\mathcal{B}$. Let $B = Dom^{\mathcal{B}}$ be the domain of $\mathcal{B}$. We write $subf^{+}(\phi)$ for the set of subformulas of ϕ and their negations:

$$subf^{+}(\phi) \;\stackrel{\text{def}}{=}\; subf(\phi) \cup \{\neg\psi : \psi \in subf(\phi)\}$$

Recall from Remark 1.7.7 that $subf(\phi)$ is countable. Hence, $subf^{+}(\phi)$ is countable too. We define inductively a sequence $B_0, B_1, B_2, \ldots$ of countable subsets of B such that B_i is a subset of B_{i+1} and such that the following conditions (i) and (ii) hold:

(i) if $\psi = \psi(x_1, \ldots, x_n, x) \in subf^{+}(\phi)$ and $b_1, \ldots, b_n \in B_i$ such that

$$(\mathcal{B}, [x_1 := b_1, \ldots, x_n := b_n]) \models \exists x.\psi$$

then there exists some $a \in B_{i+1}$ with $(\mathcal{B}, [x_1 := b_1, \ldots, x_n := b_n, x := a]) \models \psi$.

(ii) for each n-ary function symbol f of *Voc* and all elements $b_1, \ldots, b_n \in B_i$ we have:

$$f^{\mathcal{B}}(b_1, \ldots, b_n) \in B_{i+1}$$

Recall that the notation $\psi(x_1, \ldots, x_n, x)$ indicates that $x_1, \ldots, x_n, x$ are pairwise distinct variables with $Free(\psi) \subseteq \{x_1, \ldots, x_n, x\}$. By Remark 1.7.7, all formulas in $subf^{+}(\phi)$ just have finitely many free variables. Thus, they have the form $\psi(x_1, \ldots, x_n, x)$. To treat (i) and (ii) simultaneously, we consider the countable formula-set

$$\mathfrak{F} \;\stackrel{\text{def}}{=}\; subf^{+}(\phi) \cup \mathfrak{G}$$

where

$$\mathfrak{G} \;\stackrel{\text{def}}{=}\; \bigcup_{n \geqslant 0} \; \bigcup_{f \in Func_n} \{\, x = f(x_1, \ldots, x_n) \;:\; x, x_1, \ldots, x_n \in Var \text{ are pairwise distinct} \,\}$$

Note that the formulas $x = f(x_1,\ldots,x_n)$ in $\mathfrak{G}$ have the form $\psi(x_1,\ldots,x_n,x)$ and they enjoy the property $(\mathcal{B},[x_1 := b_1,\ldots,x_n := b_n]) \models \exists x.\psi(x_1,\ldots,x_n,x)$ for all $b_1,\ldots,b_n \in B$. The fact that $\mathfrak{F}$ is countable follows from the observation that *subf*$^+(\phi)$ is countable (see above) and the assumption that the vocabulary and variable-set are recursively enumerable, and therefore countable.

Let B_0 be the set of the interpretations $c^{\mathcal{B}}$ of the constant symbols c in the given vocabulary. Given B_i, the definition of B_{i+1} is as follows. Let Λ_i be the set of all tuples

$$\lambda = (\psi, x_1,\ldots,x_n,x,b_1,\ldots,b_n)$$

where $\psi = \psi(x_1,\ldots,x_n,x) \in \mathfrak{F}$ and $b_1,\ldots,b_n \in B_i$ such that:

$$(\mathcal{B},[x_1 := b_1,\ldots,x_n := b_n]) \models \exists x.\psi$$

Then, Λ_i is countable as so are $\mathfrak{F}$ and B_i. For each tuple $\lambda \in \Lambda_i$, we pick an element $a_\lambda \in B$ such that

$$(\mathcal{B},[x_1 := b_1,\ldots,x_n := b_n, x := a_\lambda]) \models \psi$$

and define

$$B_{i+1} \stackrel{\text{def}}{=} B_i \cup \{\, a_\lambda : \lambda \in \Lambda_i \,\}.$$

Then, B_{i+1} is a countable, contains B_i and enjoys the desired properties (i) and (ii) by construction. Thus, $B_0 \subseteq B_1 \subseteq B_2 \subseteq \ldots$ is an increasing sequence of countable subsets of B.

Given the sequence $B_0, B_1, \ldots$, a countable model $\mathcal{A}$ for ϕ can be derived as follows. The domain of $\mathcal{A}$ is defined by:

$$A \stackrel{\text{def}}{=} \bigcup_{i \geqslant 0} B_i$$

Clearly, A is countable as the B_i's are countable. Moreover, if $b_1,\ldots,b_n \in A$ and f is an n-ary function symbol then $f^{\mathcal{B}}(b_1,\ldots,b_n) \in A$. This statement can be seen as follows. If $b_1,\ldots,b_n \in A$, then $b_1,\ldots,b_n \in B_i$ for some i. Obviously,

$$(\mathcal{B},[x_1 := b_1,\ldots,x_n := b_n]) \models \exists x.\underbrace{x = f(x_1,\ldots,x_n)}_{=\psi(x_1,\ldots,x_n,x)}$$

and $a \stackrel{\text{def}}{=} f^{\mathcal{B}}(b_1,\ldots,b_n)$ is the unique element in B such that

$$(\mathcal{B},[x_1 := b_1,\ldots,x_n := b_n, x := a]) \models x = f(x_1,\ldots,x_n).$$

The definition of B_{i+1} yields:

$$a = f^{\mathcal{B}}(b_1,\ldots,b_n) \in B_i' \subseteq B_{i+1} \subseteq A$$

Hence, $b_1,\ldots,b_n \in A$ and $f \in \mathit{Func}_n$ implies $f^{\mathcal{B}}(b_1,\ldots,b_n) \in A$. This shows that we can define a countable substructure $\mathcal{A}$ of $\mathcal{B}$ with domain A as follows:

$$\begin{array}{lcll} P^{\mathcal{A}} & \stackrel{\text{def}}{=} & P^{\mathcal{B}} \cap A^n & \text{for each } n\text{-ary predicate symbol } P, \\ f^{\mathcal{A}} & \stackrel{\text{def}}{=} & f^{\mathcal{B}}|_{A^n} & \text{for each } n\text{-ary function symbol } f. \end{array}$$

It remains to show that $\mathcal{A}$ is a model for ϕ. This follows from the observation that for each $\psi \in \mathit{subf}^+(\phi)$ and variable valuation $\mathcal{V} : \mathit{Var} \to A$ we have:

$$(\mathcal{A}, \mathcal{V}) \models \psi \quad \text{iff} \quad (\mathcal{B}, \mathcal{V}) \models \psi$$

The proof for this statement can be provided by structural induction. We then may consider $\phi = \psi$ and get $\mathcal{A} \models \phi$, since $\mathcal{B} \models \phi$. $\square$

Note that the proof of Theorem 1.7.8 yields an alternative proof of the downward Löwenheim-Skolem theorem for FOL-formulas without equality (see Theorem 1.5.5 on page 75).