Fagin's theorem

In Theorem 2.4.4 we saw that each ESO-definable graph property is decidable in nondeterministic polynomial time. We now establish a stronger result stating that ESOL even provides a full characterization of the complexity class *NP* in the sense that the ESO-definable properties for finite structures agree with the problems in *NP*. This result is due to Fagin (1974) and states that ESOL *captures NP*, which means that the following two statements (1) and (2) hold:

(1) Given an ESOL-sentence ϕ, there is a nondeterministic algorithm that takes as input a finite structure $\mathcal{A}$ and checks whether $\mathcal{A} \models \phi$ in time $\mathcal{O}(\text{poly}(\mathit{size}(\mathcal{A})))$ (the formula ϕ is viewed to be fixed).

(2) For each class $\mathcal{C}$ of finite structures for some finite purely relational vocabulary *Voc* that is closed under isomorphism and where membership to $\mathcal{C}$ can be checked in nondeterministic polynomial time there exists an ESOL-sentence ϕ such that

$$\mathcal{A} \models \phi \quad \text{iff} \quad \mathcal{A} \in \mathcal{C}$$

for each finite structure $\mathcal{A}$.

When we consider the elements of the complexity class *NP* as structures over finite, purely relational vocabularies (rather than formal languages over finite words) then (1) and (2) state that for each finite, purely relational vocabulary *Voc* and class $\mathcal{C}$ of finite structures over *Voc* which is closed under isomorphism:

$$\begin{array}{lll} \mathcal{C} \in NP & \text{iff} & \text{there exists an ESOL-sentence } \phi \text{ over } Voc \text{ that defines } \mathcal{C}, \\ & & \text{i.e., for all finite structures } \mathcal{A} \text{ over } Voc\text{:} \\ & & \qquad \mathcal{A} \in \mathcal{C} \Longleftrightarrow \mathcal{A} \models \phi \end{array}$$

For example, consider the satisfiability problem of propositional logic (SAT). The elements of SAT are satisfiable propositional formulas. In the context of Fagin's Theorem, we identify any propositional formula with its syntax graph. The syntax graph is a rooted acyclic graph where the nodes represent the subformulas. More precisely, the nodes representing a subformula of length $\geqslant 1$ are labeled with the outermost operator ($\wedge$ or $\neg$ if just conjunction and negation are used as basic operators in the syntax of propositional formulas). The leaves (i.e., terminal nodes) stand for atoms (i.e., boolean variables) or the constants *true* or *false*. The syntax graph can be viewed as a structure over the vocabulary consisting of a binary predicate symbol E for the edge relation and several unary predicate symbols for the labeling of the nodes. E.g., one might use one predicate symbol P_{root} for the singleton set consisting of the root node, unary predicate symbols $P_\wedge$, $P_\neg$ for the boolean operators $\wedge$ and $\neg$ and unary predicate symbols for all boolean variables and the constants *true* or *false*.

For many decision problems a similar encoding of the inputs by finite structures can be provided. Indeed, any language L consisting of finite words over some alphabet Σ has a representation as a class $\mathcal{C}$ of finite structures over some finite and purely relational vocabulary. Such a representation relies on an encoding of finite words by finite labeled graphs. (This will be discussed in Section 2.6.) This permits to treat the elements of the complexity class *NP* as classes of finite structures over some finite and purely relational vocabulary.

Theorem 2.4.7 (Fagin's theorem). *ESOL captures NP.*

We do not present here the full proof for Fagin's theorem. We just mention that condition (1) can easily be shown by an argument that is roughly the same as for the special case of graphs (see Theorem 2.4.4). Given an ESOL-sentence

$$\phi = \exists X_1 \ldots \exists X_n.\psi,$$

where ψ is a FOL-sentence, and a finite structure $\mathcal{A}$ with domain A then we first guess relations $R_1, \ldots, R_n$ for $X_1, \ldots, X_n$. More precisely, $R_i \subseteq A^{k_i}$ where k_i is the arity of X_i. This can be done by guessing $|A|^{k_1} + \ldots + |A|^{k_n}$ bits. We then check in deterministic polynomial time whether ψ holds under the structure given by $\mathcal{A}$ and the interpretation R_i for X_i, $i = 1, \ldots, n$. This is possible with an algorithm that operates similar to the one proposed in the proof of Theorem 1.6.8 on page 88.

Fagin's Theorem can be understood as a generalization of Cook's Theorem stating the *NP*-completeness of SAT (the satisfiability problem of propositional logic). The fact that SAT is in *NP* is obvious. The *NP*-hardness of SAT is obtained by Fagin's Theorem as follows.

Corollary 2.4.8 (Cook's Theorem, NP-hardness of SAT). *SAT (the satisfiability problem for propositional logic) is NP-hard.*

Proof. Let $\mathcal{C}$ be a problem in *NP*. We have to show that $\mathcal{C}$ is polynomially reducible to SAT. We may regard $\mathcal{C}$ as a property of finite structures over some finite purely relational vocabulary *Voc*. By Fagin's Theorem there exists an ESOL-sentence ϕ over *Voc* such that for each finite structure $\mathcal{A}$ over *Voc* we have:

$$\mathcal{A} \in \mathcal{C} \quad \text{iff} \quad \mathcal{A} \models \phi$$

The goal is provide a *deterministic polynomial time* algorithm that takes as input a finite structure $\mathcal{A}$ over *Voc* and constructs a propositional formula $\varphi_{\mathcal{A}}$ such that

$$\varphi_{\mathcal{A}} \text{ is satisfiable} \quad \text{if and only if} \quad \mathcal{A} \models \phi.$$

The construction of $\varphi_{\mathcal{A}}$ is similar to the polynomial reduction from ϕ-GRAPH for ESOL Horn expressions to the satisfiability problem for propositional Horn formulas (see Theorem 2.4.6 on page 154). Let

$$\phi = \exists X_1 \ldots \exists X_r.\psi(X_1, \ldots, X_r)$$

where ψ is a FOL-sentence over *Voc* extended by $X_1, \ldots, X_r$ as new predicate symbols. Let m_i be the arity of X_i. Given a finite structure $\mathcal{A}$ for *Voc* with domain A, the propositional formula $\varphi_{\mathcal{A}}$ is obtained from the FOL-sentence $\psi(X_1, \ldots, X_r)$ by a two-step transformation:

$$\psi \rightsquigarrow \psi_{\mathcal{A}} \rightsquigarrow \varphi_{\mathcal{A}}$$

We first replace each first-order quantification $\forall x.\theta$ in ψ with

$$\bigwedge_{a \in A} \theta[x/a]$$

where the elements $a \in A$ are treated as constant symbols. The so obtained SOL-formula $\psi_{\mathcal{A}} = \psi_{\mathcal{A}}(X_1, \ldots, X_r)$ is quantifier-free and its atoms have the form:

- $P(\overline{a})$ for $P \in Pred_k$ and $\overline{a} \in A^k$,
- or $a_1 = a_2$ for elements $a_1, a_2 \in A$
- or $X_i(\overline{a})$ where $1 \leqslant i \leqslant r$ and $\overline{a} = (a_1, \ldots, a_{m_i})$ is a m_i-tuple of elements in A.

(Recall that m_i is the arity of X_i.) We now treat the atoms $X_i(\overline{a})$ as boolean variables. The idea is that truth value 1 for $X_i(\overline{a})$ indicates that $\overline{a}$ is a tuple in the relation R_i for X_i, while truth value 0 for $X_i(\overline{a})$ means that $\overline{a} \notin R_i$. The other atoms $P(a_1, \ldots, a_k)$ and $a_1 = a_2$ where P is a k-ary predicate symbol in *Voc* and the a_i's are elements in A can be evaluated over $\mathcal{A}$. That is, we replace

- each atomic formula $P(\overline{a})$ where P is a k-ary predicate symbol and $\overline{a} \in A^k$ by *true* or *false*, depending on whether $\overline{a} \in P^{\mathcal{A}}$,
- each formula $a_1 = a_2$ where $a_1, a_2 \in A$ with either *true* or *false*, depending on whether a_1 and a_2 agree.

The resulting formula $\varphi_{\mathcal{A}}$ can be regarded as a propositional formula over the atoms $X_i(\overline{a})$ with $1 \leqslant i \leqslant r$ and $\overline{a} \in A^{m_i}$. The length of $\varphi_{\mathcal{A}}$ is bounded by

$$|\varphi_{\mathcal{A}}| = \mathcal{O}\big(|\psi| \cdot |A|^{qr(\psi)}\big) = \mathcal{O}(poly(|A|),$$

as ϕ (and therefore ψ) is supposed to be fixed. Remind that $qr(\psi)$ denotes the quantifier rank of ψ (see Notation 1.6.20 on page 97) and notice that in the transformation $\psi \rightsquigarrow \varphi_{\mathcal{A}}$ each universal FO-quantifier $\forall x.\theta$ of ψ is replaced with a conjunction of $|A|$ subformulas $\theta[x/a]$. Thus, given $\mathcal{A}$, the propositional formula $\varphi_{\mathcal{A}}$ can be constructed from ϕ in time polynomial in $|A|$.

It remains to show that $\varphi_{\mathcal{A}}$ is satisfiable if and only if $\mathcal{A}$ belongs to $\mathcal{C}$. This relies on the fact that there is a one-to-one correspondence between the r-tuples $(R_1, \ldots, R_r)$ consisting of relations $R_i \subseteq A^{m_i}$ for the predicate variables X_i and the assignments for the boolean variables $X_i(a_1, \ldots, a_{m_i})$ that might appear in $\varphi_{\mathcal{A}}$. This yields:

$\mathcal{A} \in \mathcal{C}$

iff $\mathcal{A} \models \phi = \exists X_1 \ldots \exists X_r.\psi(X_1, \ldots, X_r)$

iff there exist relations $R_i \subseteq A^{m_i}$ such that $(\mathcal{A}, [X_1 := R_1, \ldots, X_r := R_r]) \models \psi$

iff there exist relations $R_i \subseteq A^{m_i}$ such that $(\mathcal{A}^+, [X_1 := R_1, \ldots, X_r := R_r]) \models \psi_{\mathcal{A}}$

(where $\psi_{\mathcal{A}}$ is viewed as a quantifier-free SOL-formula over *Voc* extended by new constant symbols $a \in A$ and $\mathcal{A}^+$ extends $\mathcal{A}$ by the interpretation $a^{\mathcal{A}^+} = a$ for the auxiliary constant symbol $a \in A$)

iff there exists an assignment for the boolean variables $X_i(\overline{a})$ that is satisfying for φ_A

iff $\varphi_{\mathcal{A}}$ is satisfiable

Thus, the transformation $\mathcal{A} \rightsquigarrow \varphi_{\mathcal{A}}$ yields a polynomial time-bounded reduction from $\mathcal{C}$ (viewed as a problem in *NP*) to SAT. □

2.5 Weak second-order logic

In the previous sections, we saw that second-order logic provides an elegant and very powerful framework to specify properties of structures. However, none of the "nice" properties of FOL, such as the existence of sound and complete proof systems, the compactness theorem or the Löwenheim-Skolem theorems, are preserved. This motivates the consideration of alternative semantics for SOL. That is, we keep the elegant syntax of SOL, but modify the meaning of second-order quantification by imposing certain restrictions on the range of the second-order variables. One example is weak SOL where the semantics of second-order quantification for predicate variables is restricted to *finite* relations.

The syntax of weak SOL is the same as for SOL, except that we do not allow for function variables (and function quantifiers). That is, terms are as in FOL. The syntax of weak SOL formulas extends the one for FOL by atoms $X(t_1, \ldots, t_n)$ with X is a predicate variable of arity n and by SO-quantification $\forall X.\phi$. Since we have negation, existential quantification over predicate variables can be derived in the standard way by $\exists X.\phi \stackrel{\text{def}}{=} \neg\forall X.\neg\phi$. The semantics of weak SOL is given by a satisfaction relation $\models_{\text{weak}}$ that specifies truth values for formulas over interpretations $\mathcal{I} = (\mathcal{A}, \mathcal{W})$ where $\mathcal{A}$ is a structure (as for FOL and standard SOL) and $\mathcal{W}$ a variable valuation that assigns elements of $\mathcal{A}$'s domain to the FO-variables and *finite* relations of arity k to each k-ary predicate variable. The definition of the satisfaction relation $\models_{\text{weak}}$ is by structural induction. The semantics of the FOL-fragment (atomic formulas, conjunctions $\phi_1 \wedge \phi_2$, negations $\neg\phi$ and first-order quantification $\forall x.\phi$) is as usual. The only difference between the weak satisfaction relation $\models_{\text{weak}}$ and the standard SOL-satisfaction relation $\models$ is that in weak SOL second-order predicate quantification ranges over finite relations. That is, if X is an n-ary predicate variable then:

$$\mathcal{I} \models_{\text{weak}} \forall X.\phi \quad \text{iff} \quad \begin{cases} \text{for all finite relations } R \subseteq A^n: \\ \qquad \mathcal{I}[X := R] \models_{\text{weak}} \phi \end{cases}$$

Here, A is the domain of $\mathcal{I}$, i.e., $A = \mathit{Dom}^{\mathcal{A}}$ if $\mathcal{I} = (\mathcal{A}, \mathcal{W})$. Obviously, we then have:

$$\mathcal{I} \models_{\text{weak}} \exists X.\phi \quad \text{iff} \quad \begin{cases} \text{there exists a finite relation } R \subseteq A^n \\ \text{such that } \mathcal{I}[X := R] \models_{\text{weak}} \phi \end{cases}$$

Clearly, the satisfaction relations $\models$ and $\models_{\text{weak}}$ agree for finite structures, while they differ for infinite structures. For example, consider the formula

$$\phi \stackrel{\text{def}}{=} \exists X \forall y.(\, P(y) \leftrightarrow X(y)\,)$$

where X is a unary predicate variable and P a unary predicate symbol. Then,

$$\mathcal{A} \models_{\text{weak}} \phi \quad \text{iff} \quad P^{\mathcal{A}} \text{ is finite,}$$

while ϕ is a tautology under the standard semantics. Thus, if $\mathcal{A} = (A, P^{\mathcal{A}})$ is an infinite structure where $P^{\mathcal{A}} = A$ then $\mathcal{A} \models \phi$, but $\mathcal{A} \not\models_{\text{weak}} \phi$. This yields:

Lemma 2.5.1. *There exists a SOL-sentence ϕ without function variables and a structure $\mathcal{A}$ such that $\mathcal{A} \not\models_{\text{weak}} \phi$, while $\mathcal{A} \models \phi$.*

The SOL-characterization of finite structures in Example 2.3.1 on page 145 relies on the observation that, for each nonempty set A, the finiteness of A is equivalent to the statement that each injective function $h : A \to A$ is surjective as well as to the statement that each total, transitive binary relation on A intersects the diagonal $\{(a,a) : a \in A\}$. These characterization of finite sets are no longer appropriate for weak SOL as the former requires quantification over function symbols and the latter requires quantification over binary predicate variables to be interpreted by infinite relations when the domain is infinite. However, there is even a simpler characterization of finite structures by means of the following weak SOL-sentence over the empty vocabulary:

$$\phi^{w}_{\text{fin}} \stackrel{\text{def}}{=} \exists X \forall y.\, X(y).$$

Under the weak SOL-semantics, ϕ^{w}_{fin} asserts the existence of a finite set that covers all elements of the domain. This, of course, is equivalent to the finiteness of the domain. Under the standard semantics, ϕ^{w}_{fin} is a tautology since for each structure $\mathcal{A}$ with domain A, $(\mathcal{A}, [X := A]) \models \forall y.X(y)$. Thus, for all structures $\mathcal{A}$:

$$\mathcal{A} \models_{\text{weak}} \phi^{w}_{\text{fin}} \quad \text{iff} \quad \mathcal{A} \text{ is finite,}$$

while $\mathcal{A} \models \phi^{w}_{\text{fin}}$ for all structures $\mathcal{A}$. We write $\equiv_{\text{weak}}$ to denote the equivalence of weak SOL-formulas. That is:

$$\phi \equiv_{\text{weak}} \psi \quad \text{iff} \quad \left\{ \begin{array}{c} \text{for all interpretations } \mathcal{I} = (\mathcal{A}, \mathcal{W}): \\ \mathcal{I} \models_{\text{weak}} \phi \iff \mathcal{I} \models_{\text{weak}} \psi \end{array} \right.$$

Similarly, $\Vdash_{\text{weak}}$ stands for the consequence relation induced by the weak SOL semantics. If $\mathfrak{F} \cup \{\psi\}$ is a set of weak SOL formulas then:

$$\mathfrak{F} \Vdash_{\text{weak}} \psi \quad \text{iff} \quad \left\{ \begin{array}{c} \text{for all interpretations } \mathcal{I} = (\mathcal{A}, \mathcal{W}): \\ \text{if } \mathcal{I} \models_{\text{weak}} \mathfrak{F} \text{ then } \mathcal{I} \models_{\text{weak}} \psi \end{array} \right.$$

Formula ϕ^{w}_{fin} can be used to show that $\equiv_{\text{weak}}$ and $\Vdash_{\text{weak}}$ are not the same as standard equivalence or consequence relation, since, e.g., we have:

$$\mathit{true} \equiv \exists X \forall y.\, X(y), \quad \text{while} \quad \mathit{true} \not\equiv_{\text{weak}} \exists X \forall y.\, X(y)$$

Since weak SOL is powerful enough to distinguish between infinite and finite structures (e.g., the formula ϕ^{w}_{fin} above), weak SOL does not meet the compactness property. To see this, we use the same argument as in Theorem 2.3.2 (page 146). We consider the set

$$\mathfrak{F} \stackrel{\text{def}}{=} \{\phi^{w}_{\text{fin}}\} \cup \{\psi_2, \psi_3, \psi_4, \ldots\}$$

where ψ_n is a FOL-sentence that characterizes the structures with n or more elements. Then, $\mathfrak{F}$ is finitely satisfiable, but has no model with respect to $\models_{\text{weak}}$. This shows that the compactness property does not hold for weak SOL.

Theorem 2.5.2 (Lack of compactness property for weak SOL). *There is a set of weak SOL-sentences that is finitely satisfiable, but not satisfiable.*

As a consequence we get that weak SOL does not have sound and complete proof systems. Formula ϕ^{w}_{fin} is a weak SOL-sentence that has finite models of arbitrary size, but there is no infinite model. Hence, the upward Löwenheim-Skolem theorem *"from finite to infinite models"* does not hold for weak SOL. Also the upward Löwenheim-Skolem theorem *"from infinite to larger models"* does not hold for weak SOL since the structures isomorphic to $(\mathbb{N}, \leqslant)$ have a characterization in weak SOL:

Example 2.5.3 (Characterization of the ordered natural numbers in weak SOL). The class of structures over the vocabulary consisting of a single binary predicate symbol $\sqsubseteq$ that are isomorphic to $(\mathbb{N}, \leqslant)$ is definable in weak SOL. To provide such a weak SOL-sentence that covers the essential features of $(\mathbb{N}, \leqslant)$, we reuse the characterization of $(\mathbb{N}, \leqslant)$ presented in Example 1.7.10 on page 129. That is, we use the fact that $(\mathbb{N}, \leqslant)$ is the unique linear order $\mathcal{A} = (A, \sqsubseteq)$ without maximal elements such that for each element $a \in A$ the downward closure $A \downarrow a$ is finite. (Uniqueness is up to isomorphism.) Let ϕ^{w}_{nat} be the following weak SOL-sentence:

$$\phi^{w}_{\text{nat}} \quad \stackrel{\text{def}}{=} \quad \phi_{\text{lin\_order}} \ \wedge \ \underbrace{\forall x \exists y. x < y}_{\substack{\text{no maximal} \\ \text{element}}} \ \wedge \ \underbrace{\forall x \exists Y \forall y. (y \leqslant x \to Y(y))}_{\substack{\text{the downward closure of} \\ \text{each element is finite}}}$$

Here, $\phi_{\text{lin\_order}}$ is a FOL-sentence that formalizes the conditions of a linear order (see Definition 1.4.13 on page 53). As in Example 1.7.10 on page 129 it can now be shown that for each structure $\mathcal{A} = (A, \sqsubseteq)$ we have: $\mathcal{A} \models_{\text{weak}} \phi^{w}_{\text{nat}}$ if and only if $\mathcal{A}$ is isomorphic to $(\mathbb{N}, \leqslant)$. ■

Although the weak SOL-semantics differs from the standard semantics, weak SOL can be viewed as a sublogic of standard SOL:

Theorem 2.5.4 (Embedding of weak SOL into SOL). *For each weak SOL-sentence ϕ there exists a SOL-sentence ϕ' over the same vocabulary such that ϕ and ϕ' have the same models, i.e., for all structures $\mathcal{A}$:*

$$\mathcal{A} \models_{\text{weak}} \phi \quad \textit{iff} \quad \mathcal{A} \models \phi'$$

Proof. Let X be a predicate variable of arity n and let $\psi_{\text{fin}}(X)$ be a SOL-formula stating that X is finite. I.e., for all interpretations $\mathcal{I}$:

$$\mathcal{I}[X := R] \models \psi_{\text{fin}}(X) \quad \text{iff} \quad R \text{ is finite}$$

For the definition of $\psi_{\text{fin}}(X)$ we can use the characterization of (in)finite sets as in Example 2.3.1 on page 145. If X is n-ary then

$$\psi_{\text{fin}}(X) \ \stackrel{\text{def}}{=} \ \forall Y. \big(\psi_{\text{tt}}(X,Y) \ \to \ \exists \overline{z}. Y(\overline{z},\overline{z}) \big)$$

where Y is a $2n$-ary predicate variable and $\overline{z}$ an n-tuple of pairwise distinct FO-variables. Subformula $\psi_{\text{tt}}(X,Y)$ asserts that Y stands for a binary relation on X that is total and transitive:

$$\begin{aligned} \psi_{\text{tt}}(X,Y) \ \stackrel{\text{def}}{=} \ & \forall \overline{z} \forall \overline{u}. \big(Y(\overline{z},\overline{u}) \to X(\overline{z}) \wedge X(\overline{u}) \big) \ \wedge \ \forall \overline{z}. \big(X(\overline{z}) \to \exists \overline{u}. Y(\overline{z},\overline{u}) \big) \ \wedge \\ & \forall \overline{z} \forall \overline{u} \forall \overline{v}. \big(Y(\overline{z},\overline{u}) \wedge Y(\overline{u},\overline{v}) \ \to \ Y(\overline{z},\overline{v}) \big) \end{aligned}$$

Given a weak SOL-sentence ϕ, the definition of an equivalent SOL-sentence ϕ' relies on the syntactic replacement of each second-order predicate quantification $\forall X.\psi$ in ϕ with the formula

$$\forall X. \big(\psi_{\text{fin}}(X) \to \psi\big)$$

This corresponds to the replacement of $\exists X.\psi$ with $\exists X. \big(\psi_{\text{fin}}(X) \wedge \psi\big)$. Formally we define the transformation "weak SOL-formula $\phi \rightsquigarrow$ SOL-formula ϕ'" by structural induction, see Figure 17.

$$\begin{array}{rcll}
\mathit{true}' & \stackrel{\text{def}}{=} & \mathit{true} & \\
P(\bar{t})' & \stackrel{\text{def}}{=} & P(\bar{t}) & \text{for } P \in \mathit{Pred} \\
X(\bar{t})' & \stackrel{\text{def}}{=} & X(\bar{t}) & \text{for } X \in \mathit{PVar} \\
(\phi_1 \wedge \phi_2)' & \stackrel{\text{def}}{=} & \phi_1' \wedge \phi_2' & \\
(\neg\phi)' & \stackrel{\text{def}}{=} & \neg\phi' & \\
(\forall y.\phi)' & \stackrel{\text{def}}{=} & \forall y.\phi' & \\
(\forall X.\phi)' & \stackrel{\text{def}}{=} & \forall X. (\psi_{\text{fin}}(X) \to \phi') &
\end{array}$$

Figure 17: Transformation of weak SOL formula ϕ into an equivalent SOL formula ϕ'

Recall that weak SOL does not use function quantification. Obviously, ϕ and ϕ' have the same free variables. It is now easy to see that for each structure $\mathcal{A}$ and variable valuation $\mathcal{W}$ that assigns finite relations to all predicate variables, $(\mathcal{A},\mathcal{W})$ is a model for ϕ under the weak SOL semantics if and only if $(\mathcal{A},\mathcal{W}) \models \phi'$. Thus, if ϕ is a sentence then so is ϕ'. Furthermore, ϕ and ϕ' have the same models when ϕ is interpreted with respct to $\models_{\text{weak}}$ and ϕ' with respect to the standard SOL satisfaction relation $\models$. $\square$

Unlike standard SOL, satisfiable sets of weak SOL-sentences are satisfiable over countable structures. That is, the downward Löwenheim-Skolem theorem holds for weak SOL. We show this by providing that an embedding of weak SOL into FOL with countable conjunctions (see Section 1.7.1):

Theorem 2.5.5 (Embedding of weak SOL into FOL($\bigwedge$)). *For each weak SOL-sentence ϕ there exists a* FOL($\bigwedge$)*-sentence ϕ' over the same vocabulary such that ϕ and ϕ' have the same models.*

Proof. Let ϕ be a weak SOL-sentence. For each k-ary predicate variable X that appears in ϕ we choose infinitely many k-tuples $\bar{x}^1, \bar{x}^2, \ldots$ of fresh first-order variables, i.e., $\bar{x}^i = (x_1^i, \ldots, x_k^i)$ where the x_j^i's are pairwise distinct and do not appear in ϕ. Furthermore, we require that the x_j^i's chosen for X are pairwise distinct from the corresponding variables chosen for another predicate variable Y that appears in ϕ.

The FOL($\bigwedge$)-sentence ϕ' is generated from ϕ using an inductive approach. Given a subformula ψ of ϕ and a function $\sigma : \mathit{PVar} \to \mathbb{N}$ we will define a "corresponding" FOL($\bigwedge$)-formula ψ^σ. Intuitively, ψ^σ arises from ψ when interpreting each predicate variable X with the finite set $\{\bar{x}^1, \ldots, \bar{x}^m\}$ where $m = \sigma(X)$.

If ψ is *true* or an atomic formula $P(t_1,\ldots,t_n)$ for some n-ary predicate symbol P or an atomic formula of the form $t_1 = t_2$, then the parameter σ has no effect. That is, in these cases ψ^σ is defined as follows:

$$\begin{aligned} \mathit{true}^\sigma &\stackrel{\text{def}}{=} \mathit{true} \\ P(t_1,\ldots,t_n)^\sigma &\stackrel{\text{def}}{=} P(t_1,\ldots,t_n) \\ (t_1 = t_2)^\sigma &\stackrel{\text{def}}{=} t_1 = t_2 \end{aligned}$$

where P is a n-ary predicate symbol. We now consider an atomic formula of the form $X(\bar{t})$ where X is a k-ary predicate variable and $\bar{t} = (t_1,\ldots,t_k)$ an k-tuple of terms. The intuitive meaning of the parameter σ is reflected in the definition of $X(\bar{t})^\sigma$:

$$X(\bar{t})^\sigma \stackrel{\text{def}}{=} \bigvee_{1\leqslant i\leqslant m} (\bar{t} = \bar{x}^i) \qquad \text{where } m = \sigma(X)$$

Here, $\bar{t} = \bar{x}^i$ abbreviates $\bigwedge_{1\leqslant j\leqslant k} (t_j = x^i_j)$. For the special case $\sigma(X) = 0$ we have $X(\bar{t})^\sigma \stackrel{\text{def}}{=} \mathit{false}$.

For formulas where the outermost operator is a conjunction, negation or first-order quantification, the definition is straightforward:

$$\begin{aligned} (\phi_1 \wedge \phi_2)^\sigma &\stackrel{\text{def}}{=} \phi_1^\sigma \wedge \phi_2^\sigma \\ (\neg\psi)^\sigma &\stackrel{\text{def}}{=} \neg\psi^\sigma \\ (\forall x.\psi)^\sigma &\stackrel{\text{def}}{=} \forall x.\psi^\sigma \end{aligned}$$

It remains to explain the treatment of second-order predicate quantification. If X does not appear free in ψ, then $\forall X.\psi$ is equivalent to ψ and the parameter σ has no effect on the SO-quantifier $\forall X$. Thus:

$$(\forall X.\psi)^\sigma \stackrel{\text{def}}{=} \psi^\sigma \qquad \text{if } X \notin \mathit{PFree}(\psi)$$

Suppose now that $X \in \mathit{PFree}(\psi)$. In this case, $(\forall X.\psi)^\sigma$ states that ψ holds for all $m \in \mathbb{N}$ and all sets X with m or less elements. Formally:

$$(\forall X.\psi)^\sigma \stackrel{\text{def}}{=} \psi^{\sigma[X:=0]} \wedge \bigwedge_{m\geqslant 1} \forall\bar{x}^1 \ldots \forall\bar{x}^m.\psi^{\sigma[X:=m]}$$

where $\sigma[X := m]$ denotes the function that agrees with σ on all predicate variables $Y \neq X$ and assigns m to X. It is easy to see that $\psi^\sigma = \psi^\rho$ if σ, $\rho : \mathit{PVar} \to \mathbb{N}$ agree on all predicate variables $X \in \mathit{PFree}(\psi)$. Hence, for the given weak SOL-sentence ϕ, the FOL($\bigwedge$)-sentence ϕ' corresponding to ϕ can be defined by

$$\phi' \stackrel{\text{def}}{=} \phi^\sigma \ \text{ for an arbitrary function } \sigma : \mathit{PVar} \to \mathbb{N}.$$

The equivalence of ϕ and ϕ' can be established by showing that for each subformula ψ of ϕ and each function σ as above we have:

$$(\mathcal{A}, \mathcal{W}) \models_{\text{weak}} \psi \quad \text{iff} \quad (\mathcal{A}, \mathcal{W}') \models \psi^\sigma$$

where $\mathcal{A}$ is an arbitrary structure and $\mathcal{W}$, $\mathcal{W}'$ variable valuations for the free variables of ψ and ψ^σ, respectively, such that $\mathcal{W}(x) = \mathcal{W}'(x)$ for all first-order variables $x \in \mathit{Free}(\psi)$ and

$$\mathcal{W}(X) \;=\; \left\{ \mathcal{W}'(\overline{x}^1), \ldots, \mathcal{W}'(\overline{x}^m) \right\} \quad \text{if } X \in \mathit{PFree}(\psi) \text{ and } \sigma(X) = m.$$

This can be shown by structural induction. For the given weak SOL-sentence ϕ, we obtain the desired result stating that $\mathcal{A} \models_{\text{weak}} \phi$ iff $\mathcal{A} \models \phi'$ for all structures $\mathcal{A}$. □

By the above result and Theorem 1.7.8 (page 122) we obtain:

Corollary 2.5.6 (Downward Löwenheim-Skolem Theorem for weak SOL). *Each satisfiable weak SOL-sentence has a countable model.*

As a consequence we get that weak SOL is strictly less expressive than standard SOL. For instance, the SOL-sentence $\neg\phi_{\text{ctbl}}$ of Example 2.3.3 (on page 146) that characterizes the uncountable sets has no equivalent weak SOL-sentence.