

Automated Theorem Proving

Prof. Dr. Jasmin Blanchette, Yiming Xu, PhD,
Tanguy Bozec, and Lydia Kondylidou
based on exercises by Dr. Uwe Waldmann

Winter Term 2025/26

Exercises 4: First-Order Logic

Exercise 4.1: Let $\Sigma = (\{b/0, c/0, d/0, f/1\}, \{P/1\})$. Does the formula

$$P(b) \wedge P(c) \wedge \neg P(d) \wedge \neg(\exists x P(f(f(x))))$$

have a Σ -model whose universe has exactly two elements? Give an example of such a model or show that such a model does not exist.

Proposed solution. The Σ -algebra \mathcal{A} with $U_{\mathcal{A}} = \{2, 3\}$, $b_{\mathcal{A}} = 2$, $c_{\mathcal{A}} = 2$, $d_{\mathcal{A}} = 3$, $f_{\mathcal{A}}(u) = 3$ for all $u \in U_{\mathcal{A}}$, and $P_{\mathcal{A}} = \{2\}$ is a model of the given formula. Its universe has two elements.

Exercise 4.2: Let the signature $\Sigma = (\Omega, \Pi)$ be given by $\Omega = \{+/2, s/1, 0/0\}$ and $\Pi = \emptyset$, and let

$$F_1 = \forall x (x + 0 \approx x)$$

$$F_2 = \forall x \forall y (x + s(y) \approx s(x + y))$$

$$F_3 = \forall x \forall y (x + y \approx y + x)$$

$$F_4 = \neg \forall x \forall y (x + y \approx y + x).$$

- (a) Determine a Σ -algebra \mathcal{A} with an universe of exactly two elements such that \mathcal{A} is a model of F_1, F_2, F_3 .
- (b) Determine a Σ -algebra \mathcal{A} with an universe of exactly two elements such that \mathcal{A} is a model of F_1, F_2, F_4 .

Proposed solution. (a) The Σ -algebra \mathcal{A} with $U_{\mathcal{A}} = \{0, 1\}$, $u +_{\mathcal{A}} v = u + v \bmod 2$ for all $u, v \in U_{\mathcal{A}}$, $s_{\mathcal{A}}(u) = u + 1 \bmod 2$ for all $u \in U_{\mathcal{A}}$, and $0_{\mathcal{A}} = 0$ is a model of F_1, F_2, F_3 .

(b) The Σ -algebra \mathcal{A} with $U_{\mathcal{A}} = \{0, 1\}$, $u +_{\mathcal{A}} 0 = u$ for all $u \in U_{\mathcal{A}}$, $0 +_{\mathcal{A}} 1 = 0$, $1 +_{\mathcal{A}} 1 = 0$, $s_{\mathcal{A}}(u) = u$ for all $u \in U_{\mathcal{A}}$, and $0_{\mathcal{A}} = 0$ is a model of F_1, F_2, F_4 .

Exercise 4.3: Let $\Sigma = (\Omega, \emptyset)$ with $\Omega = \{f/1, c/0\}$. Give a Σ -model \mathcal{A} of

$$\neg f(c) \approx c \quad \wedge \quad \forall x (f(f(x)) \approx x)$$

with $U_{\mathcal{A}} = \{1, 2, 3\}$.

Proposed solution. Define $f_{\mathcal{A}}(1) = 2$, $f_{\mathcal{A}}(2) = 1$, $f_{\mathcal{A}}(3) = 3$, and $c_{\mathcal{A}} = 1$.

Exercise 4.4: Let $\Sigma = (\Omega, \Pi)$ with $\Omega = \{f/2, g/2\}$ and $\Pi = \{P/2, Q/1\}$. Let

$$F = \forall x (P(x, y) \vee \exists y P(x, f(y, z)))$$

and $\sigma = \{y \mapsto g(x, z), z \mapsto g(x, y)\}$. Compute $F\sigma$.

Proposed solution. $\forall u (P(u, g(x, z)) \vee \exists v P(u, f(v, g(x, y))))$.

Exercise 4.5: Let F be a formula. Prove that $\exists x F$ is satisfiable if and only if $F\{x \mapsto b\}$ is satisfiable, where b is a constant that does not occur in F .

Proposed solution. For the “if” part, we assume that $F\{x \mapsto b\}$ is satisfiable. Then there exists an algebra \mathcal{A} and an assignment β such that $\mathcal{A}(\beta)(F\{x \mapsto b\}) = 1$. By the substitution lemma, we have $\mathcal{A}(\beta[x \mapsto e])(F) = 1$ where $e = \mathcal{A}(\beta)(b)$. Consequently, $\mathcal{A}(\beta)(\exists x F) = \max_{a \in U_{\mathcal{A}}} \{\mathcal{A}(\beta[x \mapsto a])(F)\} \geq \mathcal{A}(\beta[x \mapsto e])(F) = 1$. We have that $\exists x F$ is satisfiable.

For the “only if” part, we assume that $\exists x F$ is satisfiable. Thus there exists an algebra \mathcal{A} and an assignment β such that $\mathcal{A}(\beta)(\exists x F) = 1$. We have

$$\mathcal{A}(\beta)(\exists x F) = \max_{a \in U_{\mathcal{A}}} \{\mathcal{A}(\beta[x \mapsto a])(F)\} = 1.$$

Thus there exists $a \in U_{\mathcal{A}}$ such that $\mathcal{A}(\beta[x \mapsto a])(F) = 1$. Let b be a constant that does not appear in F . We define an algebra \mathcal{A}' , with $U_{\mathcal{A}'} = U_{\mathcal{A}}$, $P_{\mathcal{A}'} = P_{\mathcal{A}}$ for every $P/m \in \Pi$, $f_{\mathcal{A}'} = f_{\mathcal{A}}$ if $f \neq b$ for every $f/n \in \Omega$, and $b_{\mathcal{A}'} = a$.

\mathcal{A} and \mathcal{A}' differ only in their interpretations of b , consequently, for all assignments β , $\mathcal{A}(\beta)(F) = \mathcal{A}'(\beta)(F)$, because b never appears in F . By the substitution lemma, we have $\mathcal{A}'(\beta[x \mapsto a])(F) = \mathcal{A}'(\beta)(F\{x \mapsto b\})$. We can conclude that $\mathcal{A}(\beta[x \mapsto a])(F) = \mathcal{A}'(\beta)(F\{x \mapsto b\}) = 1$, which shows that $F\{x \mapsto b\}$ is satisfiable.

Exercise 4.6: Let $\Sigma = (\Omega, \Pi)$ be a signature. For every Σ -formula F without equality, let $neg(F)$ be the formula that one obtains from F by replacing every atom $P(t_1, \dots, t_n)$ in F by its negation $\neg P(t_1, \dots, t_n)$ for every $P/n \in \Pi$. Prove: If F is valid, then $neg(F)$ is valid.

Hint: Somewhere in the proof you need an induction over the structure of formulas. It is sufficient if you check the base cases and \wedge , \neg , and \exists . The other boolean connectives and quantifiers (\vee , \rightarrow , \leftrightarrow , \forall) can be handled analogously; you may omit them.

Proposed solution. We first define for every Σ -algebra \mathcal{A} an algebra \mathcal{A}' by $U_{\mathcal{A}'} = U_{\mathcal{A}}$, $f_{\mathcal{A}'} = f_{\mathcal{A}}$ for every $f/n \in \Omega$, and $P_{\mathcal{A}'} = U_{\mathcal{A}}^n \setminus P_{\mathcal{A}}$ for every $P/m \in \Pi$.

In the next step, we prove the lemma that $\mathcal{A}(\beta)(neg(F)) = \mathcal{A}'(\beta)(F)$ for every formula F and every assignment β .

We use induction over the structure of formulas. Clearly $\mathcal{A}(\beta)(neg(\perp)) = \mathcal{A}(\beta)(\perp) = 0 = \mathcal{A}'(\beta)(\perp)$ and $\mathcal{A}(\beta)(neg(\top)) = \mathcal{A}(\beta)(\top) = 1 = \mathcal{A}'(\beta)(\top)$.

Since all function symbols are interpreted in the same way in the algebras \mathcal{A} and \mathcal{A}' , we get $\mathcal{A}(\beta)(t) = \mathcal{A}'(\beta)(t)$ for every term t ; therefore $\mathcal{A}(\beta)(neg(P(t_1, \dots, t_m))) = \mathcal{A}(\beta)(\neg P(t_1, \dots, t_m)) = 1$ iff $(\mathcal{A}(\beta)(t_1), \dots, \mathcal{A}(\beta)(t_m)) \notin P_{\mathcal{A}}$ iff $(\mathcal{A}'(\beta)(t_1), \dots, \mathcal{A}'(\beta)(t_m)) \in P_{\mathcal{A}'}$ iff $\mathcal{A}'(\beta)(P(t_1, \dots, t_m)) = 1$.

By structural induction, we now obtain $\mathcal{A}(\beta)(neg(F \wedge G)) = \mathcal{A}(\beta)(neg(F) \wedge neg(G)) = \min\{\mathcal{A}(\beta)(neg(F)), \mathcal{A}(\beta)(neg(G))\} = \min\{\mathcal{A}'(\beta)(F), \mathcal{A}'(\beta)(G)\} = \mathcal{A}'(\beta)(F \wedge G)$ and $\mathcal{A}(\beta)(neg(\neg F)) = \mathcal{A}(\beta)(\neg(neg(F))) = 1 - \mathcal{A}(\beta)(neg(F)) = 1 - \mathcal{A}'(\beta)(F) = \mathcal{A}(\beta)(\neg F)$ and $\mathcal{A}(\beta)(neg(\exists x F)) = \mathcal{A}(\beta)(\exists x(neg(F))) = \max_{a \in U_{\mathcal{A}}} \{\mathcal{A}(\beta[x \mapsto a])(neg(F))\} = \max_{a \in U_{\mathcal{A}'}} \{\mathcal{A}'(\beta[x \mapsto a])(F)\} = \mathcal{A}'(\beta)(\exists x F)$.

Using the lemma, we now see that if F is valid, then for every \mathcal{A} and β we get $\mathcal{A}(\beta)(neg(F)) = \mathcal{A}'(\beta)(F) = 1$, which implies that $neg(F)$ is valid as well.

Exercise 4.7 (*): Let Π be a set of propositional variables. Let N and N' be sets of clauses over Π . Let S be a set of literals that does not contain any complementary literals. Prove: If every clause in N contains at least one literal L with $L \in S$ and if no clause in N' contains a literal L with $\bar{L} \in S$, then $N \cup N'$ is satisfiable if and only if N' is satisfiable.

Proposed solution. The “only if” part is trivial. For the “if” part, suppose that N' is satisfiable, that is, there is a valuation \mathcal{B} such that $\mathcal{B}(C) = 1$ for every $C \in N'$. Define a valuation \mathcal{A} by $\mathcal{A}(P) = 1$ if $P \in S$, $\mathcal{A}(P) = 0$ if $\neg P \in S$, and $\mathcal{A}(P) = \mathcal{B}(P)$ otherwise. Since every clause in N contains some literal of S , $\mathcal{A}(C) = 1$ for every $C \in N$. For a clause $C \in N'$ we distinguish two cases: If C contains some literal of S , then again $\mathcal{A}(C) = 1$. Otherwise C contains neither a literal in S nor the complement of a literal in S , so $\mathcal{A}(C) = \mathcal{B}(C)$. Since $\mathcal{B}(C) = 1$ for every $C \in N'$, we get $\mathcal{A}(C) = 1$ for every $C \in N'$.

Exercise 4.8: Let $\Sigma = (\Omega, \Pi)$ be a signature where Π contains two predicate symbols Q and R with the same arity n and possibly further predicate symbols. For any Σ -formula F let $\text{rep}(F)$ be the formula that one obtains by replacing every atom $Q(s_1, \dots, s_n)$ in F by the corresponding atom $R(s_1, \dots, s_n)$.

(a) Prove: If F is valid, then $\text{rep}(F)$ is valid. It is sufficient if you consider nonequational atoms, disjunctions $G \vee G'$, and negations $\neg G$; the other cases are handled analogously.

(b) Refute: If F is satisfiable, then $\text{rep}(F)$ is satisfiable.

Proposed solution. (a) Assume that the Σ -formula F is valid. Let \mathcal{A} and β be an arbitrary Σ -algebra and an assignment. We have to show that $\mathcal{A}(\beta)(\text{rep}(F)) = 1$. Define a Σ -algebra \mathcal{B} such that $U_{\mathcal{B}} = U_{\mathcal{A}}$, $f_{\mathcal{B}} = f_{\mathcal{A}}$ for every $f \in \Omega$, $Q_{\mathcal{B}} = R_{\mathcal{A}}$, and $P_{\mathcal{B}} = P_{\mathcal{A}}$ for every $P \in \Pi \setminus \{Q\}$. Obviously, $\mathcal{B}(\gamma)(t) = \mathcal{A}(\gamma)(t)$ for every assignment γ and Σ -term t . We show that $\mathcal{B}(\gamma)(G) = \mathcal{A}(\gamma)(\text{rep}(G))$ for every Σ -formula G and every γ by induction over the formula structure:

If $G = Q(s_1, \dots, s_n)$, then $\text{rep}(G) = R(s_1, \dots, s_n)$. The tuple $(\mathcal{A}(\gamma)(s_1), \dots, \mathcal{A}(\gamma)(s_n)) = (\mathcal{B}(\gamma)(s_1), \dots, \mathcal{B}(\gamma)(s_n))$ is contained in $Q_{\mathcal{B}}$ iff it is contained in $R_{\mathcal{A}}$ by definition of $Q_{\mathcal{B}}$, therefore we get $\mathcal{B}(\gamma)(Q(s_1, \dots, s_n)) = \mathcal{A}(\gamma)(R(s_1, \dots, s_n)) = \mathcal{A}(\gamma)(\text{rep}(Q(s_1, \dots, s_n)))$.

If $G = P(t_1, \dots, t_m)$ for some $P \neq Q$, then $\text{rep}(G) = P(s_1, \dots, s_n)$. Then the tuple $(\mathcal{A}(\gamma)(s_1), \dots, \mathcal{A}(\gamma)(s_n)) = (\mathcal{B}(\gamma)(s_1), \dots, \mathcal{B}(\gamma)(s_n))$ is contained in $P_{\mathcal{B}}$ iff it is contained in $P_{\mathcal{A}}$, therefore we get $\mathcal{B}(\gamma)(P(s_1, \dots, s_n)) = \mathcal{A}(\gamma)(\text{rep}(P(s_1, \dots, s_n)))$.

If $G = G' \vee G''$, then $\text{rep}(G) = \text{rep}(G') \vee \text{rep}(G'')$. By induction, $\mathcal{B}(\gamma)(G') = \mathcal{A}(\gamma)(\text{rep}(G'))$ and $\mathcal{B}(\gamma)(G'') = \mathcal{A}(\gamma)(\text{rep}(G''))$, therefore $\mathcal{B}(\gamma)(G) = \mathcal{B}(\gamma)(G' \vee G'') = \max\{\mathcal{B}(\gamma)(G'), \mathcal{B}(\gamma)(G'')\} = \max\{\mathcal{A}(\gamma)(\text{rep}(G')), \mathcal{A}(\gamma)(\text{rep}(G''))\} = \mathcal{A}(\gamma)(\text{rep}(G') \vee \text{rep}(G'')) = \mathcal{A}(\gamma)(\text{rep}(G))$.

If $G = \neg G'$, then $\text{rep}(G) = \neg \text{rep}(G')$. By induction, $\mathcal{B}(\gamma)(G') = \mathcal{A}(\gamma)(\text{rep}(G'))$, therefore $\mathcal{B}(\gamma)(G) = \mathcal{B}(\gamma)(\neg G') = 1 - \mathcal{B}(\gamma)(G') = 1 - \mathcal{A}(\gamma)(\text{rep}(G')) = \mathcal{A}(\gamma)(\neg \text{rep}(G')) = \mathcal{A}(\gamma)(\text{rep}(G))$.

The other cases are handled analogously.

Since F is supposed to be valid, we have therefore $\mathcal{A}(\beta)(rep(F)) = \mathcal{B}(\beta)(F) = 1$.

(b) Let $F = Q(b) \wedge \neg R(b)$, then $rep(F) = R(b) \wedge \neg R(b)$. Clearly, F is satisfiable, but $rep(F)$ is unsatisfiable.

Exercise 4.9 (*): Let $\Sigma = (\Omega, \Pi)$ be a signature. Let $P/1$ and $Q/0$ be predicate symbols in Π . Let N be a set of (universally quantified) clauses over Σ . Let N_0 be the set of all clauses in N that contain a literal $\neg P(t)$ for some $t \in T_\Sigma(X)$. Let N_1 be the set of all clauses in N that contain a literal $P(t')$ for some $t' \in T_\Sigma(X)$. Prove: If all clauses in $N_0 \setminus N_1$ contain also the literal $\neg Q$ and if all clauses in $N_1 \setminus N_0$ contain also the literal Q , then N and $(N \setminus N_0) \setminus N_1$ are equisatisfiable.

Proposed solution. We have to show that N has a model whenever $(N \setminus N_0) \setminus N_1$ has a model, and vice versa.

Since $(N \setminus N_0) \setminus N_1$ is a subset of N , every model of N is obviously a model of $(N \setminus N_0) \setminus N_1$.

For the reverse direction assume that the Σ -algebra \mathcal{A} is a model of $(N \setminus N_0) \setminus N_1$. We define a Σ -algebra \mathcal{B} that has the same universe as \mathcal{A} and that agrees with \mathcal{A} for all function and predicate symbols except for $P/1$.

If $Q_{\mathcal{A}} = 1$, we define $P_{\mathcal{B}} = \emptyset$. Since the predicate symbol P does not occur in $(N \setminus N_0) \setminus N_1$, \mathcal{B} agrees with \mathcal{A} for all the symbols that occur in these clauses, therefore $\mathcal{B} \models (N \setminus N_0) \setminus N_1$. Since all clauses in N_0 contain at least one negated literal $\neg P(t)$ and since $P_{\mathcal{B}}$ is false for every argument, $\mathcal{B} \models N_0$. Finally, all clauses in $N_1 \setminus N_0$ contain the positive literal Q , and since $Q_{\mathcal{B}} = Q_{\mathcal{A}} = 1$, we get $\mathcal{B} \models N_1 \setminus N_0$. Since $N = ((N \setminus N_0) \setminus N_1) \cup N_0 \cup (N_1 \setminus N_0)$, we conclude that $\mathcal{B} \models N$.

Otherwise $Q_{\mathcal{A}} = 0$, then we define $P_{\mathcal{B}} = U_{\mathcal{B}}$. Again, for all the symbols that occur in clauses in $(N \setminus N_0) \setminus N_1$, \mathcal{B} agrees with \mathcal{A} , therefore $\mathcal{B} \models (N \setminus N_0) \setminus N_1$. Since all clauses in N_1 contain at least one positive literal $P(t)$ and since $P_{\mathcal{B}}$ is true for every argument, $\mathcal{B} \models N_1$. Finally, all clauses in $N_0 \setminus N_1$ contain the negated literal $\neg Q$, and since $Q_{\mathcal{B}} = Q_{\mathcal{A}} = 0$, we get $\mathcal{B} \models N_0 \setminus N_1$. Since $N = ((N \setminus N_0) \setminus N_1) \cup N_1 \cup (N_0 \setminus N_1)$, we conclude again that $\mathcal{B} \models N$.

Exercise 4.10 (*): Let \succ be a well-founded strict partial ordering on a set M . A function $\phi : M^n \rightarrow M$ with $n \geq 1$ is called strictly monotonic in the j th argument if $a_j \succ a'_j$ implies $\phi(a_1, \dots, a_j, \dots, a_n) \succ \phi(a_1, \dots, a'_j, \dots, a_n)$ for all arguments $a_1, \dots, a_n, a'_j \in M$.

(a) Prove: If the ordering \succ on the set M is well-founded and total, and if $\phi : M^n \rightarrow M$ with $n \geq 1$ is strictly monotonic in the j th argument, then $\phi(a_1, \dots, a_j, \dots, a_n) \succeq a_j$ for all $a_1, \dots, a_n \in M$.

(b) In part (a), it was required that \succ is a *total* ordering. Give an example that shows that the property of part (a) does not hold if the ordering \succ is well-founded but not total.

(c) Use part (a) to prove the following property: Let $\Sigma = (\Omega, \Pi)$ be a signature, let \mathcal{A} be a Σ -algebra. Let \succ be a well-founded total ordering on the universe $U_{\mathcal{A}}$ of \mathcal{A} , such that $f_{\mathcal{A}} : U_{\mathcal{A}}^n \rightarrow U_{\mathcal{A}}$ is strictly monotonic in every argument for every $f/n \in \Omega$ with $n \geq 1$. Let β be an arbitrary \mathcal{A} -assignment, let $t \in T_{\Sigma}(X)$. Then $\mathcal{A}(\beta)(t) \succeq \beta(x)$ for every variable $x \in \text{var}(t)$.

Proposed solution. (a) Let \succ be a well-founded and total ordering on a set M , let $\phi : M^n \rightarrow M$ be a function that is strictly monotonic in the j th argument, where $1 \leq j \leq n$. Let $a_1, \dots, a_{j-1}, a_{j+1}, \dots, a_n$ be elements of M . We show $\phi(a_1, \dots, a_j, \dots, a_n) \succeq a_j$ for all $a_j \in M$ by well-founded induction over a_j and \succ .

Let $b := \phi(a_1, \dots, a_j, \dots, a_n)$. Assume that $b \not\succeq a_j$. Since \succ is total, we conclude that $a_j \succ b$. So by the induction hypothesis, we must have $\phi(a_1, \dots, b, \dots, a_n) \succeq b$. But this implies $\phi(a_1, \dots, a_j, \dots, a_n) = b \preceq \phi(a_1, \dots, b, \dots, a_n)$, contradicting the strict monotonicity of ϕ in the j th argument. So $\phi(a_1, \dots, a_j, \dots, a_n) = b \succeq a_j$ as required.

(b) Let $M = \{b, c\}$, let $\succ = \emptyset$, that is, the ordering in which all elements are incomparable. Now define $\phi(b) = c$ and $\phi(c) = b$. Then ϕ is trivially strictly monotonic in the first argument (since the condition $a_1 \succ a'_1$ is never satisfied), but $\phi(b) \succeq b$ does not hold.

(c) We use induction over the structure of terms. If t is a variable y , then $x \in \text{var}(t)$ implies $x = y$, so $\mathcal{A}(\beta)(y) = \beta(y)$ by definition of $\mathcal{A}(\beta)$. If t is a term $f(t_1, \dots, t_n)$, then $x \in \text{var}(t)$ implies $x \in \text{var}(t_i)$ for some i . So $\mathcal{A}(\beta)(t) = f_{\mathcal{A}}(\mathcal{A}(\beta)(t_1), \dots, \mathcal{A}(\beta)(t_n)) \succeq \mathcal{A}(\beta)(t_i)$ by strict monotonicity of $f_{\mathcal{A}}$ and part (a), and $\mathcal{A}(\beta)(t_i) \succeq \beta(x)$ by induction for t_i .