Automated Theorem Proving

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

Winter Term 2024/25

Exercises 2: Preliminaries Continued and Propositional Logic

Exercise 2.1: Determine all strict total orderings $\succ$ on the set $\{a, b, c, d, e\}$ such that the following properties hold simultaneously:

(1) $\{a, b\} \succ_{\text{mul}} \{a, a, c\}$

(2) $\{c, d\} \succ_{\text{mul}} \{b, b, b\}$

(3) $\{a, e\} \succ_{\text{mul}} \{c, e, e\}$

Proposed solution. Ineq. (1) holds if and only if $b \succ a$ and $b \succ c$. Ineq. (2) holds if and only if $c \succ b$ or $d \succ b$, but the first of the two possibilities is excluded by (1). Finally Ineq. (3) holds if and only if $a \succ c$ and $a \succ e$. There are two strict orderings that satisfy these conditions, namely $d \succ b \succ a \succ c \succ e$ and $d \succ b \succ a \succ e \succ c$.

Exercise 2.2: Let M be a set, and let $\succ$ be a strict partial ordering over M. Let $b, b_1, b_2 \in M$, and let S, S_1, S_2 be finite multisets over M.

(a) Prove or refute: If $\{b\} \succ_{\text{mul}} S_1$ and $\{b\} \succ_{\text{mul}} S_2$, then $\{b\} \succ_{\text{mul}} S_1 \cup S_2$.

(b) Prove or refute: If $S \succ_{\text{mul}} \{b_1\}$ and $S \succ_{\text{mul}} \{b_2\}$, then $S \succ_{\text{mul}} \{b_1, b_2\}$.

Proposed solution. (a) The statement holds: If $\{b\} \succ_{\text{mul}} S_1$, then by definition $S_1 = (\{b\} - X) \cup Y$ for multisets X and Y such that $\emptyset \neq X \subseteq \{b\}$ and such that for each $y \in Y$ there is an $x \in X$ with $x \succ y$. Clearly X must equal $\{b\}$, and therefore $Y = S_1$. Thus we have $b \succ y$ for each $y \in S_1$. Analogously, we can show that $b \succ y$ for each $y \in S_2$. Therefore $b \succ y$ for each $y \in S_1 \cup S_2$, which implies $\{b\} \succ_{\text{mul}} S_1 \cup S_2$.

(b) The statement does not hold: Let $S = \{b_1, b_2\}$, then obviously $\{b_1, b_2\} \succ_{\text{mul}} \{b_1\}$ and $\{b_1, b_2\} \succ_{\text{mul}} \{b_2\}$, but not $\{b_1, b_2\} \succ_{\text{mul}} \{b_1, b_2\}$.

Exercise 2.3: (a) Let $M = \{a, b, c, d\}$. Suppose that the binary relation $\to$ over multisets over M is defined by the rules (1)–(3):

(1) $\quad S \cup \{b, c\} \to S \cup \{a, a, a\}$

(2) $\quad S \cup \{b, a\} \to S \cup \{b, c, c\}$

(3) $\quad S \cup \{c\} \to S \cup \{d\}$

Then $\to$ can be shown to be terminating using the multiset extension $\succ_{\text{mul}}$ of an appropriate well-founded ordering on M. What does $\succ$ look like?

(b) If the binary relation $\to$ is defined by the rules (4)–(6),

(4) $\quad S \cup \{a, a\} \to S \cup \{b, c\}$

(5) $\quad S \cup \{b, b\} \to S \cup \{a, c\}$

(6) $\quad S \cup \{b, c\} \to S \cup \{a, d, c, c\}$

then there is no well-founded ordering on M such that $\to$ is contained in $\succ_{\text{mul}}$. Why? Give a short explanation.

(c) Nevertheless, the relation $\to$ defined by the rules (4)–(6) is terminating. Prove it. (Hint: Think about lexicographic combinations.)

Proposed solution. (a) The only possible ordering on M is $b \succ a \succ c \succ d$.

(b) For rule (4), we need $\{a, a\} \succ_{\text{mul}} \{b, c\}$, therefore $a \succ b$ and $a \succ c$. For rule (5), we need $\{b, b\} \succ_{\text{mul}} \{a, c\}$, therefore $b \succ a$ and $b \succ c$. From $a \succ b$ and $b \succ a$, it follows that $a \succ a$, contradicting irreflexivity.

(c) We map every multiset S over M to a pair of two natural numbers, where the first one is $S(a) + S(b)$ (that is, the sum of the numbers of occurrences of a and b in S), and the second one is $S(b)$, and compare these pairs of natural numbers lexicographically. In rule (4), the first component decreases, in rule (5), the first component decreases, in rule (6), the first component remains constant and the second component decreases, therefore the lexicographic combination decreases for all rules (4)–(6).

Alternatively, we can map every multiset S to the natural number $2 \cdot S(a) + 3 \cdot S(b)$. This number also decreases for all rules (4)–(6).

Exercise 2.4 $(*)$**:** Prove: If S and S' are finite multisets over a set M, and $S \succ_{\text{mul}} S'$ holds for *every* strict partial ordering $\succ$ over M, then $S' \subset S$ (that is, $S' \subseteq S$ and $S' \neq S$).

Proposed solution. Suppose that S and S' are finite multisets over a set M, and that $S \succ_{\mathrm{mul}} S'$ holds for *every* strict partial ordering $\succ$ over M. The empty relation $\succ_0$, for which $x \succ_0 y$ is false for all elements x and y, is a strict partial ordering (it is trivially irreflexive and transitive). So the property holds in particular for $\succ_0$. By the definition of the multiset extension, $S\ (\succ_0)_{\mathrm{mul}}\ S'$ if and only if there are multisets X and Y such that $\emptyset \neq X \subseteq S$ and $S' = (S - X) \cup Y$ and for every $y \in Y$ there is an $x \in X$ such that $x \succ_0 y$. Since $x \succ_0 y$ is false for all x and y, Y must be empty. So S' equals $S - X$, which is a subset of S, and since X is nonempty, we obtain $S' \subset S$.

Exercise 2.5: Which of the following propositional formulas are valid? Which are satisfiable?

(1) $\neg P$

(2) $P \rightarrow \bot$

(3) $\bot \rightarrow P$

(4) $(P \vee Q) \rightarrow P$

(5) $P \rightarrow (Q \rightarrow P)$

(6) $Q \rightarrow \neg Q$

(7) $Q \wedge \neg Q$

(8) $\neg(\neg P \wedge \neg\neg P)$

Proposed solution. (1) invalid (with P interpreted as $\top$) but satisfiable (with P interpreted as $\bot$)

(2) invalid (with P interpreted as $\top$) but satisfiable (with P interpreted as $\bot$)

(3) valid and hence satisfiable

(4) invalid (with P interpreted as $\bot$ and Q interpreted as $\top$) but satisfiable (with P and Q interpreted as $\bot$)

(5) valid and hence satisfiable

(6) invalid (with P interpreted as $\top$) but satisfiable (with P interpreted as $\bot$)

(7) invalid and in fact unsatisfiable

(8) valid and hence satisfiable

Exercise 2.6 ($*$)**:** Let $N = \{C_1, \dots, C_n\}$ be a finite set of propositional clauses without duplicated literals or complementary literals such that for every $i \in \{1, \dots, n\}$ the clause C_i has exactly i literals. Prove or refute: N is satisfiable.

Proposed solution. The property holds. We show by induction that for every $j \in \{0, \dots, n\}$ there is a partial valuation $\mathcal{A}_j$ that satisfies $C_1, \dots, C_j$ and in which exactly j atoms are defined.

If $j = 0$ the statement is trivial: Define $\mathcal{A}_0$ as the valuation that is undefined for every propositional variable. If $0 < j \leq n$, we assume by induction that the statement holds for $j - 1$; so there exists a partial valuation $\mathcal{A}_{j-1}$ that satisfies $C_1, \dots, C_{j-1}$ and in which exactly $j - 1$ atoms are defined. As C_j contains j literals that are different and noncomplementary, C_j must contain j different atoms. Since only $j-1$ atoms are defined in $\mathcal{A}_{j-1}$, there exists at least one atom P in C_j that is undefined in $\mathcal{A}_{j-1}$. Now define $\mathcal{A}_j$ as the valuation that maps P to 1 if P occurs positively in C_j, or to 0 if P occurs negatively in C_j, and that interprets every other atom Q in the same way as $\mathcal{A}_{j-1}$. Since all atoms that are defined in $\mathcal{A}_{j-1}$ are defined in the same way in $\mathcal{A}_j$, $\mathcal{A}_j$ satisfies $C_1, \dots, C_{j-1}$; moreover $\mathcal{A}$ satisfies C_j since it interprets P appropriately.

Exercise 2.7: Let F, G, H be propositional formulas, let p be a position of H. Prove or refute: If $H[F]_p$ is valid and $H[G]_p$ is valid, then $H[F \vee G]_p$ is valid.

Proposed solution. Proof: Suppose that $H[F]_p$ and $H[G]_p$ are valid. Let $\mathcal{A}$ be any valuation. By assumption, $\mathcal{A}(H[F]_p) = \mathcal{A}(H[G]_p) = 1$. If $\mathcal{A}(F) = 1$, then $\mathcal{A}(F \vee G) = \mathcal{A}(F)$, therefore, by Prop. 2.3.6 $\mathcal{A}(H[F \vee G]_p) = \mathcal{A}(H[F]_p) = 1$. Otherwise $\mathcal{A}(F) = 0$, then $\mathcal{A}(F \vee G) = \mathcal{A}(G)$, therefore. by Prop. 2.3.6 $\mathcal{A}(H[F \vee G]_p) = \mathcal{A}(H[G]_p) = 1$. So $\mathcal{A}(H[F \vee G]_p) = 1$ for every valuation $\mathcal{A}$.

Exercise 2.8: Let F, G, H be propositional formulas, let p be a position of H. Prove or refute: If $H[F \wedge G]_p$ is valid, then $H[F]_p$ and $H[G]_p$ are valid.

Proposed solution. Counterexample: Let $F = P$, $G = \neg P$, and $H = \neg Q$. Then $H[F \wedge G]_1 = \neg(F \wedge G) = \neg(P \wedge \neg P)$ is valid, but $H[F]_1 = \neg F = \neg P$ and $H[G]_1 = \neg G = \neg\neg P$ are not valid.

Exercise 2.9: Let Π be a set of propositional variables with $P, Q \in \Pi$. For every propositional formula F over Π, let $\phi(F)$ be the formula that one obtains from F by replacing every occurrence of P by $P \vee Q$. For instance, if $F = ((R \vee \neg P) \wedge (Q \vee P))$, then $\phi(F) = ((R \vee \neg(P \vee Q)) \wedge (Q \vee (P \vee Q)))$, and if $F = R$, then $\phi(F) = R$.

(a) Prove: If $\phi(F)$ is satisfiable, then F is satisfiable. (Note: It is sufficient if you consider propositional variables, negations, and conjunctions; the other cases are treated analogously.)

(b) Refute: If $\phi(F)$ is valid, then F is valid.

Proposed solution. (a) Assume that $\phi(F)$ is satisfiable. Let the valuation $\mathcal{A}$ be a model of $\phi(F)$. We define a valuation $\mathcal{A}'$ by $\mathcal{A}'(P) = \mathcal{A}(P \vee Q)$ and $\mathcal{A}'(R) = \mathcal{A}(R)$ for every propositional variable $R \in \Pi$ different from P.

Now we can show by induction over the structure of formulas that $\mathcal{A}'(G) = \mathcal{A}(\phi(G))$ for every Π-formula G:

Case 1: G is a propositional variable. If $G = P$, then $\mathcal{A}'(P) = \mathcal{A}(P \vee Q)$ by definition of $\mathcal{A}'$ and $\mathcal{A}(\phi(P)) = \mathcal{A}(P \vee Q)$ by definition of ϕ; if G is a propositional variable R different from P, then $\mathcal{A}'(R) = \mathcal{A}(R)$ and $\mathcal{A}(\phi(R)) = \mathcal{A}(R)$.

Case 2: G is a negation $\neg G_1$. We must show $\mathcal{A}'(\neg G_1) = \mathcal{A}(\phi(\neg G_1))$. Then $\mathcal{A}'(\neg G_1) = 1 - \mathcal{A}'(G_1) = 1 - \mathcal{A}(\phi(G_1))$ by induction and $\mathcal{A}(\phi(\neg G_1)) = \mathcal{A}(\neg\phi(G_1)) = 1 - \mathcal{A}(\phi(G_1))$.

Case 3: G is a conjunctive formula $G_1 \wedge G_2$. We must show $\mathcal{A}'(G_1 \wedge G_2) = \mathcal{A}(\phi(G_1 \wedge G_2))$. Then $\mathcal{A}'(G_1 \wedge G_2) = \min\{\mathcal{A}'(G_1), \mathcal{A}'(G_2)\} = \min\{\mathcal{A}(\phi(G_1)), \mathcal{A}(\phi(G_2))\}$ by induction and $\mathcal{A}(\phi(G_1 \wedge G_2)) = \mathcal{A}(\phi(G_1) \wedge \phi(G_2)) = \min\{\mathcal{A}(\phi(G)), \mathcal{A}(\phi(G_2))\}$.

The remaining cases are handled analogously.

Since $\mathcal{A}(\phi(F)) = 1$, we conclude that $\mathcal{A}'(F) = 1$, so $\mathcal{A}'$ is a model of F.

(b) Let $F = P \vee \neg Q$, then F is not valid, but $\phi(F) = (P \vee Q) \vee \neg Q$ is valid.

Exercise 2.10: Let Π be a set of propositional variables. Let Q and R be two propositional variables in Π. For any Π-formula F let $\phi(F)$ be the formula that one obtains by replacing every occurrence of Q in F by R.

Prove: If $\phi(F)$ is satisfiable, then F is satisfiable. (It is sufficient if you consider propositional variables, conjunctions, and negations; the other cases are handled analogously.)

Proposed solution. Assume that $\phi(F)$ is satisfiable. Then there exists a valuation $\mathcal{A}$ such that $\mathcal{A}(\phi(F)) = 1$. We have to show that there exists a valuation $\mathcal{A}'$ such that $\mathcal{A}'(F) = 1$. Define $\mathcal{A}'$ by $\mathcal{A}'(Q) = \mathcal{A}(R)$ and $\mathcal{A}'(P) = \mathcal{A}(P)$ for every propositional variable $P \in \Pi \setminus \{Q\}$.

We show by induction over the formula structure that $\mathcal{A}'(G) = \mathcal{A}(\phi(G))$ for every Π-formula G.

Case 1: G is a propositional variable. If $G = Q$, then $\phi(Q) = R$. Therefore $\mathcal{A}'(Q) = \mathcal{A}(R) = \mathcal{A}(\phi(Q))$ by definition of $\mathcal{A}'(Q)$. Otherwise $G = P$ for some $P \in \Pi \setminus \{Q\}$, then $\phi(P) = P$. Therefore $\mathcal{A}'(P) = \mathcal{A}(P) = \mathcal{A}(\phi(P))$ by definition of $\mathcal{A}'(P)$.

Case 2: G is a conjunctive formula $G_1 \wedge G_2$. Using the induction hypothesis for G_1 and G_2 we get $\mathcal{A}'(G) = \mathcal{A}'(G_1 \wedge G_2) = \min(\mathcal{A}'(G_1), \mathcal{A}'(G_2)) = \min(\mathcal{A}(\phi(G_1)), \mathcal{A}(\phi(G_2))) = \mathcal{A}(\phi(G_1) \wedge \phi(G_2)) = \mathcal{A}(\phi(G_1 \wedge G_2))$.

Case 3: G is a negation $\neg G_1$. We use the induction hypothesis for G_1 and obtain $\mathcal{A}'(G) = \mathcal{A}'(\neg G_1) = 1 - \mathcal{A}'(G_1) = 1 - \mathcal{A}(\phi(G_1)) = \mathcal{A}(\neg\phi(G_1)) = \mathcal{A}(\phi(\neg G_1))$.

The remaining cases are handled analogously.

Since $\mathcal{A}(\phi(F)) = 1$ by assumption and $\mathcal{A}'(G) = \mathcal{A}(\phi(G))$ for every Π-formula G, we obtain $\mathcal{A}'(F) = 1$, so F is satisfiable.

Exercise 2.11: Let N be a set of propositional clauses. Prove or refute the following statement: If N contains clauses $C_i \vee D_i$ ($i \in \{1, \ldots, n\}$) such that $\{C_i \mid i \in \{1, \ldots, n\}\} \models \bot$, then $N \models \bigvee_{i\in\{1,\ldots,n\}} D_i$.

Proposed solution. The statement holds. Proof: Suppose that N contains clauses $C_i \vee D_i$ ($i \in \{1, \ldots, n\}$) such that $\{C_i \mid i \in \{1, \ldots, n\}\} \models \bot$. Let $\mathcal{A}$ be an arbitrary model of N. We have to show that $\mathcal{A} \models \bigvee_{i\in\{1,\ldots,n\}} D_i$. Assume otherwise. Then $\mathcal{A}(\bigvee_{i\in\{1,\ldots,n\}} D_i) = 0$ and therefore $\mathcal{A}(D_i) = 0$ for every $i \in \{1, \ldots, n\}$. On the other hand, $\mathcal{A}$ is a model of every clause in N, and therefore $\mathcal{A}(C_i \vee D_i) = 1$ for every $i \in \{1, \ldots, n\}$. Consequently, $\mathcal{A}(C_i) = 1$ for every $i \in \{1, \ldots, n\}$. This is impossible, however, since $\{C_i \mid i \in \{1, \ldots, n\}\}$ is unsatisfiable.