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 7: Resolution Continued

Exercise 7.1: Find a strict total ordering $\succ$ on the ground atoms $P(b), P(c), Q, R$ such that

$$P(b) \vee \neg P(c) \succ_{\mathrm{C}} \neg P(b) \vee P(c) \tag{1}$$

$$P(b) \vee P(b) \vee P(b) \vee R \succ_{\mathrm{C}} P(b) \vee R \vee R \tag{2}$$

$$\neg P(b) \vee Q \succ_{\mathrm{C}} P(c) \vee R \tag{3}$$

Proposed solution. Since $P(b) \prec_{\mathrm{L}} \neg P(b)$, (1) implies $\neg P(c) \succ_{\mathrm{L}} \neg P(b)$ and thus $P(c) \succ P(b)$. From (2) we conclude $P(b) \succ_{\mathrm{L}} R$ and thus $P(b) \succ R$. Since we already know that $\neg P(b) \prec_{\mathrm{L}} P(c)$, (3) implies $Q \succ_{\mathrm{L}} P(c)$ and thus $Q \succ P(c)$. Combining all properties, we obtain $Q \succ P(c) \succ P(b) \succ R$.

Exercise 7.2: Consider the following formulas:

$$F_1 = \forall x\,(S(x) \rightarrow \exists y\,(R(x,y) \wedge P(y)))$$

$$F_2 = \forall x\,(P(x) \rightarrow Q(x))$$

$$F_3 = \exists x\, S(x)$$

$$G = \exists x\,\exists y\,(R(x,y) \wedge Q(y))$$

Use ordered resolution to prove that $\{F_1, F_2, F_3\} \models G$. You may choose the selection function and the ordering on atoms.

Hint: You will need some preprocessing.

Proposed solution. We first need to clausify $F_1 \wedge F_2 \wedge F_3 \wedge \neg G$. The formula is already a conjunction, so we can clausify its components separately and take the union of the resulting clause sets.

For F_1:

$$\begin{aligned}
& \forall x\,(S(x) \rightarrow \exists y\,(R(x,y) \wedge P(y))) \\
\Rightarrow_P\ & \forall x\,\exists x_1\,(S(x) \rightarrow (R(x,x_1) \wedge P(x_1))) \\
\Rightarrow_S\ & \forall x\,(S(x) \rightarrow (R(x,f_1(x)) \wedge P(f_1(x)))) \\
\Rightarrow_{CNF}\ & \forall x\,(\neg S(x) \vee (R(x,f_1(x)) \wedge P(f_1(x)))) \\
\Rightarrow_{CNF}\ & \forall x\,((\neg S(x) \vee R(x,f_1(x))) \wedge (\neg S(x) \vee P(f_1(x))))
\end{aligned}$$

Resulting clauses: $\neg S(x) \vee R(x, f_1(x))$ and $\neg S(x) \vee P(f_1(x))$.

For F_2:

$$\begin{aligned}
& \forall x\,(P(x) \rightarrow Q(x)) \\
\Rightarrow_{CNF}\ & \forall x\,(\neg P(x) \vee Q(x))
\end{aligned}$$

Resulting clause: $\neg P(x) \vee Q(x)$.

For F_3:

$$\begin{aligned}
& \exists x\, S(x) \\
\Rightarrow_S\ & S(f_2)
\end{aligned}$$

Resulting clause: $S(f_2)$.

For $\neg G$:

$$\begin{aligned}
& \neg\exists x\,\exists y\,(R(x,y) \wedge Q(y)) \\
\Rightarrow_P\ & \forall x\,\neg\exists y\,(R(x,y) \wedge Q(y)) \\
\Rightarrow_P\ & \forall x\,\forall y\,\neg(R(x,y) \wedge Q(y)) \\
\Rightarrow_{CNF}\ & \forall x\,\forall y\,(\neg R(x,y) \vee \neg Q(y))
\end{aligned}$$

Resulting clause: $\neg R(x,y) \vee \neg Q(y)$.

Putting all of this together, we obtain the following clause set:

$$\begin{aligned}
\neg S(x) \vee R(x, f_1(x)) \quad & (1) \\
\neg S(x) \vee P(f_1(x)) \quad & (2) \\
\neg P(x) \vee Q(x) \quad & (3) \\
S(f_2) \quad & (4) \\
\neg R(x,y) \vee \neg Q(y) \quad & (5)
\end{aligned}$$

We take the empty selection function and an atom ordering such that $S(\_) \succ R(\_) \succ Q(\_) \succ P(\_)$. (The exact ordering of atoms with the same predicate symbol but different arguments is irrelevant for this exercise.)

From (1) and (4), by "Ordered Resolution with Selection" we obtain

$$R(f_2, f_1(f_2)) \quad (6)$$

From (5) and (6), by "Ordered Resolution with Selection" we obtain

$$\neg Q(f_1(f_2)) \quad (7)$$

From (7) and (3), by "Ordered Resolution with Selection" we obtain

$$\neg P(f_1(f_2)) \quad (8)$$

From (2) and (4), by "Ordered Resolution with Selection" we obtain

$$P(f_1(f_2)) \quad (9)$$

From (8) and (9), by "Ordered Resolution with Selection" we obtain the empty clause.

Exercise 7.3: Let $\Sigma = (\Omega, \Pi)$ be a signature with $\Omega = \{b/0, f/1\}$ and $\Pi = \{P/2, Q/1, R/2\}$. Suppose that the atom ordering $\succ$ compares ground atoms by comparing lexicographically first the predicate symbols ($P \succ Q \succ R$), then the size of the first argument, then the size of the second argument (if present). If at least one of the two atoms to be compared is nonground, $\succ$ compares only the predicate symbols.

Let N be the following set of clauses:

$$P(f(x), x) \lor R(b, b) \quad (1)$$
$$\neg P(b, x) \lor \neg P(x, b) \lor Q(x) \quad (2)$$
$$Q(f(b)) \lor \neg Q(b) \lor R(f(x), b) \quad (3)$$
$$Q(b) \lor \neg R(f(x), f(x)) \quad (4)$$
$$\neg Q(x) \lor R(x, x) \quad (5)$$

(a) Which literals are strictly maximal in the clauses of N?

(b) Which literals are maximal in the clauses of N?

(c) Which $Res^{\succ}_{sel}$-inferences are possible if *sel* selects no literals? What are their conclusions?

(d) Is there a $Res^{\succ}_{sel}$-inference between the clause

$$P(x, f(x)) \lor R(b, b) \qquad (1')$$

and clause (2) if *sel* selects no literals? Justify your answer.

(e) Define a selection function *sel* such that N is saturated under $Res^{\succ}_{sel}$.

Proposed solution. (a) The following literals are maximal in their respective clauses:
$P(f(x), x)$ in (1)
$\neg P(b, x)$ and $\neg P(x, b)$ in (2)
$Q(f(b))$ in (3)
$Q(b)$ in (4)
$\neg Q(x)$ in (5)

(b) Same as part (a), since there are no duplicate literals in any of the clauses (1)–(5).

(c) The following inferences are possible:

- An "Ordered Resolution with Selection" inference from (1) and (a renamed copy of) (2) with $\neg P(b, f(b)) \lor Q(f(b)) \lor R(b, b)$ as the conclusion.
- An "Ordered Factorization with Selection" inference from (2) with $\neg P(b, b) \lor Q(b)$ as the conclusion.
- An "Ordered Resolution with Selection" inference from (3) and (a renamed copy of) (5) with $\neg Q(b) \lor R(f(x), b) \lor R(f(b), f(b))$ as the conclusion.
- An "Ordered Resolution with Selection" inference from (4) and (a renamed copy of) (5) with $\neg R(f(x), f(x)) \lor R((b), b)$ as the conclusion.

(d) No. Unification is possible only with the first literal of (2). Clause (2) after substitution is $\neg P(b, f(b)) \lor \neg P(f(b), b) \lor Q(f(b))$. The first literal is not maximal after the substitution is applied, so no inference is possible.

(e) Let *sel* be the selection function that selects no literal in (1), $\neg P(b, x)$ in (2), $\neg Q(b)$ in (3), $\neg R(f(x), f(x))$ in (4), and $\neg Q(x)$ in (5).

Exercise 7.4: In Sect. 3.12 of the lecture notes, the inference rules for ground resolution with ordering restrictions (without selection functions) are given by

(Ground) Ordered Resolution:

$$\frac{D \lor A \qquad C \lor \neg A}{D \lor C} \quad \text{if } A \succ L \text{ for all } L \text{ in } D \text{ and } \neg A \succeq L \text{ for all } L \text{ in } C.$$

(Ground) Ordered Factorization:

$$\frac{C \vee A \vee A}{C \vee A} \quad \text{if } A \succeq L \text{ for all } L \text{ in } C.$$

This calculus is sound and refutationally complete for sets of ground clauses.

Suppose that we replace the ordering restriction for the first inference rule by "if $A \succ L$ for all L in D and $A \succeq L$ for all L in C."

(a) Is the calculus with this modification still sound? If yes, give a short explanation; if no, give a counterexample.

(b) Is the calculus with this modification still refutationally complete? If yes, give a short explanation; if no, give a counterexample.

Proposed solution. (a) Yes. Every inference of the modified calculus is also an inference of the unrestricted resolution calculus. Since the inference rules of the unrestricted resolution calculus are sound, the inference rules of the modified calculus are sound as well.

(b) No. The modified calculus is not refutationally complete. In particular, it is impossible to derive the empty clause from the clauses P and $\neg P \vee \neg P$, since the necessary resolution step violates the modified ordering restriction.

Exercise 7.5: Determine all strict total orderings $\succ$ on the atomic formulas P, Q, R, S such that the associated clause ordering $\succ_C$ satisfies the properties (1)–(3) simultaneously:

$$P \vee Q \succ_C \neg Q \qquad (1)$$

$$R \vee Q \succ_C \neg P \vee \neg P \qquad (2)$$

$$\neg R \vee \neg R \succ_C S \qquad (3)$$

Proposed solution. Ineq. (1) holds if and only if $P \succ Q$. Ineq. (2) holds if and only if $R \succ P$ or $Q \succ P$, but the second of the two possibilities is excluded by (1). Ineq. (3) holds if and only if $R \succ S$. There are three strict orderings that satisfy these conditions, namely $R \succ S \succ P \succ Q$, $R \succ P \succ S \succ Q$, and $R \succ P \succ Q \succ S$.

Exercise 7.6: Let $\Sigma = (\Omega, \Pi)$ be a signature with $\Omega = \{b/0, f/1\}$ and $\Pi = \{P/1, Q/1\}$. Suppose that the atom ordering $\succ$ compares ground atoms by comparing lexicographically first the size of the argument and then the predicate symbols ($P \succ Q$). Let N be the following set of clauses:

$$\neg P(x) \lor P(f(x)) \qquad (1)$$
$$\neg Q(f(b)) \lor P(f^3(b)) \qquad (2)$$
$$Q(b) \lor Q(f(b)) \qquad (3)$$

where $f^0(b)$ stands for b and $f^{i+1}(b)$ stands for $f(f^i(b))$.

(a) Sketch what the set $G_\Sigma(N)$ of all ground instances of clauses in N looks like. How is it ordered w.r.t. the clause ordering $\succ_C$?

(b) Construct the candidate interpretation $I^{\succ}_{G_\Sigma(N)}$ of the set of all ground instances of clauses in N. Is it a model of $G_\Sigma(N)$?

Proposed solution. (a) The clauses (2) and (3) are ground, so the only ground instance of (2) is (2) itself, and the only ground instance of (3) is (3) itself. The ground instances of (1) are those clauses that we obtain from $\neg P(x) \lor P(f(x))$ by replacing x by a ground term. The set of ground terms is $\{b, f(b), f^2(b), \dots\}$, so the ground instances of (1) are $\{\neg P(f^i(b)) \lor P(f^{i+1}(b)) \mid i \in \mathbb{N}\}$. Comparing the largest literals in these clauses, we find that the ground instances of (1) are ordered as follows:

$$\begin{array}{lll}
 & \neg P(b) \lor P(f(b)) & (1.1)\\
\prec_C & \neg P(f(b)) \lor P(f^2(b)) & (1.2)\\
\prec_C & \neg P(f^2(b)) \lor P(f^3(b)) & (1.3)\\
\prec_C & \neg P(f^3(b)) \lor P(f^4(b)) & (1.4)\\
\prec_C & \cdots &
\end{array}$$

We still need to figure out where to put clauses (2) and (3) in the clause ordering: $Q(b) \lor Q(f(b))$ is smaller than $\neg P(b) \lor P(f(b))$, and $\neg Q(f(b)) \lor P(f^3(b))$ comes between $\neg P(f(b)) \lor P(f^2(b))$ and $\neg P(f^2(b)) \lor P(f^3(b))$. So the ordering is

$$\begin{array}{lll}
 & Q(b) \lor Q(f(b)) & (3)\\
\prec_C & \neg P(b) \lor P(f(b)) & (1.1)\\
\prec_C & \neg P(f(b)) \lor P(f^2(b)) & (1.2)\\
\prec_C & \neg Q(f(b)) \lor P(f^3(b)) & (2)\\
\prec_C & \neg P(f^2(b)) \lor P(f^3(b)) & (1.3)\\
\prec_C & \neg P(f^3(b)) \lor P(f^4(b)) & (1.4)\\
\prec_C & \neg P(f^4(b)) \lor P(f^5(b)) & (1.5)\\
\prec_C & \neg P(f^5(b)) \lor P(f^6(b)) & (1.6)\\
\prec_C & \cdots &
\end{array}$$

(b) Clause (3) produces $Q(f(b))$. Clauses (1.1) and (1.2) are true in their own interpretations and produce nothing. Clause (2) produces $P(f^3(b))$. Clause (1.3) is true in its own interpretation and produces nothing. All further clauses are productive and produce $P(f^4(b))$, $P(f^5(b))$, $P(f^6(b))$, ... Since the construction succeeds, the limit $I^{\succ}_{G_\Sigma(N)} = \{Q(f(b))\} \cup \{P(f^i(b)) \mid i \geq 3\}$ is a model of $G_\Sigma(N)$.

The following table summarizes the candidate interpretation construction:

Iteration	Clause C	R_C	E_C
0	$Q(b) \vee Q(f(b))$	$\emptyset$	$\{Q(f(b))\}$
1	$\neg P(b) \vee P(f(b))$	$\{Q(f(b))\}$	$\emptyset$
2	$\neg P(f(b)) \vee P(f^2(b))$	$\{Q(f(b))\}$	$\emptyset$
3	$\neg Q(f(b)) \vee P(f^3(b))$	$\{Q(f(b))\}$	$\{P(f^3(b))\}$
4	$\neg P(f^2(b)) \vee P(f^3(b))$	$\{Q(f(b)), P(f^3(b))\}$	$\emptyset$
5	$\neg P(f^3(b)) \vee P(f^4(b))$	$\{Q(f(b)), P(f^3(b))\}$	$\{P(f^4(b))\}$
6	$\neg P(f^4(b)) \vee P(f^5(b))$	$\{Q(f(b)), P(f^3(b)), P(f^4(b))\}$	$\{P(f^5(b))\}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Exercise 7.7: Let $\Sigma = (\Omega, \Pi)$ be a signature with $\Omega = \{b/0, f/1\}$ and $\Pi = \{P/1, Q/1\}$. Suppose that the atom ordering $\succ$ compares ground atoms by comparing lexicographically first the predicate symbols ($P \succ Q$) and then the size of the argument. Let N be the following set of clauses:

$$\neg Q(y) \vee P(y) \qquad (1)$$

$$Q(x) \vee Q(f(x)) \qquad (2)$$

(a) Sketch what the set $G_\Sigma(N)$ of all ground instances of clauses in N looks like. How is it ordered w.r.t. the clause ordering $\succ_C$?

(b) Construct the candidate interpretation $I^{\succ}_{G_\Sigma(N)}$ of the set of all ground instances of clauses in N.

Proposed solution. (a) The ground instances of (1) are those clauses that we obtain from $\neg Q(y) \vee P(y)$ by replacing y by a ground term. The set of ground terms is $\{b, f(b), f(f(b)), \dots\}$, so the ground instances of (1) are $\{\neg Q(f^i(b)) \vee P(f^i(b)) \mid i \in \mathbb{N}\}$, where $f^0(b)$ stands for b and $f^{i+1}(b)$ stands for $f(f^i(b))$. Comparing the largest literals in these clauses, we find that the ground instances of (1) are ordered as follows:

$$\begin{array}{lll} & \neg Q(b) \vee P(b) & (1.1) \\ \prec_C & \neg Q(f(b)) \vee P(f(b)) & (1.2) \\ \prec_C & Q(f(f(b))) \vee P(f(f(b))) & (1.3) \\ \prec_C & \cdots & \end{array}$$

The ground instances of (2) are those clauses that we obtain from $Q(x) \vee Q(f(x))$ by replacing x by a ground term. The set of ground terms is $\{b, f(b), f(f(b)), \dots\}$, so the ground instances of (2) are $\{Q(f^i(b)) \vee Q(f^{i+1}(b)) \mid i \in \mathbb{N}\}$. Comparing the largest literals in these clauses, we find that the ground instances of (2) are ordered as follows:

$$
\begin{array}{lll}
 & Q(b) \vee Q(f(b)) & (2.1) \\
\prec_C & Q(f(b)) \vee Q(f(f(b))) & (2.2) \\
\prec_C & Q(f(f(b))) \vee Q(f(f(f(b)))) & (2.3) \\
\prec_C & \cdots &
\end{array}
$$

Since $P \succ Q$ the set $G_\Sigma(N)$ looks like this:

$$
\begin{array}{lll}
 & Q(b) \vee Q(f(b)) & (2.1) \\
\prec_C & Q(f(b)) \vee Q(f(f(b))) & (2.2) \\
\prec_C & Q(f(f(b))) \vee Q(f(f(f(b)))) & (2.3) \\
\prec_C & \cdots & \\
\prec_C & \neg Q(b) \vee P(b) & (1.1) \\
\prec_C & \neg Q(f(b)) \vee P(f(b)) & (1.2) \\
\prec_C & Q(f(f(b))) \vee P(f(f(b))) & (1.3) \\
\prec_C & \cdots &
\end{array}
$$

(b) Clause (2.1) produces $Q(f(b))$. Clause (2.2) is true in its own interpretation and produces nothing. Clause (2.3) produces $Q(f(f(f(b))))$.

Clause (1.3) is true in its own interpretation and produces nothing. Clause (1.2) produces $P(f(b))$. Clause (1.2) is true in its own interpretation and produces nothing.

From this, we infer that $I_N^{\succ} = \{Q(f^{2i+1}(b)) \mid i \in \mathbb{N}\} \cup \{P(f^{2i+1}(b)) \mid i \in \mathbb{N}\}$.

The following table summarizes the candidate interpretation construction:

Iter.	Clause C	R_C	E_C
0	$Q(b) \vee Q(f(b))$	$\emptyset$	$\{Q(f(b))\}$
1	$Q(f(b)) \vee Q(f(f(b)))$	$\{Q(f(b))\}$	$\emptyset$
2	$Q(f(f(b))) \vee Q(f(f(f(b))))$	$\{Q(f(b))\}$	$\{Q(f(f(f(b))))\}$
3	$Q(f(f(f(b)))) \vee Q(f(f(f(f(b)))))$	$\{Q(f(b)), Q(f(f(f(b))))\}$	$\emptyset$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
ω	$\neg Q(b) \vee P(b)$	$\{Q(f^{2i+1}(b)) \mid i \in \mathbb{N}\}$	$\emptyset$
$\omega+1$	$\neg Q(f(b)) \vee P(f(b))$	$\{Q(f^{2i+1}(b)) \mid i \in \mathbb{N}\}$	$\{P(f(b))\}$
$\omega+2$	$\neg Q(f(f(b))) \vee P(f(f(b)))$	$\{Q(f^{2i+1}(b)) \mid i \in \mathbb{N}\} \cup \{P(f(b))\}$	$\emptyset$
$\omega+3$	$\neg Q(f(f(f(b)))) \vee P(f(f(f(b))))$	$\{Q(f^{2i+1}(b)) \mid i \in \mathbb{N}\} \cup \{P(f(b))\}$	$\{P(f(f(f(b))))\}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$

Exercise 7.8: Let N be a set of ground clauses, and let $\succ$ be a total and well-founded atom ordering. Prove or refute: If every clause in N is redundant w.r.t. N, then every clause in N is a tautology.

Proposed solution. The claim holds. We present two proofs: a proof by contradiction relying on the minimality of a counterexample and a direct proof by well-founded induction. Both approaches are equally valid.

The first proof is by contradiction. Assume that there exists a clause $C \in N$ that is redundant but not a tautology. Without loss of generality, we can even assume that C is minimal w.r.t. $\succ$. Since C is redundant, there exist $C_1, \ldots, C_n \in N$ ($n \geq 0$) such that $C_i \prec C$ for every i and $C_1, \ldots, C_n \models C$. Since C is the minimal counterexample, $C_1, \ldots, C_n$, which are smaller than C, must all be tautologies. By definition of entailment, C must be a tautology as well. Contradiction.

The second proof is by induction. We will prove that every clause $C \in N$ is a tautology by well-founded induction on C w.r.t. $\prec$. The induction hypothesis tells us that D is a tautology for all $D \in N$ such that $D \prec C$. If C is redundant, then there exist $C_1, \ldots, C_n \in N$ ($n \geq 0$) such that $C_i \prec C$ for every i and $C_1, \ldots, C_n \models C$. By the induction hypothesis, $C_1, \ldots, C_n$, which are smaller than C, must all be tautologies. By definition of entailment, C must be a tautology as well.

Exercise 7.9 ($*$)**:** Give an example of two different first-order clauses F and G such that F entails G and G is not redundant w.r.t. $\{F\}$. If necessary, specify the atom ordering used by the redundancy criterion.

Proposed solution. We take $F := P(x)$ and $G := P(a)$. Clearly, $P(x)$ entails $P(a)$, yet there is no clause in $G_\Sigma(\{P(x)\})$ that is smaller than $P(a)$ and that entails $P(a)$. The only clause in $G_\Sigma(\{P(x)\}$ that entails $P(a)$ is $P(a)$ itself, and it is not smaller, regardless of the atom ordering.

Exercise 7.10 ($*$)**:** Give a clause C such that an "Ordered Resolution with Selection" inference is possible from C and C and the inference is not redundant w.r.t. $\{C\}$.

Proposed solution. We take $C := \neg p(x, y, z) \vee p(y, z, x)$. The conclusion of resolving C with a renamed copy of C is $\neg p(x, y, z) \vee p(z, x, y)$.

In fact, self-inferences (inferences from two copies of the same premise) are necessary for refutational completeness. Consider the clause set $N = \{C, D, E\}$, where $D = p(a, c, b)$ and $E = \neg p(b, a, c)$, and an atom ordering with $p(c, b, a) \succ p(b, a, c) \succ p(a, c, b)$. Inferences between C and D or between C and E are impossible due to ordering restrictions;

and inferences between D and E are clearly impossible as well. The only possible inference is the one from two copies of C that we described above. From its conclusion and E, we obtain $F = \neg p(a, c, b)$ and from F and D, we obtain $\perp$.