Mockup Examination in the Course Automated Theorem Proving

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

Winter Term 2025/26

For convenience, a handout is provided with the definitions of the main calculi and concepts covered in the course.

Last name (in CAPITAL LETTERS):

First name (in CAPITAL LETTERS):

Matriculation number:

Program of study:

Hereby I confirm the correctness of the above information:

Signature

With your signature, you confirm that you are in sufficiently good health at the beginning of the examination and that you accept this examination bindingly.

Please leave the following table blank:

Question	1	2	3	4	5	6	$\sum$
Points	16	20	10	20	16	18	100
Score							

Instructions

You have **120 minutes** at your disposal. Written or electronic aids are not permitted except for normal watches. Carrying forbidden devices, even turned off, will be considered a cheating attempt.

Write your full name and matriculation number clearly legible on this cover sheet, as well as your name in the header on each sheet. Hand in all sheets. Leave them stapled together. Use only **pens** and **neither** the color **red nor green**.

Check that you have received all the sheets and the handout. The questions can be found on **pages 3–10**. You may use the back of the sheets for secondary calculations. If you use the back of a sheet to answer, clearly mark what belongs to which question, and indicate in the corresponding question where all parts of your answer can be found. Cross out everything that should not be graded.

There are 6 questions for a total of 100 points.

Answer. This version of the exam contains model answers in blocks like this one.

Grading. And blocks like this one specify the grading scheme.

Question 1 (16 points)**:** Let F be the first-order formula

$$\exists z\, \forall x\, ((\exists y\, P(z,y) \to \exists y\, Q(x,y)))$$

(a) Transform F into clausal normal form. Proceed one step at a time and record each step.

Answer.

$$\begin{array}{lll} & \exists z\, \forall x\, ((\exists y\, P(z,y)) \to (\exists y\, Q(x,y))) & (1) \\ \Rightarrow_P & \exists z\, \forall x\, \forall y_1\, (P(z,y_1) \to (\exists y\, Q(x,y))) & (2) \\ \Rightarrow_P & \exists z\, \forall x\, \forall y_1\, \exists y_2\, (P(z,y_1) \to Q(x,y_2)) & (3) \\ \Rightarrow_S & \forall x\, \forall y_1\, \exists y_2\, (P(f_1,y_1) \to Q(x,y_2)) & (4) \\ \Rightarrow_S & \forall x\, \forall y_1\, (P(f_1,y_1) \to Q(x,f_2(x,y_1))) & (5) \\ \Rightarrow_{CNF} & \forall x\, \forall y_1\, (\neg P(f_1,y_1) \vee Q(x,f_2(x,y_1))) & (6) \end{array}$$

Thus $N = \{\neg P(f_1,y_1) \vee Q(x,f_2(x,y_1))\}$ is the clausal normal form.

Grading. 12 points:

- 4 points for the result
- 2 points for each of the four intermediate steps

(b) Are the formula F and its clausal normal form equivalent? Explain your answer.

Answer. No, they are not equivalent. This is because Skolemization does not yield an equivalent formula. For example, consider the algebra $\mathcal{A}$ such that

$$\begin{aligned} U_{\mathcal{A}} &= \{0,1\} \\ P_{\mathcal{A}} &= \{(0,0),(0,1),(1,0),(1,1)\} \\ Q_{\mathcal{A}} &= \{(0,0),(1,1)\} \\ (f_1)_{\mathcal{A}} &= 0 \\ (f_2)_{\mathcal{A}}(0,0) = (f_2)_{\mathcal{A}}(0,1) = (f_2)_{\mathcal{A}}(1,0) &= (f_2)_{\mathcal{A}}(1,1) = 0 \end{aligned}$$

Then formula (4) (which is equivalent to (1) via (3) and (2)) is true in $\mathcal{A}$—the witness for z is irrelevant, and the witness for y_2 is x. In contrast, formula (5) (which is equivalent to the clause set N via (5) and (6)) is false in $\mathcal{A}$—essentially because f_2 provides the wrong witness when $x = 1$.

Grading. 4 points:

1. 2 points for "No"
2. 2 points for the explanation

Question 2 (20 points)**:** Let N be the following set of propositional clauses over $\Pi = \{P, Q, R\}$:

$$
\begin{array}{lcccl}
P & \vee & \neg Q & & (1) \\
P & & \vee & \neg R & (2) \\
\neg P & & \vee & R & (3)
\end{array}
$$

(a) Use the DPLL procedure to find a (total) model of N. Document each step of the procedure.

Answer. We start with the literal set $M := \emptyset$ and the clause set $N := \{(1), (2), (3)\}$. We notice that $\neg Q$ is a pure literal, so we set $M := \{\neg Q\}$. Then no more pure literals are present and no unit propagation is possible. We make an arbitrary decision and make P false, so we set $M := \{\neg Q, \neg P\}$. Next, clause (2) contains the unit literal $\neg R$, so we set $M := \{\neg Q, \neg P, \neg R\}$. At this point, all clauses in N are true in M, so we stop with $M = \{\neg Q, \neg P, \not{R}\}$ as a partial model, which is a total model as well.

Grading. 10 points:

- 3 points if the given interpretation is a total model
  - 1 point subtracted if the model is partial
- 1 point for the initial state $M = \emptyset$
- 2 points for pure literal
- 2 points for decision
- 2 points for unit propagation

(b) Use the DPLL procedure to prove that $N \models Q \rightarrow R$. You will first need to transform the entailment into a formula to refute. Document each step of the procedure.

Answer. We use the fact that $N \models Q \rightarrow R$ if and only if $N' = N \cup \{\neg(Q \rightarrow R)\}$ is unsatisfiable. To use the DPLL procedure, we transform N' into a set of clauses and obtain $N \cup \{(4), (5)\}$, with Q (4) and $\neg R$ (5). Since (4) contains the unit literal Q, we set $M := \{Q\}$. Since (5) contains the unit literal $\neg R$, we set $M := \{Q, \neg R\}$. Since (1) contains the unit literal P, we set $M := \{Q, \neg R, P\}$. At this point, (3) is false in M, and there is nowhere to backtrack to, so the clause set N' is unsatisfiable.

Grading. 10 points:

- 2 points for N'
- 2 points for clausification of N'

- 2 points for each of the three unit propagations

Question 3 (10 points)**:** Let $\Sigma = (\Omega, \emptyset)$ with $\Omega = \{f/1,\ g/1,\ h/2,\ b/0,\ c/0\}$. Find a total precedence $\succ$ on Ω such that the lexicographic path ordering $\succ_{\text{lpo}}$ that is induced by $\succ$ satisfies the following three properties simultaneously.

$$h(x, f(y)) \succ_{\text{lpo}} h(g(y), y) \qquad (1)$$
$$h(x, c) \succ_{\text{lpo}} g(h(x, b)) \qquad (2)$$
$$g(b) \succ_{\text{lpo}} c \qquad (3)$$

There is no need to explain your answer.

Answer. $f \succ h \succ g \succ c \succ b$.

Even though no explanation was asked for, we will provide one.

Property (1) could be a consequence of Case (2a) or Case (2c) of the lpo definition. Case (2c) can be ruled out, though, since x cannot be larger than $g(y)$. So Case (2a) of the lpo definition must hold, and therefore $f(y) \succ_{\text{lpo}} h(g(y), y)$, which implies $f \succ h$ and $f \succ g$.

Property (2) could be a consequence of Case (2a) or Case (2b) of the lpo definition. Case (2a) can be ruled out, since neither x nor c can be larger than $g(h(x, b))$. (If $x \succ g(h(x, b))$, then since we also have $g(h(x, b)) \succ x$ by Case (1), by transitivity $x \succ x$, violating irreflexivity. If $c \succ g(h(x, b))$, since $g(h(x, b)) \succ x$ by Case (1) we obtain $c \succ x$ by transitivity, which is impossible since $c \succ x$ follows from none of the cases of the lpo definition.) So Case (2b) of the lpo definition must hold, and therefore $h \succ g$ and $c \succ b$.

From property (3), we conclude that $g \succ c$ or $b \succ c$. We can rule out $b \succ c$, since we already know $c \succ b$. Therefore $g \succ c$.

Combining these results, we obtain the precedence specified above.

Grading. 10 points:

- 10 points for correct answer
  - 5 points subtracted if only two out of three equations are oriented correctly

Question 4 (20 points)**:** Let $\Sigma = (\{f/1,\, g/2,\, b/0,\, c/0\},\, \{P/2,\, Q/1,\, R/2\})$. Let the atom ordering $\succ$ be the kbo with weight 1 for all symbols and variables and the precedence $P \succ Q \succ R \succ f \succ g \succ b \succ c$. Let N be the following set of clauses over Σ:

$$\neg R(b, x) \quad (1)$$
$$Q(g(x, z)) \vee Q(g(b, f(y))) \vee \neg R(x, z) \quad (2)$$
$$P(f(x), x) \vee \neg Q(f(c)) \vee \neg R(x, c) \quad (3)$$
$$\neg P(x, c) \vee Q(x) \quad (4)$$
$$\neg P(x, b) \vee R(c, y) \quad (5)$$

(a) Suppose that the selection function *sel* selects no literals. Compute all $\mathrm{Res}^{\succ}_{sel}$ inferences between the clauses (1)–(5). Do not compute inferences between derived clauses.

Answer. The following literals are maximal in clauses (1)–(5):
(1): literal 1;
(2): literal 1 and 2 (literal 3 is smaller than 1);
(3): literal 1 (literals 2 and 3 are smaller than 1);
(4): literal 1 (literal 2 is smaller than 1);
(5): literal 1 and 2.
We get the following three $\mathrm{Res}^{\succ}_{sel}$ inferences:

Factorization from (2) literals 1 and 2:
mgu $\{x \mapsto b,\, z \mapsto f(y)\}$,
conclusion $Q(g(b, f(y))) \vee \neg R(b, f(y))$ (6).

Resolution from (3) literal 1 and (4) literal 1 (after renaming x in (4) to x'):
mgu $\{x' \mapsto f(c),\, x \mapsto c\}$,
conclusion $\neg Q(f(c)) \vee \neg R(c, c) \vee Q(f(c))$ (7).

Resolution from (3) literal 1 and (5) literal 1 (after renaming x in (5) to x'):
mgu $\{x' \mapsto f(b),\, x \mapsto b\}$,
conclusion $R(c, y) \vee \neg Q(f(c)) \vee \neg R(b, c)$ (8).

Grading. 16 points:

- 5 points for each of the three correct inferences
  - 2 points subtracted if variables were not renamed apart
- 1 point for absence of extraneous inferences

(b) Is the set N saturated up to redundancy? Explain your answer.

Answer. Yes. All conclusions of $\mathrm{Res}^{\succ}_{sel}$ inferences between the clauses in N are redundant w.r.t. N: Clauses (6) and (8) are subsumed by (1). Clause (7) contains complementary literals and is therefore a tautology. So N is saturated up to redundancy.

Grading. 4 points:

- 2 points for "Yes"
- 2 points for explanation

Question 5 (16 points)**:** Let R be the following set of rewrite rules over $\Sigma = (\{f/1, g/2, h/1, c/0\}, \emptyset)$:

$$\begin{aligned} f(f(x)) &\to h(h(x)) \quad (1)\\ g(f(y),x) &\to g(y,x) \quad (2)\\ h(g(z,f(c))) &\to f(z) \quad (3) \end{aligned}$$

There are three critical pairs between the three rules. Find them.

Answer. Between (1) at position 1 and a renamed copy of (1):
$\sigma = \{x \mapsto f(x')\}$,
$h(h(f(x'))) \leftarrow f(f(f(x'))) \to f(h(h(x')))$,
critical pair: $\langle h(h(f(x'))), f(h(h(x')))\rangle$.

Between (2) at position 1 and a renamed copy of (1):
$\sigma = \{y \mapsto f(x')\}$,
$g(f(x'),x) \leftarrow g(f(f(x')),x) \to g(h(h(x')),x)$,
critical pair: $\langle g(f(x'),x), g(h(h(x')),x)\rangle$.

Between (3) at position 1 and (2):
$\sigma = \{z \mapsto f(y), x \mapsto f(c)\}$,
$f(f(y)) \leftarrow h(g(f(y),f(c))) \to h(g(y,f(c)))$,
critical pair: $\langle f(f(y)), h(g(y,f(c)))\rangle$.

Since rules (1) and (2) are not variable-disjoint, to compute the critical pair between (2) at position 1 and (1), it is critical to rename the variable x in either (1) or (2), even though the term $f(f(x))$ and the subterm $f(y)$ of $g(f(y),x)$ are unifiable without the renaming.

Grading. 16 points:

- 5 points for each of the three critical pairs
  - 2 points subtracted if variables are not renamed apart
- 1 point for absence of extraneous critical pairs

Question 6 (18 points)**:** Let $\Sigma = (\{f/1,\ g/1,\ b/0\}, \emptyset)$. Consider the set N consisting of the following of equational clauses over Σ:

$$g(b) \approx b$$
$$g(x) \not\approx b \lor g(f(x)) \approx b$$

We use the superposition calculus without selection and the kbo with $g \succ f \succ b$ and weight 1 for symbols and variables as the term ordering.

(a) Is the set N saturated up to redundancy? If yes, briefly explain why. If no, give an inference from N that is not redundant w.r.t. N.

Answer. Yes, N is saturated. The last literals of both clauses are maximal, and the left-hand sides of each maximal literal is maximal. For a "Positive Superposition" inference to be applicable, we would have to unify $g(b)$ with a nonvariable subterm of $g(f(x))$, which is clearly impossible.

Grading. 4 points:

- 2 points for "Yes"
- 2 points for explanation

(b) Complete the following table documenting the first four iterations of the candidate interpretation construction leading to the rewrite system R_∞ for the set of ground clauses $G_\Sigma(N)$.

Iter.	Clause C	R_C	E_C
0	$g(b) \approx b$	$\emptyset$	
1	$g(b) \not\approx b \lor g(f(b)) \approx b$		
2	$g(f(b)) \not\approx b \lor g(f(f(b))) \approx b$		
3	$g(g(b)) \not\approx b \lor g(f(g(b))) \approx b$		

Answer.

Iter.	Clause C	R_C	E_C
0	$g(b) \approx b$	$\emptyset$	$\{g(b) \to b\}$
1	$g(b) \not\approx b \lor g(f(b)) \approx b$	$\{g(b) \to b\}$	$\{g(f(b)) \to b\}$
2	$g(f(b)) \not\approx b \lor g(f(f(b))) \approx b$	$\{g(b) \to b,\ g(f(b)) \to b\}$	$\{g(f(f(b))) \to b\}$
3	$g(g(b)) \not\approx b \lor g(f(g(b))) \approx b$	$\{g(b) \to b,\ g(f(b)) \to b,\ g(f(f(b))) \to b\}$	$\emptyset$

Notice that iteration 3 is not productive: Using $g(b) \to b$ and $g(f(b)) \to b$, we can rewrite $g(f(g(b)))$ to b, making the second literal true.

Grading. 14 points:

- 2 point for each of the seven table entries, acknowledging Folgefehler