SAARLAND UNIVERSITY ARTIFICIAL INTELLIGENCE
Prof. Jörg Hoffmann and Prof. Wolfgang Wahlster
Dr. Álvaro Torralba, Daniel Gnad, Marcel Steinmetz
Christian Bohnenberger, Cosmina Croitoru, Akram Elkorashy, Sophian Guidara,
Daniel Heller, Björn Mathis, Lukas Schaal, Julia Wichlacz

Exercise Sheet 6.
Solutions due Tuesday, **June 14**, 16:00 – 16:15, in the lecture hall.[1]

Exercise 21. (2.5 Points)

Run the unification algorithm from the lecture on each of these sets $\{E_i\}$. Here, v, w, x, y, z are *variables* while a, b are *constants*. Write down the contents of $D_0, s_1, T_1.D_1, \ldots$ until the algorithm terminates (cf. Chapter 12, slide 26). In particular, list the *entire* content of the substitutions $s_1, s_2, \ldots$.

a) $\{P(f(z), f(x), y), P(y, z, z\}$

b) $\{P(a, f(x, b), g(z)), P(z, f(a, y), g(x))\}$

c) $\{P(x, b, f(a)), P(g(a), f(x), y)\}$

d) $\{P(g(z, b), y, x), P(w, g(a, x), f(z)), P(g(a, b), g(z, f(a)), f(z))\}$

(Solution)

a)
- $T_0 = \{P(f(z), f(x), y), P(y, z, z\}$
  $D_0 = \{f(z), y)\}$
  $s_1 = \{\frac{y}{f(z)}\}$
- $T_1 = \{P(f(z), f(x), f(z)), P(f(z), z, z)\}$
  $D_1 = \{f(x), z\}$
  $s_2 = s_1\{\frac{z}{f(x)}\} = \{\frac{y}{f(f(x))}, \frac{z}{f(x)}\}$
- $T_2 = \{P(f(f(x)), f(x), f(f(x))), P(f(f(x)), f(x), f(x))\}$
  $\lightning$ The term $f(x)$ contains the variable x. $\Rightarrow$ output failure

b)
- $T_0 = \{P(a, f(x, b), g(z)), P(z, f(a, y), g(x))\}$
  $D_0 = \{a, z\}$
  $s_1 = \{\frac{z}{a}\}$

[1] Solutions in paper form only, and solution submission only at the stated time at the stated place. At most 3 authors per solution. All authors must be in the same tutorial group. All sheets of your solution must be stapled together. At the top of the first sheet, you must write the names of the authors and the name of your tutor. Your solution must be placed into the correct box for your tutorial group. If you don't comply with these rules, 3 points will be subtracted from your score for this sheet.

- $T_1 = \{P(a, f(x, b), g(a)), P(a, f(a, y), g(x))\}$
  $D_1 = \{a, x\}$
  $s_2 = s_1\{\frac{x}{a}\} = \{\frac{z}{a}, \frac{x}{a}\}$
- $T_2 = \{P(a, f(a, b), g(a)), P(a, f(a, y), g(a))\}$
  $D_1 = \{b, y\}$
  $s_3 = s_2\{\frac{y}{b}\} = \{\frac{z}{a}, \frac{x}{a}, \frac{y}{b}\}$
- $T_3 = \{P(a, f(a, b), g(a))\}$
  $\rightarrow$ singleton $\Rightarrow$ return s_3

c)
- $T_0 = \{P(x, b, f(a)), P(g(a), f(x), y)\}$
  $D_0 = \{x, g(a)\}$
  $s_1 = \{\frac{x}{g(a)}\}$
- $T_1 = \{P(g(a), b, f(a)), P(g(a), f(g(a)), y)\}$
  $D_1 = \{b, f(g(a))\}$
  $\lightning$ No variable is contained in D_1 $\Rightarrow$ output failure

d)
- $T_0 = \{P(g(z, b), y, x), P(w, g(a, x), f(z)), P(g(a, b), g(z, f(a)), f(z))\}$
  $D_0 = \{g(z, b), w, g(a, b)\}$
  $s_1 = \{\frac{w}{g(z,b)}\}$
- $T_1 = \{P(g(z, b), y, x), P(g(z, b), g(a, x), f(z)), P(g(a, b), g(z, f(a)), f(z))\}$
  $D_1 = \{z, a\}$
  $s_2 = s_1\{\frac{z}{a}\} = \{\frac{w}{g(a,b)}, \frac{z}{a}\}$
- $T_2 = \{P(g(a, b), y, x), P(g(a, b), g(a, x), f(a)), P(g(a, b), g(a, f(a)), f(a))\}$
  $D_2 = \{y, g(a, x), g(a, f(a))\}$
  $s_3 = s_2\{\frac{y}{g(a,x)}\} = \{\frac{w}{g(a,b)}, \frac{z}{a}, \frac{y}{g(a,x)}\}$
- $T_3 = \{P(g(a, b), g(a, x), x), P(g(a, b), g(a, x), f(a)), P(g(a, b), g(a, f(a)), f(a))\}$
  $D_3 = \{x, f(a)\}$
  $s_4 = s_3\{\frac{x}{f(a)}\} = \{\frac{w}{g(a,b)}, \frac{z}{a}, \frac{y}{g(a,f(a))}, \frac{x}{f(a)}\}$
- $T_4 = \{P(g(a, b), g(a, f(a)), f(a))\}$
  $\rightarrow$ singleton $\Rightarrow$ return s_4

(/Solution)

Exercise 22. (2.5 Points)

Consider the following statements and the set $\theta^* = \{\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5\}$ of corresponding predicate logic formulas in Skolem normal form:

1. Peter plays volleyball.
   $\varphi_1 = \mathit{plays}(\mathit{Peter}, \mathit{volleyball})$

2. Zoe knows everyone who plays in Peter's team
   $\varphi_2 = \forall x[\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(x, \mathit{Peter}) \vee \mathit{knows}(\mathit{Zoe}, x)]$

3. All good players play in Peter's team.
   $\varphi_3 = \forall x[\neg \mathit{good\text{-}player}(x) \vee \mathit{plays\text{-}in\text{-}team\text{-}of}(x, \mathit{Peter})]$

4. Ron is a good volleyball player.
   $\varphi_4 = \mathit{good\text{-}player}(\mathit{Ron})$

5. Zoe does not know Ron.
   $\varphi_5 = \neg \mathit{knows}(\mathit{Zoe}, \mathit{Ron})$

(a) Write down $CF(\theta^*)$ (compare Chapter 12 slide 9) and the resulting Herbrand universe $HU(\theta^*)$.

(b) Write down the Herbrand expansion $HE(\theta^*)$. Bring each of the formulas in $HE(\theta^*)$ into CNF, resulting in a set Δ of clauses.

(c) Use propositional resolution to prove that Δ is unsatisfiable. Assuming that statements 1.–4. are true, what does that tell us about the relationship between Zoe and Ron?

(Solution)

(a) $CF(\theta^*) = \{\mathit{Peter}, \mathit{Ron}, \mathit{Zoe}, \mathit{volleyball}\}$. $HU(\theta^*) = \{\mathit{Peter}, \mathit{Ron}, \mathit{Zoe}, \mathit{volleyball}\}$. (As there are no functions here, the Herbrand universe is equal to the set of constants appearing in θ^*.)

(b)

$$
\begin{aligned}
HE(\theta^*) \quad = \quad & \{(\mathit{plays}(\mathit{Peter}, \mathit{volleyball})),\\
& (\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Peter}, \mathit{Peter}) \vee \mathit{knows}(\mathit{Zoe}, \mathit{Peter})),\\
& (\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Ron}, \mathit{Peter}) \vee \mathit{knows}(\mathit{Zoe}, \mathit{Ron})),\\
& (\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Zoe}, \mathit{Peter}) \vee \mathit{knows}(\mathit{Zoe}, \mathit{Zoe})),\\
& (\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{volleyball}, \mathit{Peter}) \vee \mathit{knows}(\mathit{Zoe}, \mathit{volleyball})),\\
& (\neg \mathit{good\text{-}player}(\mathit{Peter}) \vee \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Peter}, \mathit{Peter})),\\
& (\neg \mathit{good\text{-}player}(\mathit{Ron}) \vee \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Ron}, \mathit{Peter})),\\
& (\neg \mathit{good\text{-}player}(\mathit{Zoe}) \vee \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Zoe}, \mathit{Peter})),\\
& (\neg \mathit{good\text{-}player}(\mathit{volleyball}) \vee \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{volleyball}, \mathit{Peter})),\\
& (\mathit{good\text{-}player}(\mathit{Ron})),\\
& (\neg \mathit{knows}(\mathit{Zoe}, \mathit{Ron}))\}
\end{aligned}
$$

All these formulas are already in CNF, so the resulting set of clauses is as follows:

$$
\begin{aligned}
\Delta \;=\; & \{\{\mathit{plays}(\mathit{Peter}, \mathit{volleyball})\}, \\
& \{\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Peter}, \mathit{Peter}), \mathit{knows}(\mathit{Zoe}, \mathit{Peter})\}, \\
& \{\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Ron}, \mathit{Peter}), \mathit{knows}(\mathit{Zoe}, \mathit{Ron})\}, \\
& \{\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Zoe}, \mathit{Peter}), \mathit{knows}(\mathit{Zoe}, \mathit{Zoe})\}, \\
& \{\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{volleyball}, \mathit{Peter}), \mathit{knows}(\mathit{Zoe}, \mathit{volleyball})\}, \\
& \{\neg \mathit{good\text{-}player}(\mathit{Peter}), \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Peter}, \mathit{Peter})\}, \\
& \{\neg \mathit{good\text{-}player}(\mathit{Ron}), \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Ron}, \mathit{Peter})\}, \\
& \{\neg \mathit{good\text{-}player}(\mathit{Zoe}), \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Zoe}, \mathit{Peter})\}, \\
& \{\neg \mathit{good\text{-}player}(\mathit{volleyball}), \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{volleyball}, \mathit{Peter})\}, \\
& \{\mathit{good\text{-}player}(\mathit{Ron})\}, \\
& \{\neg \mathit{knows}(\mathit{Zoe}, \mathit{Ron})\}\}
\end{aligned}
$$

(c) A suitable resolution tree is shown in figure c:

$$
\begin{tikzcd}
& \{\neg \mathit{good\text{-}player}(\mathit{Ron}), \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Ron}, \mathit{Peter})\} \arrow[dr] & & \{\mathit{good\text{-}player}(\mathit{Ron})\} \arrow[dl] \\
\{\neg \mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Ron}, \mathit{Peter}), \mathit{knows}(\mathit{Zoe}, \mathit{Ron})\} \arrow[dr] & & \{\mathit{plays\text{-}in\text{-}team\text{-}of}(\mathit{Ron}, \mathit{Peter})\} \arrow[dl] & \\
& \{\mathit{knows}(\mathit{Zoe}, \mathit{Ron})\} \arrow[dr] & & \{\neg \mathit{knows}(\mathit{Zoe}, \mathit{Ron})\} \arrow[dl] \\
& & \square &
\end{tikzcd}
$$

If we set the first 4 formulas as our knowledge base KB, our result takes the form "KB$\cup\{\varphi_5\}$ is unsatisfiable". By the Contradiction Theorem, we thus have KB $\models \neg\varphi_5$. In other words, assuming that the first 4 formulas are true, it follows that Zoe knows Ron.

(/Solution)

Exercise 23. (2.5 Points)

Consider the following set of predicate logic formulas, given in Skolem normal form:

1. $\varphi_1 = Q(b) \land Q(c) \land U(d, a)$
2. $\varphi_2 = \forall x[\neg S(x) \lor \neg T(d, x)]$
3. $\varphi_3 = \forall x \forall y[\neg P(x, y) \lor \neg Q(y) \lor R(x)]$
4. $\varphi_4 = \forall x[\neg R(x) \lor (S(f(x)) \land T(x, f(x)))]$
5. $\varphi_5 = \forall x[\neg U(x, a) \lor P(x, b) \lor P(x, c)]$

Perform the following operations:

(a) Write down the set Δ of clauses corresponding to the clausal normal forms of these formulas.

(b) Use binary PL1 resolution to show that Δ is unsatisfiable. Note: If you need to use a clause more than once, make a copy of the clause and rename the variables to avoid name collisions.

(Solution)

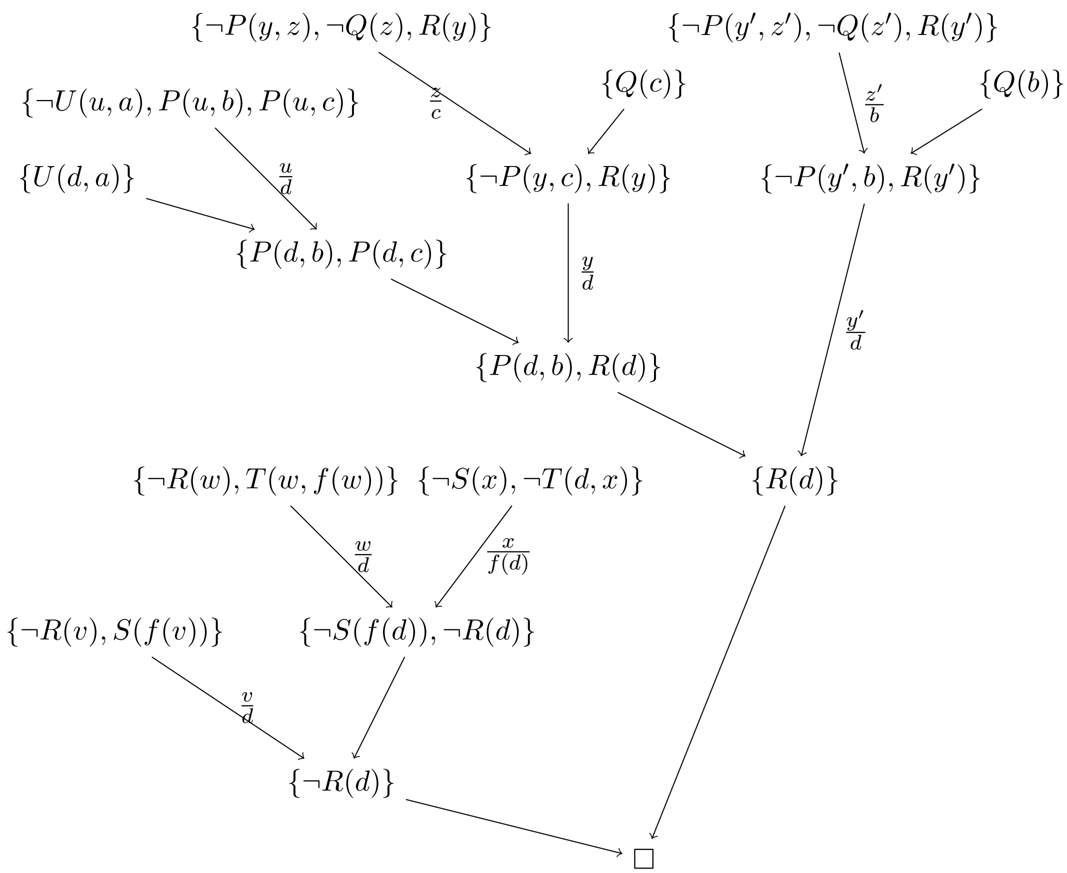
(a) The clauses for each formula are:

1. φ_1: $\{Q(b)\}, \{Q(c)\}, \{U(d, a)\}$
2. φ_2: $\{\neg S(x), \neg T(d, x)\}$
3. φ_3: $\{\neg P(x, y), \neg Q(y), R(x)\}$
4. φ_4: $\{\neg R(x), S(f(x))\}, \{\neg R(x), T(x, f(x))\}$
5. φ_5: $\{\neg U(x, a), P(x, b), P(x, c)\}$

When creating Δ, do not forget to standardize the variables apart!

1. φ_1: $\{Q(b)\}, \{Q(c)\}, \{U(d, a)\}$
2. φ_2: $\{\neg S(x), \neg T(d, x)\}$
3. φ_3: $\{\neg P(y, z), \neg Q(z), R(y)\}$
4. φ_4: $\{\neg R(v), S(f(v))\}, \{\neg R(w), T(w, f(w))\}$
5. φ_5: $\{\neg U(u, a), P(u, b), P(u, c)\}$

(b) Resolution tree:



(/Solution)

Exercise 24. (2.5 Points)

Say that φ is a predicate logic formula taking the form $\varphi = \forall x_1 \ldots \forall x_i \exists y \psi$ where all quantified variables are pairwise distinct. Say that f is a function symbol that does not occur in φ and a is a constant that does not occur in φ. Denote $\varphi^* = \forall x_1 \cdots \forall x_i \psi \frac{y}{f(x_1,\ldots,x_i)}$ and $\varphi' = \forall x_1 \cdots \forall x_i \psi \frac{y}{a}$.

(a) Prove that, if φ is satisfiable, then φ^* is satisfiable as well.

(b) Prove that the same is not true for φ', i.e., there are cases where φ is satisfiable but φ' is not. (Tip: Remember here that predicate and function symbols can be interpreted in arbitrary ways, and that our language does not include "=" to directly compare constants.)