Universität Siegen
Lehrstuhl Theoretische Informatik
Markus Lohrey

Exercise 4

Task 1

Consider the structure $(\mathbb{N}, +, \cdot, s, 0)$. Use Gödel's β-function in order to formalize the following statements in predicate logic:

(a) $x^y = z$ (use free variables x, y and z),

(b) Fermat's Last Theorem,

(c) Collatz conjecture.

Solution:

We give the main ideas:

(a) We express $x^y = z$ as: There is a sequence $(a_1, \ldots, a_y, a_{y+1})$ with $a_1 = 1$, $a_{i+1} = a_i \cdot x$ for $1 \leq i \leq y$ and $a_{y+1} = z$. This holds if there are $t, p \in \mathbb{N}$ with $\beta(t, p, 1) = 1$, $\beta(t, p, i + 1) = \beta(t, p, i) \cdot x$ for every $1 \leq i \leq y$ and $\beta(t, p, y + 1) = z$.

(b) Fermat's Last Theorem states the following: For all natural numbers $a, b, c \geq 1$ and $n \geq 3$ we have $a^n + b^n \neq c^n$. We already know how to formalize $x^y = z$. From Exercise 2, Task 2, we know how to formalize the numbers 1 and 2 and the relations $\geq$ and $>$ in $(\mathbb{N}, +, \cdot, s, 0)$. We can thus formalize: If $a, b, c \geq 1$ and $n > 2$ and $a^n = a'$, $b^n = b'$ and $c^n = c'$, then $a' + b' \neq c'$.

(c) Let $f \colon \mathbb{N} \to \mathbb{N}$ be defined as $f(2n) = n$ and $f(2n + 1) = 3(2n + 1) + 1$. Let C_n be the sequence $(n, f(n), f(f(n)), \ldots)$. We write $C_n[i]$ for the ith element of the sequence. The Collatz conjecture is the following question: Is there for every n an integer j, such that $C_n[j] = 1$? The function f can be formalized by distinguishing between odd and even numbers and by defining the numbers 2 and 3 (Exercise 2, Task 2). Using the β-function, we can formalize the Collatz conjecture as follows: For every $n \in \mathbb{N}$ there are $t, p \in \mathbb{N}$, such that $\beta(t, p, 1) = n$, $\beta(t, p, i + 1) = f(\beta(t, p, i))$ and there is $j \in \mathbb{N}$ such that $\beta(t, p, j) = 1$.

Task 2
Show that the set of valid formulas of predicate logic is undecidable. Use a reduction to the Post correspondence problem for the proof.

Post correspondence problem (PCP)
Input: A sequence of pairs $(s_1, t_1), \ldots, (s_k, t_k)$, such that $k \geq 0, s_i, t_i \in \{0,1\}^*$ $(1 \leq i \leq k)$.
Question: Are there indices $i_1, \ldots, i_m \in \{1, \ldots, k\}$ $(m \geq 1)$ such that $s_{i_1} \cdots s_{i_m} = t_{i_1} \cdots t_{i_m}$?

Hint: Let e be a 0-ary function symbol, f_0 and f_1 be unary function symbols and P denote a binary predicate symbol. Let $z = c_1 \ldots c_\ell \in \{0,1\}^\ell$ $(\ell \geq 0)$ be a string and let t be a term. We define $f_z(t)$ as $f_{c_\ell}(\ldots(f_{c_1}(t))\ldots)$. Let

$$\phi_1 = \bigwedge_{i=1}^{k} P(f_{s_i}(e), f_{t_i}(e)),$$

$$\phi_2 = \forall v \forall w (P(v, w) \to \bigwedge_{i=1}^{k} P(f_{s_i}(v), f_{t_i}(w))),$$

$$\phi_3 = \exists z (P(z, z))$$

and let $\phi = \phi_1 \wedge \phi_2 \to \phi_3$. Show that ϕ is valid if and only if the corresponding instance $(s_1, t_1), \ldots, (s_k, t_k)$ of the post correspondence problem has a solution.

Solution:
We have to show two directions. In order to prove the first direction, we assume that ϕ is valid. The main idea is to define a structure $\mathcal{A}$ in such a way that it yields the existance of a solution to the corresponding instance $(s_1, t_1), \ldots, (s_k, t_k)$ of the post correspondence problem. We define $\mathcal{U}_\mathcal{A} = \{0,1\}^*$, $e^\mathcal{A} = \varepsilon$, $f_0^\mathcal{A}(s) = s0$, $f_1^\mathcal{A}(s) = s1$ and

$$P^\mathcal{A} = \{(s, t) \mid \exists (i_1, \ldots, i_m) \in \mathbb{N}^m : s = s_{i_1} \ldots s_{i_m} \wedge t = t_{i_1} \ldots t_{i_m}\}.$$

That is, $e^\mathcal{A}$ is the empty string ε and $f_0^\mathcal{A}$, $f_1^\mathcal{A}$ (and also $f_w^\mathcal{A}$ for $w \in \{0,1\}^*$) concatenate strings (for example, $f_v^\mathcal{A}(u) = uv$). If $(s,t) \in \{0,1\}^* \times \{0,1\}^*$ satisfies $(s,t) \in P^\mathcal{A}$, then $s = s_{i_1} \ldots s_{i_m}$ and $t = t_{i_1} \ldots t_{i_m}$ for some indices $i_1, \ldots, i_m \in \mathbb{N}$. As ϕ is valid, we have $\mathcal{A} \models \phi$. If $\mathcal{A} \models \phi_1 \wedge \phi_2$, then it follows that $\mathcal{A} \models \phi_3$, that is, there is a solution to the instance $(s_1, t_1), \ldots, (s_k, t_k)$ of the post correspondence problem: If $\mathcal{A} \models \phi_3$ then there exists $z \in \{0,1\}^*$, such that there are indices $i_1, \ldots, i_m \in \mathbb{N}$ with $z = s_{i_1} \ldots s_{i_m} = t_{i_1} \ldots t_{i_m}$. We have $\mathcal{A} \models \phi_1$, as $f_{s_i}^\mathcal{A}(e^\mathcal{A}) = s_i$, $f_{t_i}^\mathcal{A}(e^\mathcal{A}) = t_i$ and $(s_i, t_i) \in P^\mathcal{A}$ for every $1 \leq i \leq k$. It remains to show that $\mathcal{A} \models \phi_2$: Let $s, t \in \{0,1\}^*$ with $(s,t) \in P^\mathcal{A}$. Then there is a sequence of indices $(i_1, \ldots, i_m)$ with $s = s_{i_1} \ldots s_{i_m}$ and $t = t_{i_1} \ldots t_{i_m}$. Let $1 \leq j \leq k$. We have to show that $(f_{s_j}^\mathcal{A}(s), f_{t_j}^\mathcal{A}(t)) \in P^\mathcal{A}$. This holds as there are indices $(i_1, \ldots, i_m, j)$ with $f_{s_j}^\mathcal{A}(s) = s_{i_1} \ldots s_{i_m} s_j$ and $f_{t_j}^\mathcal{A}(t) = t_{i_1} \ldots t_{i_m} t_j$.
In order to prove the other direction, we have to show that if the instance $(s_1, t_1), \ldots, (s_k, t_k)$ of the post correspondence problem has a solution, then ϕ is valid. Let $(i_1, \ldots, i_m)$ be the solution to the instance $(s_1, t_1), \ldots, (s_k, t_k)$ of the post correspondence problem. If $\mathcal{A} \not\models \phi_1$

or $\mathcal{A} \not\models \phi_2$, then $\mathcal{A} \models \phi$ trivially holds. It remains to consider the case that $\mathcal{A} \models \phi_1$ and $\mathcal{A} \models \phi_2$: we then have to show that $\mathcal{A} \models \phi_3$ holds as well. Define $\psi\colon \{0,1\}^* \to \mathcal{U}_{\mathcal{A}}$ by $\psi(\varepsilon) = e^{\mathcal{A}}$, $\psi(s0) = f_0^{\mathcal{A}}(\psi(s))$ and $\psi(s1) = f_1^{\mathcal{A}}(\psi(s))$. As $\mathcal{A} \models \phi_1$, we find that $(\psi(s_i), \psi(t_i)) \in P^{\mathcal{A}}$ for every $1 \leq i \leq k$. In particular, we have $(\psi(s_{i_1}), \psi(t_{i_1})) \in P^{\mathcal{A}}$. We show inductively, that $(\psi(s_{i_1} \ldots s_{i_m}), \psi(t_{i_1} \ldots t_{i_m})) \in P^{\mathcal{A}}$ holds as well:
Let $(\psi(s_{i_1} \ldots s_{i_\ell}), \psi(t_{i_1} \ldots t_{i_\ell})) \in P^{\mathcal{A}}$ for $1 \leq \ell < m$. As $\mathcal{A} \models \phi_1$, we have $(\psi(s_{i_{\ell+1}}), \psi(t_{i_{\ell+1}})) \in P^{\mathcal{A}}$. By definition of ψ, we find that for all $v, w \in \{0,1\}^*$, it holds that $\psi(vw) = f_w^{\mathcal{A}}(\psi(v))$. In particular, we find $\psi(s_{i_1} \ldots s_{i_{\ell+1}}) = f^{\mathcal{A}}_{s_{i_{\ell+1}}}(\psi(s_{i_1} \ldots s_{i_\ell}))$ and $\psi(t_{i_1} \ldots t_{i_{\ell+1}}) = f^{\mathcal{A}}_{t_{i_{\ell+1}}}(\psi(t_{i_1} \ldots t_{i_\ell}))$. As $\mathcal{A} \models \phi_2$ we find that $(\psi(s_{i_1} \ldots s_{i_\ell+1}), \psi(t_{i_1} \ldots t_{i_{\ell+1}})) \in P^{\mathcal{A}}$: this concludes the induction.
As $(\psi(s_{i_1} \ldots s_{i_m}), \psi(t_{i_1} \ldots t_{i_m})) \in P^{\mathcal{A}}$ and as $s_{i_1} \ldots s_{i_m} = t_{i_1} \ldots t_{i_m}$, as $(i_1, \ldots, i_m)$ is a solution to the instance of the post correspondence problem, we find that $\mathcal{A} \models \phi_3$ and hence $\mathcal{A} \models \phi$.