Homework 7

An instance of PCP is a finite set of two-part tiles $\frac{s_1}{t_1}, \frac{s_2}{t_2}, \ldots, \frac{s_k}{t_k}$ where each tile is from $\frac{\{0,1\}^*}{\{0,1\}^*}$. There exists a solution to a PCP instance if there is a finite indexing $i_1, i_2, \ldots, i_l \in 1, \ldots, k$ such that $s_{i_1} s_{i_2} \ldots s_{i_l} = t_{i_1} t_{i_2} \ldots t_{i_l}$.

Theorem 1. Solving PCP is undecidable.

By reducing the problem of PCP to a statement of first order logic, we prove its undecidability. For a PCP instance, let us define a signature:

- Constant symbol e.
- Unary function symbols f_0 and f_1.
- Binary predicate symbol p.

We will use unary function symbols $s_1, \ldots, s_k, t_1, \ldots, t_k$ as abbreviations for appending a tile to a string (e.g., if $s_i = 010$, then $s_i = f_0(f_1(f_0)))$. Let us define the following axioms:

$$\varphi_1 := \bigwedge_{i \in 1, \ldots, k} p(s_i(e), t_i(e))$$

$$\varphi_2 := \forall u, v \left(p(u, v) \Rightarrow \bigwedge_{i \in 1, \ldots, k} p(s_i(u), t_i(v)) \right)$$

$$\varphi_3 := \exists z \, p(z, z)$$

Theorem 2. Deciding validity of a first order formula is undecidable.

Proof. We reduce a PCP instance to deciding validity of the statement $(\varphi_1 \wedge \varphi_2) \Rightarrow \varphi_3$. We need to prove that the formula is valid iff there is a solution to the PCP instance.

$\Leftarrow$ Let $i_1, i_2, \ldots, i_l \in 1, \ldots, k$ be indices of the solution of a PCP instance. Hence, $s_{i_1} s_{i_2} \ldots s_{i_l} = t_{i_1} t_{i_2} \ldots t_{i_l}$. We need to show the validity of the formula, i.e., that it holds in any interpretation. The above fact implies that

$$u_s = s_k(s_{i_{k-1}}(\cdots(s_{i_1}(e))\cdots)) = t_{i_k}(t_{i_{k-1}}(\cdots(t_{i_1}(e))\cdots)) = u_t$$

(note that this holds syntactically because these functions are just abbreviations for f_0, f_1).

Inductively, on the length of k, we prove that $p(u_s, u_t)$ assuming φ_1 and φ_2.

Base case: If $k = 1$. Using the statement φ_1, we know that $p(s_{i_1}(e), t_{i_1}(e))$ which is what we want.

Inductive case: We know that $p(s_{k-1}(\cdots(s_{i_1}(e))\cdots), t_{k-1}(\cdots(t_{i_1}(e))\cdots))$. Then, using φ_2, we know that this implies $p(s_k(s_{i_{k-1}}(\cdots(s_{i_1}(e))\cdots)), t_{i_k}(t_{i_{k-1}}(\cdots(t_{i_1}(e))\cdots)))$.

Hence, we proved $p(u_s, u_t)$ and since $u_s = u_t$ we get that φ_3. Therefore, the formula is valid.

$\Rightarrow$ We assume validity of the formula and find a solution for the PCP instance. Therefore, we may choose any interpretation.

Let us choose the following: Domain is the set of finite strings over $\{0,1\}^*$. Symbol e is the empty string. Unary functions f_0 and f_1 represent appending 0 or 1 to a string, respectively. Finally, the subset of $\{0,1\}^* \times \{0,1\}^*$ for which the predicate p is valid is chosen to be minimal in terms of inclusion while preserving validity of φ_1 and φ_2, i.e., we interpret $p(u,v)$ as there is a sequence of tiles such that u is made up of the upper parts of the tiles and v is made up of the bottom parts of the tiles.

Such intepretation exists. Therefore, the statement φ_3 is true which means there is a finite string z such that $p(z,z)$. By well-founded induction on the length of z, we find the solution to the PCP instance.

Since p is minimal up to inclusion closed under φ_1 and φ_2, then there are u,v for each $p(u,v)$ such that u,v either match the formula φ_1 or right-hand side of the formula φ_2.

Let us assume that z matches the first axiom, i.e., there is $i \in 1,\dots,k$ such that $z = s_i(e) = t_i(e)$. Then, we found a solution to the problem which is using the tile i.

Let us now assume that z matches the second axiom, i.e., there is $u,v \in \{0,1\}^*$ and $i \in 1,\dots,k$ such that $z = s_i(u) = t_i(v)$ and $p(u,v)$. This means the last tile used was i and the rest of the tiling is obtained using the induction hypothesis from u,v.

Therefore, we obtained a finite solution to the PCP instance.

□

Theorem 3. A theory T is consistent iff T does not contain all wffs over the T-signature.

Proof. If T does contain all wffs over the T-signature, then there is no model of T since for every model $\mathcal{I}$, either φ or $\neg\varphi$ is not true (where φ is any wff), i.e., either $\mathcal{I} \not\models \varphi$ or $\mathcal{I} \not\models \neg\varphi$ which means $\mathcal{I} \not\models T$. Non-existence of a model implies inconsistency directly from the definition.

Let us assume T is inconsistent and let ψ be any wff. Let $\mathcal{I}$ be a model. Since T is inconsistent, then $\mathcal{I} \not\models T$ which means there is a formula $\varphi \in T$ such that $\mathcal{I} \not\models \varphi$. Therefore, the statement $\forall\mathcal{I}((\forall\xi \in T, \mathcal{I} \models \xi) \Rightarrow \mathcal{I} \models \psi)$ is true since the implication is vacuously true. Hence, directly from the definition, ψ is in T which means any wff is in T. □