Solution Sheet n°10

Solution of exercise 1:

1. Not necessarily. By definition of 'true', if X is printable then $P(X)$ is true. However, we are not given that the machine is capable of printing all true sentences, but only that the machine never prints false ones.

2. No. To see this, consider the sentence $\neg PN(\neg PN)$. By definition of *true*, this sentence is true if and only if the norm of $\neg PN$ is not printable. But the norm of $\neg PN$ is the very sentence $\neg PN(\neg PN)$. Consequently, the sentence $\neg PN(\neg PN)$ is true if and only if the sentence $\neg PN(\neg PN)$ is not printable. We face two mutually exclusive alternatives

   (a) $\neg PN(\neg PN)$ is true and not printable;

   (b) $\neg PN(\neg PN)$ is not true and printable.

   The second alternative is not plausible since it violates the given hypothesis according to which the machine never prints sentences that are not true. Hence the sentence $\neg PN(\neg PN)$ must be true, but the machine cannot print it.

3. For the same reasons as in 2., the sentence $\neg PN101001000$ is true and not printable.

4. The sentence $PN101001000$ is not printable, since $\neg PN101001000$ is true and the machine never prints false sentence. Its negation is not printable either by the previous point.

Solution of exercise 2:

Theorem (After Gödel with shades of Tarski)**.** *If the set $d^{-1}(P^{\complement})$ is expressible in $\mathcal{L}$ and $\mathcal{L}$ is correct, then there is a true sentence of $\mathcal{L}$ not provable in $\mathcal{L}$.*

Proof. Suppose $\mathcal{L}$ is correct and $d^{-1}(P^{\complement})$ is expressible in $\mathcal{L}$. Let H be a predicate which expresses $d^{-1}(P^{\complement})$ in $\mathcal{L}$ and let h be the Gödel number of H. We call G the diagonalisation of H, that is, the sentence $H(h)$. We show that G is true ($G \in \mathcal{T}$) but not provable in $\mathcal{L}$ ($G \notin \mathcal{P}$).

Since H expresses $d^{-1}(P^{\complement})$ in $\mathcal{L}$, for any natural number n, $H(n) \in \mathcal{T}$ iff $n \in d^{-1}(P^{\complement})$ iff $d(n) \notin P$. In particular for h the Gödel number H, we have

$$H(h) \in \mathcal{T} \quad \text{iff} \quad d(h) \notin P.$$

Now by definition of the diagonal function, $d(h)$ is the Gödel number of the diagonalisation of $E_h = H$, that is, of the very sentence $H(h)$. Since P is by definition the set of all Gödel numbers of provable formulas ($\mathcal{P}$), we have $d(h) \notin P$ iff $H(h) \notin \mathcal{P}$ and thus

$$H(h) \in \mathcal{T} \quad \text{iff} \quad H(h) \notin \mathcal{P}.$$

Hence, the sentence $H(h)$ is true in $\mathcal{L}$ iff $H(h)$ is not provable in $\mathcal{L}$. Finally, since $\mathcal{L}$ is supposed correct, we have necessarily that $H(h)$ is true in $\mathcal{L}$ but not provable in $\mathcal{L}$. $\square$

Lemma (Diagonal lemma)**.** *For any set A of natural numbers, if $d^{-1}(A)$ is expressible in $\mathcal{L}$, then there is a Gödel sentence for A.*

Proof. Suppose that H is a predicate which expresses $d^{-1}(A)$ in $\mathcal{L}$. Let h be its Gödel number. Then by definition of the diagonal function, $d(h)$ is the Gödel number of $H(h)$. For any natural number n, $H(n) \in \mathcal{T}$ iff $n \in d^{-1}(A)$ iff $d(n) \in A$. In particular, for h, we have

$$H(h) \in \mathcal{T} \quad \text{iff} \quad d(h) \in A.$$

Since $d(h) = g(H(h))$, it follows that the sentence $G = H(h)$ is a Gödel sentence for the set A. $\square$

We now give another proof of the previous theorem by using this diagonal lemma:

Since $d^{-1}(P^{\complement})$ is expressible in $\mathcal{L}$, there exists by the diagonal lemma a Gödel sentence G for $P^{\complement}$. A Gödel sentence for $P^{\complement}$ is precisely a sentence which is true if and only if it is not provable in $\mathcal{L}$. Hence, under the hypothesis that the system $\mathcal{L}$ is correct, it follows that G is a true but not provable sentence in $\mathcal{L}$.