Gödel's Fixed Point Lemma in the Land of Sets

We work in **HF** and assume that all formulas of our language are objects in **HF**, i.e., $\mathrm{Fml} \subseteq \mathbf{HF}$. As usual, if $a \in \mathbf{HF}$ and φ is a formula in one free variable x, we write

$$\mathbf{HF} \models \varphi(a) :\Longleftrightarrow \mathbf{HF}\frac{a}{x} \models \varphi,$$

i.e., for all valuables v such that $v(x) = a$, we have that $\mathbf{HF}, v \models \varphi$.

Lemma (Gödel). If φ is a formula in one free variable x, then there is a sentence σ such that

$$\mathbf{HF} \models \sigma \iff \mathbf{HF} \models \varphi(\sigma).$$

Proof. If x is a variable and $w \in \mathbf{HF}$, there is a concrete formula that states that x is equal to w, defined by recursion on the rank of w in the von Neumann hierarchy. Let us write $\mathsf{is}_w(x)$ for that formula. Now for any formula ξ with one free variable x and $w \in \mathbf{HF}$, we let $\xi[w] := \exists x(\xi \wedge \mathsf{is}_w(x))$. Clearly,

$$\mathbf{HF} \models \xi[w] \iff \mathbf{HF} \models \xi(w) \iff \mathbf{HF}\frac{w}{x} \models \xi. \tag{*}$$

This operation is called the *substitution operation*: we are syntactically substituting the set w for the variable x in the formula ξ.

Since $\mathrm{Fml} \subseteq \mathbf{HF}$, we can substitute any formula η for the variable x; in particular, we can substitute the formula ξ itself, i.e., form $\xi[\xi]$. We have that $\mathbf{HF} \models \xi[\xi]$ if and only if the formula ξ applies to the object $\xi \in \mathbf{HF}$.[1]

Let φ be the formula in the statement of the lemma (in one free variable x) and define a formula $\widehat{\varphi}$, again in one free variable x, expressing "x is a formula in one free variable and $\varphi[x[x]]$". So, $\widehat{\varphi}$ holds of a formula ξ if and only if the sentence $\xi[\xi]$ satisfies the formula φ.

Finally, we now form the sentence $\sigma := \widehat{\varphi}[\widehat{\varphi}]$ and prove that it satisfies the statement of the lemma:

$$\begin{aligned}
\mathbf{HF} \models \sigma &\iff \mathbf{HF} \models \widehat{\varphi}[\widehat{\varphi}] \\
&\iff \mathbf{HF} \models \widehat{\varphi}(\widehat{\varphi}) \quad \text{by (*)} \\
&\iff \mathbf{HF} \models \varphi[\widehat{\varphi}[\widehat{\varphi}]] \quad \text{by definition of } \widehat{\varphi} \\
&\iff \mathbf{HF} \models \varphi(\widehat{\varphi}[\widehat{\varphi}]) \quad \text{by (*)} \\
&\iff \mathbf{HF} \models \varphi(\sigma).
\end{aligned}$$

q.e.d.

Applying Gödel's Fixed Point Lemma to the formula $\varphi(x) := \neg\exists v(\mathrm{O}(x, v) \wedge v \in \mathrm{Proof}(\mathcal{R}, S))$ gives us the *Gödel sentence* ψ with the property that

$$\mathbf{HF} \models \psi \iff \mathbf{HF} \models \text{“}\psi \text{ is not provable”}. \tag{†}$$

The Gödel sentence cannot be false since this would imply that it's provable, hence true. So it must be true. But by (†), it is not provable.

B.L.
6 July 2023

[1] As an example in the style of Frege and Russell, think of the formula ξ expressing "has an even number of symbols"; then $\xi[\xi]$ is the sentence stating that the formula describing the property of having an even number of symbols has itself an even number of symbols.