108 (Gödel/Turing incompleteness) Prove that we cannot consistently and completely define a total, deterministic interpreter. An interpreter is a predicate $\mathtt{I}$ that applies to texts; when applied to a text representing a binary expression, its result is equal to the represented expression. For example,

$$\mathtt{I}\text{“}\forall s\text{: [*}char\text{]}\cdot \#s \ge 0\text{”} \;=\; \forall s\text{: [*}char\text{]}\cdot \#s \ge 0$$

After trying the question, scroll down to the solution.

§ Let $Q = $ "$\neg\ \mathtt{I}\ Q$" . Now

$$\begin{array}{lll} & \mathtt{I}\ Q & \text{replace } Q \text{ with its equal} \\ = & \mathtt{I}\ \text{“}\neg\ \mathtt{I}\ Q\text{”} & \text{If } \mathtt{I} \text{ is a complete interpreter as described in the question, then} \\ = & \neg\ \mathtt{I}\ Q & \end{array}$$

If $\mathtt{I}$ is a complete interpreter, we have an inconsistency. To save ourselves we can leave the interpreter incomplete. In particular,

$$\mathtt{I}\ \text{“}\neg\ \mathtt{I}\ Q\text{”} \ = \ \neg\ \mathtt{I}\ Q$$

must not be a theorem. If it is an antitheorem, then $\mathtt{I}$ is not an interpreter. So leave it unclassified. Alternatively, we could let $\mathtt{I}$ be partial so that $\mathtt{I}\ Q = null$, or nondeterministic so that $\mathtt{I}\ Q = bin$. Then $\mathtt{I}\ Q = \neg\ \mathtt{I}\ Q$ is a theorem, but we cannot use the Completion Rule to prove it is an antitheorem because $\mathtt{I}\ Q$ is not elementary. So we do not have an inconsistency, but we also do not have a total, deterministic interpreter. As any programmer can see, applying $\mathtt{I}$ to "$\neg\ \mathtt{I}\ Q$" will cause an infinite execution, and produce no answer.

Although the question does not ask for this, here is how you define an interpreter. Start with

$$\mathtt{I}\ \text{“}\top\text{”} \ = \ \top$$
$$\mathtt{I}\ \text{“}\bot\text{”} \ = \ \bot$$

Now, for texts that represent negations, we want to say something like

$$\mathtt{I}\ (\text{“}\neg\text{”};\ s) \ = \ \neg\ \mathtt{I}\ s$$

It says: to apply $\mathtt{I}$ to a text that starts with "$\neg$" , just apply $\mathtt{I}$ to the text after the $\neg$, and then negate the result. For texts that represent conjunctions, we want to say something like

$$\mathtt{I}\ (s;\ \text{“}\wedge\text{”};\ t) \ = \ \mathtt{I}\ s\ \wedge\ \mathtt{I}\ t$$

And so on for all operators of the theory we are interpreting. The trouble is precedence. For example, the expression

$$\neg\top \wedge \bot$$

starts with $\neg$, but it's not negating $\top \wedge \bot$. One solution is to insist that all expressions be fully parenthesized. Another solution is to use Polish prefix notation (see Subsection 3.2.2 on page 31.)