Math 160B - Winter 2022 - Homework 8
Due Sunday, March 13 11:00pm
(Note revised due date — Hand in by uploading to Gradescope)

Problems 1 and 5 depend on the proof of the Gödel Second Incompleteness Theorem which will be covered in class on Tuesday (March 8).

1. Let A be an arbitrary L_{PA}-sentence. Prove that PA does not prove $\neg \mathrm{THM}_{\mathrm{PA}}(\underline{\ulcorner A \urcorner})$.

2. Let E be a self-referential formula such that R proves $E \leftrightarrow \mathrm{THM}_{\mathrm{PA}}(\underline{\ulcorner \neg E \urcorner})$. Is E true or false? Justify your answer.

3. Let F be a self-referential formula such that R proves $F \leftrightarrow \neg \mathrm{THM}_{\mathrm{PA}}(\underline{\ulcorner \neg F \urcorner})$. Is F true or false? Justify your answer.

4. Let $T \supseteq \mathrm{R}$ be a consistent, axiomatizable theory. This exercise asks you to carry Rosser's construction of an independent sentence. A *Rosser proof* of a formula A is defined to be a proof P of A such that there is no proof P' of $\neg A$ with $\ulcorner P' \urcorner < \ulcorner P \urcorner$. More formally, let $\mathrm{NEG}(x_1, x_2)$ represent the mapping $\ulcorner A \urcorner \mapsto \ulcorner \neg A \urcorner$ in R and define $\mathrm{RPRF}_T(w, x)$ to be the formula

$$\mathrm{RPRF}_T(w, x) \;:=\; \mathrm{PRF}_T(w, x) \wedge \forall v \, \forall y \, (v < w \wedge \mathrm{NEG}(x, y) \rightarrow \neg \mathrm{PRF}_T(v, y)).$$

Let $\mathrm{RTHM}_T(x)$ be the formula $\neg \exists w \, \mathrm{RPRF}_T(w, x)$. Finally, let D_R be a self-referential formula such that R proves $D_R \leftrightarrow \neg \mathrm{RTHM}_T(\underline{\ulcorner D_R \urcorner})$. Also, let A be an arbitrary formula.

(a) There is a Rosser T-proof of A if and only if there is a T-proof of A.
(b) If $T \vdash A$, then $T \vdash \mathrm{RTHM}_T(\underline{\ulcorner A \urcorner})$.
[Hint: You will need to use $T \supseteq R$ and Axiom $\mathrm{R}'_{\leq}$.]
(c) If $T \vdash \neg A$, then $T \vdash \neg \mathrm{RTHM}_T(\underline{\ulcorner A \urcorner})$.
(d) Prove that $T \nvdash D_R$.
(e) Prove that $T \nvdash \neg D_R$.
(f) Conclude that D_R is independent of T.

5. Suppose $T \supseteq R$ is consistent, axiomatizable and satisfies the Hilbert-Bernays-Löb conditions. Let D be a self-referential formula such that $\mathrm{R} \vdash D \leftrightarrow \neg \mathrm{THM}_T(\underline{\ulcorner D \urcorner})$.

(a) Prove $T \vdash D \leftrightarrow \mathrm{CON}_T$.
(b) Suppose that E is another sentence such that $\mathrm{R} \vdash E \leftrightarrow \neg \mathrm{THM}_T(\underline{\ulcorner E \urcorner})$. Prove that $T \vdash D \leftrightarrow E$.
(c) Show that there is a theory T such that $T \vdash \neg D$. (This shows that the assumption of ω-consistency in Theorem IX.31 cannot be weakened to the assumption of consistency.)