Phil 140b	**Spring 2016**

Assignment #6

Nikhil Unni

1. *Suppose that $B(y)$ is a provability predicate for T, where T is consistent. Let $D(y)$ be the formula:*

$$B(y) \land y \neq \ulcorner 0 = 1 \urcorner$$

Show that $D(y)$ meets the third condition of the definition of a provability predicate but not the second.

The third provability condition of a provability predicate $D(y)$ is:

$$\models_T D(\ulcorner A \urcorner) \implies D(\ulcorner D(\ulcorner A \urcorner) \urcorner)$$

$$\models_T B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \implies D(\ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner)$$

$$\models_T B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \implies B(\ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner) \land \ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner \neq \ulcorner 0 = 1 \urcorner$$

Because of (P1), we know that:

$$\models_T B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \implies B(\ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner)$$

And using the coding scheme discussed in class (and chapter 15), we know that since $\ulcorner 0 = 1 \urcorner$ is inside the sentence $B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner$, we know that the number $\ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner$ must be some number where there are a sequence of digits on either side of $\ulcorner 0 = 1 \urcorner$, so we know that $\ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner$ can't be equal to $\ulcorner 0 = 1 \urcorner$. With this we have the third provability condition:

$$\models_T B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \implies B(\ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner) \land \ulcorner B(\ulcorner A \urcorner) \land \ulcorner A \urcorner \neq \ulcorner 0 = 1 \urcorner \urcorner \neq \ulcorner 0 = 1 \urcorner$$

For the second provability condition of $D(y)$ to be incorrect would mean:

$$\models_T D(\ulcorner A_1 \implies A_2 \urcorner) \not\Rightarrow (D(\ulcorner A_1 \urcorner) \implies D(\ulcorner A_2 \urcorner))$$

Replacing all of the negations:

$$\models_T B(\ulcorner A_1 \implies A_2 \urcorner) \land \ulcorner A_1 \implies A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner \not\Rightarrow (B(\ulcorner A_1 \urcorner) \land \ulcorner A_1 \urcorner \neq \ulcorner 0 = 1 \urcorner \implies B(\ulcorner A_2 \urcorner) \land \ulcorner A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner)$$

$$\models_T B(\ulcorner A_1 \implies A_2 \urcorner) \land \ulcorner A_1 \implies A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner \land \neg (B(\ulcorner A_1 \urcorner) \land \ulcorner A_1 \urcorner \neq \ulcorner 0 = 1 \urcorner \implies B(\ulcorner A_2 \urcorner) \land \ulcorner A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner)$$

$$\models_T B(\ulcorner A_1 \implies A_2 \urcorner) \land \ulcorner A_1 \implies A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner \land (B(\ulcorner A_1 \urcorner) \land \ulcorner A_1 \urcorner \neq \ulcorner 0 = 1 \urcorner \land \neg (B(\ulcorner A_2 \urcorner) \land \ulcorner A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner))$$

$$\models_T B(\ulcorner A_1 \implies A_2 \urcorner) \land \ulcorner A_1 \implies A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner \land B(\ulcorner A_1 \urcorner) \land \ulcorner A_1 \urcorner \neq \ulcorner 0 = 1 \urcorner \land (\neg B(\ulcorner A_2 \urcorner) \lor \ulcorner A_2 \urcorner = \ulcorner 0 = 1 \urcorner)$$

As discussed above, it should be clear that $\ulcorner A_1 \implies A_2 \urcorner \neq \ulcorner 0 = 1 \urcorner$ is true by the construction of codes. So instead we can show the truth of:

$$\models_T B(\ulcorner A_1 \implies A_2 \urcorner) \land B(\ulcorner A_1 \urcorner) \land \ulcorner A_1 \urcorner \neq \ulcorner 0 = 1 \urcorner \land (\neg B(\ulcorner A_2 \urcorner) \lor \ulcorner A_2 \urcorner = \ulcorner 0 = 1 \urcorner)$$

If we have a consistent system and if A_1 is not $0 = 1$, which would imply **any** sentence, from (P2) we have:

$$\models_T B(\ulcorner A_1 \urcorner) \land B(\ulcorner A_1 \implies A_2 \urcorner) \implies B(\ulcorner A_2 \urcorner)$$

However, if $B(\ulcorner A_1 \urcorner) \land B(\ulcorner A_1 \implies A_2 \urcorner) \land \neg B(\ulcorner A_2 \urcorner)$, we have an inconsistency with our last property, and so we can prove anything, namely $\ulcorner A_2 \urcorner = \ulcorner 0 = 1 \urcorner$. So the statement **must** be true, meaning that the second provability predicate condition is false.

2. *Let $B(y)$ be a provability predicate for T (extending Q). Show that T proves the following:*

(i) $B(\ulcorner A \wedge C \urcorner) \iff (B(\ulcorner A \urcorner) \wedge B(\ulcorner C \urcorner))$

First we can show:

$$B(\ulcorner \neg A \urcorner) \iff \neg B(\ulcorner A \urcorner)$$

We know that if you can prove $\neg A$, then there is no way to prove A, since this would introduce an inconsistency to the system. So we have:

$$B(\ulcorner A \wedge C \urcorner)$$
$$\iff B(\ulcorner \neg(\neg A \vee \neg C) \urcorner)$$
$$\iff B(\ulcorner \neg(A \implies \neg C) \urcorner)$$

Using what we proved above:

$$\iff \neg B(\ulcorner A \implies \neg C \urcorner)$$

And using (P2):

$$\iff \neg(B(\ulcorner A \urcorner) \implies B(\ulcorner \neg C \urcorner))$$
$$\iff \neg(B(\ulcorner A \urcorner) \implies \neg B(\ulcorner C \urcorner))$$
$$\iff \neg(\neg B(\ulcorner A \urcorner) \vee \neg B(\ulcorner C \urcorner))$$
$$\iff B(\ulcorner A \urcorner) \wedge B(\ulcorner C \urcorner)$$

(ii) $B(\ulcorner 0 = 1 \urcorner) \implies B(\ulcorner A \urcorner)$

Since Q has $0 \neq 1$ as an axiom, T is consistent if and only if $\not\models_T B(\ulcorner 0 = 1 \urcorner)$. So if $\models_T B(\ulcorner 0 = 1 \urcorner)$, then T is inconsistent, and so for any sentence A:

$$\models_T A$$

And from (P1), we have $\models_T B(\ulcorner A \urcorner)$.

(iii) $B(\ulcorner A \urcorner) \implies (B(\ulcorner \neg A \urcorner) \implies B(\ulcorner 0 = 1 \urcorner))$

This is equivalent to:

$$\neg B(\ulcorner A \urcorner) \vee (\neg B(\ulcorner \neg A \urcorner) \vee B(\ulcorner 0 = 1 \urcorner))$$

And using the negation property we showed in (i):

$$\neg B(\ulcorner A \urcorner) \vee B(\ulcorner A \urcorner) \vee B(\ulcorner 0 = 1 \urcorner)$$

This is true simply because if we have a consistent system, then either $B(\ulcorner A \urcorner)$ or $\neg B(\ulcorner A \urcorner)$, and if we don't have a consistent system, then $B(\ulcorner 0 = 1 \urcorner)$.

3. *Suppose $B(y)$ is a provability predicate for T. Use the existence of a sentence G such that:*

$$\models_T G \iff B(\ulcorner G \urcorner)$$

to construct an "alternative" proof that if T is consistent, then not:

$$\models_T \neg B(\ulcorner 0 = 1 \urcorner)$$

Suggestion : show

$$\models_T B(\ulcorner B(\ulcorner G \urcorner) \urcorner) \implies B(\ulcorner \neg G \urcorner),$$
$$\models_T B(\ulcorner G \urcorner) \implies B(\ulcorner \neg G \urcorner),$$

and

$$\models_T B(\ulcorner G \urcorner) \implies (B(\ulcorner \neg G \urcorner) \implies B(\ulcorner 0 = 1 \urcorner))$$

Conclude that if $\models_T \neg B(\ulcorner 0 = 1 \urcorner)$, *then:*

$$\models_T \neg B(\ulcorner G \urcorner),$$

$$\models_T G,$$

and

$$\models_T B(\ulcorner G \urcorner)$$

Assume:

$$(1) \models_T B(\ulcorner B(\ulcorner G \urcorner) \urcorner) \implies B(\ulcorner \neg G \urcorner)$$

Then, from (P3), we have:

$$(2) \models_T B(\ulcorner G \urcorner) \implies B(\ulcorner B(\ulcorner G \urcorner) \urcorner) \implies B(\ulcorner \neg G \urcorner)$$

Since T is consistent, $\neg B(\ulcorner G \urcorner) \vee \neg B(\ulcorner \neg G \urcorner) \vee B(\ulcorner 0 = 1 \urcorner)$ must be true, since you cannot prove both G and $\neg G$ while still maintaining a consistent system. This is equivalent to:

$$(3) \models_T B(\ulcorner G \urcorner) \implies (B(\ulcorner \neg G \urcorner) \implies B(\ulcorner 0 = 1 \urcorner))$$

With this all in mind, suppose that $\models_T \neg B(\ulcorner 0 = 1 \urcorner)$. Then if $\models_T B(\ulcorner G \urcorner)$, then from (3), we know that $\models_T \neg B(\ulcorner \neg G \urcorner)$. And from (2), we know that $\models_T \neg B(\ulcorner G \urcorner)$, which is an inconsistency. But if $\models_T \neg B(\ulcorner G \urcorner)$, then we get $\models_T G$ and $\models_T B(\ulcorner G \urcorner)$, which is also inconsistent.

Since our assumption lead to an inconsistency, we know $\not\models_T \neg B(\ulcorner 0 = 1 \urcorner)$

4. *Suppose that* $B(y)$ *is a provability predicate for* T *and that* $\models_T B(\ulcorner A \urcorner) \implies C$ *and that* $\models_T B(\ulcorner C \urcorner) \implies A$. *Show that* $\models_T A$ *and* $\models_T C$.

Let $D(y)$ be the formula $B(y) \implies A$, and let $E(y)$ be the formula $B(y) \implies C$ and apply the diagonal lemma to obtain sentences D, E such that:

$$(1) \models_T D \iff B(\ulcorner D \urcorner) \implies A$$

$$(2) \models_T E \iff B(\ulcorner E \urcorner) \implies C$$

By (P1), we have:

$$(3) \models_T B(\ulcorner D \iff B(\ulcorner D \urcorner) \implies A \urcorner)$$

$$(4) \models_T B(\ulcorner E \iff B(\ulcorner E \urcorner) \implies C \urcorner)$$

By (P2), we have:

$$(5) \models_T B(\ulcorner D \urcorner) \implies B(\ulcorner B(\ulcorner D \urcorner) \implies A \urcorner)$$

$$(6) \models_T B(\ulcorner E \urcorner) \implies B(\ulcorner B(\ulcorner E \urcorner) \implies C \urcorner)$$

$$(7) \models_T B(\ulcorner D \urcorner) \implies (B(\ulcorner B(\ulcorner D \urcorner) \urcorner) \implies B(\ulcorner A \urcorner))$$

$$(8) \models_T B(\ulcorner E \urcorner) \implies (B(\ulcorner B(\ulcorner E \urcorner) \urcorner) \implies B(\ulcorner C \urcorner))$$

From (P3) we have:

$$(9) \models_T B(\ulcorner D \urcorner) \implies B(\ulcorner B(\ulcorner D \urcorner) \urcorner)$$

$$(10) \models_T B(\ulcorner E \urcorner) \implies B(\ulcorner B(\ulcorner E \urcorner) \urcorner)$$

Combining (7) and (9) and (8) and (10) we get:

$$(11) \models_T B(\ulcorner D \urcorner) \implies B(\ulcorner A \urcorner)$$

$$(12) \models_T B(\ulcorner E \urcorner) \implies B(\ulcorner C \urcorner)$$

Combining (11) and (12) with the formulas stated in the problem:

$$(13) \models_T B(\ulcorner D \urcorner) \implies C$$

$$(14) \models_T B(\ulcorner E \urcorner) \implies A$$

Then combining (1) with (13) and (2) with (14) we get:

$$(15) \models_T D$$

$$(16) \models_T E$$

And with (P1):

$$(17) \models_T B(\ulcorner D \urcorner)$$

$$(18) \models_T B(\ulcorner E \urcorner)$$

And combining (13) with (17) and (14) with (18) we finally get:

$$\models_T C$$

$$\models_T A$$