The Agreement Theorem and the Substitution Theorems

I. Introduction.

In §3.4, the Hamilton text develops some crucial facts about valuations for the Predicate Calculus. The text does not state or prove these results as cleanly as I'd like; there is a even a flaw in one of the induction arguments. The purpose of this handout is to rewrite those parts of the section that seem to need it. Below, I establish three propositions. The first (which I call the "Agreement Theorem") is Proposition 3.33 of the text. I prove that *before* proving Proposition 3.23 (which I dub the "Formula Substitution Theorem"), because I need the Agreement Theorem to push through the proof of the Formula Substitution Theorem. (So does Hamilton, but he does not realize it.) In between these two propositions, I insert a third one (which I call the "Term Substitution Theorem"), whose main purpose is to help prove the Formula Substitution Theorem. (The Term Substitution Theorem *does* appear in Hamilton, but statement and proof are buried in the proof of Proposition 3.23, and I prefer to handle them separately.) Throughout, I will be dealing with valuations in one fixed interpretation.

II. The Agreement Theorem.

Definition. Let v and w be valuations, and let $\mathcal{A}$ be a *wf*. Say that v and w *agree* on $\mathcal{A}$ if $f_v(\mathcal{A}) = f_w(\mathcal{A})$;[1] that is, if

$$\begin{cases} \text{either both } v \text{ and } w \text{ satisfy } \mathcal{A} \\ \text{or neither } v \text{ nor } w \text{ satisfies } \mathcal{A}. \end{cases}$$

Proposition 3.33 (Agreement Theorem). *For all valuations v and w, and for every wf $\mathcal{A}$: if*

$$v(x) = w(x) \text{ for every variable that occurs free in } \mathcal{A}, \qquad (*)$$

then $f_v(\mathcal{A}) = f_w(\mathcal{A})$.

Proof. The first step is to observe that if any two valuations v and w have the property that

$$v\left(x_{i_j}\right) = w\left(x_{i_j}\right), 1 \leq j \leq n,$$

then it must also be the case that $v(t) = w(t)$ for all terms t that contain no variables other than (at most) $\{x_{i_1}, x_{i_2}, \ldots, x_{i_n}\}$. This is established by an easy induction argument.

Exercise 1. Supply the details of this argument.

The second step is to prove the assertion of the theorem by induction on the number k of logical symbols in $\mathcal{A}$.[2]

Basis ($k = 0$). In this case, $\mathcal{A}$ is an atomic *wf*, say $A_i^\ell(t_1, \ldots, t_\ell)$, so

$$\begin{aligned} v \text{ satisfies } \mathcal{A} &\iff \overline{A}_i^\ell(v(t_1), \ldots, v(t_\ell)) \text{ is true} \\ [\text{because, by Exercise } \mathbf{1}, v(t_i) = w(t_i), 1 \leq i \leq \ell] &\iff \overline{A}_i^\ell(w(t_1), \ldots, w(t_\ell)) \text{ is true} \\ &\iff w \text{ satisfies } \mathcal{A}. \end{aligned}$$

Inductive step. Suppose the statement of the theorem holds that for all *wf*s with $\leq (k-1)$ logical symbols (and for all valuations), and let $\mathcal{A}$ have $k > 0$ logical symbols.

Case 1. Say $\mathcal{A}$ is $\sim\mathcal{B}$. Then $\mathcal{A}$ and $\mathcal{B}$ have exactly the same free variables; so any valuations v and w that fulfill condition $(*)$ for $\mathcal{A}$ must also do so for $\mathcal{B}$. By induction, then, v and w must agree on $\mathcal{B}$. Thus we have:

$$\begin{aligned} v \text{ satisfies } \mathcal{A} &\iff v \text{ does not satisfy } \mathcal{B} \\ &\iff w \text{ does not satisfy } \mathcal{B} \\ &\iff w \text{ satisfies } \mathcal{A}. \end{aligned}$$

---

[1] Recall from class that any valuation v induces a $\{T, F\}$–valued function f_v defined on the set of *wf*s.

[2] That is, the number of occurrences of "$\sim$", "$\rightarrow$" and "$\forall$" in $\mathcal{A}$.

Case 2. Say $\mathcal{A}$ is $\mathcal{B} \to \mathcal{C}$. The set of free variables in $\mathcal{A}$ comprises precisely those that appear in $\mathcal{B}$ or $\mathcal{C}$ or both, so any valuations v and w that fulfill condition $(*)$ for $\mathcal{A}$ must also do so for $\mathcal{B}$ and $\mathcal{C}$. By induction, then, v and w must agree on $\mathcal{B}$ and $\mathcal{C}$, so we get

$$\begin{aligned} v \text{ does not satisfy } \mathcal{A} &\iff v \text{ satisfies } \mathcal{B} \text{ but does not satisfy } \mathcal{C} \\ &\iff w \text{ satisfies } \mathcal{B} \text{ but does not satisfy } \mathcal{C} \\ &\iff w \text{ does not satisfy } \mathcal{A}. \end{aligned}$$

Case 3 (trickier). Say $\mathcal{A}$ is $(\forall x_i)\mathcal{B}$. Let v and w fulfill condition $(*)$ for $\mathcal{A}$. I will argue that

$$v \text{ does not satisfy } \mathcal{A} \Longrightarrow w \text{ does not satisfy } \mathcal{A}; \tag{$\dagger$}$$

then, reversing the rôles of v and w will give

$$w \text{ does not satisfy } \mathcal{A} \Longrightarrow v \text{ does not satisfy } \mathcal{A},$$

and the induction will be complete.

So—to prove (†)—suppose that v does not satisfy $\mathcal{A}$. Then there is some valuation v' such that $v' \mathrel{\tilde{i}} v$ and such that v' does not satisfy $\mathcal{B}$. From v', manufacture a valuation w' such that $w' \mathrel{\tilde{i}} w$ and such that $w'(x_i) = v'(x_i)$. That is, set

$$\begin{cases} w'(x_i) := v'(x_i), & \text{and} \\ w'(x_j) := w(x_j), & j \neq i. \end{cases}$$

We are thus considering four valuations at the moment; the relationships among them are summarized in the following table. In it, $\{x_{j_1}, x_{j_2}, \ldots, x_{j_n}\}$ are the free variables of $\mathcal{A}$; the free variables of $\mathcal{B}$ comprise these n variables—and possibly x_i as well, if x_i appears free in $\mathcal{B}$.[3]

Variable	v	v'	w'	w
—	—	—	—	—
x_{j_1}	d_1	d_1	d_1	d_1
x_{j_2}	d_2	d_2	d_2	d_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_{j_n}	d_n	d_n	d_n	d_n
—	—	—	—	—
x_i	b_1	b_2	b_2	b_3

In the first n lines of the table:

column #1 matches column #4	because v and w agree on the free variables of $\mathcal{A}$;
column #1 matches column #2	because $v' \mathrel{\tilde{i}} v$; and
column #3 matches column #4	by the the construction of w'.

In the last line of the table: $v'(x_i) = w'(x_i)$ by construction of w'.

Now, since the free variables of $\mathcal{B}$ are at most $\{x_{j_1}, x_{j_2}, \ldots, x_{j_n}, x_i\}$, v' and w' fulfill condition $(*)$ for $\mathcal{B}$, and so by induction, they must agree on $\mathcal{B}$. But v' does not satisfy $\mathcal{B}$; so w' does not satisfy $\mathcal{B}$ either. Finally, since $w' \mathrel{\tilde{i}} w$, w does not satisfy $\mathcal{A}$. ∎

---

[3] If $\mathcal{B}$ does not contain x_1 free, then one can take $v' := v$ and $w' := w$. Note also that $\mathcal{B}$ might have *more* free variables than $\mathcal{A}$, but it has *fewer* logical symbols, so that the induction will be valid.

III. The Substitution Theorems.

Let t be a fixed term, and let x_i be a fixed variable. The rest of the handout deals with the effects of replacing free occurrences of x_i with t in terms, in *wf*s, and in valuations. The precise definitions are these.

Definitions.

[a]: For any term u, let $\text{sub}(u, t \to x_i)$ be the term that results from replacing each occurrence of x_i in u with term t.

[b]: For any *wf* $\mathcal{A}$, let $\text{sub}(\mathcal{A}, t \to x_i)$ be the term that results from replacing each *free* occurrence[4] of x_i in $\mathcal{A}$ with term t.

[c]: For any valuation v, let $\text{sub}(v, t \to x_i)$ be the valuation v' such that $v' \tilde{i} v$ and $v'(x_i) = v(t)$. That is,

$$\begin{cases} v'(x_i) := v(t), & \text{and} \\ v'(x_k) := v(x_k), & k \neq i. \end{cases}$$

The intuition behind the Term Substitution Theorem (below) is this. For any term u and any valuation v: if you plan to replace x_i with t, it should not matter whether you (a), replace x_i with t in u and then use v to evaluate, or (b), leave u unchanged but instead make the corresponding change to v.

Proposition 3.225 (Term Substitution Theorem). *For any term u and any valuation v, let $u' = \text{sub}(u, t \to x_i)$ and let $v' = \text{sub}(v, t \to x_i)$. Then $v'(u) = v(u')$.*

The diagram below should help you get the statement of this theorem into focus:

$$\begin{tikzcd} u \arrow[rr, "\text{sub}"] \arrow[ddrr, "v'"'] & & u' \arrow[dd, "v"] \\ & & \\ & & \text{same} \end{tikzcd}$$

Proof. **Exercise 2.** (*Hints*: Argue by induction on the number of function letters in u. The basis step has three cases; argue each one carefully.) ∎

The idea behind the Formula Substitution Theorem is similar but subtler; we need to add a restriction. Let $\mathcal{A}$ be a *wf*. *If t is free for x_i in $\mathcal{A}$*, then it should not matter whether you (a), replace x_i with t in $\mathcal{A}$ and then apply f_v; or (b), leave $\mathcal{A}$ unchanged but instead make the corresponding change to v. The following example illustrates the need for italicized restriction.

Example. We will consider an interpretation I for a first-order language in which D_I is the set of positive real numbers. Let x, y and z be variables of the language; assume that in the interpretation, the function letter f_1^1 names the function $\overline{f}_1^1(r) = r^3$; and let $\mathcal{A}(x)$ be interpreted by the assertion,

$$\lim_{y \to 0^+} \left(\frac{y}{x} \right) \text{ exists.}$$

Exercise 3. Let t be the term $f_1^1(z)$, let $v(z) = 4$, and let $v' = \text{sub}(v, t \to x)$. (Observe that t is free for x in $\mathcal{A}(x)$.) Write down the assertion that interprets $\mathcal{A}(t)$; find the statement whose truth or falsity determines $f_v(\mathcal{A}(t))$; and find the statement whose truth or falsity determines $f_{v'}(\mathcal{A}(x))$. Do the truth values match?

Exercise 4. Let t be the term $f_1^1(y)$, let $v(y) = 4$, and let $v' = \text{sub}(v, t \to x)$. (Observe that t is *not* free for x in $\mathcal{A}(x)$.) Write down the assertion that interprets $\mathcal{A}(t)$; find the statement whose truth or falsity determines $f_v(\mathcal{A}(t))$; and find the statement whose truth or falsity determines $f_{v'}(\mathcal{A}(x))$. Do the truth values match?

Proposition 3.23 (Formula Substitution Theorem). *Let $\mathcal{A} = \mathcal{A}(x_i)$ be a wf, and let $\mathcal{A}(t) = \text{sub}(\mathcal{A}(x_i), t \to x_i)$. If t is free for x_i in $\mathcal{A}(x_i)$, then for any valuation v,*

$$f_{v'}(\mathcal{A}(x_i)) = f_v(\mathcal{A}(t)),$$

where $v' = \text{sub}(v, t \to x_i)$.

---

[4] In definition **[a]**, *all* occurrences of x_i in u are free. Why?

A diagram similar to the one for the Term Substitution Theorem should help you to get this statement into focus:

$$\begin{tikzcd}
\mathcal{A}(x_i) \arrow[r, "\text{sub}"] \arrow[rrd, "f_{v'}"'] & {} \arrow[r] & \mathcal{A}(t) \arrow[d, "f_v"] \\
 & & \text{same}
\end{tikzcd}$$

Proof by induction on the number k of logical symbols in $\mathcal{A}(x_i)$.

Basis ($k = 0$). In this case, there is a predicate letter A^n_ℓ for which $A(x_i) = A^n_\ell(u_1, \dots, u_n)$ and $A(t) = A^n_\ell(u'_1, \dots, u'_n)$, where $u'_j = \text{sub}(u_j, t \to x_i)$, for $1 \le j \le n$. Thus

$$\begin{aligned}
v \text{ satisfies } \mathcal{A}(t) &\iff (v(u'_1), \dots, v(u'_n)) \in \overline{A^n_\ell} \\
\text{(by the Term Substitution Theorem)} &\iff (v'(u_1), \dots, v'(u_n)) \in \overline{A^n_\ell} \\
&\iff v' \text{ satisfies } \mathcal{A}(x_i).
\end{aligned}$$

Inductive step. Suppose the statement of the theorem holds that for all *wf*s with $\le (k-1)$ logical symbols, and let $\mathcal{A}(x_i)$ have $k > 0$ logical symbols.

Case 1. Say $\mathcal{A}(x_i)$ is $(\sim\mathcal{B})(x_i)$.

Exercise 5. Complete the proof of this case.

Case 2. Say $\mathcal{A}(x_i)$ is $(\mathcal{B} \to \mathcal{C})(x_i)$.

Exercise 6. Complete the proof of this case.

Case 3. Say $\mathcal{A}(x_i)$ is $(\forall x_i)\mathcal{B}(x_i)$. Then $\mathcal{A}$ contains no free occurrences of x_i, so that $\mathcal{A}(x_i)$ and $\mathcal{A}(t)$ are the same *wf* $\mathcal{A}$. Moreover, v and v' agree on every free variable of $\mathcal{A}$, so by the Agreement Theorem,

$$f_v(\mathcal{A}) = f_{v'}(\mathcal{A}).$$

Case 4 (the interesting case). Say $\mathcal{A}(x_i)$ is $(\forall x_j)\mathcal{B}(x_i)$, where $j \ne i$; In this case, $\mathcal{A}(t)$ is $(\forall x_j)\mathcal{B}(t)$. Since t is free for x_i in $\mathcal{A}(x_i)$, either

4(α) : $\mathcal{A}(x_i)$ contains no free occurrences of x_i, or
4(β): t contains no occurrences of x_j.

Case 4(α) is exactly like case 3: because $\mathcal{A}(x_i)$ contains no free occurrences of x_i, the two formulas $\mathcal{A}(x_i)$ and $\mathcal{A}(t)$ are the same formula, and v and v' agree on it. So we need to argue case 4(β): t has no occurrence of x_j. I will prove both directions of the implication

$$f_{v'}(\mathcal{A}(x_i)) = F \iff f_v(\mathcal{A}(t)) = F; \tag{††}$$

this will establish 4(β) and complete the induction. The diagram below will be useful for understanding the proofs in both directions.

$$\begin{tikzcd}
v \arrow[r, leftrightarrow, "\underset{j}{\sim}"] \arrow[d, leftrightarrow, "\underset{i}{\sim}"'] & w \arrow[d, leftrightarrow, "\underset{i}{\sim}"'] \\
v' \arrow[r, leftrightarrow, "\underset{j}{\sim}"] & w'
\end{tikzcd}$$

$\Longleftarrow$: Say $f_v(\mathcal{A}(t)) = F$. Then there is a valuation $w \underset{j}{\sim} v$ for which $f_w(\mathcal{B}(t)) = F$. For this w, put $w' := \text{sub}(w, t \to x_i)$; that is, put

$$\begin{cases} w'(x_i) := w(t), & \text{and} \\ w'(x_k) := w(x_k), & k \ne i. \end{cases}$$

By the inductive step, we know that $f_{w'}(\mathcal{B}(x_i)) = f_w(\mathcal{B}(t))$, so $f_{w'}(\mathcal{B}(x_i)) = F$. Moreover,

I CLAIM THAT $w' \stackrel{\sim}{j} v'$,

from which it will follow that $f_{v'}(\mathcal{A}(x_i)) = F$, and the proof of $\Longleftarrow$ will be complete.

Proof of claim. Because of the equivalences indicated in the diagram,

$$w'(x_k) = w(x_k) = v(x_k) = v'(x_k) \text{ for all } k \notin \{i, j\}; \tag{‡}$$

it remains to show that $w'(x_i) = v'(x_i)$. The crucial observation here is this: since $w \stackrel{\sim}{j} v$ and t contains no occurrences of x_j, the Agreement Theorem implies that $w(t) = v(t)$. So in this case—see diagram—we have

$$w'(x_i) = w(t) = v(t) = v'(x_i).$$

Thus $w' \stackrel{\sim}{j} v'$, and the proof of the claim is complete.

$\Longrightarrow$: Say $f_{v'}(\mathcal{A}(x_i)) = F$. Then there is a valuation $w' \stackrel{\sim}{j} v'$ for which $f_{w'}(\mathcal{B}(x_i)) = F$. From this valuation w', define a valuation w by putting $w := \text{sub}(w', v(x_i) \to x_i)$; that is, put

$$\begin{cases} w(x_i) := v(x_i), & \text{and} \\ w(x_k) := w'(x_k), & k \neq i. \end{cases}$$

Thus, by construction, $w \stackrel{\sim}{j} v$. Moreover,

I CLAIM THAT $w' = \text{sub}(w, t \to x_i)$.

Proof of claim. By construction of w, $w' \stackrel{\sim}{i} w$; it remains to establish that $w'(x_i) = w(t)$. One argues this as follows (see diagram):

$$\begin{array}{rl} & w'(x_i) \\ \text{because } w' \stackrel{\sim}{j} v' \to & = v'(x_i) \\ \text{because } v' = \text{sub}(v, t \to x_i) \text{ by hypothesis} \to & = v(t) \\ \text{argued in proof of first claim} \to & = w(t). \end{array}$$

Thus $w' = \text{sub}(w, t \to x_i)$, so he claim has been proved.

The claim now permits us to apply the inductive hypothesis to w, w', and $\mathcal{B}$, so that

$$f_w(\mathcal{B}(t)) = f_{w'}(\mathcal{B}(x_i)) = F;$$

and, since $w \stackrel{\sim}{j} v$, this in turn implies that $f_w(\mathcal{B}(t)) = F$. The proof of $\Longrightarrow$, of (††) and of the theorem are now complete. ∎