§1. Semantic theory and the paradoxes. Frege’s attempts to formulate a theory of properties to serve as a foundation for logic, mathematics and semantics all dissolved under the weight of the logical paradoxes. The language of Frege’s theory permitted the representation of the property which holds of everything which does not hold of itself. Minimal logic, plus Frege’s principle of abstraction, leads immediately to a contradiction. The subsequent history of foundational studies was dominated by attempts to formulate theories of properties and sets which would not succumb to the Russell argument. Among such are Russell’s simple theory of types and the development of various iterative conceptions of set. All of these theories ban, in one way or another, the self-reference responsible for the paradoxes; in this sense they are all “typed” theories. The semantical paradoxes, involving the concept of truth, induced similar nightmares among philosophers and logicians involved in semantic theory. The early work of Tarski demonstrated that no language that contained enough formal machinery to represent the various versions of the Liar could contain a truth-predicate satisfying all the Tarski biconditionals. However, recent work in both disciplines has led to a re-evaluation of the limitations imposed by the paradoxes.

In the foundations of set theory, the work of Gilmore [1974], Feferman [1975], [1979], [1984], and Aczel [1980] has clearly demonstrated that elegant and useful type-free theories of classes are feasible. Work on the semantic paradoxes was given new life by Kripke’s contribution (Kripke [1975]). This inspired the recent work of Gupta [1982] and Herzberger [1982]. These papers demonstrate that much room is available for the development of theories of truth which meet almost all of Tarski’s desiderata.

The primary motivation for the present theory stems from the desire to develop a new foundation for natural language semantics. Recent work in semantics and philosophical logic (Chierchia [1985]; Turner [1984]; Barwise and Perry [1983]; Kamp [1983]) has stressed the need for theories of properties, relations and propositions in which type-freeness is a major desiderata. Apart from the obvious need for semantic theory to come to terms with the paradoxes there are more specific issues (Chierchia [1985]) to do with nominalization, the semantics of
perception (Kamp [1983]), and the analysis of belief (Barwise and Perry [1983]). Chierchia, for example, has quite persuasively argued that semantic theory requires a theory of properties which is not hampered by the strong type discipline imposed by Montague semantics. A second problem with traditional foundations relates to the analysis of intensional notions. Much recent work in semantics employs the notion of possible world: properties are represented as functions from worlds into sets and propositions as functions from worlds into truth values. Despite the success of this regime it suffers from many well-known difficulties. There is a growing consensus that many of these difficulties can be blamed on the employment of a too-extensional notion of property. Current theories seem at odds with the basic intuition that two properties can be logically equivalent without being identical.

The present paper presents a theory of properties, relations and propositions which is type-free and intensional (i.e. the axiom of extensionality fails). The intuitions which drive the theory are based on those model-theoretic intuitions which underpin the theories of truth developed by Gupta and Herzberger. This paper is largely concerned with the development of the underlying theory. The theory has some intrinsic interest with respect to the foundations of mathematics. In this regard, we shall, throughout, exploit connections with the theories of Gilmore, Aczel and Feferman. The actual application of the present theory to natural language semantics is detailed in Chierchia and Turner [1985].

§2. A theory of properties, relations and propositions. In this section we present a theory of properties, relations and propositions which is motivated by reflection upon Frege's principle of abstraction. We shall, for reasons of generality, assume an underlying first-order language $L_0$, and a first-order theory $T_0$. Initially, at least, we only assume that $L_0$ and $T_0$ have the following content. $L_0$ is to contain a constant symbol 0 (also written 0), a binary operation symbol $d_-$, and two unary operation symbols $q_1$ and $q_2$. We also assume that the underlying theory ($T_0$) has the following formulae provable:

(i) $d_-(x, y) \neq 0,$
(ii) $q_1(x, y) = x \& q_2(x, y) = y.$

Consequently, $d_-(x, y)$ acts as a pairing operation, where we have written $(x, y)$ for $d_-(x, y).$ We shall further unpack the content of $L_0$ and $T_0$ as we proceed. In $T_0$ we can define $x' = (x, 0).$ Thus $q_1$ acts as a predecessor operation, and from the axioms of $T_0$ we can derive:

(iii) $x' \neq 0,$
(iv) $x' = y' \rightarrow x = y.$

Tuples $(t_1, \ldots, t_n)$ can be introduced recursively: $(t) = t$ and $(t_1, \ldots, t_k, t_{k+1}) = ((t_1, \ldots, t_k), t_{k+1}).$ Moreover, we have projection operations $q_i(t_1, \ldots, t_k) = t_i,$ etc.

The language of the theory is essentially that of Bealer [1982]. It is a first-order language which, in addition to the apparatus of $L_0$ and equality, contains a family of relation symbols $\langle p_k \rangle_{k \geq 0}.$ The basic vocabulary of $L$ also includes the usual connectives ($\&$, $\lor$, $\neg$, $\forall$, $\exists$, $\rightarrow$, $\leftrightarrow$) together with denumerable sets of individual variables ($\{x_i : i \geq 0\}$) and individual constants ($\{c_i : i \geq 0\}$) (we shall often drop the subscripts). We define the terms and wffs of $L$ by (simultaneous) recursion as follows:

(i) every variable, constant or term of $L_0$ is a term of $L$; (ii) if $t_0, \ldots, t_k$ are terms then
A THEORY OF PROPERTIES

$p_k(t_0, \ldots, t_k)$ is an atomic wff; (iii) if $A$ is a wff and $x_1, \ldots, x_k, y_1, \ldots, y_m$ include all the variables of $A$, then $\hat{x}_1, \ldots, \hat{x}_k[A]$ is a term in which $x_1, \ldots, x_k$ are bound and $y_1, \ldots, y_m$ are free. (iv) if $A$ is a wff and $x$ is a variable then $\forall x A$ is a wff; (v) if $s$ and $t$ are terms then $s = t$ is a wff; (vi) if $A$ and $B$ are wffs then so are $\neg A$ and $A \land B$; (vii) every wff of $L_0$ is a wff of $L$. Other wff are obtained by definition in the conventional way.

We adopt the usual conventions regarding bondage and freedom. Moreover, we shall write $A(t)$ and $s(t)$ for the result of substituting $t$ for $x$ in $A$ and $s$, respectively. We shall often write $t$ for the tuple of terms $t_1, \ldots, t_k$ where $k$ is understood. $A(t)$ then represents the result of simultaneously substituting $t_i$ for $x_i$, $1 \leq i \leq k$. We shall be most interested in two special cases. When $k = 1$, $p_k$ is to act as the membership relation, and when $k = 0$, $p_k$ is to act as the truth predicate. We shall write $p$ for $p_1$ and $T$ for $p_0$. When the case $k = 1$ is considered we shall frequently drop the subscript—this applies to all operations we define where the subscripting is absent; the default case is $k = 1$.

We could actually introduce the terms of $L$ as terms of $L_0$ (following Feferman [1984]) as follows. Using the apparatus of $L_0$ let $r_A$ be the Gödel number of $A$. If $A$ has variables among $x_1, \ldots, x_k, y_1, \ldots, y_n$ then $(r_A, y_1, \ldots, y_n)$ serves as an operation which “abstracts” $x_1, \ldots, x_k$, treating $y_1, \ldots, y_n$ as parameters. We could then introduce $\hat{x}_1, \ldots, \hat{x}_k[A(x_1, \ldots, x_k, y_1, \ldots, y_n)]$, by definition, as $(r_A, y_1, \ldots, y_n)$. We shall appeal to this representation in the next section.

The intended interpretation of $p_k(\hat{x}[A], t)$ is supplied by the axiom schema of abstraction:

(A) $p_k(\hat{x}[A], t) \leftrightarrow A(t)$.

Of course, (A) is inconsistent with classical logic, since Russell’s property can be represented by $r = \hat{x}[(\neg p(x, x)]$, and (A) applied to $p(r, r)$ yields an immediate contradiction. Thus we cannot adopt (A) in full generality. Nevertheless, the intuitions which underpin (A) are extremely compelling; they only need to be modified a little. Our intention is to unpack the content of (A) in an effort to see which aspects can be salvaged and which have to be discarded. This is not the usual route taken. More often one searches for philosophical scruples which might be employed to rule out certain instances. In contrast, our initial approach will be more formal. We shall employ (A) as a guiding principle in the development of the theory. Shortly, we shall offer a more conventional model-theoretic approach to justifying the resulting theory.

A central observation in the development concerns the dual role played by wffs. In their primary or “external” role they occur as the wffs themselves; in their secondary or “internal” role they are located withing the scope of an abstraction operator. In their primary role, we assume that wffs are to be governed by the laws of classical logic. This leaves us to spell out the logic of their secondary or internal role. For example, is the following principle to be adopted:

$(p_k(\hat{x}[A], t) \land p_k(\hat{x}[A \rightarrow B], t)) \rightarrow p_k(\hat{x}[B], t)$?

More generally, is this internal logic also to be taken as classical? Certainly, (A) guarantees this since it reduces the internal logic to the external one. Therefore, to remain in spirit with (A), we shall also adopt a classical logic internally.
For convenience, we first present a standard axiomatization of classical logic. In addition to the axioms of equality we adopt the following axiomatization.

(i) The “external” logic of wffs. AXIOMS.

(E1) \( A \rightarrow (B \rightarrow A) \).

(E2) \( (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)) \).

(E3) \( (\neg A \rightarrow \neg B) \rightarrow ((\neg A \rightarrow B) \rightarrow A) \).

(E4) \( \forall x A \rightarrow A(t) \) (\( t \) free for \( x \) in \( A \)).

(E5) \( \forall x(A \rightarrow B) \rightarrow (A \rightarrow \forall x B) \) (\( A \) contains no free occurrences of \( x \)).

RULES. MP. \( \vdash A \) and \( \vdash A \rightarrow B \), then \( \vdash B \).

GEN. \( \vdash A \), then \( \vdash \forall x A \).

We shall write \( T_0 \vdash A \) for \( A \) derivable from \( T_0 \).

(ii) The “internal” logic of wffs.

(11)–(15) \( p_k(\hat{x}[A], t) \) for \( A \) any instance of (E1)–(E5), and \( t \) any terms.

(16) \( p_k(\hat{x}[A], t) \) & \( p_k(\hat{x}[A \rightarrow B], t) \rightarrow p_k(\hat{x}[B], t) \).

(17) \( (\forall y)[p_k(\hat{x}[A], t)] \rightarrow p_k(\hat{x}[\forall y A], t) \).

These principles govern the internal logic of the connectives \( \lor, \land, \neg, \exists, \forall \), etc.: they guarantee that the logic is classical.

We shall have occasion to consider, not just the bare logics, but also certain first-order theories. When some first-order theory \( Z \) is considered we shall supplement the internal theory by:

\[
(IZ) \quad T([A])
\]

where \( A \) is any wff of \( Z \). Thus the internal logic, IL, has the form (11)–(17). If IL is supplemented by \( Z \) we shall write \( IL_Z \), to indicate the inclusion of (IZ) as part of the internal theory. We shall assume a basic internal theory \( Z_0 \) which includes the usual axioms of equality, together with the following schema which relates the various instances of predication:

\[
(S) \quad p_k(\hat{x}[A], t) \leftrightarrow T([A(t)])
\]

This captures the obvious intuition that \( k \)-place predication is definable in terms of substitution and the truth-predicate.

THEOREM 2.1. If \( Z \vdash A \) then \( IL_Z \vdash T([A]) \).

PROOF. By induction on the proof of \( A \). If \( A \) is an axiom of \( Z \) then employ (IZ). If \( A \) is an instance of (E1)–(E5), then the result follows from (11)–(15). If \( A \) follows by (MP) we employ (16); if it follows by (GEN) we employ (17). \( \blacksquare \)

Observe that we cannot derive

\[
\text{If } IL_Z \vdash A \text{ then } IL_Z \vdash T([A]).
\]

The addition of such a rule will be considered shortly when we develop a modal version of the theory.

The remainder of the content of (A) can be expressed in terms of certain principles of connection which relate the two logics. For pedagogical reasons we divide the
schema (A) into its components:

(E) \[ A(\xi) \rightarrow p_k(\hat{x}[A], t), \]
(R) \[ p_k(\hat{x}[A], t) \rightarrow A(\xi). \]

We shall refer to (E) as Expansion and (R) as Reduction. The conjunction of (E) and (R) renders the theory inconsistent: one of them must be given up, but which? There appears to be no obvious intuitive grounds for choosing between them. However, the model-theoretic account which follows suggests a modal interpretation of predication whereby \( p_k(\hat{x}[A], t) \) is best unpacked as \( A(\xi) \) is, in some sense, “necessarily” true. This suggests the adoption of (R) and the rejection of (E)—at least in general.

The rest of the theory is concerned with (E). In this connection, the inspiration stems from the fact that every wff can be equivalently expressed in prenex disjunctive normal form. In this set-up negations only govern atomic wffs. This motivates the following schemata: For atomic \( A \),

(A1) \[ A(\xi) \rightarrow p_k(\hat{x}[A], t), \]
(A2) \[ \neg A(\xi) \rightarrow p_k(\hat{x}[\neg A], t). \]

In other words, to derive all the instances of expansion, it is sufficient to add expansion for atomic and negative atomic wffs. Unfortunately the theory will again be inconsistent. In fact, under the assumption of (R), (A2) is precisely the instance of (E) which enables the derivation of Russell’s paradox. Thus, under the assumption of (R), we have located the aspect of (A) from which the inconsistency stems, namely (A2). This is the principle we must give up—but the only one.

However, although we cannot maintain (A2) in full generality, we can replace it by a slight weakening without destroying the inconsistency. In this regard, observe that, under the assumptions (R) and (A1), the schema (A2) is equivalent to:

(LEMP) For each atomic wff \( A \), \( \forall z (p_k(\hat{x}[A], z) \lor p_k(\hat{x}[\neg A], z)) \)

—a law of excluded middle for properties. So, although we always have \( p_k(\hat{x}[A], z) \lor \neg p_k(\hat{x}[A], z) \), (LEMP), in general, fails. This failure highlights the fact that the “internal negation” of properties is not always definable in terms of sentential negation. The theory thus employs two negations which are both classical but which are not interdefinable. We can institutionalize this by introducing “negative” predication relations. The easiest way to achieve is as follows.

Define \( \bar{p}_k(t) =_{\text{def}} T([\neg p_k(t)]) \).

To ensure that \( \bar{p}_k \) acts as a negative predication relation we insist on

(IN) \[ \bar{p}_k(\hat{x}[A], t) \leftrightarrow p_k(\hat{x}[\neg A], t). \]

It follows directly from (R) that \( p_k \) and \( \bar{p}_k \) are mutually exclusive, i.e.

(DIS) \[ \forall x (\neg (p_k(x) \& \bar{p}_k(x))). \]

We can now state our analogue of (A2) which forces \( p_k(\hat{x}[\neg A], t) \) (for atomic \( A \)) to be equivalent to the “internal negation” of \( A(\xi) \):

(A2') \[ \bar{p}_k(\hat{x}[A], t) \leftrightarrow \bar{A}(\xi), \]

where \( \bar{p}_k(t) = \bar{p}_k(t) \).
As a direct consequence of \((R), (A1), (IN)\) and \((A2')\) we obtain
\[
(\text{NEG}) \quad p_k(\bar{x}[^0 p_k(\bar{y}[^A], t)], s) \leftrightarrow p_k(\bar{x}[^p_k(\bar{y}[^A], t)], s),
\]
which allows negations to be pushed through to the innermost atomic wff. Indeed, in the presence of \((R)\) and \((A1)\), \((IN)\) plus \((A2')\) is equivalent to \((\text{NEG})\).

Let \(T\) be the theory consisting of the inner logic plus \(Z_0, (R), (A1), (IN)\) and \((A2')\).

A more elegant formulation of the theory, which brings out its modal character, can be obtained by defining \(\Box A =_{\text{def}} T([A])\). The theory then takes on the following form.

\textbf{Modal Version \(M\).}

\[
\begin{align*}
\Box A & \to A && \text{(T)}, \\
A & \to \Box A && \text{(S4) for atomic \(A\),} \\
(\Box A & \land \Box(A \to B)) & \to \Box B && \text{(Imp),} \\
\forall x \Box A & \to \Box \forall x A && \text{(Barcan),} \\
p_k(\bar{x}[^A], t) & \leftrightarrow \Box(A(t)) && \text{(S),} \\
\Box \neg \Box A & \leftrightarrow \Box(\neg \Box A) && \text{(Neg),} \\
Z_0 & \vdash A \text{ then } M & \vdash \Box A && \text{(Nec').}
\end{align*}
\]

\textbf{THEOREM 2.2.} \(T\) and \(M\) are equivalent.

\textbf{PROOF.} Assume \(T\). Then axiom \((T)\) is a special case of \((R)\) where \(k = 0\). Again, for \(k = 0\), \((A1)\) takes on the form \(B \to T([B])\) for \(B\) atomic. This yields \((S4)\). \((\text{Imp})\) is a special case of \((16)\) for \(k = 0\), and \((\text{Barcan})\) an instance of \((17)\) for \(k = 0\). \((\text{Nec'})\) follows from Theorem 2.1. \((S)\) is just a reformulation of the original.

Conversely, assume \(M\). We have to check that the more general schemata follow. \((11)-(15)\) follows from \((\text{Nec'})\) and \((S)\). \((16)\) and \((17)\) follow from \((S)\) and \((\text{Imp})\) and \((S)\) and \((\text{Barcan})\), respectively. \((R)\) follows from \((S)\) and \((T)\). \((A1)\) follows from \((S)\) and \((S4)\), and \((\text{NEG})\) follows from several applications of \((S)\) (part of the underlying theory \(Z_0\)) and \((\text{Neg})\) in the modal version. \(\blacksquare\)

This modal version is "almost" \(\text{S4} + \text{Barcan}\). The difference concerns the weakened version of the rule of necessitation. We have only the special case of necessitation where \(A\) is a consequence of the underlying theory \(Z_0\). Indeed, full necessitation permits the use of \((R)\), internally, and this allows the derivation of Russell's paradox. Similarly, the addition of the \((S5)\) axiom \(\sim \Box \sim A \to \Box(\sim \Box \sim A)\) leads to Russell's paradox. No obvious modal strengthening of the theory seems possible.

\textbf{§3. Models of \(T\).} Our model-theoretic account of \(T\) draws its inspiration from the recent work of Gupta [1982] and Herzberger [1982] on the theory of truth. Essentially, we shall construct a theory of predication which is analogous to the Gupta-Herzberger theory of truth. For ease of exposition we shall only consider the case where \(k = 1\).

In general, models of \(L\) have the form \(M = \langle M, P \rangle\), where \(M\) is a nonempty set and \(P\) is a binary relation on \(M\). We here assume (as outlined in §2) that the abstractions are terms of \(L_0\). We also assume \(M\) is a model of \(T_0\).

For convenience, we shall further assume that \(L\) is expanded to \(L^M\) which is to include a symbol \(m'\) for each \(m \in M\), and that \(M\) contains a subset generated by the
successor operation. Let \textbf{Form} be the subset of $M$ consisting of all $(r_A, m_1, \ldots, m_n)$, where $A = A(x, m'_1, \ldots, m'_n)$. Thus each such term $\hat{x}[A(x, m'_1, \ldots, m'_n)]$ denotes an element of $M$. We can than provide the truth-conditions for the sentences of $L$ as follows.

(i) If $A$ is an atomic sentence of $L_0^M$ then

$$[[A]]^M = \begin{cases} t & \text{if } M \models A, \\ f & \text{otherwise.} \end{cases}$$

(ii) $[[p(t_1, t_2)]]^M = 1$ iff $[[t_1]]^M, [[t_2]]^M \in P$, where $t_1$ and $t_2$ are closed terms of $L_0^M$.

(iii) $[[A \& B]]^M = 1$ iff $[[A]]^M = 1$ and $[[B]]^M = 1$.

(iv) $[[\neg A]]^M = 1$ iff $[[A]]^M \neq 1$.

(v) $[[\forall x A]]^M = 1$ iff for each $m \in M$, $[[A(m)]]^M = 1$.

The idea behind the construction of models for $T$ is elegantly simple and is based upon those intuitions which motivate the axiom schema of abstraction. Our aim is to construct a model for $L$ which validates as many instances of (A) as possible. We shall construct a sequence of models in which the relation of predication is continually modified.

Initially, the relation of predication only gives the appropriate interpretation for those elements of $M$ which are not elements of \textbf{Form}. For elements of \textbf{Form} the interpretation is arbitrary, and hence the truth of

$$(A) \quad p(\hat{x}[A], t) \leftrightarrow A(t)$$

is not guaranteed.

We wish to force (A) wherever possible. To this end we construct a new model, $J(M) = \langle M, J(P) \rangle$, where

$$\langle k, m \rangle \in J(P) \iff [[A(m', m'_1, \ldots, m'_n)]]^M = 1,$$

where $k = (r_A, m_1, \ldots, m_n) \in \textbf{Form}$, and otherwise the extension of $J(P)$ is the same as that of $P$.

In other words, we revise $P$ in an effort to validate (A). Certain instances of (A) will certainly follow from this revision, namely those which involve wffs which contain no elements of \textbf{Form} in their atomic subformulae. But this is only a small step forward. For those wffs which do make reference to elements of \textbf{Form}, through their atomic subformulae, the revision guarantees nothing. Subsequently, we repeat the process and form a second revision, $J(J(M))$. At this stage certain instances of (A), which involve wffs with at most one level of nested elements of \textbf{Form}, will be validated. Hence by repeating this process more and more instances of (A) get validated.

This is a crude account of the process of revision. Actually, things are a good deal more involved than this simple picture implies. Certain instances of predication (e.g. $p(\hat{x}[\sim p(x, x)], \hat{x}[\sim p(x, x)])$ can never satisfy the schema. In fact, as the process of model-revision is carried through, such instances continually oscillate in truth-value. For example, if $p(\hat{x}[\sim p(x, x)], \hat{x}[\sim p(x, x)])$ is true at the initial model, it will be false at the second, true at the third, and so on.
The phenomenon is relevant to the extension of the process to transfinite stages—something which will be surely necessary if the process is to "terminate", in some intuitive sense, i.e. where all consistently possible instances of (A) get validated. Here we follow the lead of Herzberger rather than Gupta.

At limit ordinals we place in the extension of the predication relation just those instances which have assumed a constant value. In other words, for a limit ordinal $\lambda$, we place in the extension of $A$ those pairs $\langle m, m' \rangle$ which are in the the extension of $P_\beta$ for all large enough $\beta < \lambda$. More precisely, we can summarize the process of model-revision as follows:

**Definition 3.1.** Let $M$ be any model of $T_0$. We define an ordinal sequence of models $\langle M_\alpha \rangle_{\alpha \geq 0}$, where $M_0 = M$ and $M_{\alpha+1} = J(M_\alpha)$, and for limit ordinals $\lambda$, $M_\lambda = \langle M, P_\lambda \rangle$, where

$$\langle e, d \rangle \in P_\lambda \text{ iff } (\exists x < \lambda)(\forall \beta)(x \leq \beta < \lambda)(\langle e, d \rangle \in P_\beta),$$

so that $\langle e, d \rangle \in P_\lambda$ iff $\langle e, d \rangle$ is in the extension of all $P_\beta$ for $\beta < \lambda$ with $\beta$ large enough.

Our theory relates to those instances of predication which eventually assume a constant value in the revision process. Hence,

**Definition 3.2.** An element $\langle e, d \rangle$ of $M \times M$ is positively stable iff $(\exists x < \lambda)(\forall \beta \geq x)(\langle e, d \rangle \in P_\beta)$; it is negatively stable iff $(\exists x)(\forall \beta \geq x)(\langle e, d \rangle \notin P_\beta)$. It is stable if it is either positively or negatively stable. If $(\forall \beta \geq x)(\langle e, d \rangle \in P_\beta)$ we say $\langle e, d \rangle$ is positively stable from $x$; if $(\forall \beta \geq x)(\langle e, d \rangle \notin P_\beta)$ we say it is negatively stable from $x$.

Indeed, our theory is meant to be a theory about just those instances of predication which are a stable in the above sense. To validate the theory we thus require a model with validates precisely the stable instances.

**Definition 3.3.** An ordinal $\sigma$ is a stabilization ordinal iff for each $m, m' \in M$ we have that $\langle m, m' \rangle$ is positively stable iff $\langle m, m' \rangle \in P_\sigma$.

The following result is taken from Herzberger [1982].

**Theorem (Herzberger).** For any model $M$ there exists an ordinal $\sigma$, in the revision process based on $M$, such that $\sigma$ is a stabilization ordinal. Moreover, there exists an ordinal $\rho$ such that stabilization ordinals occur with periodicity $\rho$. Furthermore, if $\langle e, d \rangle$ is positively stable it is positively stable from $\sigma$.

The result holds for any "semi-inductive process". Such a process is given by a family of sets $B$ together with a function $J: B \rightarrow B$. The semi-inductive process generated by $\langle B, b, J \rangle$, where $b \in B$, is given as follows.

$$f(0) = b,$$
$$f(\alpha + 1) = J(f(\alpha)),$$
$$f(\lambda) = \{b' \in B : (\exists x < \lambda)(\forall \beta)(x \leq \beta < \lambda)(b \in f(\beta))\}, \text{ for limit ordinals } \lambda.$$  

Under such general conditions Herzberger establishes the existence of a stabilization ordinal. Our process is clearly a special case where $B = P(M \times M)$ and $J$ is given as above.

**Theorem 3.4.** For any starting model $M$ of $T_0$ and stabilization ordinal $\sigma$, $M_\sigma$ is a model of $T$. 


PROOF. Clearly all properties defined by logical truths, expressible in \( L \), are positively stable since they are true in all models. This applies to the laws of equality. For (A1), suppose \( \alpha \) is a stabilization ordinal and \( \left[ t(p(t_1, t_2)(\xi)) \right]^{M_{\alpha}} = 1 \). Then \( \forall x \geq \alpha, \left[ t(p(t_1, t_2)(\xi)) \right]^{M_{\alpha}} = 1 \). By definition of \( J \),

\[
\left[ p(\hat{x}[p(t_1, t_2)], s) \right]^{M_{\beta}} = 1
\]

for \( \beta \geq \alpha + 1 \). But then, by definition, \( p(\hat{x}[p(t_1, t_2)], s) \) is positively stable. Hence, since \( \alpha \) is a stabilization ordinal, \( \left[ p(\hat{x}[p(t_1, t_2)], s) \right]^{M_{\alpha}} = 1 \). For (R), we have only to note that if \( \left[ p(\hat{x}[A], t) \right]^{M_{\alpha}} = 1 \), then \( \left[ p(\hat{x}[A], t) \right]^{M_{\alpha + 1}} = 1 \) and so, by the definition of \( J \), \( \left[ A(\xi) \right]^{M_{\alpha}} = 1 \). This leaves us to deal with (NEG), since (S) is true in all revised models by the definition of \( J \). For (NEG) we need only observe that, for any ordinal \( \alpha \),

\[
\left[ p(\hat{x}[p(\hat{y}[B], s)], t) \right]^{M_{\beta + 1}} = \left[ B(\xi)(\xi) \right]^{M_{\beta}} = \left[ p(\hat{x}[p(\hat{y}[\sim B], s), t) \right]^{M_{\beta + 2}}.
\]

§4. Principles of abstraction. The theory does not support the full abstraction principle (A). In this section we investigate the circumstances under which full abstraction can be maintained. There are some immediate answers.

**Theorem 4.1.** For every wff \( A(t) \) of \( L \) such that \( Z_0 \vdash A(t) \), we have

(i) \( T \vdash p_k(\hat{x}[A], t) \leftrightarrow A(t) \),

(ii) \( T \vdash p_k(\hat{x}[\sim A], t) \leftrightarrow \sim A(t) \).

**Proof.** By reduction, the right-to-left directions are immediate. By (S) and Theorem 2.1, \( Z_0 \vdash A(t) \) implies \( \forall \alpha \vdash p_k(\hat{x}[A], t) \). Hence (i) follows. Since \( Z_0 \vdash A(t) \), (ii) is immediate. 

Thus, the theorems of the underlying theory induce properties for which full abstraction is guaranteed. This gives us a sufficient criterion to ensure full abstraction. We now turn to a second sufficient condition. First, however, we require the following:

**Theorem 4.2.** The following are derivable in \( T \):

(F1) \( p_k(\hat{x}[A], t) \& p_k(\hat{x}[B], t) \rightarrow p_k(\hat{x}[A \& B], t) \); 
(F2) \( p_k(\hat{x}[A], t) \lor p_k(\hat{x}[B], t) \rightarrow p_k(\hat{x}[A \lor B], t) \); 
(F3) \( p_k(\hat{x}[\sim A \lor B], t) \rightarrow p_k(\hat{x}[\sim (A \& B)], t) \); 
(F4) \( p_k(\hat{x}[\sim A \& B], t) \rightarrow p_k(\hat{x}[\sim (A \& B)], t) \); 
(F5) \( p_k(\hat{x}[A], t) \rightarrow p_k(\hat{x}[\sim A], t) \); 
(F6) \( \exists y p_k(\hat{x}[A], t) \rightarrow p_k(\hat{x}[\exists y A], t) \); 
(F7) \( \exists y p_k(\hat{x}[A], t) \rightarrow p_k(\hat{x}[\exists y A], t) \); 
(F8) \( p_k(\hat{x}[\exists y A], t) \rightarrow p_k(\hat{x}[\sim \exists y A], t) \); 
(F9) \( p_k(\hat{x}[\forall y A], t) \rightarrow p_k(\hat{x}[\sim \forall y A], t) \); 
(F10) \( p_k(\hat{x}[A \& B], t) \rightarrow p_k(\hat{x}[A], t) \& p_k(\hat{x}[B], t) \); 
(F11) \( p_k(\hat{x}[\sim (A \& B)], t) \rightarrow p_k(\hat{x}[\sim A \lor \sim B], t) \); 
(F12) \( p_k(\hat{x}[\sim (A \lor B)], t) \rightarrow p_k(\hat{x}[\sim A \& \sim B], t) \); 
(F13) \( p_k(\hat{x}[\forall y A], t) \rightarrow \forall y p_k(\hat{x}[A], t) \); 
(F14) \( p_k(\hat{x}[\sim A], t) \rightarrow p_k(\hat{x}[A], t) \); 
(F15) \( p_k(\hat{x}[\sim \forall y A], t) \rightarrow p_k(\hat{x}[\exists y \sim A], t) \); 
(F16) \( p_k(\hat{x}[\sim \exists y A], t) \rightarrow p_k(\hat{x}[\forall y \sim A], t) \).
PROOF. These are all easy to establish and follow directly from the classical nature of the internal logic. (F3), for example follows from the fact that $\vdash \sim A \lor \sim B \rightarrow \sim (A \land B)$. Similarly, (F7) follows from the classically valid

\begin{align*}
(1) & \quad \forall s \rightarrow \exists y A(y), \\
(2) & \quad \text{if } A(y) \rightarrow B \text{ then } \exists y A \rightarrow B.
\end{align*}

We employ (1) internally to obtain $p_k(\hat{x}[A(s) \rightarrow \exists y A], t)$. Under the assumption $p_k(\hat{x}[\exists y A], t)$, by (I6), we obtain $p_k(\hat{x}[\exists y A], t)$. Then (F7) follows from (2).

To state our second class of sufficient conditions we introduce the following notions, familiar from Gilmore [1974].

DEFINITION by recursion as follows:

4.3. Let $A$ be any wff of $L$. Define $A^+$ and $A^-$ by recursion as follows:

(0) If $A$ is $L_0$-atomic, $A^+ = A$ and $A^- = A$.

(1) If $A$ is $L$-atomic (i.e. $p_k(t)$), then $p_k(t)^+ = p_k(t)$ and $p_k(t)^- = \bar{p}_k(t)$.

(2) $(\sim B)^+ = B^-$ and $(\sim B)^- = B^+$.

(3) $(B \land C)^+ = B^+ \land C^+$ and $(B \land C)^- = B^- \lor C^-$.

(4) $(B \lor C)^+ = B^+ \lor C^+$ and $(B \lor C)^- = B^- \land C^-$.

(5) $(\forall x B)^+ = \forall x B^+$ and $(\forall x B)^- = \exists x B^-$. 

(6) $(\exists x B)^+ = \exists x B^+$ and $(\exists x B)^- = \forall x B^-$. 

LEMMA 4.4. In $T$ we have $A^+ \rightarrow A$ and $A^- \rightarrow \sim A$.

PROOF. Trivial by induction (the base step employs (DIS)).

Thus $A^+$ and $A^-$ are, in the above sense, "approximations" to $A$ and $\sim A$ respectively.

LEMMA 4.5. For each wff $A$ of $L$ we have:

(i) $T \vdash A(t)^+ \rightarrow p_k(\hat{x}[A], t)$,

(ii) $T \vdash A(t)^- \rightarrow p_k(\hat{x}[\sim A], t)$.

PROOF. The proof is routine and is established by simultaneous induction, on $A$, for (i) and (ii). We therefore omit it.

It is interesting to note that the proof only employs (F1)–(F9). Each of these principles follows from an internal logic which is intuitionistic. Full classical internal reasoning is not required. Principles (F11), (F14), and (F15) are, on the other hand, not intuitionistically sound.

THEOREM 4.6. For each wff $A$ of $L$,

(i) $T \vdash A(t) \leftrightarrow A(t)^+$ implies $T \vdash p_k(\hat{x}[A], t) \leftrightarrow A(t)$;

(ii) $T \vdash A(t) \leftrightarrow A(t)^-$ implies $T \vdash p_k(\hat{x}[\sim A], t) \leftrightarrow \sim A(t)$.

PROOF. This is immediate from (R) and 4.5.

However, according to 4.1, these conditions, although sufficient, are not necessary. For example, consider $A(x) = p(x, x) \lor \sim p(x, x)$. By (4.1), $p_k(\hat{x}[A], t) \leftrightarrow A(t)$; however, $A(t)^+ \leftrightarrow A(t)$, where $t = \hat{x}[A]$. Indeed, if $A(t)^+ \leftrightarrow A(t)$ then substitution yields $p(t, t) \lor \sim p(t, t)$. Both alternatives lead, via (R), to a contradiction.

Principles which cannot be generally deduced in $T$ are the inverses of (F2) and (F7). Indeed, if they were (R) would be a consequence, by a simple inductive argument. This fact, however, does raise an interesting question: what theory results from replacing the classical internal logic (plus (R)) by (F1)–(F13) plus (F17) and (F18)?

(F17) $p_k(\hat{x}[A \lor B], t) \rightarrow p_k(\hat{x}[A], t) \lor p_k(\hat{x}[B], t)$.

(F18) $p_k(\hat{x}[\exists y A], t) \rightarrow \exists y p_k(\hat{x}[A], t)$.
In fact, as we shall shortly demonstrate, the theory which results (call it $F$) is precisely the theory developed in Gilmore [1974] and Feferman [1984]. The theory consists of (DIS) & (IN), together with the following schemata:

\[(C1) \ p_k(\tilde{x}[A], t) \leftrightarrow A(t)^+ , \]
\[(C2) \ p_k(\tilde{x}[\sim A], t) \leftrightarrow A(t)^- . \]

**Theorem 4.7.** (IN) + (DIS) + (C1) + (C2) is equivalent to $F$.

**Proof.** Given (C1), (C2), (IN) and (DIS), it is tedious but easy to check that each schema of $F$ is validated. Moreover, (A1) and (A2') are instances of (C1) and (C2) respectively. We illustrate the induction step by reference to disjunction, negation and existential quantification. Suppose, inductively, $p_k(\tilde{x}[A], t) \leftarrow A(t)^+$ and $p_k(\tilde{x}[B], t) \leftarrow B(t)^+$. Then by (F2) and (F17), $p_k(\tilde{x}[A \lor B], t) \leftarrow (A \lor B)(t)^+$. Moreover,

\[p_k(\tilde{x}[A \lor B], t) \leftrightarrow p(\tilde{x}[\sim (A \lor B)], t) \] (by (IN)). This is equivalent to $p_k(\tilde{x}[\sim A \& \sim B], t)$ (by (F12) and (F4)). By (F1) and (F10) the latter is equivalent to $p_k(\tilde{x}[\sim A], t) \& p_k(\tilde{x}[\sim B], t)$. The result now follows by induction and the definition $(A \lor B)^- = A^- \& B^-$. For negation, half the result is immediate and the other half requires (F7) and (F14). For existential quantification we spell it out more fully: by (F7) and (F16),

\[p_k(\tilde{x}[\exists y A], t) \leftrightarrow \exists y p_k(\tilde{x}[A], t). \] By induction, the latter is equivalent to $(\exists y A(t))^+$, i.e. $(\exists y A(t))^+$. Moreover,

\[p_k(\tilde{x}[\sim \exists y A], t) \leftrightarrow p(\tilde{x}[\forall y \sim A], t) \] by (F9) and (F16). By (F6) and (F13) we obtain $\forall y p_k(\tilde{x}[\sim A], t)$ which, by induction, is equivalent to $(\forall y A(t))^-$ i.e. $(\exists y A(t))^-$.

The theory $F$ is orthogonal to $T$ in that many nonclassical equivalences, such as those induced by (F17) and (F18), hold in $F$ but are not derivable in $T$. However, as (4.6) illustrates, the theory $F$ yields no more instances of abstraction or (LEMP). Indeed, by adding (F17) and (F18), the underlying logic is no longer classical and (4.1) fails. In particular, $p(\tilde{x}[A \rightarrow A], t) \leftrightarrow (A \rightarrow A)(t)$ is not even derivable. As we shall later see, this nonclassical nature of $F$ blocks the development of arithmetic in the theory. This is one of the advantages of $T$ over $F$.

### §5. Total properties, relations and propositions.

Certain PRP’s will satisfy abstraction but not (LEMP). In this section we investigate the PRP’s which are completely well-behaved in the sense that (LEMP) will hold generally. We can probe this issue a little more systematically by introducing the following concept.

**Definition 5.1.** A PRP $a$ is $k$-total iff $\forall x(p_k(a, x) \lor p_k(a, x))$; otherwise $a$ is $k$-partial.

The cases of special interest ($k = 1$ and $k = 0$) we shall refer to as total properties and total propositions, respectively. Total PRP’s are precisely those for which every instance of (LEMP) is validated.

**Definition 5.2.** (i) $\text{Tot}_k(a) = \text{def} \, \forall x(p_k(a, x) \lor p_k(a, x)).$
(ii) $\text{Tot}_k = \text{def} \, \tilde{y}[\forall x(p_k(y, x) \lor p_k(y, x))].$
Since the matrix of Tot\(_k\) satisfies (i) (of 4.5), we have

**Theorem 5.3.** \(T \vdash \phi(Tot_k, t) \iff \phi(Tot_t(t)).\)

Thus, Tot\(_k\) satisfies abstraction for all its instances. However, it is not itself total.

**Theorem 5.4.** \(T \vdash \neg \phi(Tot_k, Tot_t).\)

**Proof.** Assume \(\phi(Tot_k, Tot_t).\) Let \(a = \tilde{a}[\phi(Tot_k, x) \& \neg \phi(x, x)].\) By 4.6, (i) and (ii), we obtain

(iii) \(\phi(a, a) \iff \phi(Tot_k, a) \& \phi(x, a),\)

(iv) \(\neg \phi(a, a) \iff \phi(Tot_k, a) \lor \phi(x, a).\)

By (iii) and (iv), \(\phi(a, a) \iff \phi(a, a)\) and therefore \(\phi(Tot_k, a).\) Moreover, by (iv), \(\phi(a, a) \iff \phi(Tot_k, a) \lor \phi(a, a),\) and since \(\phi(Tot_k, a), \phi(a, a) \iff \phi(a, a)\)—a contradiction.

There are some immediate closure conditions for total PRP's, which follow from 4.6. The following is taken from Feferman [1984].

**Definition 5.5.** Let \(A(x, y)\) be a wff of \(L.\) \(A\) is in \((a, \ldots, a_n)-\)form if each atomic subformula of \(A\) (excluding equality) is of the form \(p_i(a_i, s)\) for some \(i, 1 \leq i \leq n.\)

**Theorem 5.6.** If \(A\) is in \((a_1, \ldots, a_n)-\)form then, if \(\phi(Tot_k(a), a)\) for \(1 \leq i \leq n,\) then

\(\phi(Tot_k(\tilde{A}[x_1, \ldots, x_m, a_1, \ldots, a_n])).\)

**Proof.** By a simple inductive argument we can establish \(A^+ \iff A\) and \(A^- \iff \neg A.\) We then employ 4.6.

The various Boolean operations have their analogues for PRP's. Here we investigate their behavior as regards totality.

**Definition 5.7.** (i) \(a \land b = \tilde{a}[p(a, x)\& p(b, x)].\)

(ii) \(a \lor b = \tilde{a}[p(a, x)\lor p(b, x)].\)

(iii) \(a \land b = \tilde{a}[p(a, x)\& \neg p(b, x)].\)

(iv) \(\land_{\phi} a = \tilde{a}[\forall y(p(a, y) \land p_k(y, x))].\)

(v) \(\lor_{\phi} a = \tilde{a}[\exists y(p(a, y) \land p_k(y, x))].\)

**Theorem 5.8.** (i) \(\phi(Tot_k(a)\land \phi(Tot_k(b))\) together imply

\(\phi(Tot_k(a \land b)), \phi(Tot_k(a \lor b)), \phi(Tot_k(a \land b)).\)

(ii) \(\phi(Tot_1(a)) and a \subseteq Tot_k\) (i.e. \(\forall x(p(a, x) \rightarrow \phi(Tot_k(x)))\) imply \(\phi(Tot_k(\lor a))\) and \(\phi(Tot_k(\land a)).\)

**Proof.** Part (i) follows from 5.6. For part (ii) we employ 4.6. First observe that \(\exists y(p_1(a, y) \land p_k(y, x))\) is positive and hence half the condition for totality is satisfied. For the other half of the requirement observe that, under the assumption that \(a \subseteq Tot_k,\)

\((\forall y(p(a, y) \rightarrow \neg p_k(y, x))) \iff (\forall y(p(a, y) \rightarrow \tilde{p}_k(y, x))).\)

Since \(\phi(Tot_1(a)),\) this is equivalent to \(\forall y(\tilde{p}(a, y) \lor \tilde{p}_k(y, x)),\) which is in turn equivalent to

\((\forall y(\tilde{p}(a, y) \& \tilde{p}_k(y, x))^-).\)

So the other half of 4.6 is satisfied. A similar argument applies to \(\lor_k.\)

Where \(k\) is understood, we shall frequently drop the subscripts and superscripts.

**Corollary 5.9.** (i) \(\phi(Tot_k(a)\land \phi(Tot_k(b))\) together imply

\(\phi(Tot_k(a \land b), \phi(Tot_k(a \land b), \phi(Tot_k(a \land b)).\)

\(P_k(a \land b, x) \iff P_k(a, x) \& P_k(b, x),\)

\(P_k(a \lor b, x) \iff P_k(a, x) \lor P_k(b, x),\)

\(P_k(a \land b, x) \iff P_k(a, x) \& \neg P_k(b, x).\)
(ii) \(\text{Tot}_1(a) \& a \subseteq \text{Tot}_k\) implies

\[
P_k(\bigwedge_k a, x) \iff \forall y(P(a, y) \rightarrow P_k(y, x)),
\]

\[
P_k(\bigvee_k a, x) \iff \exists y(P(a, y) \& P_k(y, x)).
\]

Notice that for \(k = 1\) we obtain closure of \(\text{Tot}_1\) under the normal set-theoretic operations of union, intersection and difference. Moreover,

\[
\bigvee_1 a = \exists x(\exists y(p(a, y) \& p(y, x))) \quad \text{and} \quad \bigwedge_1 a = \exists x(\forall y(p(a, y) \rightarrow p(y, x))],
\]

which are generalised union and intersection.

The existence of universal PRP’s is sanctioned by the theory. Indeed, any classically valid wff defines a universal PRP. In particular we have

**Definition 5.10.** \(\bigvee_k = \exists x(x = x)\) and \(\bigwedge_k = \exists x(x \neq x)\).

**Theorem 5.11.** For each \(k\), \(\bigvee_k\) and \(\bigwedge_k\) are \(k\)-total.

**Proof.** The usual rules for equality form part of the underlying internal theory \(Z_0\). ■

We shall write \(\bar{a}^k\) for \(\bigvee k - k\ a\).

The usual definitions of power-set and function spaces are also available to us. First we require the notions of extensional containment and extensional equivalence:

\[
a \subseteq_k b \iff \forall x(p(a, x) \rightarrow p(b, x)),
\]

\[
a \equiv_k b \iff a \subseteq_k b \& b \subseteq_k a.
\]

For pedagogical reasons we shall now write \(p(a, t)\) as \(t \in a\) and \(\exists A\) as \(\{x: A\}\).

**Definition 5.12.** \(P(a) = \{x: x \subseteq_1 a\}\).

**Theorem 5.13.** For \(\text{Tot}_1(a)\) and \(\text{Tot}_1(x)\),

\[
x \subseteq_1 a \leftrightarrow x \in P(a), \quad x \not\subseteq_1 a \leftrightarrow x \not\in P(a).
\]

**Proof.** Use 4.6 and the totality of \(a\) and \(x\). ■

This result can be put in a slightly different form, namely: \(\text{Tot}_1(a)\) and \(\text{Tot}_1(b)\) together imply \(\text{Tot}_0([b \in P(a)])\). What we cannot derive is the totality of \(P(a)\) itself—even when \(a\) is total. We shall return to this point in §6 when we shall consider a notion of totality which is preserved by the power-set construction. Similar remarks apply to function spaces. For function spaces we again proceed in the standard way. First, we define cartesian products.

**Definition 5.14.** \(a \times b = \{z: \exists x \exists y(x \in a \& y \in b \& z = (x, y)\}\}.

**Theorem 5.15.** For \(\text{Tot}_1(a)\) and \(\text{Tot}_1(b)\), \(\text{Tot}(a \times b)\).

**Proof.** Use 5.6. ■

**Definition 5.16.** (i) \(\text{Fun}(f, a, b) = f \subseteq_1 a \times b \& (\forall x \in a \exists y \in b)((x, y) \in f)\).

(ii) \((a \rightarrow b) = \{f: \text{Fun}(f, a, b)\}\).

**Theorem 5.17.** For \(\text{Tot}(a)\), \(\text{Tot}(b)\) and \(\text{Tot}(f)\),

(i) \(f \in (a \rightarrow b) \leftrightarrow \text{Fun}(f, a, b)\),

(ii) \(f \in (a \rightarrow b) \leftrightarrow \sim \text{Fun}(f, a, b)\).

**Proof.** Similar to Theorem 5.13: use totality of \(f, a, b\) and 4.6. ■

The specialization of these results to the case \(k = 0\) yields some interesting consequences relating to the theory of truth and propositions. We shall write \(\text{Prop}(a)\) for \(\text{Tot}_0(a)\) (i.e. \(T(a) \vee \overline{T(a)}\)).
THEOREM 5.18. (i) Prop(a) implies (Prop(\(\overline{a}\)) and \(T(\overline{a}) \leftrightarrow \sim T(a)\)).
(ii) Prop(a) and Prop(b) together imply (Prop(a \sqcap b) and T(a \sqcap b) iff T(a) and T(b)).
(iii) Prop(a) and Prop(b) together imply (Prop(a \sqcup b) and T(a \sqcup b) iff T(a) or T(b)).

PROOF. These are special cases of 5.8(i) and 5.9. ■

We can also introduce generalized quantifiers in the standard way:

DEFINITION 5.19.

\[
\begin{align*}
\text{Every} & =_{\text{def}} \exists \exists \forall z (z \in x \rightarrow z \in y], \\
\text{Some} & =_{\text{def}} \exists \exists \exists z (z \in x \& z \in y], \\
\text{Not Every} & =_{\text{def}} \exists \exists \forall z (z \in x \& z \notin y], \\
\text{No} & =_{\text{def}} \exists \exists \forall z (z \in x \rightarrow z \notin y].
\end{align*}
\]

THEOREM 5.20. Tot(a) and Tot(b) together imply
(i) Prop([p_2(\text{Every}, a, b)]),
(ii) Prop([p_2(\text{Some}, a, b)]),
(iii) Prop([p_2(\text{Not Every}, a, b)]),
(iv) Prop([p_2(\text{No}, a, b))].

PROOF. Use 4.6 and totality of a and b. ■

COROLLARY 5.21. (i) p_2 (\text{Every}, a, b) \rightarrow a \sqcap b = A.
(ii) p_2 (\text{Some}, a, b) \rightarrow a \sqcap b \neq A.
(iii) p_2 (\text{Not Every}, a, b) \rightarrow a \sqcap b \neq A.
(iv) Prop([p_2(\text{No}, a, b)]).

In essence, 5.18, 5.20 and 5.21 furnish the logical schema of a Frege structure. ■

§6. Inductive definitions. The internal logic has played a central role in the development of the theory but, so far, the theory has yielded little more than the theory of Feferman and Gilmore. In this section, we investigate how the internal logic can be employed to derive a certain theory of inductive definitions. We begin with the simplest example and then derive a more general characterisation.

(i) Arithmetic in T. The first topic we address concerns the development of arithmetic in T. One approach to the natural numbers views them as the smallest class containing 0 and closed under successor. Recall that the successor operation is represented x' = (x, 0); d then acts as predecessor and from the axioms of d, q, and q_2, we have x' \neq 0, and also x' = y' \rightarrow x = y. The natural numbers can then be defined as follows:

N = \{x: \forall z[(0 \in z \& \forall y(y \in z \rightarrow y' \in z)) \rightarrow x \in z]\}.

We require N to satisfy the normal closure and induction principles:

\text{Closure.} \ 0 \in N \& \forall y(y \in N \rightarrow y' \in N).

\text{Induction.} \ [A(0) \& \forall y(A(y) \rightarrow A(y'))] \rightarrow \forall x(x \in N \rightarrow A(x)).

We first deal with closure.

THEOREM 6.1. N satisfies closure.

PROOF. This requires essential use of the internal logic. Let A(x, z) = (0 \in z \& \forall y(y \in z \rightarrow y' \in z)) \rightarrow x \in z. First claim 0 \in N. Clearly, \neg \forall z A(0, z); thus by 2.1,
A 'THEORY OF PROPERTIES

0 ∈ \{x: \forall z A(x, z)\}. Assume \(x \in N\). We must establish \(x' \in N\). Claim \(\vdash \forall z A(x, z) \rightarrow \forall z A(x', z)\). Assume

\[
\begin{align*}
(1) & \quad \forall z A(x, z), \\
(2) & \quad 0 \in z \& \forall y (y \in z \rightarrow y' \in z).
\end{align*}
\]

By (1) and (2), \(x \in z\). By (2) again, \(x' \in z\). Hence (1), (2) \(\vdash x' \in z\). By the deduction theorem, (1) \(\vdash (2) \rightarrow x' \in z\). By (GEN), (1) \(\vdash \forall z ((2) \rightarrow x' \in z)\). Hence, by the deduction theorem again, \(\vdash (1) \rightarrow \forall z A(x', z)\), and hence the claim. By (2), \(x \in \{u: \forall z A(u, z) \rightarrow \forall z A(u', z)\}\). By assumption \(x \in N\), i.e. \(x \in \{u: \forall z A(u, z)\}\). By (I6), we obtain \(x \in \{u: \forall z A(u', z)\}\). By two applications of (S), \(x' \in \{u: \forall z A(u, z)\}\), i.e. \(x' \in N\), as required.

The induction schema as stated is not derivable from the definition of \(N\). All we can derive is the schema

\[
\forall z \{ (0 \in z \& \forall y (y \in z \rightarrow y' \in z) ) \rightarrow \forall x (x \in N \rightarrow x \in z) \}.
\]

This much follows from closure and reduction (R). However, for those wffs \(A\) which define total properties the normal schema holds.

(ii) General inductive definitions. The technique employed to obtain closure and induction for \(N\) can be employed to furnish generalized inductive definitions. Let \(A(x, y)\) be any wff of \(L\). Consider

\[
I_A = \{u: \forall z (\forall x (A(x, y) \rightarrow y' \in z) \rightarrow x \in z) \rightarrow u \in z\}.
\]

Under what circumstances does \(I_A\) satisfy closure and induction?

**Closure.** \(\forall x [\forall y (A(x, y) \rightarrow y \in I_A) \rightarrow x \in I_A]\).

**Induction.** \(\forall x (\forall y (A(x, y) \rightarrow B(y)) \rightarrow B(x)) \rightarrow (\forall u \in I_A) (B(u))\).

We first establish a general result.

**Theorem 6.2.** For each wff \(A(x, y)\) such that \(\text{Tot}(\{x: A(x, y)\})\), \(I_A\) satisfies closure.

**Proof.** Let \(C(z) = \forall x (\forall w (A(x, w) \rightarrow w \in z) \rightarrow x \in z)\) and \(B(u, z) = C(z) \rightarrow u \in z\). Claim \(\vdash \forall y (A(u, y) \rightarrow \forall z B(y, z)) \rightarrow \forall z B(u, z)\).

Assume

\[
\begin{align*}
(1) & \quad \forall y (A(u, y) \rightarrow \forall z B(y, z)), \\
(2) & \quad C(z).
\end{align*}
\]

We need to conclude \(u \in z\). Assume \(A(u, y)\). By (1), \(\forall z B(y, z)\). From (2) and \(B(y, z)\), we have \(y \in z\). Hence, \(A(u, y) \rightarrow y \in z\), and so, by (GEN), \(\forall y (A(u, y) \rightarrow y \in z)\). Hence, (1), (2) \(\vdash u \in z\). The deduction theorem yields the claim.

Now by 2.1,

\[
\begin{align*}
(3) & \quad u \in \{v: \forall y (A(v, y) \rightarrow \forall z B(y, z)) \rightarrow \forall z B(v, z)\}.
\end{align*}
\]
Assume $\forall y(A(u, y) \rightarrow y \in I_a)$. By two applications of (S), $\forall y(A(u, y) \rightarrow u \in \{v: \forall z B(y, z)\})$. By the totality of $\{x: A(x, y)\}$, we obtain

$$u \in \{v: \forall y(A(x, y) \rightarrow \forall z B(y, z))\}. \tag{4}$$

By (3), (4) and (I6) we obtain $u \in \{v: \forall z B(v, z)\}$, i.e. $u \in I_a$.  

A useful extension to 6.2 is the following:

**COROLLARY 6.3.** For each wff $A(x, y)$ such that $\text{Tot}(a) \rightarrow u \in \{x: A(x, y)\} \leftrightarrow A(u, y)$, $I_a$ satisfies $\forall x(\forall y((\text{Tot}(x) \& A(x, y)) \rightarrow y \in I) \rightarrow x \in I_a)$. 

**PROOF.** This follows by a slight variation on the last part of the argument of 6.2.  

**EXAMPLES.** (A) As one special case we obtain induction on the well-founded part of a relation. Let $A(x, y) = x \in a \& (y, x) \in r$. Then for $\text{Tot}(a) \& \text{Tot}(r)$ we obtain (via 4.6) $\text{Tot}(\{x: A(x, y)\})$. Thus the theorem yields a class $I$ such that

$\text{Closure. } \forall x \in a(\forall y(y < x \rightarrow y \in I) \rightarrow x \in I)$,

where $y < x$ is written for $(y, x) \in r$.

(B) Induction on $N$ is a special case where $a = \{x: x = 0 \lor \exists y\{x = y'\}\}$ and $r = \{z: \exists x \exists y(z = (y, x) \& x = y')\}$.

(iii) A universe of inductively total objects. Let $A(x, y) = y \in x \& \text{Tot}(x)$. We then obtain from 6.3 a class $I$ with the following closure and induction principles:

$\text{Closure. } \text{Tot}(x) \& x \subseteq I \rightarrow x \in I$.

$\text{Induction. } \forall x(\text{Tot}(x) \& x \subseteq z \rightarrow x \in z) \rightarrow I \subseteq z$.

**THEOREM 6.4.** $x \in I \rightarrow \text{Tot}(x) \& x \subseteq I$.

**PROOF.** Let $Z = \{y: \text{Tot}(y) \& y \subseteq I\}$. Suppose $\text{Tot}(x) \& x \subseteq Z$. By reduction, $\forall y(y \in x \rightarrow \text{Tot}(y) \& y \subseteq I)$. By closure, $\forall y(y \in x \rightarrow y \subseteq I)$, i.e. $x \subseteq I$. Since $\text{Tot}(x)$, 4.6 yields $x \in Z$. Therefore, $\forall x(\text{Tot}(x) \& x \subseteq Z \rightarrow x \in Z)$. By induction, $I \subseteq Z$. Thus, $x \in I$ implies, by reduction, $\text{Tot}(x) \& x \subseteq I$.

**COROLLARY 6.5.** $x \in I \leftrightarrow x \subseteq I \& \text{Tot}(x)$.

$I$ thus consists of a class whose elements are inductively total. We now explore some of the closure properties of $I$.

**THEOREM 6.6.** For $a, b \in I$ we have:

(i) $a \uplus b \in I, a \sqcup b \in I, a - b \in I$;

(ii) $\forall a \in I, \exists a \in I$.

**PROOF.** Part (i) is straightforward given the previous corollary. We illustrate part (ii) with respect to union. We employ the previous corollary: we have to show $\text{Tot}(\exists a)$ and $\exists a \subseteq I$. By Theorem 5.8, $\text{Tot}(\sqrt{a})$ follows since, by transitivity of $I$, $a \in I$ implies $a \in \text{Tot}$ and $a \subseteq \text{Tot}$. To show containment first recall $\sqrt{a} = \{z: \exists y(y \in a \& z \in y)\}$. If $x \in \sqrt{a}$, by reduction, $\exists y(y \in a \& x \in y)$. Since $y \in a$ and $a \in I$ we know, by transitivity of $I$, that $y \in I$. Since $x \in y \in I, x \in I$. Hence, $\sqrt{a} \subseteq I$.

We can now introduce a notion of totality which is preserved by power-set and function space.

**DEFINITION 6.7.** $\text{CL}(a) =_{\text{def}} (\forall x \in I)(x \in a \lor x \in a)$, and $\text{CL} = \{x: \text{CL}(x)\}$.

In essence, $\text{CL}(a)$ states that $a$ is total relative to the universe of inductively total PRP's. Those PRP's which satisfy CL have some rather strong closure properties.
Theorem 6.8. (i) \(\text{CL}(a)\) implies \(\text{CL}(\bar{a})\).
(ii) \(\text{CL}(a)\) and \(\text{CL}(b)\) imply \(\text{CL}(a \cap b)\) and \(\text{CL}(a \cup b)\).
(iii) \(\text{CL}(a)\) implies \(\text{CL}(P(ai))\).
(iv) \(\text{CL}(a)\) and \(\text{CL}(b)\) imply \(\text{CL}(a \Rightarrow b)\).
(v) \(\text{CL}(a)\) and \(\text{CL}(b)\) imply \(\text{CL}(a \times b)\).
(vi) \(\text{CL}(I), \text{CL}(CL), \text{CL}(\sqrt{1}), \text{CL}(\setminus)\).

Proof. The only nontrivial verification involves power-set and function space. We illustrate with the former. Suppose \(\text{CL}(a)\). First suppose \(x \in I\) and \(x \not\in a\). Then \(\text{Tot}(x)\). Thus, by Theorem 5.13, \(x \in I\) implies \(x \subseteq a \iff x \in P(a)\). Next assume \(x \not\in a\) for \(x \in I\). Then \(\exists y(y \in x \land y \not\in a)\). Thus we have \(y \in I\) \& \(y \not\in a\). Since \(\text{CL}(a)\) we have \(y \in a\). Hence, by \((A2')\), \(x \in \{u: y \in a\}\).

By the definition, \((A1)\) and \((R)\), \(x \in \{u: y \in u\} \iff y \in x\). Thus,

\[
\exists y(y \in x \land y \not\in a) \rightarrow x \in \{u: \exists y(y \in u \land y \not\in a)\},
\]
i.e. \(x \in P(a)\). By \((R)\), we obtain \(x \not\in a \iff x \in P(a)\). Thus, \(x \in I\) implies \(x \in P(a) \lor x \in P(a)\). ■

§7. First-order theories. In this section we shall briefly consider possible extensions to the theory \(T\). We shall only consider extensions which involve the underlying theory \(T_0\).

One extension of some interest involves a theory of operations. Following Feferman [1979] we could include the theory of partial operations \(\text{APP}\). Essentially we extend \(L_0\) by the addition of new individual constants \(k, s, d\) and a new atomic formulae \(\text{app}(a, b, c)\) where \(a, b, c\) are terms of this new extended language. The operation \(\text{app}\) is partial and asserts that the first argument, when applied to the second, yields the third. It is not assumed that every operator is total. The axioms for \(\text{APP}\) are the normal ones for the combinators but where partiality is taken into account. Natural models of this theory can be constructed from the partial recursive functions. Details can be found in Feferman [1979] or in a more elegant form in Beeson [1985]. Such an extension facilitates the definition of a new notion of function space as follows:

\[
\{a \Rightarrow b\} = \hat{f}[\forall x(x \in a \rightarrow f(x) \in b)]
\]
Since the defining formula is positive the totality of \(a\) and \(b\) guarantees, by 4.6, that \(\{a \Rightarrow b\}\) will be total. In this way the "intensional" mathematics of Feferman [1979] can be formulated.

An alternative approach might involve the lambda calculus. We could extend \(T_0\) by the addition of the axioms of the lambda calculus. Such an approach would bring us closer to that of Aczel [1980]. The underlying models of the theory \(T_0\) would then have to be models of the lambda calculus.

As regards the consistency of these extensions, one has only to observe that the construction of §3 is left intact by such extensions: these theories make no reference to the relations of predication and so, in the stabilization process, every model will remain a model of \(T_0\).

More interesting extensions involve the internal logic. Let \(Z\) be any first-order theory in the language \(L\). Under what circumstances can \(Z\) be added to the internal
logic? Which first order theories are stable; i.e., does there exist an ordinal from which every model is a model of Z? We shall not pursue this topic any further, but it seems like a fruitful line given the results of Flagg and Myhill [1985].

Acknowledgements. I would like to thank Barbara H. Partee both for encouragement and for making this work possible through support on her SDF grant. Among others, I have had useful comments from Johan van Benthem, Peter Aczel, George Bealer, Hans Kamp, and Gennaro Chierchia.

REFERENCES

G. Bealer [1982], Quality and concept, Oxford University Press, Oxford.
——— [1979], Constructive theories of functions and classes, Logic Colloquium '78 (M. Boffa et al., editors), North-Holland, Amsterdam, pp. 159–224.
B. Flagg and J. Myhill [1985], Notes on a type-free system extending ZFC (manuscript).
H. Kamp [1983], A scenic tour through the land of naked infinitives (manuscript).

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF ESSEX
COLCHESTER CO4 3SQ, ENGLAND