

5

Philosophical Studies published by the Philosophical Society and the
Department of Philosophy, University of Uppsala, Sweden.

No. 38

Changing Positions

Essays Dedicated to

Lars Lindahl

on the Occasion of His Fiftieth Birthday

Edited by

Paul Needham and Jan Odelstad



Uppsala 1986

A SEMANTIC ANALYSIS OF RUSSELLIAN SIMPLE TYPE THEORY*

Sten Lindström

As emphasized in recent articles by Alonzo Church and David Kaplan (Church 1974, Kaplan 1975), the philosophies of language of Frege and Russell incorporate quite different methods of semantic analysis with different basic concepts and different ontologies. Accordingly we distinguish between a Fregean and a Russellian tradition in intensional logic. Semantic analysis in the tradition which started with Frege's article 'Über Sinn und Bedeutung' (1892) and which makes a distinction between sense (or intension) and denotation (or extension) will be referred to as Fregean semantics. Semantic analysis in the tradition of Russell's articles 'On Denoting' (1905), 'Mathematical Logic as based on the Theory of Types' (1908) and Principia Mathematica we refer to as Russellian semantics.

Both Fregean and Russellian semantics are faced with solving the following closely related problems:

(i) Frege's puzzle concerning the "cognitive significance" of identity statements: how can 'a = b', if true, be an informative statement differing in cognitive significance from 'a = a'?

(ii) The problem of oblique or nonextensional contexts: how can two well-formed expressions with the same denotation ever fail to be interchangeable *salva veritate*?

(iii) The problem of providing an adequate truth-conditional semantics for propositional attitude reports. Fregean solutions to these problems essentially involve the distinction between sense and denotation. The appearance of oblique contexts in natural languages was interpreted by Frege to indicate a certain kind of systematic ambiguity rather than a failure of extensionality. According to Frege's doctrine of indirect denotation, expressions denote in

(unembedded) oblique contexts what is ordinarily their sense. Frege's extensional point of view has been advocated and developed further in this century by Alonzo Church (1951, 1973, 1974a).

Carnap and Montague, on the other hand, while still working within the Fregean tradition, saw the occurrence of oblique contexts in natural languages as genuine counterexamples to the principle of extensionality, the principle that the denotation of a well-formed expression is always a function of the denotations of its semantically relevant parts (Carnap 1947, Montague 1974).

According to the Carnap-Montague view, each well-formed expression of a language has both an extension (corresponding to Frege's denotation) and an intension (roughly corresponding to Frege's sense). The intension of an expression is identified with a function from possible worlds (or models or state-descriptions, representing possible worlds) to appropriate extensions. Hence, the intension of a singular term, an individual concept, is a function from possible worlds to (possible) individuals; the intension of an n-ary predicate of individuals is a function from possible worlds to sets of ordered n-tuples of (possible) individuals; the intension of a sentence, a (Carnapian) proposition, is a function from possible worlds to the truth-values T (true) and F (false). If I(E) is the intension of the expression E and w is a possible world, then I(E)(w) is the extension of E in the world w. Thus, within the Carnap-Montague framework, the notion of extension is relativized to possible worlds. By the extension of E, simpliciter, we then understand the extension of E in the actual world.

The Carnap-Montague program, while rejecting the principle of extensionality, subscribes instead to the principle of intensionality, i.e., the principle that the intension of a complex expression is a function of the intensions of its semantically relevant parts. Modal contexts, although apparently nonextensional, satisfy the principle of intensionality. However, as was already pointed out by Carnap in Meaning and Necessity (1947), §13, the principle seems to fail

for propositional attitude reports. Consider, for example, statements of elementary number theory. It seems that any true number-theoretic statement must be true by logical necessity and hence true in every possible world. Hence, all true number-theoretic statements have the same intension, namely the constant function mapping each possible world to the truth-value T. Now, assume that George, who is a rational person and a competent mathematician, has just proved a certain number-theoretic statement '...'. George, therefore, holds the statement '...' to be true. It seems that we are then justified in affirming:

(1) George believes that

We also assume that there is another statement '___' of elementary number theory which George has failed both to prove and to disprove. Now, the statement '___' is in fact true, although George does not think so. He rather suspects that it is false. Under these circumstances, we seem justified in asserting:

(2) George does not believe that ___.

In inferring (1) and (2), we are presupposing a disquotation principle connecting the attitude of holding a sentence true (acceptance) with belief:

(DP) If x is a competent English speaker, then x accepts (holds true) '---' if and only if he believes that ---.

Here, '---' is to be replaced by any normal English sentence not containing ambiguities, indexicals or pronominal devices.

Since the two number-theoretic statements '...' and '___' are both true, they have the same intension, i.e.,

(3) $I('...') = I('___')$.

In case we doubt that any two true number-theoretic statements have the same intension, we may make the stronger assumption that '...' and '___' are provably equivalent in elementary number-theory or even in first-order logic. We may, for example, assume that '...' is a simple number-theoretic statement and that '___' is a logically equivalent statement of considerably higher quantifier complexity. It still seems

possible that George should hold '...' true while denying or doubting the truth of '___'. Presumably, it could even be the case that a rational person would accept one statement while not accepting or even rejecting a tautologically equivalent statement.

Now, from (1) and (2), we may conclude

(4) I('George believes that ...') ≠ I('George believes that ___').

However, (3) and (4) are in apparent contradiction with the principle of intensionality.

One may object to the above argument that the principle (DP) is unreasonably strong. Perhaps we only have the weaker principle:

(WDP) If x is a competent English speaker and x accepts (holds true) '---', then x believes that ---.

This principle corresponds to the weaker of Kripke's (1979) principles connecting sincere assent (rather than acceptance) and belief. Using (WDP), we cannot infer (2) from George's failure to accept '___'. However, it seems quite plausible that there should be some number-theoretic statement '___' such that:

- (i) '...' and '___' have the same intension;
- (ii) George has good reasons to reject '___', i.e., to accept 'not-___'. We may imagine that George has constructed a "proof" of 'not-___' containing some subtle error;
- (iii) George accepts 'not-___'.

Using (WDP) we can then conclude:

(5) George believes that not-___.

From (3) and (5) we get by the principle of intensionality:

(6) George believes that not-... .

Hence, we describe George as both believing that ... and believing that not-... , i.e., as holding explicitly contradictory beliefs. To do so seems to misrepresent George's frame of mind.

Furthermore, since George accepts both '...' and 'not-___', it is quite likely that he also accepts '... and not-___'. However, then we may apply (WDP) to infer:

(7) George believes that (... and not-___).

Using (7), (3) and the principle of intensionality, we then conclude:

(8) George believes that (... and not-...),

i.e., we are describing George as believing an explicit contradiction. This seems to contradict the assumption that George is a rational person. The semantic analysis of propositional attitude reports thus presents an apparent obstacle to the Carnap-Montague program.

The purpose of this paper is to pursue the Russellian alternative in intensional logic, where a sense/denotation distinction is avoided. Russellian semantics in contrast to Fregean semantics assigns only one kind of semantic value, most naturally thought of as denotation, to the well-formed expressions of a language. Following Quine (1969), Church (1974), and Kaplan (1975), we interpret Russell's logical doctrines in the first edition of Principia Mathematica (1910) and in the early writings (Russell 1903, 1905, 1908) intensionally. The values of the propositional variables of Russellian type theory are taken to be abstract propositions rather than sentences or truth-values. Sentences, being substitutable for propositional variables, are taken to denote propositions. Predicate terms are assigned propositional functions, i.e., functions from entities of the appropriate kind to propositions, as their semantic values. Individual terms, finally, refer to individuals.

Russellian semantics is not extensional in the usual sense, since the principles:

(1) $(A \leftrightarrow B) \rightarrow (...A... \leftrightarrow ...B...)$

and

(2) $\forall x_1 \dots \forall x_n (Fx_1 \dots x_n \leftrightarrow Gx_1 \dots x_n) \rightarrow (...F... \leftrightarrow ...G...)$

fail, for example, in propositional attitude contexts.

Instead of (1) and (2), we have the principles:

$$(3) (A = B) \rightarrow (A \leftrightarrow B)$$

$$(4) (A = B) \rightarrow (...A... = ...B...)$$

$$(5) \forall x_1 \dots \forall x_n (Fx_1 \dots x_n = Gx_1 \dots x_n) \rightarrow (...F... = ...G...),$$

where '=' stands for the relation of identity between propositions. Hence, Russellian semantics is extensional in the following nonstandard sense: the Russellian denotation of an expression is always a function of the Russellian denotations of its semantically relevant parts.

It seems natural, both on a Fregean and on a Russellian approach, to interpret propositional attitudes, like beliefs, fears, and desires, as relations between a person and a proposition (or between a person, a time, and a proposition). The two approaches differ, however, in their methods of semantic analysis and in their views of the nature of propositions. Consider the example:

(1) Othello believes that Desdemona loves Cassio.

On the Fregean analysis, propositional attitude verbs, like 'believes' apply syntactically to noun phrases which are interpreted to denote propositions. Hence, the logical form of (1) can on the Fregean analysis be symbolized as:

(2) Believes (Othello, that Desdemona loves Cassio), where the that-clause denotes the proposition that Desdemona loves Cassio. According to Frege's doctrine of indirect denotation, the expressions 'Desdemona', 'loves', 'Cassio', will in the context:

(3) that Desdemona loves Cassio

refer not to their ordinary denotations but rather to their ordinary senses. The Fregean proposition denoted by the that-clause (3) seems to be a complex built up from the (ordinary) senses of the expressions 'Desdemona', 'loves', and 'Cassio'. In general, propositions for Frege appear to have been sentence-like structures having senses as their constituents.

On the Russellian approach, propositional attitude verbs

are analyzed as sentential operators that apply syntactically to sentences rather than to noun phrases. Assuming that 'Othello', 'Desdemona', and 'Cassio' are genuine proper names and not abbreviations of definite descriptions, the logical form of (1) can on the Russellian analysis be represented as:

(4) Believes (Othello, Desdemona loves Cassio).

Belief is still analyzed as a relation between a person and a proposition, since the Russellian denotation of a sentence is a proposition rather than a truth-value. However, within the Russellian framework there are no senses available and Russellian propositions are not like Fregean propositions constructed out of senses. Atomic Russellian propositions are instead complexes built up from properties (or relations) and objects. Assuming that 'Desdemona loves Cassio' refers to an atomic Russellian proposition, this proposition is a complex:

(5) [loves; Desdemona, Cassio],

having as its constituents the loving-relation and the two persons Desdemona and Cassio. In general, an atomic Russellian proposition has the form $[R; a_1, \dots, a_n]$, where R is an n-ary relation (in intension) and a_1, a_2, \dots, a_n are any entities (of appropriate logical types). The proposition $[R; a_1, \dots, a_n]$ is true iff the objects a_1, \dots, a_n stand in the relation R to each other. In Principia Mathematica, complex propositions are constructed from atomic ones by logical operations:

- (i) if p is a proposition, then there is a proposition $\sim p$ (the negation of p) which is true iff p is not true;
- (ii) if p and q are propositions, then there is a proposition (p & q) (the conjunction of p and q) which is true iff p and q are both true;
- (iii) if f is a propositional function having as its domain the collection D of all entities of a certain logical type, then there is a proposition $(\forall x)f(x)$ (the universal generalization of f) which is true iff f(a) is a true proposition for every entity a in D.

Russellian semantics differs from Fregean semantics in not viewing definite descriptions "the + common noun phrase" as genuine singular terms. Instead they are treated syntactically

and semantically as quantifier phrases on the analogy of phrases like "a + common noun phrase" and "every + common noun phrase". Within a Russellian framework, quantifier expressions (or determiners) like "all", "most", "many", "no", "some", etc., may be viewed as denoting relations between properties.¹⁾ In other words, quantifier expressions denote higher-order propositional functions. Consider, for example, the sentences:

- (1) All kings of France are spies.
- (2) Some kings of France are spies.

These sentences may be analyzed as being of the forms:

- (3) All(F,S)
- (4) Some(F,S),

where F and S denote the properties (propositional functions) of being a king of France and being a spy, respectively. All and some denote binary propositional functions taking propositional functions of individuals as arguments. Similarly, on the Russellian view,

- (5) The king of France is a spy.

may be analyzed as having the logical form

- (6) The(F,S),

where the quantifier expression the denotes a relation (propositional function) Q such that Q(X,Y) holds if and only if

- (i) there is precisely one individual x such that X(x) is a true proposition; and
- (ii) for every individual x, if X(x) is true, then Y(x) is also true.

Hence, the following is a logical truth:

- (7) The(X,Y) ↔ (∃x)[(∀y)(Xy ↔ y = x) & Yx].

The sentence

- (8) Ralph believes that the king of France is a spy.

has, on the Russellian analysis, two readings, namely, one wide scope or de re reading:

- (9) The(F, (λx.Ralph believes that Sx));

and one narrow scope or de dicto reading:

- (10) Ralph believes that the(F,S).

From (8) and

- (11) The king of France is Ortcutt.

we may conclude

- (12) Ralph believes that Ortcutt is a spy.

only in case (8) is given the wide scope reading. Hence, failure at the level of surface syntax of the principle:

$$a = b \rightarrow (A(a/x) \leftrightarrow A(b/x))$$

is easily explained, on the Russellian approach, for those cases in which either a or b is a definite description or can be analyzed as standing for one.

Apparent failures of substitutivity of genuine singular terms are difficult to handle on the Russellian approach. The linguistic function of a meaningful expression is, on the Russellian view, entirely exhausted by the fact that it has a certain denotation. Accordingly, substitution of coreferential singular terms in a sentence can never change its content, i.e., the proposition it expresses. Russell writes in 'The Philosophy of Logical Atomism' (1918): "...if one thing has two names, you make exactly the same assertion whichever of the two names you use..."²⁾

It follows that if a and b are genuine singular terms denoting the same object, then the two sentences 'a = a' and 'a = b' express the same Russellian proposition. However, we are reluctant to describe someone who accepts the sentence 'not a = b', where a and b are coreferential singular terms, as believing that not a = a.

Failures of substitutivity of coextensional sentences and predicate terms are explained by Russell's doctrine that sentences and predicate terms denote propositions and propositional functions, respectively. Also logically equivalent sentences may fail to correspond to the same Russellian proposition and hence fail to be substitutable salva veritate, for example, in propositional attitude contexts.

1. Simple versus Ramified Types

Type distinctions were introduced into the Russellian ontology in order to avoid antinomies. In this paper, we shall modify Russell's approach by considering the simple theory of types rather than Russell's ramified theory. The simple theory is sufficient for avoiding Russell's paradox of predication, i.e., the contradiction resulting from the assumption:

$$(\exists f)(\forall x)(fx \leftrightarrow (\exists g)(g = x \ \& \ \sim g(g))),$$

where 'f' and 'g' range over predicates (propositional functions) and the variable 'x' ranges over any entities whatsoever (including predicates). For the resolution of semantic paradoxes involving notions like satisfaction, truth, and definition, we rely on Tarski's levels of language doctrine. The division of propositions and propositional functions into levels (or orders) which is involved in the ramified theory of types seems to be useful for the analysis of those self-referential paradoxes that involve propositional attitudes like assertion, belief, and knowledge. For example, if we add a primitive predicate symbol 'A' ("asserts") to the simple theory of types, we can reproduce the Liar antinomy, applied to propositions rather than to sentences. To derive this version of the Liar antinomy, we simply have to make the intuitively coherent assumption:

$$(\forall p)(A(e,p) \leftrightarrow p = (\forall q)(A(e,q) \rightarrow \sim q)),$$

i.e., that Epimenides makes precisely one assertion (a certain day) namely that all assertions made by Epimenides (on that day) are false. One possible way of avoiding the paradox is to separate the notion of an assertion into an infinite hierarchy of predicate symbols 'Aⁿ' of levels 1, 2, 3, ... However, this type of solution, although syntactically simpler than Russell's, seems to presuppose a division of propositions and propositional functions into levels which is similar to Russell's ramified type hierarchy.

According to the indicated resolution of the Liar antinomy, we can let Aⁿ(x,p) be a proposition of level n which is true iff p is a proposition of level < n and x asserts p.

Thus, if p is a proposition of level $\geq n$, then Aⁿ(x,p) is simply false. Assume now:

$$(1) (\forall p)(A^n(e,p) \leftrightarrow p = (\forall q)(A^m(e,q) \rightarrow \sim q))$$

and let 'P₀' abbreviate '($\forall q)(A^m(e,q) \rightarrow \sim q)$ '. In order for (1) to be true, we must have $n > m$. Next, we attempt to derive a contradiction. Hence, we assume P₀, i.e.,

$$(2) (\forall q)(A^m(e,q) \rightarrow \sim q) \quad \text{(by definition of } P_0)$$

$$(3) A^n(e,P_0) \quad \text{(from (1))}$$

However, from (2) and (3) we cannot derive $\sim P_0$, since A^m(e,P₀) is false. In fact, P₀ will be true in the described situation.

As was noted already by Russell in The Principles of Mathematics, appendix B (1903) and also by Myhill (1958), the simple theory of types is not sufficient to avoid antinomies if we assume very strict principles of individuation for propositions. In fact, the following principle is sufficient to generate a contradiction in the simple theory of types:

$$(P) (\forall f)(\forall g)([(\forall p)f(p) = (\forall p)g(p)] \rightarrow f = g),$$

where 'f' and 'g' range over propositional functions taking propositions as arguments. This principle seems intuitively plausible. If $f \neq g$, then it seems possible to assert $(\forall p)f(p)$ without asserting $(\forall p)g(p)$ and vice versa. Hence, $(\forall p)f(p)$ and $(\forall p)g(p)$ appear to be different propositions.

Let us now see how to derive a contradiction from the principle (P) within simple type theory. The argument is a formalized version of Russell's argumentation in appendix B of Russell (1903). It also has a formal similarity to the semantical antinomy of Tarski (1944), note 11.

We say that a proposition p is self-applicable if

$$(\exists f)([p = (\forall q)f(q)] \ \& \ f(p))$$

and non-self-applicable if

$$(\exists f)([p = (\forall q)f(q)] \ \& \ \sim f(p)).$$

By the comprehension principle of simple type theory there is a propositional function h such that

$$(1) (\forall p)(h(p) \leftrightarrow (\exists f)([p = (\forall q)f(q)] \ \& \ \sim f(p))).$$

That is, h is the property of being non-self-applicable. Consider now the proposition $(\forall p)h(p)$, i.e., the proposition that all propositions are non-self-applicable. Let p_0 be this proposition. Assume that the proposition p_0 is non-self-applicable, i.e.,

- (2) $h(p_0)$ (assumption)
- (3) $(\exists f)([p_0 = (\forall q)f(q)] \ \& \ \sim f(p_0))$ (from (1),(2))
- (4) $[p_0 = (\forall q)f(q)] \ \& \ \sim f(p_0)$ ((3) ES)
- (5) $(\forall p)h(p) = (\forall q)f(q)$ (from (4))
- (6) $h = f$ (from (5) by (P))
- (7) $\sim h(p_0)$ (from (4),(6))
- (8) $\sim h(p_0)$ ((2),(7) RAA)
- (9) $[p_0 = (\forall p)h(p)] \ \& \ \sim h(p_0)$ (from (8))
- (10) $(\exists f)([p_0 = (\forall p)f(p)] \ \& \ \sim f(p_0))$ ((9) EG)
- (11) $h(p_0)$ ((10) Df. of p_0)
- (12) $h(p_0) \ \& \ \sim h(p_0)$ (from (8),(11))

Thus, we have shown that too strict principles of individuation for propositions are incompatible with the simple theory of types. We leave it to the reader to verify that the above argument cannot be carried through within the ramified theory of types.

2. The Language of Simple Type Theory

In the rest of this paper we shall focus on a system STT of Russellian simple type theory. First, we shall give a detailed presentation of the syntax and semantics of STT. Then, we provide STT with a natural axiomatization and finally we state the appropriate Henkin-type generalized completeness theorem for the system STT.

Simple Types. Let i be any object which is not a finite sequence. The set ST of simple types (s-types) is the intersection of all sets T satisfying the conditions:

- (i) $i \in T$
- (ii) if $m \geq 0$ and $t_1, \dots, t_m \in T$, then the finite sequence

$\langle t_1, \dots, t_m \rangle$ belongs to T .

That is, ST is the smallest set which contains i and is closed under the formation of arbitrary finite sequences. It follows from the definition that the empty sequence \emptyset is in ST . i and \emptyset are the basic s-types. Types of the form $\langle t_1, \dots, t_m \rangle$ for $m \geq 1$ are called functional s-types.

The intuitive interpretation of the simple types is this: The objects of type i are individuals and objects of type \emptyset are propositions. Objects of type $\langle t_1, \dots, t_m \rangle$, where $m \geq 1$, are functions f such that the k th domain of f ($1 \leq k \leq m$) is the collection of all entities of type t_k and the range of f is included in the collection of all propositions. Such functions are called propositional functions of type $\langle t_1, \dots, t_m \rangle$.

We shall usually write 0 instead of \emptyset and (t_1, \dots, t_m) instead of $\langle t_1, \dots, t_m \rangle$.

Primitive Symbols. The language of STT contains (primitive) symbols of the following kinds:

- (i) for each s-type t , a denumerable sequence x_t^0, x_t^1, \dots of variables of type t ;
- (ii) for each s-type t , a denumerable sequence c_t^0, c_t^1, \dots of (non-logical) constants of type t ;
- (iii) truth-functional connectives \perp and \rightarrow ;
- (iv) the universal quantification symbol Π ;
- (v) the logical necessity symbol \Box ;
- (vi) the identity symbol $=$;
- (vii) the lambda operator λ ;
- (viii) parentheses and the comma sign.

Terms. We define inductively the set Tm_t of terms of STT of type t , as follows:

- (i) every variable of type t belongs to Tm_t ;
- (ii) every constant of type t belongs to Tm_t ;
- (iii) if $A \in Tm_{(t_1, \dots, t_n)}$ and $B_{t_1} \in Tm_{t_1}, \dots, B_{t_n} \in Tm_{t_n}$, then $A(B_{t_1}, \dots, B_{t_n}) \in Tm_t$;
- (iv) if $A, B \in Tm_0$, then $\Box A \in Tm_0$ and $(A \rightarrow B) \in Tm_0$;
- (v) $\perp \in Tm_0$;
- (vi) if $A \in Tm_{(t)}$, then $\Pi A \in Tm_0$;

- (vii) if $A, B \in Tm_t$, then $(A = B) \in Tm_0$;
- (viii) if $A \in Tm_0$ and x_{t_1}, \dots, x_{t_n} are distinct variables of respective types t_1, \dots, t_n , then $(\lambda x_{t_1} \dots x_{t_n} A) \in Tm_{(t_1, \dots, t_n)}$.

A closed term is a term that does not contain any free variables.

A formula is a term of type 0. A sentence is a closed formula.

We introduce the following metalinguistic abbreviations:

$$(\forall x_t)A =_{df.} \Pi(\lambda x_t A)$$

$$\sim A =_{df.} (A \rightarrow \perp).$$

$(\exists x_t)$, $\&$, \vee , \leftrightarrow are defined in the usual way.

Let A_u and B_t be terms of type u and t , respectively. We say that B_t is free for the variable x_t in A_u if there is no term $(\lambda x_1 \dots x_n C)$ such that:

- (i) some free occurrence of x_t in A_u lies within an occurrence of $(\lambda x_1 \dots x_n C)$; and
- (ii) one or several of the variables x_1, \dots, x_n occur free in B_t .

We write $A_t[B_{t_1}, \dots, B_{t_n} / x_{t_1}, \dots, x_{t_n}]$ for the result of (simultaneously) substituting B_{t_1}, \dots, B_{t_n} for all free occurrences of the variables x_{t_1}, \dots, x_{t_n} in A_t . Whenever this notation is used, x_{t_1}, \dots, x_{t_n} are assumed to be distinct variables.

3. Semantics for Simple Type Theory

A frame is a structure

$$F = \langle W, D_t, E, \underline{f}, \text{imp}, \text{nec}, \text{uq}_t, \text{id}_t \rangle_t \in \text{ST}$$

where

- (1) W is a non-empty set (of possible worlds).
- (2) D_i is a non-empty set (of individuals).
- (3) D_0 is a set having at least two elements. The elements of D_0 are called propositions.
- (4) For all s-types t_1, \dots, t_n ,

$$\emptyset \neq D_{(t_1, \dots, t_n)} \subseteq \left[D_0^{(D_{t_1} \times \dots \times D_{t_n})} \right],$$

i.e., $D_{(t_1, \dots, t_n)}$ is a non-empty set of functions from $(D_{t_1} \times \dots \times D_{t_n})$ into D_0 . The elements of $D_{(t_1, \dots, t_n)}$ are called propositional functions of type (t_1, \dots, t_n) .

- (5) $E: D_0 \times W \rightarrow \{0, 1\}$. For each proposition $p \in D_0$ and each world $w \in W$, $E(p, w)$ is the truth-value (or extension) of p in the world w . 0 and 1 represent the truth-values false and true, respectively. We define the intension $I(p)$ of a proposition p as the set of all worlds in which p is true, i.e.,

$$I(p) = \{w \in W : E(p, w) = 1\}.$$

- (6) $\underline{f} \in D_0$, $\text{imp}: D_0 \times D_0 \rightarrow D_0$, $\text{nec}: D_0 \rightarrow D_0$, $\text{uq}_t: D_{(t)} \rightarrow D_0$, $\text{id}_t: D_t \times D_t \rightarrow D_0$.

- (7) The function E satisfies the following conditions:
 - (a) $E(\underline{f}, w) = 0$, for all $w \in W$;
 - (b) for all $p, q \in D_0$, $w \in W$, $E(\text{imp}(p, q), w) = 1$ iff either $E(p, w) = 0$ or $E(q, w) = 1$;
 - (c) for all $p \in D_0$, $w \in W$, $E(\text{nec}(p), w) = 1$ iff for all $w' \in W$, $E(p, w') = 1$;
 - (d) for all $f \in D_{(t)}$, $w \in W$, $E(\text{uq}_t(f), w) = 1$ iff for all $x \in D_t$, $E(f(x), w) = 1$;
 - (e) for all $x, y \in D_t$, $w \in W$, $E(\text{id}_t(x, y), w) = 1$ iff $x = y$.

It follows from clauses (7). (a) - (e) that:

- (i) $I(\underline{f}) = \emptyset$;
- (ii) $I(\text{imp}(p, q)) = (W - I(p)) \cup I(q)$;
- (iii)
$$I(\text{nec}(p)) = \begin{cases} W, & \text{if } I(p) = W \\ \emptyset, & \text{if } I(p) \neq W \end{cases}$$
- (iv) $I(\text{uq}_t(f)) = \{w \in W : (\forall x \in D_t)(E(f(x), w) = 1)\}$;
- (v)
$$I(\text{id}_t(x, y)) = \begin{cases} W, & \text{if } x = y \\ \emptyset, & \text{if } x \neq y. \end{cases}$$

A structure for STT is an ordered triple $\underline{M} = \langle \underline{F}, w_0, V \rangle$, where

- (i) $\underline{F} = \langle W, D_t, E, \underline{f}, \text{imp}, \text{nec}, \text{uq}_t, \text{id}_t \rangle_{t \in ST}$ is a frame;
- (ii) $w_0 \in W$ (the actual world);
- (iii) V is a function which assigns a value $V(c_t) \in D_t$ to every (non-logical) constant of type t in STT.

Let $\underline{M} = \langle \underline{F}, w_0, V \rangle$ be a structure for STT. An M-assignment is a function g from the set of all variables such that for every $t \in ST$ and every variable x_t of type t , $g(x_t) \in D_t$. Given an assignment g , a variable x_t of type t , and an element $a \in D_t$, we define the assignment:

$$g(a/x_t) = (g - \{ \langle x_t, g(x_t) \rangle \}) \cup \{ \langle x_t, a \rangle \}.$$

The assignment $g(a_1, \dots, a_n/x_1, \dots, x_n)$, where x_1, \dots, x_n are distinct variables of respective types t_1, \dots, t_n and $a_1 \in D_{t_1}, \dots, a_n \in D_{t_n}$, is defined analogously.

General versus Standard Models for STT. The distinction between general and standard models was introduced by Leon Henkin in (1950). There he proved the completeness of a system of extensional simple type theory with respect to general models. Our goal here is to prove the corresponding result for Russellian simple type theory.

A general model (G-model) for STT is a structure $\underline{M} = \langle \underline{F}, w_0, V \rangle$ for STT such that there exists a function (the value function) which assigns, to each M-assignment g and each term $A_t \in Tm_t$, a value $|A_t[g]|^{\underline{M}} \in D_t$, in such a way as to satisfy the conditions (i) to (ix) below:

- (i) if c_t is a constant of type t , then $|c_t[g]|^{\underline{M}} = V(c_t)$;
- (ii) if x_t is a variable of type t , then $|x_t[g]|^{\underline{M}} = g(x_t)$;
- (iii) if $A \in Tm_{(t_1, \dots, t_n)}$ and $B_1 \in Tm_{t_1}, \dots, B_n \in Tm_{t_n}$, then $|A(B_1, \dots, B_n)[g]|^{\underline{M}} = |A[g]|(|B_1[g]|, \dots, |B_n[g]|)$ (we often omit the superscript \underline{M});
- (iv) $|I[g]| = \underline{f}$;

- (v) $|(A \rightarrow B)[g]| = \text{imp}(|A[g]|, |B[g]|)$;
- (vi) $|\Box A[g]| = \text{nec}(|A[g]|)$;
- (vii) $|\Pi A_t[g]| = \text{uq}_t(|A_t[g]|)$;
- (viii) $|(\Lambda_t = B_t)[g]| = \text{id}_t(|A_t[g]|, |B_t[g]|)$;
- (ix) if A is a formula and x_1, \dots, x_n are distinct variables of types t_1, \dots, t_n , respectively, then $|(\lambda x_1 \dots \lambda x_n A)[g]|$ equals the function F from $D_{t_1} \times \dots \times D_{t_n}$ to D_C such that for all $a_1 \in D_{t_1}, \dots, a_n \in D_{t_n}$,

$$F(a_1, \dots, a_n) = |A[g(a_1, \dots, a_n/x_1, \dots, x_n)]|.$$

A standard model for STT is a structure $\underline{M} = \langle \underline{F}, w_0, V \rangle$ for STT such that for all types t_1, \dots, t_n ,

$$D_{(t_1, \dots, t_n)} = \left[\begin{array}{c} (D_{t_1} \times \dots \times D_{t_n}) \\ D_0 \end{array} \right].$$

Clearly, if \underline{M} is a standard model, then it is also a general model. Later we shall see that there exist general models that are not standard models.

Truth in a Model. Let $\underline{M} = \langle \underline{F}, w_0, V \rangle$ be a general model for STT, where $\underline{F} = \langle W, D_t, E, \underline{f}, \text{imp}, \text{nec}, \text{uq}_t, \text{id}_t \rangle_{t \in ST}$. We say that the formula A is true at the world $w \in W$ relative to the model \underline{M} and the M-assignment g (in symbols: $\underline{M} \models_w A[g]$) iff

$$E(|A[g]|^{\underline{M}}, w) = 1.$$

The formula A is true in the model \underline{M} relative to the M-assignment g (in symbols: $\underline{M} \models A[g]$) iff $E(|A[g]|^{\underline{M}}, w_0) = 1$.

We have the following consequences of these definitions:

- (i) $\underline{M} \models_w A(B_{t_1}, \dots, B_{t_n})[g]$ iff $E(|A[g]|(|B_{t_1}[g]|, \dots, |B_{t_n}[g]|), w) = 1$.
- (ii) $\underline{M} \not\models_w I[g]$.
- (iii) $\underline{M} \models_w (A \rightarrow B)[g]$ iff $\underline{M} \not\models_w A[g]$ or $\underline{M} \models_w B[g]$.
- (iv) $\underline{M} \models_w \sim A[g]$ iff $\underline{M} \not\models_w A[g]$.

- (v) $M \models_w \Pi A_{(t)}[g]$ iff for all $a \in D_t$, $E(|A_{(t)}[g]|(a), w) = 1$.
- (vi) $M \models_w (A_t = B_t)[g]$ iff $|A_t[g]| = |B_t[g]|$.
- (vii) $M \models_w \Box A[g]$ iff for all $w' \in W$, $M \models_{w'} A[g]$.

Logical Truth and Logical Consequence. There are two notions of logical truth corresponding to the two concepts of a model. Hence, we define:

A formula A is G-valid in STT (in symbols: $\models_G A$) iff for each general model M and each M -assignment g , $M \models A[g]$. A is S-valid (in symbols: $\models_S A$) iff for each standard model M and each M -assignment g , $M \models A[g]$.

Similarly, we have two notions of logical consequence:

A formula A is a G-semantic consequence of a set Γ of formulas (in symbols: $\Gamma \models_G A$) iff for each general model M and each M -assignment g , if $M \models B[g]$ for every $B \in \Gamma$, then $M \models A[g]$. A is an S-semantic consequence of Γ (in symbols: $\Gamma \models_S A$) iff for every standard model M and every M -assignment g , if $M \models B[g]$ for every $B \in \Gamma$, then $M \models A[g]$.

We have, of course, the following connections between the G-notions and the S-notions:

- (i) $\models_G A \Rightarrow \models_S A$.
- (ii) $\Gamma \models_G A \Rightarrow \Gamma \models_S A$.

4. The Theory STT

We specify a recursive set of axioms and inference rules for the system STT of Russellian simple type theory. We then define a theorem of STT to be any formula obtainable from the axioms by repeated application of the rules of inference.

Axioms of STT. An axiom of STT is any instance of one of the following schemata:

- A1. Any formula A which is a tautology in \perp and \rightarrow .
- A2. $x_t = x_t$
- A3. $(x_t = y_t) \rightarrow (A_{t_1}[x_t/z_t] = A_{t_1}[y_t/z_t])$, where x_t and y_t are free for z_t in A_{t_1} .
- A4. $(\forall x_{t_1}) \dots (\forall x_{t_n})(f(x_{t_1}, \dots, x_{t_n}) = g(x_{t_1}, \dots, x_{t_n})) \rightarrow (f = g)$.
- A5. $(\lambda x_{t_1} \dots x_{t_n} A_0)(B_{t_1}, \dots, B_{t_n}) = A_0[B_{t_1}, \dots, B_{t_n}/x_{t_1}, \dots, x_{t_n}]$, where B_{t_i} ($1 \leq i \leq n$) is free for x_{t_i} in A_0 .

- A6. $\Pi A_{(t)} \rightarrow A_{(t)} B_t$.
- A7. $(\forall x_t)(A_0 \rightarrow B_{(t)}(x_t)) \rightarrow (A_0 \rightarrow \Pi B_{(t)})$, provided that x_t is not a free variable of A_0 or $B_{(t)}$.
- A8. $\Box A \rightarrow A$.
- A9. $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$.
- A10. $\sim \Box A \rightarrow \Box \sim A$.
- A11. $(A_0 = B_0) \rightarrow \Box(A \rightarrow B)$.

Rules of Inference.

- R1. From A_0 and $(A_0 \rightarrow B_0)$ to infer B_0 .
- R2. From $A_{(t)}(x_t)$ to infer $\Pi A_{(t)}$, if x_t is not free in $A_{(t)}$.
- R3. From A to infer $\Box A$.

We obtain a system of normal intensional type theory, if we add the schema

$$NI. \Box(A \leftrightarrow B) \rightarrow (A = B)$$

to the system STT. If we add instead

$$E. (A \leftrightarrow B) \rightarrow (A = B),$$

we obtain a system of extensional simple type theory. In this system we have as a theorem:

$$\Box A = A,$$

i.e., the operator \Box becomes superfluous.

Semantically, NI means that propositions having the same intension are identical. Hence, NI is equivalent to the assumption that propositions may be represented by sets of possible worlds. E means semantically that there are exactly two propositions which may be identified with the truth-values true and false.

A proof in STT is a finite sequence of formulas each of which is either an axiom or else is obtainable from earlier formulas by one of the inference rules R1 - R3. A formula A is provable in STT, or a theorem of STT, and we write

$$\vdash A \quad (\text{or to be more exact: } \vdash_{\text{STT}} A),$$

if it is the last line of a proof in STT. A derivation in STT of a formula A from a set of formulas Γ is a finite sequence of formulas A_1, \dots, A_n ($n \geq 1$) such that A_n is the

formula A and for each i ($1 \leq i \leq n$) either:

- (i) A_i is an axiom of STT; or
- (ii) $A_i \in \Gamma$; or
- (iii) A_i may be inferred from two previous members of the sequence by rule R1 (modus ponens); or
- (iv) for some $k < i$, A_k is the last member of a subsequence of A_1, \dots, A_n which is a proof of A_k in STT and A_i may be inferred from A_k either by rule R2 (generalization) or by rule R3 (necessitation).

We say that A is derivable from Γ in STT (in symbols: $\Gamma \vdash_{STT} A$) if there exists a derivation in STT of A from Γ . We state without proof:

Theorem. (Deduction Theorem for STT)

If $\Gamma, A \vdash_{STT} B$, then $\Gamma \vdash_{STT} A \rightarrow B$.

5. Soundness and Generalized Completeness of the System STT

In this section we shall state the appropriate soundness and completeness theorems for the system STT. First, however, we introduce some terminology. A set of formulas Γ is STT-consistent iff $\Gamma \not\vdash_{STT} \perp$. A set of formulas Γ is G-satisfiable [S-satisfiable] iff there exists a general [standard] model \underline{M} and an \underline{M} -assignment g such that $\underline{M} \models A[g]$ for all $A \in \Gamma$.

Theorem. (Soundness Theorem for STT)

- (i) $\vdash_{STT} A$ implies $\models_G A$;
- (ii) $\Gamma \vdash_{STT} A$ implies $\Gamma \models_G A$;
- (iii) if Γ is G-satisfiable, then Γ is STT-consistent.

Proof: (i) - (iii) are equivalent. We sketch the proof of (i). The axioms of STT are easily seen to be G-valid and the rules of inference clearly preserve G-validity. Hence, all theorems of STT are G-valid. Q.E.D.

It follows from classical results of Gödel (1931) that the set of S-valid formulas is not recursively enumerable. Hence, there is no hope of proving a completeness theorem with respect to S-validity. However, we have the following result with respect to the general semantics:

Theorem. (Generalized Completeness Theorem for STT)

- (i) $\models_G A$ implies $\vdash_{STT} A$;
- (ii) $\Gamma \models_G A$ implies $\Gamma \vdash_{STT} A$;
- (iii) if Γ is an STT-consistent set of formulas, then Γ is G-satisfiable.

We omit the long and fairly standard Henkin-type proof of this theorem. Our proof uses the same general techniques as Gallin's (1975, pp. 25-37) proof of strong completeness for Montague's higher-order intensional logic.

NOTES

*) Previous versions of this paper have been presented to the seminar in Theoretical Philosophy at Uppsala University. I am grateful to the participants, especially Thorild Dahlquist, Stig Kanger, Jan Odelstad, Włodzimierz Rabinowicz and Richard Sliwinski, for criticism and valuable suggestions.

- 1) Compare the treatment of quantifiers in van Benthem (1984).
- 2) Marsh (ed.), Logic and Knowledge (1956), p. 245.

REFERENCES

van Benthem, J., 1984. 'Questions about Quantifiers', The Journal of Symbolic Logic, vol. 49, pp. 443-466.

Carnap, R., 1947. Meaning and Necessity: A Study in Semantics and Modal Logic. University of Chicago Press, Chicago.

Church, A., 1951. 'A Formulation of the Logic of Sense and Denotation', in Structure, Method and Meaning: Essays in honor of H. M. Sheffer, Liberal Arts Press, New York.

Church, A., 1973. 'Outline of a Revised Formulation of the Logic of Sense and Denotation (Part I)', Nous 7, pp. 24-33.

Church, A., 1974, 'Russellian Simple Type Theory', in Proceedings and Addresses of the American Philosophical Association, vol. 47, pp. 21-33.

Church, A., 1974a, 'Outline of a Revised Formulation of the Logic of Sense and Denotation (Part II)', Nous 8, pp. 135-156.

- Church, A., 1976. 'Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski', The Journal of Symbolic Logic, vol 41, pp. 747-760.
- Frege, G., 1892. 'Über Sinn und Bedeutung', Zeitschrift für Philosophie und Philosophische Kritik 100, pp. 25-50.
- Gallin, D., 1975. Intensional and Higher-order Modal Logic, Mathematics Studies 19, North-Holland, Amsterdam.
- Gödel, K., 1931. 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I', Monatshefte für Mathematik und Physik, vol. 38, pp. 173-198.
- Henkin, L., 1950. 'Completeness in the Theory of Types', The Journal of Symbolic Logic, vol. 15, pp. 81-91.
- Kaplan, D., 1975. 'How to Russell a Frege-Church', The Journal of Philosophy, vol 72, pp. 716-729.
- Montague, R., 1974. Formal Philosophy. Selected papers of Richard Montague. Edited and with an introduction by Richmond H. Thomason. Yale University Press. New Haven.
- Myhill, J., 1958. 'Problems arising in the Formalization of Intensional Logic', Logique et Analyse 1, pp. 78-83.
- Quine, W. V., 1969. Set Theory and its Logic, Revised Edition, Harvard University Press, Cambridge, Mass.
- Russell, B., 1903. The Principles of Mathematics, Cambridge University Press, Cambridge.
- Russell, B., 1905. 'On Denoting', Mind, vol. 14, pp. 479-493.
- Russell, B., 1908. 'Mathematical Logic as based on the Theory of Types', American Journal of Mathematics 30, pp. 222-262.
- Russell, B. and Whitehead, A. N., 1910. Principia Mathematica, vol. I, Cambridge University Press, Cambridge.
- Russell, B., 1918. 'The Philosophy of Logical Atomism', The Monist 28, pp. 495-527; 29, pp. 33-63, 190-222, 344-380. Lectures delivered in London 1918. (Reprinted in Marsh (ed.) Logic and Knowledge (1956)).
- Tarski, A., 1944. 'The Semantic Conception of Truth and the Foundations of Semantics', Philosophy and Phenomenological Research 4, pp. 341-375.