Notes on intensional theories

Notas sobre teorías intensionales

Joseph Sneed
Colorado School of Mines, USA. discufilo@ucaldas.edu.co

Recibido el 11 de enero de 2011 y aprobado el 1 de febrero de 2011

Resumen

La cuestión de si los lenguajes intensionales son más expresivos que los lenguajes no-intensionales surge en el marco de una perspectiva semántica de las teorías. Desde esta perspectiva, la cuestión es esta. ¿Hay clases modelo que se pueden caracterizar mediante teorías que usan conceptos intensionales que no se pueden caracterizar mediante teorías que no usan conceptos intensionales? Se sugiere una formulación precisa de esta cuestión, pero no se ofrece una respuesta.

Para aproximarse a esta cuestión, se resume la teoría de modelos de primer orden [II] y se revisa el enfoque semántico de las teorías que emplea incrementos teóricos, no intensionales, de primer orden [III].

Los incrementos teóricos de primer orden se bosquejan pero no se definen rigurosamente [IV]. Este lenguaje intensional proporciona el aparato para atribuir uso del lenguaje y actitudes intensionales a individuos cuyo comportamiento es el objeto de investigación. También proporciona el aparato para hablar sobre traducción del lenguaje atribuido al lenguaje del investigador.

La cuestión inicial se convierte entonces en si hay clases modelo que se puedan caracterizar mediante incrementos intencionales de lógica de primer orden que no pueden ser capturados por incrementos teóricos no-intensionales [V].

Palabras clave

Sintaxis de primer orden -FOS, intensional, leyes, modelos, no-intensional, ley psico-física, functor de Ramsey.

Abstract

The question of whether intensional languages are more expressive than non-intensional languages is raised within the framework of a semantic view of theories. From this perspective, the question is this. Are there model classes that can be characterized by theories using intensional concepts that cannot be characterized by theories that do not use intensional concepts? A precise formulation of this question is suggested, but no answer is given.

To approach this question, model theory of first order theories is summarized [II] and the semantic approach to theories using non-intensional, theoretical augmentations of first order theories is reviewed [III].

Intensional augmentations of first order theories are sketched [IV] but not rigorously defined. This intensional language provides the apparatus for attributing language use and intensional attitudes to individuals whose behavior is the object investigation. It also provides apparatus for talking about translation from the attributed language to the investigator's language.

The initial question then becomes whether there are model classes that can be characterized by intensional augmentations of first order logic that cannot be captured by non-intensional theoretical augmentations [V].

Key words

First Order Syntax-FOS, intensional, laws, models, non-intensional, psycho-physical law, Ramsey functor.

I

1. Introduction

First we consider the purpose and then the approach to the present enterprise.

1.1. Purpose

What is an intensional theory of behavior? Are intensional theories of behavior more powerful than non-intensional theories? Is there a kind of behavior that can only be "explained" by intensional theories? Here I try to address these questions using a semantic (structuralist) conception of theory (Wolfgang Moulines Sneed 1987).

Consider a situation in which an "external observer" is trying to understand the behavior of a some number of individuals — human beings, animals, black boxes and possibly other kinds of things as well. Intuitively, the observer sees only overt behavior of these individuals — how they move about in space relative to each other, change color, produces sounds, etc. There is no *a priori* reason to believe that the individuals communicate with each other (or the external observer), use language, have beliefs, desires, etc. Such "intensional attributes" may be imputed to individuals in an effort to explain their behavior, but they are not a part of the behavior to be explained. 'Understanding' or 'explaining' the behavior of these individuals is taken in a minimal sense of distinguishing, in a general way, kinds of behavior that may occur from those that may not. This kind of understanding may lead to an ability to predict and/or control behavior, but it need not.

A bit more precisely, the observer uses some fixed vocabulary to describe the behavior of the individuals. Her task is to distinguish in some general way the kinds of behavior she observes (or countenances as possible) of these individuals from kinds of behavior he does not observe (and countenances as impossible). That is she wants to have a theory about the behavior of these individuals. Are there kinds of behavior that could *only* be characterized by attributing intensional states to some of the individuals... or are they, in principle, *eliminable*?

These questions can be given precise formulation by:

i) viewing the enterprise of producing a theory as characterizing a class of non-theoretical models using theoretical models;
ii) distinguishing intensional theoretical concepts from non-intensional theoretical concepts.

On my account of the matter, intensional concepts are essentially related to language. So we need to talk both about the language used by the observer to construct his theory as well as whatever language she might attribute to individuals about which she is theorizing.

A "side benefit" of raising the question of eleminability of intensional concepts in a semantic framework is that the "descriptive import" of laws involving intensional attitudes and linguistic concepts and the degree to which intensional attitudes determined by behavorial data is also illuminated.

For present purposes, languages are essentially formal devices (of a specific kind) that characterize classes of models (in a specific way). Since we know a good bit about the model theory of first order logic, it is expedient to begin thinking about our question in terms of languages conceived as syntax for first order logic (FOS).

Restricting our attention to FOS *de facto* excludes form consideration one important feature of the syntactic view of theories –"laws" that operate across different models of the theory– socalled "constraints". I say *'de facto'* because I don't know how to formulate constraints in FOS. But, there may be a way. Should the formulation of the question presented here lead us to conclude that intensional theories are no more expressive than non-intensional, theoretical theories; the question of whether consideration of constraints would make a difference should be considered.

1.2. Approach

We will begin by considering a descriptive language L_D -some specific instance of a first order syntax (FOS) —

$$L_D = \langle P_D, V, Q, C, concat_D, F_D, S_D \rangle$$

[E-I-1]

its interpretations with finite domains drawn from an "ur-domain" H— $\mathbb{I}[H, L_D]$ and the set of subsets of $\mathbb{I}[H, L_D]$, M_D, that can be characterized by finite sets of sentences of L_D — L_D theories. Here, P_D is a finite set of predicate types and S_D is the set of sentence types of L_D. The sets V, Q, and C are respectively variables, quantifiers and sentential connectives of L_D. F_D is the set of formulas and $concat_D$ is a concatenation relation used to define F_D recursively on the basis of P_D, V, Q, and C.

We will then consider theoretical, but non-intensional augmentations of L_D of the form

$$L_T = \langle P_D, P_T, V, Q, K, concat_T, F_T, S_T \rangle$$

[E-I-2]

where members of P_T are theoretical predicates. We will consider interpretations of L_T with domains drawn from a domain K — H emended with theoretical individuals — $\mathbb{I}[K, L_T]$ and M_T — the class of sub-sets of $\mathbb{I}[K, L_T]$ that can be characterized by finite sets of sentences of L_T. Each member t of $\mathbb{I}[K, L_T]$ has a descriptive fragment Ram**(t)** which is a member of $\mathbb{I}[H, L_D]$. Thus, sub-sets of $\mathbb{I}[K, L_T]$ determined by sentences of L_T correspond, via Ram, to subsets of $\mathbb{I}[H, L_D]$. In some cases, the Ram images of sub-sets of $\mathbb{I}[K, L_T]$ can not be characterized by any finite set of L_D sentences. In these cases, L_T is stronger than L_D.

Finally, we will consider an intensional, theoretical augmentation of L_D -- $L_{\mathbb{I}}$. The intensional language $L_{\mathbb{I}}$ will contain sufficient syntactical apparatus to permit the attribution of language use and intensional attitudes to some individuals. The objects of intensional attitudes are taken to be sets of non-theoretical models — sub-sets of $\mathbb{I}[H, L_D]$. Intensional theoretical augmentations are distinguished from non-intensional theoretical augmentations of L_D essentially in that the former have singular terms that denote sets of non-theoretical models. The formal apparatus used to do this is somewhat baroque. Many, I suspect, would deny that my $L_{\mathbb{I}}$ is an intuitively adequate, and even internally coherent, rendition of an intensional language. Some effort is devoted to anticipating these objections. The predicates of a first order, intensional syntax (FOIS) $L_{\mathbb{I}}$ will be:

$$P_{\mathbb{I}} = \langle P_D, \langle P_{LD}, P_{LA}, P_{QN}, trans, token, concat_{token}, that, P_A \rangle\rangle$$

[E-I-3]

where P_D are the predicates of the "underlying" descriptive language. The remainders are theoretical predicates analogous to the theoretical predicates P_T in the non-intensional theoretical augmentation.

Some idea of the intended interpretations of these predicates is required to understand why $L_{\mathbb{I}}$ might plausibly be viewed as an intensional language.

> • P_{LA} and P_{LD} are the predicates required to characterize the syntax of the attributed language L_A and the descriptive language L_D.

Consonant with the semantic conception of theory, these syntaxes are viewed as kinds of settheoretic structures and specific languages are viewed as instances (models) of these structures. The individuals in these structures are taken to be symbol types. They are regarded as theoretical individuals.

There must be singular terms in $L_{\mathbb{I}}$ that can be interpreted as referring to at least some symbol types in L_A and L_D. These are needed for $L_{\mathbb{I}}$-sentences which both *use* and *mention* L_D sentences. For example,

Whatever Bill wants he gets.

∀(x)(y)[prefer(b, that(❛y❜), that (❛¬y❜)) → y]

[E-I-4]

If Bill believes that pizza tastes good then pizza tastes good.

believe (b, that(❛p❜)) → p.

[E-I-5]

Apparently, we can get by with quote-names for sentences only. But including quote-names for other symbol types appears to be cost free. Thus:

> • P_{QN} is a set of $L_{\mathbb{I}}$-singular term types which will be interpreted as quote-names of L_D and L_A symbol-types, including sentences.

Flanking, black single quote marks '❛_❜' are symbol types used in constructing (via concat$_{\mathbb{I}}$) quote-names for L_D and L_A symbol types. They are $L_{\mathbb{I}}$-logical symbols analogous to quantifiers.

See [N-1].

- trans(⌜s_A⌝, ⌜s_D⌝) means that L_A-sentence type s_A is a translation of L_D-sentence type s_D.

I will describe below an $L_{\mathbb{I}}$-interpretation relative sense of translation.

- token(x, ⌜s_A⌝) means that non-theoretical individual x is a token for the attributed sentence type s_A.

- $concat_{token}$ is required to formulate "laws" requiring that the concatenation structure of the non-theoretical individuals identified as tokens for linguistic symbols have a set-theoretic structure isomorphic to a finite fragment of the structure of the abstract (theoretical) symbol types providing the interpretation for P_{LA} and P_{LD}.

Strictly speaking, $concat_{token}$ is a non-intensional theoretical predicate.

- that is a unary operation interpreted so that('s_A') denotes the class of models determined by the L_A-sentence s_A.

To interpret the that-operation in this way, we include the set of all sub-sets of $\mathbb{I}$[K, L_A] in the domains of all interpretations of $L_{\mathbb{I}}$. These we regard as theoretical individuals. Other theoretical individuals are required to provide interpretations for sets of symbol types.

- P_A is a set of intensional attitude predicate types.

Members of P_A will be interpreted with sets of tuples whose first member is a non-theoretical individual and whose other members are sub-sets of $\mathbb{I}$[K, L_A]. Thus, the objects of intensional attitudes are taken to be sets of non-theoretical models.

This intuitive sketch of interpretations of $L_{\mathbb{I}}$ will be filled out to characterize a set of interpretations $\mathbb{I}$[$K_{\mathbb{I}}$, $L_{\mathbb{I}}$]. Finite sets of $L_{\mathbb{I}}$-sentences may — intensional theories — will be regarded as characterizing sub-sets of $\mathbb{I}$[$K_{\mathbb{I}}$,$L_{\mathbb{I}}$].

Examples, of sentences that might be in intensional theories are:

$$\text{assert}(x, \text{that}(\ulcorner s_A \urcorner)) \leftrightarrow \exists\,(y)\,[\text{token}(y, \ulcorner s_A \urcorner) \wedge D(x,y) \wedge \ldots$$

[E-I-6]

where $D(x,y)$ is some purely descriptive L_D-predicate and '**...**' indicates more $L_{\mathbb{I}}$-predicates, either descriptive or intensional;

$$\text{prefer}(x, \text{that}(\ulcorner s_A \urcorner), \text{that}(\ulcorner \neg s_A \urcorner)) \;\&\; D(x,\ldots.) \;\&\; \text{trans}(\ulcorner s_A \urcorner, \ulcorner s_D \urcorner) \rightarrow s_D.$$

[E-I-7]

The second [E-I-7] is an example of a putative "psycho-physical law". It purports to provide descriptive conditions under which "preference leads to action". Providing an empirically acceptable intensional theory of some body of behavior is no part of the present enterprise. No claim is made that examples considered would be a part of such a theory. It is, however, claimed that a plausible account of the ontology and logical structure of intensional theories has been provided.

$L_{\mathbb{I}}$-models have purely descriptive fragments just as do L_T-models. Thus, it is possible, to regard the model classes determined by intensional theories — sub-sets of $\mathbb{I}[K_{\mathbb{I}}, L_{\mathbb{I}}]$ — as determining purely descriptive model classes — sub-sets of $\mathbb{I}[H, L_D]$, via a Ram-functor, in just the same way as it is possible to regard L_T-theories as determining purely descriptive model classes.

Are intensional theories essential to characterizing some kinds of behavior [V]? Using the apparatus sketched above, the question is roughly this.

- Is there some descriptive language L_D such that there are purely descriptive model classes determined by a theory in some intensional theoretical augmentation of L_D that can not be determined by a theory any nonintensional theoretical augmentation L_D?

It should be noted that the question is not whether the content of intensional theories can be reproduced by purely descriptive theories. Rather, it is whether the content of intensional theories can be reproduced by non-intensional, but still non-purely-descriptive, theories.

At this point, I don't have an answer to this question. But, I think the apparatus sketched above and described more fully below formulates the question with sufficient precision to admit of a rigorous answer. That is, there is a theorem to be proved — but I can't prove it.

II

2. Descriptive Language L_D

First, we describe as set-theoretic structures the syntax and then the semantics of a purely descriptive language.

2.1. Syntax

The non-theoretical, or descriptive language L_D is some specific instance of the syntax of first order logic (FOS) with a finite number of individual constants and a finite number of predicates of order's less than some fixed n.

For our purposes, it is convenient to think of instances of FOS as set-theoretic structures with the usual formation rules being part of the definition of a set-theoretic predicate that characterizes these structures. The sets appearing in these structures are to be interpreted sets of "symbol types". Symbol types are abstract entities whose "instances" are "symbol tokens". We will interpret symbol tokens to be individual physical objects. How symbol types are related to their tokens will be explained below.

More precisely, we may think of FOS's as set-theoretic structures of the following form:

$$L = \langle P, V, Q, C, concat, F, S \rangle$$

[E-II-1]

where:

$$P = \langle P^0, \ldots, P^{m+1} \rangle$$

is an m+1-tuple consisting of n-tuples

$$P^i = \langle P^i_1, \ldots, P^i_n \rangle$$

whose elements are predicate types of order i; 0 ≤ i ≤ m (Constants are 0-order predicate types.); V is a p-tuple of variable types; Q a 2-tuple of quantifier types; C is q-tuple of sentential connective types (all the usual ones), concat is a tertiary relation (binary operation) on the set of all symbol types appearing in P, V, Q, and C that characterizes the way symbol types in these sets are concatenated to form members of F — the set of formula types. S, a sub-set of F— is the set of sentence types. Different formula types in F (and sentence types in S) are distinguished by the way they are constructed by iteration of concatenations. Sentence types are distinguished from other formula types in the usual way via the concept of "bound variable".

Note that no delimiting symbols like "(" are used here. I assume these can be avoided by use of Polish notation and formation rules that attend to the arity of predicates and connectives. One could, as well, introduce set of delimiting symbols into the tuple L.

On this view, the class of set-theoretic structures that are FOS's is determined by defining a set-theoretic predicate 'is an FOS'. The usual "formation rules" for formulas and sentences in FOS appear as clauses in this definition.

Thus, we may think of L_D as some specific set-theoretic structure

$$L_D = < P_D, V, Q, C, concat_D, F_D, S_D >$$
[E-II-2]

in which P_D is a tuple consisting of some small number of predicate types, F_D and S_D are sets consisting respectively the formula types and sentence types constructed from the members of P_D using $concat_D$.

The motivation for this is that we will need to provide a formulation of the first order syntax (FOS) within a first order syntax (FOS) — provide a first order syntactical theory of first order syntax — and to speak about different "models" for this syntax in which different physical objects are taken to be the symbol tokens corresponding to the FOS-symbol-types.

2.2. Semantics

First, we describe interpretations of L_D, then an interpretation relative concept of truth for sentences in L_D, and finally L_D-theories.

2.2.1. Interpretation

Interpretations for L_D all have finite domains — h — of individuals drawn from some (possibly) infinite set of "ur-individuals" — H. Intuitively, H is just the set of all individuals — say, plovers and their predators — whose behavior interests "the observer". Specific H-interpretations are specific instances of this behavior in which a few individuals participate — say, plovers and their predators in front of my beach house on July 4, 1992.

We may consider the infinite set $\mathbb{I}[H,L_D]$ of all interpretations of L_D constructed in this way. Intuitively, this is the set of all possible data the observer might have about the behavior of the kinds of individuals that interest her. Note that the set $\mathbb{I}[H,L_D]$ will generally be infinite, though each member of it is a set-theoretic structure over a finite domain.

In the usual formulations of FOS predicate types are assigned set of tuples from h. So it is here too. An H-interpretation is just a 2-tuple

$$\mathbb{I} = < h, f^P >; f^P = < f^0, ..., f^n >$$
$$f^j \in SET\ (P^j, POT(h^j).$$

[E-II-3]

Notation is explained in [N-2]. Note that a semantic interpretation of L does *not* characterize a model for the set-theoretic structure L. Rather; it assumes that such a model is at hand and characterizes a semantic interpretation for this model.

2.2.2. Truth and models

Members of S_D — sentence types — are assigned interpretation relative truth ($\mathbb{I}$-truth) in the usual way. Thus, sentences of L_D characterize sub-sets of $\mathbb{I}[H,L_D]$ — the sub-sets of $\mathbb{I}[H,L_D]$ in which they are true. Members or these sub-sets are the models for sentences. Intuitively, we may think of sentences of L_D as "denoting" their model classes.

2.2.3. Descriptive theories

Sets of sentences T_D of L_D are linguistic expressions of descriptive theories. From a semantic point of view, the "theories" *are* the intersection of the classes of models characterized by (denoted by) the sentences. The model class of T_D is $M[T_D]$.

Let M_D be the set of all sub-sets of $\mathbb{I}[H,L_D]$ that the observer can characterize using L_D. Intuitively, M_D is the set of all possible, purely descriptive theories. These theories make use of no conceptual apparatus beyond the non-theoretical, descriptive vocabulary.

III

3. Theoretical augmentations of the descriptive language L_T

First, we describe the syntax and the semantics of the language L_T. Then we consider -theories.

3.1. Syntax

The simplest way to augment L_D is simply to add additional predicate types to those already appearing in L_D to produce L_T.

$$L_T = \langle \langle P_D, P_T \rangle V, Q, C, concat_T, F_T, S_T \rangle$$

[E-III-1]

Note that the "logical symbol types" remain unchanged. The relation $concat_T$ must be different from $concat_D$ simply because it has a bigger domain.

2.2. Semantics

First, we describe interpretations of L_T, then an interpretation relative concept of truth for sentences in L_T.

2.2.1. Interpretation

Intuitively, we want to allow for the possibility that some new kinds of individuals will be needed to satisfy some of our new, theoretical predicates. So consider a set of individuals K so that $H \subseteq K$. Then K-interpretations of L_T — theoretical interpretations — will look like:

$$\mathbb{I}_T = \langle k, f^P{}_D, f^P{}_T \rangle.$$

[E-III-2]

We allow that k may be infinite, but require that k intersect H be non-empty and finite:

$$K \cap H \neq \wedge, \text{ finite.}$$
[E-III-3]

Intuitively, we may employ an infinite number of theoretical individuals, but always in connection with some finite number of non-theoretical individuals.

3.2.2. Truth and models

This is no different from L_D except that models for sentences in L_T in have the set-theoretic structure of $\mathbb{I}_T$.

3.3. Theoretical theories

The additional theoretical apparatus may be used to construct sentences T_T that characterize classes of theoretical models $M[T_T]$ for the theoretical language –- sub-sets of $\mathbb{I}[K, L_T]$. From $M[T_T]$ we may obtain a sub-set of $\mathbb{I}[H, L_T]$ by doing two things:

i) from the members of $M[T_T]$ delete all the pairs containing P_{Tj}^{i}'s;

ii) from the sets of tuples of individuals paired with P_{Di}^{j}'s delete all tuples containing members of K - H.

Intuitively, i) eliminates all interpretations of theoretical predicates; ii) eliminates theoretical individuals from interpretations of descriptive predicates. Call the sub-set of $\mathbb{I}[H, L_D]$ obtained in the way 'Ram(M[T_T]])'. As above, Ram(M[T_T]]) is a theory (in the semantic sense) about H-P_D -behavior — behavior of individuals in H described with predicates P_D. 'Ram' is technically a "forgetful functor" sometimes called the 'Ramsey functor' to suggest the historical origin (Ramsey, 1960), of its use in explaining the logical form of empirical theories.

IV

4. Intensional augmentations of the descriptive language - $L_{\square}$

Intentional theories, on the account offered here, essentially involve the attribution of "language use" to some individuals and the attribution of "sentence tokenhood" to some other individuals. They also involve attribution of "intensional attitudes" to the same individuals to which language use is attributed. Intensional theories *may* (but do not essentially) involve the attribution of intensional attitudes to observed individuals that are "shared" by the external observer and linguistic communication between observer and observed. For the moment, I ignore this latter aspect of intensional theories. A somewhat different formulation of the view that intensional theories have this holistic character may (I think) be attributed to Davidson.

First we consider the syntax and then the semantics of an intensional language. Then we consider intensional theories.

4.1. Intensional First Order Syntax (▯FOS)

Intensional augmentations of the descriptive language add intensional attitude predicates together with the requisite linguistic apparatus to make them work. The linguistic apparatus permits the observer to talk about the syntactic structure of the attributed language and identify some observed individuals as symbol tokens in this language. In addition it provides a means for describing translation between the observer's descriptive language and the language whose use she attributes to some individuals. The key feature of this linguistic apparatus is an FOS characterization of FOS — a FOS theory whose models are FOS's. Both the attributed language and the descriptive language L_D are required to be models for this theory. We consider first the syntax needed for this theory.

Viewed as a set-theoretic structure,

$$L_{\square} = < P_{\square}, V, Q, C, ⌜\_⌝, concat_{\square}, F_{\square}, S_{\square} >.$$

[E-IV-1]

In addition to the logical symbols V, Q and C, $L_{\square}$ contains ⌜_⌝ which will be used to construct quote names.

The predicates of a first order intensional syntax ($\mathbb{I}$FOS), $L_{\mathbb{I}}$, will be:

$$P_{\mathbb{I}} = < P_D, < P_{LD}, P_{LA}, P_{QN}, \text{trans}, \text{token}, \text{concat}_{token}, \text{that}, P_A > >$$

[E-IV-2]

where P_D are the predicates of the "underlying" descriptive language. The remainders are theoretical predicates analogous to the theoretical predicates P_T in the non-intensional theoretical augmentation. These are discussed in more detail below.

4.1.1. First Order Syntax Predicates

The essential feature of intensional theories is a "language" (call it 'L_A') whose use is attributed to individuals. It may be FOS or some other formal structure like FOS. This language must have two essential features:

1. it must consist of an infinite set of "sentences" recursively definable over a finite "alphabet".

2. the sentences must provide a way of characterizing (denoting) some subsets of $\mathbb{I}[H, L_A]$ — the *same* set of interpretations that the observer works with.

The attributed language L_A may be some specific instance of FOS — the observer's or some other. Essentially, L_A-sentences function as $L_{\mathbb{I}}$-names for possible states of affairs the observer can describe in L_D only via ***use*** of L_D-sentences. That sentences of L_A have this property will be a formal requirement on the interpretation of an intensional, theoretical augmentation of L_D.

Clearly, we can attribute languages to individuals that are both stronger and weaker than the observer's L_D — in terms of the model classes they can characterize. For the purpose of considering the "theoretical power" of intensional languages it seems natural to require that the language attributed to individuals be no stronger than the observer's language.

My discussion will be restricted to attributed languages that are instances of FOS, though there may well be other formal structures that satisfy the two conditions above.

To use $L_{\square}$ to attribute use of some FOS — L_A — we must first provide $L_{\square}$ with predicates suitable for describing the set-theoretic structure of FOS. Ultimately we will use these predicates to produce an FOS-theory whose models are these set-theoretic structures. To do this we consider those FOS's in which only predicates needed for our immediate purpose appear.

Thus, we suppose that there are one-place predicates for all the symbol types appear in the tuple that is an FOS together with a 3-place concatenation relation. That is, we have

$$P_{FOS} = < \hat{P}, \hat{V}, \hat{Q}, \widehat{concat}, \hat{F}, \hat{S} >$$

[E- IV-3]

where:

$$\hat{P} = < \hat{P}^0, \ldots, \hat{P}^{m+1} >$$

is an m+1-tuple consisting of n-tuples

$$\hat{P}^i = < \hat{P}^i_1 \ldots, \hat{P}^i_n$$

and $\hat{P}^i_j$ is simply a j-place predicate; $\hat{V}$, is a p-tuple of one-place predicates; $\hat{Q}$ a q-tuple of one place predicates; $\widehat{concat}$ a 3-place predicate; $\hat{F}$ and $\hat{S}$ are one-place predicates.

Interpretations of these P_{FOS} are the sorts of things that could be FOS's in the set-theoretic sense — provided they are models for FOS-sentences T_{FOS} that provide a theory for FOS structures. They are "potential models" for an FOS theory of FOS.

So that we can talk about translation between the descriptive and attributed languages, we need to equip $L_{\square}$ with two instances of P_{FOS} — one for the attributed language L_A and one for the observer's language L_D. Call these, respectively,

$$P_{LA} \text{ and } P_{LD}$$

Intuitively, these predicates will be true of sentence types and other symbol types in these languages. Together, they will be required (by any intensional theory) to be models for T_{FOS}.

4.1.2. Quote names

$L_{\square}$ contains apparatus for forming quote-names of symbol types in L_A and L_D. This is needed to talk about attributed language use and translation between L_A and L_D. Quote-names, rather than simple constants, are needed because we must be able to read the intended referent of the name from the syntactic form of the name. Why this is so will become evident when we consider intended interpretations for $L_{\square}$.

The apparatus we employ consists of a predicate P_{QN} interpreted as a set of $L_{\square}$-singular term types the form 'x' together with symbol type '_'. The meta-linguistic formation rules for $L_{\square}$ will assure that

$$concat_{\square} ('\_', x) = 'x'$$

appears in P_{QN} iff x is a symbol type of L_A or L_D. See below [4.1.7.].

4.1.3. Translation predicate

The syntax of $L_{\square}$ will contain a predicate trans. Intuitively, trans ('s_A', 's_D') means sentence type s_A in L_A is a translation of sentence type s_D in L_D. Just how we construe 'translation' will be explained below [4.2.2.3.].

4.1.4. Token predicate

The syntax of $L_{\square}$ will also contain a predicate token. Intuitively, token (i,'s_A') means that individual i is a token for sentence type 's_A' in the attributed language L_A.

4.1.5. Token concatenation predicate

In attributing language use, an intensional theory will identify some non-theoretical individuals as tokens for symbol types in the attributed language. In any model for the theory, there will be at most a finite number of symbol tokens. In contrast, there will be an infinite number for formula and sentence types in F_A and S_A. "Laws" of the intensional theory will require that these symbol tokens have the same set-theoretic structure as some finite fragment of L_A. More precisely, token will be required to be a homomorphism between the interpretation of $concat_{token}$ and the interpretation of $concat_A$.

Intuitively, this is a part of the way an intensional theory “connects” abstract linguistic structures with infinite numbers of symbol types to observable behavior of a finite number of individuals. The rest of the way involves saying how observable behavior involving putative symbol tokens is related to intensional attitudes — i. e. characterizing linguistic behavior in intensional terms.

4.1.6. Intensional abstraction operator

The syntax of $L_{\mathbb{I}}$ will contain a unary operation symbol that. Intuitively, when s_A is a sentence of the attributed language that('s_A') denotes the class of models for s_A. Thus, that denotes a function the set of L_A-sentences S_A into the power set of H-interpretations of L_A — POT($\mathbb{I}$[H, L_A]).

4.1.7. Intensional Attitude Predicates

In addition to predicates intended to describe the syntax of L_A, L_D, their semantic relations and physical representations, an intensional language augments L_D with predicates P_A intended to attribute intensional attitudes to some individuals. Intuitively, these predicates describe relations between some individuals to whom language use is attributed and other abstract (theoretical) individuals which are classes of H-interpretations— sub-sets of $\mathbb{I}$[H, L_A] — denoted by sentences of the attributed language L_A.

Thus,

$$a(x, \text{that}('s_A'), \text{that}('\neg s_A'))$$

might be intuitively interpreted as

x prefers that(s) to that(not-s).

To this end, we add to L_D, P_A an m-1-tuple of predicate types of orders between 2 and m. Intuitively, we intend the first place in these predicates to be occupied by a non-linguistic individual and the remainder of the places to be occupied by quote-names of sentence tokens of L_A. We allow for multiple intensional objects, but only one bearer of these objects — no group minds.

For simplicity, Iterations of intensional attitudes, e.g.

Sam believes that Sue prefers wine to beer

are not considered here. More formally, they are not syntactically well formed. However, it appears that they could be treated by iterated levels of "intensional theorization".

4.1.8. Formation rules for $L_{▯}$

The formation rules for $L_{▯}$ work to characterize it in much the same way that they work in any FOS to obtain $F_{▯}$ and $S_{▯}$.

The major exception is the formation of quote-names for L_D and L_A symbol types. To do this, we need, for each predicate P in P_{LD} and P_{LA} (except the concat predicate), a clause of the form

For all X, if P(X) then P_{QN}($concat_{▯}$(▘, X, ▘)).

This rather liberal attitude to what is to counts as a sentence in $L_{▯}$ means that any restrictions on what is "meaningful" will be made in the semantics for $L_{▯}$.

Since sentence types in L_A are effectively treated as singular terms in $L_{▯}$, it may appear that $L_{▯}$-quantification into intensional contexts is ruled out. However, this need not be the case. Consider;

A) There is someone whom Bill believes to have killed Cockrobin.

which one might render in $L_{▯}$ as:

A')∃(x) [person(x) ∧ believes(b,that(▘kill(x,c)▘)]

The syntax of $L_{▯}$ apparently can be chosen to admit such a rendition. If there is a problem, it comes with specifying interpretation relative truth conditions for sentences like A'.

4.2. Semantics

First, we describe interpretations of $L_{▯}$ [4.2.1.], then an interpretation relative concept of truth for sentences in $L_{▯}$ [4.2.2.].

4.2.1. Interpretation

Interpretation is analogous to that provided for theoretical augmentations for L_D above [3.2.1.]. All $L_{\mathbb{I}}$-predicates except P_D will be treated as theoretical predicates. A domain of "urindividuals" K ($H \subseteq K$) provides for theoretical individuals. There are two kinds of theoretical individuals. First, there are those to provide interpretations for symbol-type predicates in P_{LA} and P_{LD}. Second, the interpretation of intensional abstraction and intensional attitude predicates (see sec. [4.2.1.8.] and [4.2.1.9.] below) requires enlarging the domain k of *every* interpretation $L_{\mathbb{I}}$ to include POT($\mathbb{I}$[H, L_D]). We regard these as *theoretical* individuals –- members of K-H. For those who might have ontological scruples about this enlargement, I suggest restricting the discussion to finite H's. Intuitively, it would not be too interesting to discover that the need for an intensional vocabulary hinged on wanting to talk about infinite sets.

Unlike interpretations for simple theoretical augmentations, the interpretations of some predicates will have restrictions on them that go beyond those of set-theoretic type [4.2.1.2.1.]. In most cases, these restrictions *partially*, but not completely, specify the meaning of these predicates. One might avoid these restrictions by including sentences in intensional theories whose models are restricted in these ways. However, it is not immediately evident that all restrictions we impose can be replicated syntactically in this way.

4.2.1.1. Descriptive predicates (P_D)

Interpretations may be provided for the descriptive fragments of $\mathbb{I}$FOS's, in the usual way. These interpretations will be restricted to H intersect k and simply have the form:

$$\mathbb{I}_D = < h, f^P >$$

Intuitively, descriptive predicates are required to be interpreted with sets of tuples of nontheoretical individuals.

4.2.1.2. First order syntax predicates

First, we consider attributed language predicates, P_{LA}, [4.2.1.2.1.] and then descriptive language predicates, P_{LD}, [4.2.1.2.2.].

4.2.1.2.1. Attributed language predicates

An interpretation $\mathbb{I}_{LA}$ of P_{LA} consists of functions assigning members of members of POT(k-(h ∪ POT($\mathbb{I}$[H, L_D)) to the one-place predicates in P_{LA} and some sub-set of POT((k-(h ∪ Pot($\mathbb{I}$[-H, L_D))^3) to $concat_A$. Thus interpretations of these predicates are restricted to theoretical (abstract) individuals which are not sets of H-interpretations for L_D. These theoretical individuals are introduced just for the purpose of providing interpretations for symbol types. The only interesting thing about them is the set-theoretic structure that will be imposed on them by the "laws" of T_{LA}. Depending on our T_{LA}, there may or may not be non-isomorphic interpretations of P_{LA}.

4.2.1.2.2. Descriptive language predicates

The interpretation $\mathbb{I}_{LD}$ of $L_{\mathbb{I}}$-predicates intended to describe the syntax of L_D is structurally the same as $\mathbb{I}_{LA}$.

Intuitively, however, this interpretation should be considered as "fixed". This means the observer considers only one syntactic representation of his language even though her theory of FOS syntax might allow for multiple models. The observer countenances possibly multiple interpretations of attributed language because he has no preconceived idea about which of the possibly multiple models for T_{LA} observed individuals might be using. But it's simply hard to see what intuitive sense could be made of letting in multiple interpretations of the observer's syntax.

Formally, this means that as we consider model classes determined by $L_{\mathbb{I}}$-sentences we require the interpretation of P_{LD} to be the same in all these. These considerations become otiose when T_{FOS} is categorical.

4.2.1.3. Quote names

The functions f^P_{QN}, $f^{'\_'}$ assign disjoint sub-sets of POT(k-(h ∪ POT($\mathbb{I}$[H, L_D)) to P_{QN}, '_' respectively. The interpretations of '_' are required to be distinct from interpretations of anything else. The interpretation of P_{QN} will depend on the interpretation already given for the attributed language predicates P_{LA}. The formation rules for $L_{\mathbb{I}}$ assure that, for each predicate P in P_{LD} and P_{LA}

$$\text{For all x, if } P(x) \text{ then } P_{QN}(concat_i(\text{"\_"}, x))$$
[E-IV-4]

Thus, we need only to stipulate further that

$$f^P_{QN}(\text{"x"}) = x$$
[E-IV-5]

4.2.1.4. Semantic interpretation of L_A

In addition to interpretation for the descriptive and linguistic predicates of $L_{\mathbb{I}}$, a K-interpretation for $L_{\mathbb{I}}$ must also provide a *semantic* interpretation for the attributed language L_A. Interpretation of the attributed linguistic predicates P_{LA} provides a syntactic interpretation. Intuitively, it attributes the use of an FOS to some individuals (at least in $M[T_{LA}]$ — models for T_{LA} the FOS theory of L_A). But attribution of full language use requires as well the attribution of “meaning” to this syntax.

This suggests that K-interpretations for intensional languages $L_{\mathbb{I}}$ have as component parts H-interpretations of the attributed language L_A. Formally, this just amounts to functions that map the linguistic predicates P_A into the appropriate types of sets of the domain $h \subset H$ of the interpretation $\mathbb{I}_{LA}$ of the descriptive predicates. Thus,

$$\mathbb{I} = < h, f^P{}_{LA} >$$

Note that the $f^P{}_{LA}$'s that appear in the semantic interpretation $\mathbb{I}_*$ of the attributed language L_A are different from the $f^P{}_{LA}$'s that appear in $\mathbb{I}_{\mathbf{LA}}$. The latter simply assign sub-sets of k-(h ∪ POT($\mathbb{I}[H, L_D)$ to all the predicates regardless of arity. The intended interpretation is sets of symbol types. The former assign sub-sets of h^n depending on the arity n. The intended interpretation is the “meaning” of the symbol types assigned to the latter. In $L_{\mathbb{I}}$-models for T_{LA} where the interpretations for P_{LA} are FOS's, the semantic interpretation $\mathbb{I}_*$ will provide “denotations” for members of S_A — the sentence types of L_A — via the usual definition of interpretation relative truth. They may be viewed as denoting their model classes $M[S_A] \subset \mathbb{I}[H, L_{LA}]$. Outside $M[T_{LA}]$ we may still assign semantic interpretations to P_{LA}, but lacking the structure of FOS, the definition of truth will generally lead to nonsense. More precisely, recursive definitions will not be able to move away from their basic cases for lack of structures that satisfy their conditions.

Note that including $\square_*$ in $\square_\square$ is a departure from the usual way of interpreting FOS. At this point, and only at this point, we depart from the usual practice of simply assigning set-theoretic objects to *predicates*. However, $\square_*$ is described in the meta-language for $L_\square$ in just the same way as the rest of $\square_\square$ so that semantic paradox is apparently avoided.

4.2.1.5. Translation predicate

Now that we have agreed that an interpretation of $L_\square$ must include an interpretation of the descriptive predicates -- $\square_D$ — as well as an interpretation of the attributed language — $\square_*$ — we can explain how to interpret trans.

Note first that the arguments of an atomic $L_\square$-sentence trans(a, b) will be $L_\square$-singular-terms. They will not *be* L_A- and L_D-sentences. Our intention is to interpret trans so that trans(a,b) will be true only if the singular terms a and b refer to sentence types. To this end, we begin by interpreting 'trans' as a sub-set of:

$$f^P_{LA}(S_A) \text{ X } f^P_{LD}(S_D)$$

That is, it is interpreted as a set of ordered pairs of L_A-L_D-sentence types. This interpretation of trans assumes we have already assigned the interpretations to the sentence-type predicates in P_{LA} and P_{LA}.

Intuitively, we want to impose further conditions on the interpretation of trans that capture the idea of *interpretation relative* sameness of meaning. Having already assigned interpretations to L_D and L_A what (if any) L_D and L_A-sentence types have the same meaning?

For example, we could interpret trans(a,b) to mean something like 'a has the same syntactic structure as b and the same interpretation of all predicates'. More precisely,

$$\text{trans}(\ulcorner s_A \urcorner, \ulcorner s_D \urcorner) \text{ is true in } < \square_*, \square_D >$$

$$\Leftrightarrow$$

there is a one-one mapping from symbols in s_A to symbols in s_D which preserves syntactic structure and corresponding predicate symbols are assigned the same interpretation by both $\square_*$ and $\square_D$.

According to this interpretation, trans entails material equivalence — i.e. trans(a,b) is ▯–true only if a and b are both ▯-true or both ▯-false. But, it is stronger than material equivalence. One can think of weaker requirements for the truth of trans-sentences that would still be plausible and still entail material equivalence. One might call this interpretation of trans ‘literal translation’.

On any plausible weaker interpretation, the truth of trans-sentences depends on the syntactic structure of its arguments and on specific pairs of interpretations.

Note that a more expressive $L_{▯}$ could be obtained by replacing trans with a one-way translation predicate include and defining trans(a,b) as include(a,b) & include(b,a).

4.2.1.6. Token predicate

Intuitively, domains of non-theoretical individuals consist of things that can *be* symbol tokens — including sentence tokens in L_A and L_D — as well as things that can have intensional attitudes to model classes denoted by sentence types and other things as well that have descriptive properties. The interpretation of token is thus simply a sub set of h X (k-h). Typically, we expect it to be a many-one mapping *into* k-h. That is, many physical objects may count as tokens of the same symbol type. And some symbol types will not have corresponding tokens. For example, only some small number of the infinite number of sentence tokens will be “represented” by tokens in any given interpretation.

4.2.1.7. Token concatenation predicate

The predicate ‘$concat_{token}$ is to be interpreted with a set of 3-tuples from H ∩ k — that is with 3-tuples of non-theoretical individuals. Thus it is a non-intensional, theoretical predicate. Intuitively, it is theoretical because “tokenhood” and what counts as concatenated tokens is something that is *imputed* by the theory — not something that is a part of the behavioral data for the theory. One can imagine there being several ways of imputing tokenhood and concatenation among tokens that would be compatible with the same behavioral data.

4.2.1.8. Intensional abstraction operator

The unary operator that is interpreted so that, in the case that s_A is interpreted in $\mathbb{I}_{\mathbb{I}}$ as denoting an L_A sentence type, that("s_A") denotes $M[s_A]$ — the H-model class of s_A. In all other cases, we simply let that("s_A") denote the null-set.

4.2.1.9. Intensional attitude predicates

First, we describe the interpretation of intensional altitude predicates [4.2.1.9.1.], then we consider intensional abstraction [4.2.1.9.2.].

4.2.1.9.1. Interpretation

Interpretations for the intensional attitude predicates

$$A = \langle A^2, A^3, \ldots, A^m \rangle$$

are more subtle. Intuitively, we take the *objects* of individual i's intensional attitudes to be sets of H-interpretations of the purely descriptive, non-linguistic and non-intensional, part of the intensional language. Thus, a K-interpretation with descriptive domain h assigns to predicates in A^{n+1} some sub-set of

$$h \text{ X } (POT(\mathbb{I}[H,L_A]))^n$$

Each n+1-tuple in this set consists of an individual member of h plus an n-tuple of *sets* of H-interpretations for the attributed language L_A.

More formally, the interpretation of $L_{\mathbb{I}}$ will contain

$$f^A = \langle f^2, f^3, \ldots, f^m \rangle$$

so that

$$f^{n+1} \in SET(A^{n+1}, h \text{ X } (POT(\mathbb{I}[H,L_A]))^n).$$

4.2.1.9.2. Intensional attitudes and intensional abstraction

Consider the intensional attitude $L_{\mathbb{I}}$-sentence

$$a(c, t)$$

where a is in the set of intensional attitude predicates A^2, c in h, and t is a singular term (either a member of P^0 or of the form that(▪s_A▪). Intuitively, the idea is that intensional attitude $L_{▯}$-sentences like this one are i-true only if the singular term denotes (in interpretation ▯) the model class of some L_A-sentence type.

So far, we have effectively stipulated that the singular term t denotes a model class of an L_A-sentence if it is of the form that(▪s_A▪). It still remains open that other $L_{▯}$-constants might be ▯ interpreted as denoting model classes for L_A-sentences or, indeed, other sub-sets of ▯[H, L_D].

For our purposes, it seems clear that this should be ruled out. That is, the only singular terms in $L_{▯}$ that denote sub-sets of ▯[H, L_D] are of the form that(▪s_A▪).

Intuitively, this means that the only apparatus in $L_{▯}$ that can "directly" refer to these model classes is that provided by L_A. It is just here that the potential for increased strength in determining models classes in ▯[H, L_D] could arise.

It should also be noted here that the "observer", the user of $L_{▯}$, does not *use* sentence types in L_A, she only *mentions* them via singular terms of $L_{▯}$ that denote them. The observer can also talk about the model classes characterized by these L_A sentences both in attributions of intensional attitudes and — so far as the preceding discussion has taken us — in other contexts as well. We have said nothing yet that rules out attributing descriptive (L_D) predicates to model classes. Thus, we might say something like

heavier_than(george, that(it's raining))

It is not completely obvious that we want to rule this out. In general, we do not want to preclude attributing descriptive properties to theoretical individuals (see [3.3.] above). For example, we attribute (descriptive) spatial properties to (arguably, theoretical) electrons. However, for present purposes, it appears natural to regard a predicate's attributability to model classes as a sufficient condition for taking it to be intensional. Thus, we require descriptive predicates to be interpreted with sets tuples of non-intensional individuals — either non-theoretical or theoretical, but non-intensional. An intensional individual is just a member of POT(▯[H, L_D]).

4.2.2. Truth and models

An interpretation of $L_{▯}$ will have the form:

$$▯_{▯} = < k, ▯_{D}, ▯_{LA}, ▯_{LD}, ▯_{*}, f^{P}{}_{QN}, f^{'\_'}, f^{trans}, f^{token}, f^{concat}{}_{token}, f^{A} >$$
[E-IV-6]

Each member of the tuple $L_{▯}$ will be described below.

4.2.2.1. Truth definition

A definition of 'true in interpretation $▯_{▯}$' for $L_{▯}$-sentences can be provided in the something like the usual way. The interpretation of predicates and singular terms leads in the obvious way to ▯▯-truth definitions for atomic sentences. Once this is done, sentential connectives and quantifiers work as they usually do.

4.2.2.2. Opacity of intensional contexts

It should be noted that referential opacity of intensional contexts will fall out of this truth definition in a natural way. Suppose that

$$a = b$$

and

attitude(x, that(ˈP(a)ˈ))

are both $▯_{▯}$-true. It will not then generally be the case that

attitude(x, that(ˈP(b)ˈ))

is also $▯_{▯}$-true. For,

that(ˈP(a)ˈ) ≠ that(ˈP(b)ˈ)

that(ˈP(a)ˈ) denotes the set of interpretations in which the denotation of a is in the sub-set of h denoted by P, while that(ˈP(b)ˈ) denotes the set of interpretations in which the denotation of b is in the sub-set of h denoted by P. These two sets of interpretations are isomorphic under permutation of tuples <a, _> and <b, _> in the functions that comprise the interpretations. But, they are *not* identical.

This may be clarified by making explicit one feature of our concept of interpretation. Our interpretations are tuples of functions

$$f^{i} \in SET(\{P^{i}\}, POT(k^{i}))$$

These functions are sets of ordered pairs of the form

$$<p, M>$$

where p is a linguistic symbol type (predicate or constant) and M is a set of tuples from k. In semantic formulations of theories (for example, those given by informal definition of a settheoretic predicate), it is usually just the *values* of the f^{i}'s that appear in the theory's "models" — the M's. These values appear a members of ordered tuples. Their position in these tuples serves to identify and distinguish them — e. g. to say which set of tuples is the "heavier-than" relation and which is the "longer-than". Here, it is the arguments of the f^{i}'s that identify these sets of tuples.

This entails that there is some ambiguity involved in talking about L_D-theories "determining a model classes". On one hand, it is perfectly clear to say that L_D-sentences determine sub-sets of $\mathbb{I}[H,L_D]$. Members of these subs-sets are *interpretations* of L_D in the sense just described. But these sets are not *exactly* the same as sets of models for a "corresponding" theory provided by an informal definition of a set-theoretic predicate. In fact, there will generally be a manyone correspondence between the set of interpretations of L_D determined by an L_D-theory and the set of models determined by a "corresponding" informal definition of a set-theoretic predicate. L_D-interpretations that differ "trivially" in that different constants and/or predicates are assigned to the same tuples will all correspond to the same model for the set-theoretic predicate.

Intuitively, our concept of interpretation makes explicit *exactly how* linguistic symbol-types are (literally) mapped onto (small parts of) the world. It is just this explicitness that makes intensional contexts opaque. The singular term that(⌜P(a)⌝) denotes a different set of interpretations than the singular term that(⌜P(b)⌝) just because the constants a and b are mapped onto the world in different ways.

Note, as well, that essential to referential opacity of intensional contexts is the fact that intensional objects — that(⌜sa⌝) — are, in interpretation $\mathbb{I}_I$, set-theoretic objects constructed from individuals outside the domain h of non-theoretical individuals. Aside from the L_D symbol types, members of H-h appear as well. On one view of intensional objects, these individuals H-h are "possible individuals", while those in h are "actual individuals". On the present view, members of H-h are no less real than members of h. Theories — whether expressed in formal or informal languages — generally have multiple models. All these models consist of real, actual individuals. The purpose of theorizing is not to characterize the world as a whole, but rather a number of small, possibly overlapping, fragments of the world.

4.2.2.3. Theory of meaning?

One might expect, having provided a semantics for $L_{\square}$, one would be able to say something about the logical consequence relation among intensional members of $S_{\square}$. Are there interesting, general things to note about when $s_{\square}$ is true in all interpretations in which $s_{\square}'$ is true? For example, can we define 'logical consequence' in such a way that

Bill believes someone killed Cockrobin

turns out to be a logical consequence of

Bill believes Socks killed Cockrobin?

Clearly, we can not. The reason is evident. We have placed no limitations at all on how the sets of H-interpretations assigned to intensional attitude predicates are to be related. This is rather like failing to place conditions on truth value assignments that make them "normal" with respect to the truth functional connectives. "Somehow" configurations of interpretations of intensional attitude predicates must be constrained by properties of the objects of these attitudes. To show "just how" is to provide a "theory of meaning" for intensional sentences.

There are basically two ways to proceed. One way is to enrich the concept of $L_{\square}$-interpretation in such a way that limitations on how sub-sets of $\square[H, L_{\square}]$ are assigned to intensional attitude predicates are built-in to the concept of interpretation. The other way is to leave the concept of $L_{\square}$-interpretation relatively weak and allow the meaning of intensional concepts to be further constrained by the laws of intensional theories. Following the second line, one might (at most) regard the concept of $L_{\square}$-interpretation provided here as a preliminary step toward an interesting theory of meaning.

Returning to our initial question [1.1.], is there anything interesting that can be said about the expressive power of intensional languages without saying more about a theory of meaning for such languages?

4.3. Intensional theories

Intensional theories are simply sets of $L_{\square}$-sentences, $T_{\square}$ [4.3.2.], with some interesting distinguished sub-sets [4.3.1.]. Models for intensional theores [4.3.3.] are described and indetermininacy of intensional concepts considered [4.3.4.].

4.3.1. Partial intensional theories

It is useful to distinguish three parts of full intensional theories $T_{\square}$ [4.3.2.]: an attributed language theory T_{LA} [4.3.1.1.], a purely intensional theory T_A [4.3.1.2.], and a purely descriptive theory T_D [4.3.1.3.].

4.3.1.1. Attributed language theory: T_{LA}

T_{LA} is a set of FOS sentences characterizing the syntax of L_A. In the case that L_A is an FOS, it is apparent what these sentences must like. Their models must have the set-theoretic structure characteristic of the syntax of FOS. T_{LA} will also require that the token relation be a homeomorphism between $concat_{token}$ and a fragment of $concat_A$.

4.3.1.2. Purely intensional theory: T_A

T_A is the "purely intensional" part of $T_{\square}$. It characterizes the structure required of the entities that are models for intensional attitude predicates.

An example of a plausible T_A is the purely qualitative fragment of Jeffrey decision theory (the Jeffrey-Bolker axioms) (Richard, 1965) — more precisely an FOS axiomatization of this theory slightly modified to accommodate the present model theoretic conception of the objects of the intensional attitudes. This theory deals with the intensional attitudes:

x believes a is at least as likely as b = more_likely(x ,a, b)

x weakly prefers a to b = pref(x, a, b).

Recalling that intensional objects (a and b) are *sets* of interpretations, two requirements of this theory can be rendered as:

A) If b ⊆ a then more_likely(x ,a, b).

B) If more_likely(x ,a, b) & more_likely(x ,b, c) then more_likely(x ,a, c)

To see how such a theory might reproduce plausible inferences about beliefs, note that belief simpliciter is rendered in this theory as:

C) believe(x ,a) iff more_likely(x ,a, that(⌜P v ¬P⌝)).

Now note that

D) that(⌜P(a)⌝) ⊆ that(⌜∀xP(x)⌝).

Thus, using A), B), C) and D), we may infer

believe(x, that(⌜P(a)⌝))

from

believe(x, that(⌜∀xP(x)⌝))

Transitivity is also required of the pref relation:

If pref(x, a, b) & pref(x, b, c) then pref(x, a, c)

but no plausible conditions connecting set-theoretic properties of a and b with pref (analogous to A) above) are readily apparent. This suggests that interesting inferences involving pref alone are not likely to be found.

4.3.1.3. Purely descriptive theory: T_D

T_D is the "purely descriptive" part of T_I. It is what the observer holding T_I believes about the situations in question that can be expressed in the purely descriptive vocabulary.

4.3.2. Full intensional theories: T_I

Clearly, no plausible T_I can be just the union (conjunction) of T_{LA}, T_A and T_D. Nor can it be any purely set-theoretic (truth functional) combination of them. There has to be some quantificational link among the components.

$T_{\mathbb{I}}$ must be require some kind of connection between (some of) the sentence tokens whose intensional abstractions fill the intensional object places in T_A and the predicates appearing in T_D. Crudely, the attribution of intensional attitudes must have some "descriptive import".

How should this work? The full intensional theory $T_{\mathbb{I}}$ may contain sentences describing linguistic actions [4.3.2.1.] and psyco-physical laws [4.3.2.2.]. However, holistic theories [4.3.2.3.] may have descriptive import without containing such sentences.

4.3.2.1. Linguistic actions

First, consider "linguistic actions" like asserting, questioning, commanding, etc. Within the present framework, it seems natural to regard these as *manifested by* "descriptive" or "observable" relations between non-linguistic individuals and sentence tokens. But, these descriptive relations are connected by $T_{\mathbb{I}}$ to attributions of intensional attitudes. That is, Hans shouting 'Tur schliessen!' (a relation between non-theoretical individual Hans and a disturbance in the ambient atmosphere — a token for the sentence type "Tur schliessen!" and also a nontheoretical individual) counts as Hans commanding that the door be shut only in models where certain intensional relations are also attributed to Hans, 'Tur schliessen' and perhaps other sentence tokens as well.

Thus, in $T_{\mathbb{I}}$ one might expect to find sentences like:

command(x, that(▘s_A▝) iff ∃(y) [token(y, ▘s_A▝) & D(x,y) &...

where D(x,y) is some purely descriptive relation and '...' indicates more conditions, either descriptive or intensional.

4.3.2.2. Psyco-physical laws

What about other, non-linguistic, actions? Intuitively, the obvious tack here is to suppose that $T_{\mathbb{I}}$ contains "psycho-physical laws". That is, $T_{\mathbb{I}}$ requires that some configurations of intensional attitude attributions entail the "truth" of some of the sentence tokens appearing as objects in these attitudes. For example, if T_A is Jeffrey decision theory then for *some* sentences s among the sentences of L_A it is plausible to suppose $T_▯$ contains something roughly like:

$$\text{pref}(x, \text{that}(⌜s_A⌝), \text{that}(⌜\neg s_A⌝)) \,\&\, D(x,...) \,\&\, \text{trans}(⌜s_A⌝, ⌜s_D⌝) \Rightarrow s_D.$$

That is, under certain descriptive conditions described by ‘D(x,...)’ x’s preferring that(⌜s_A⌝) to that(⌜$\neg s_A$⌝) entails s_D when ‘s_D’ is a translation of ‘s_A’ — e.g. when s_D describes something that x can do in circumstances D(x,...).

Similarly, one might expect that for some perception predicates like “sees” laws of the following form might appear:

$$s_D \,\&\, ▯(⌜s_D⌝,...) \,\&\, D(x,...) \,\&\, \text{trans}(⌜s_A⌝, ⌜s_D⌝) \;-\!> \text{sees}(x, \text{that}(⌜s_A⌝))$$

That is, under certain descriptive conditions described by ‘D(x,...)’ and for certain kinds of sentences described by ▯(⌜s_D⌝,...), whenever s_D (is true), x sees that s_D. A “causal” theory of perception might be formulated in this way.

4.3.2.3. Holistic theories

Having considered the possibility that $T_▯$ might contain psycho-physical (and pysio-psychological) laws to clarify our conception of the apparatus permitted in $L_▯$, we may consider not the possibility that $T_▯$ can have “descriptive import” without sentences of these forms. That is, $T_▯$ might contain no sentences that had purely descriptive or purely intensional sentences connected to others as consequents in universally quantified implications. Roughly, there are no sentences in $T_▯$ that might count (even as conditional, partial) definitions of intensional predicates in terms of descriptive (or conversely).

This kind of holism is commonplace in theories from physical science. In such theories various theoretical concepts are so tightly interwoven with each other and with non-theoretical concepts that only in a few, very special models of the theory can one make inferences from fully non-theoretical sentences to fully theoretical (and conversely). Nevertheless, such theories do have descriptive, non-theoretical import. That is, they serve to characterize an non-trivial class of non-theoretical models.

4.3.3. Models for intensional theories

We consider theoretical [4.3.3.1.] and non-theoretical [4.3.3.2.] models for intensional theories $T_{\mathbb{I}}$.

4.3.3.1. Theoretical models

Recalling [E-IV-6], interpretations of $L_{\mathbb{I}}$ have the form:

$$i_{\mathbb{I}} = < k,\ i_D,\ i_{LA},\ i_{LD},\ i_*,\ f^P{}_{QN},\ f^{\{',\}},\ f^{trans},\ f^{token},\ f^{concat}{}_{token},\ f^A >$$

The set of all such interpretations — relative to a fixed ur-domains H, and K of non-theoretical and theoretical individuals (H ⊆ K) is $\mathbb{I}$[H, K, $L_{\mathbb{I}}$]. Each set of $L_{\mathbb{I}}$-sentences $T_{\mathbb{I}}$ — an $L_{\mathbb{I}}$-theory — determines of sub-set of $\mathbb{I}$[H, K, $L_{\mathbb{I}}$], M[$T_{\mathbb{I}}$]. The set of all sub-sets of $T_{\mathbb{I}}$ that can be determined in this way is **M**[$L_{\mathbb{I}}$].

4.3.3.2. Non-theoretical models

Each $i_{\mathbb{I}}$ in $\mathbb{I}$[H, K, $L_{\mathbb{I}}$] corresponds to exactly one member of $\mathbb{I}$[H, L_D] via a functor Ram such that

$$\text{Ram}(i_{\mathbb{I}}) = < h,\ i_D >$$

where h = k ∩ H and k and i_D are respectively the first an second members of $i_{\mathbb{I}}$. Intuitively, Ram just wipes out everything but the first and second members of $i_{\mathbb{I}}$.

Extending Ram to operate on sets, we note that each $L_{\mathbb{I}}$-theory, $T_{\mathbb{I}}$, determines a sub-set of $\mathbb{I}$[H, L_D], $\overline{\text{Ram}}$(M[$T_{\mathbb{I}}$]). This we call 'the descriptive content' of $T_{\mathbb{I}}$. The "empirical claim" of $T_{\mathbb{I}}$ is that all L_D-descriptions of observed behavior — $L_{\mathbb{I}}$-descriptions of behavioral data — are to be found in $\overline{\text{Ram}}$(M[$T_{\mathbb{I}}$]).

4.3.4. Indeterminacy of intensional concepts

The question of indeterminacy of intensional concepts is considered generally [4.3.4.1.] and specifically with respect to traslation [4.3.4.2.].

4.3.4.1. General

The question of whether some specific description of putative behavioral data — some specific member of $\square[H, L_D]$, i_D — is in the content of $T_{\square}$ is essentially this. Is there *some* theoretical augmentation if i_D, $i_{\square}$ that is in $M[T_{\square}]$? Except in very special cases, when the answer to this question is affirmative, there will be multiple theoretical augmentations of i_D to models for $T_{\square}$. That is, the intensional theoretical concepts required to demonstrate that i_D is in the content of $T_{\square}$ will not be uniquely determined. More intuitively, there may be a variety of ways to impute linguistic behavior and intensional attitudes to members of **h** that satisfy the laws of the intensional theory $T_{\square}$. This may be so even though the theory $T_{\square}$ is non-trivial — in the sense that $\overline{\text{Ram}}(M[T_{\square}])$ is a proper sub-set of $\square[H, L_D]$.

This kind of indeterminacy of theoretical concepts is common in theories from the physical sciences. Indeed, it remains even when these theories are strengthened by conditions -- socalled 'constraints' — that operate across different models for the theory. Thus, there is every reason to expect that intensional theories will exhibit the same kind of indeterminacy.

4.3.4.2. Translation

Indeterminacy of attributions of "meaning" to attributed language is one aspect of the indeterminacy of intensional concepts has received considerable attention within the framework of somewhat different formulations of the issues at hand (Quine 1960).

Intuitively, we may regard the triple

$$< f^{trans}, i_D, i_* >$$

as a meta-linguistic "translation manual" (in the Quinean sense) between the observer's descriptive language L_D and the language L_A which he attributes to some of the individuals he observes. At least we may make this intuitive identification in $M[T_{LA}]$ — those models for the descriptive-linguistic part of the language $L_{\square}$ in which interpretation of the symbol-type predicates of the attributed language have the formal properties of an FOS.

In the absence of further restriction on the models of interest, it is clear that there will be a multiplicity of possible translation manuals. Further restriction on the models is provided by intensional theories $T_{\square}$ can not be counted on to completely eliminate this. There will still be a multiplicity of translation manual triples intensional augmentations of an i_D that are in $\mathbf{M}[T_{\square}]$. That is, there will generally be a multiplicity of translation manuals compatible with the behavioral data.

Some features of this multiplicity are worth noting. First, there may be multiple possibilities for the i_* associated with some fixed i_D. Clearly, our intensional theory of L_D-described, H-behavior will allow for different instances of this behavior, i. e. different i_D's. It could be the case that each of these i_D's had associated with it (in $\mathbf{M}[T_{\square}]$) exactly one corresponding i_*. In this case we would say that the theory $T_{\square}$ uniquely determined the interpretation of the attributed language. In the case that there were multiple i_*'s associated with the same i_D we would say that $T_{\square}$ countenanced an "indeterminacy of translation". This appears to be the kind of *semantic* translation indeterminacy discussed by Quine (196).

There is, however, a further possibility for *syntactic* translation indeterminacy that becomes explicit in this formulation. For a fixed $< i_D, i_* >$, there may be multiple possibilities for interpreting 'trans'. Whether there are depends (in part) on how strong a notion of "syntactic translation" we build into the interpretation of 'trans'. Intuitively, this indeterminacy appears to be identifiable as that commonly encountered (even by true bilinguals) in rendering text of one language into that of another.

It should also be noted that some semantic translation pairs might be compatible only with a null-set interpretation of 'trans'. That is, on some acceptable semantic "translations" there might be no way to identify sentences as having the same meaning.

V

5. Comparison of theories

Our question is roughly this. Are there any descriptive model classes that can be characterized by an intensional theory that can not be characterized by a non-intensional theory? More precisely, are there any descriptive model classes that can be characterized by an intensional theory than can not be characterized by a theory using only non-intensional theoretical concepts?

There are at least two interesting ways to make this question precise. One way is to take the descriptive language L_D and ur-domain H to be fixed; the other is to consider all possible descriptive languages and ur-domains.

First, consider the case of a fixed L_D and H. Here the question is:

Is it the case that: Given L_D and H, for all intensional theoretical augmentations of L_D, $L_{\square}$, and all L_{T-}theories $T_{\square}$, there is some non-intensional theoretical augmentation of L_D, L_T, and L_T-theory T_T such that

$$\overline{\text{Ram}}(M[T_{\square}]).= \overline{\text{Ram}}(M[T_{\square}]) \ ?$$

Intuitively, we have settled on the kind of behavior to be theorized about by fixing L_D and H. We simply want to know wether there is anything we can say about this kind of behavior using intensional concepts that we could not say using non-intensional, theoretical concepts.

Next, consider the more sweeping question: is there any kind of behavior that demands intensional concepts for its characterization?

Is it the case that: For all L_D and H, and for all intensional theoretical augmentations of L_D, $L_{\square}$, and all $L_{\square}$-theories, $T_{\square}$, there is some non-intensional theoretical augmentation of L_D, L_T, and L_T-theory such that

$$\overline{\text{Ram}}(M[T_{\square}]).= \overline{\text{Ram}}(M[T_{\square}]) \ ?$$

I confess that, at this point, I have no idea how to answer either of these questions. Supposing you conjecture that the answer to the second question is negative, the natural strategy is to try to produce a counter example. But, even at the intuitive level, it's not clear to me what kind of use of an intensional language might provide counter examples. Some kinds of potential counter examples would clearly be unconvincing — e. g. those that depended on things like the cardinality of domains and arity of predicates. Should one be able to produce them, reformulating the question to rule them out would appear to be in order.

6. Notes

[N-1]: Use of single quotes in the author's meta-language will be governed by the usual conventions. Meta-linguistic names used to describe the formal languages (L_I, L_D, L_A) and their component parts will ***not*** be enclosed in single quotes unless the meta-linguistic name is mentioned, rather than used.

[N-2]: Here and in what follows, 'SET(A, B)' denotes the set of all functions from set A to set B. 'POT(X)' denotes the power set of X; X^0 = X. 'h^j' denotes the set of all j-tuples formed from members of h.

References

Balzer, W., Moulines, C. U. and Sneed, J. *An Architectonic for Science: The Structuralist Program*. D. Reidel, Dordrecht, 1987. Print.

Davidson, D. *Essays on Actions and Events*. Oxford: Clarendon Press, 1980. Print.

- - -. *Inquiries into Truth and Meaning*. Oxford: Clarendon Press, 1984. Print.

Jeffrey, R. *The Logic of Decision*. New York: McGraw-Hill, 1965. Print.

Quine, Willard V. O. *Word and Object*. New York: Wiley, 1960. Print.

Ramsey, Frank P. "Theories". *The Foundations of Mathematics.* New Jersey: Littlefield, Adams & Co. Patterson, 1960. Print.