APOHA

BUDDHIST NOMINALISM AND HUMAN COGNITION

EDITED BY

Mark Siderits

Tom Tillemans

Arindam Chakrabarti

Columbia University Press   New York

13

Classical Semantics and Apoha Semantics

Brendan S. Gillon

As is well known, Indian Buddhist thinkers found universals (*sāmānya*) to be metaphysically repugnant. Yet, many Indian thinkers considered universals necessary to account for how general expressions, such as common nouns, manage to apply to an unbounded set of individuals. Indian Buddhist thinkers, however, believed that a satisfactory account of this linguistic phenomenon could be provided without appeal to universals. In particular, such thinkers, called *apoha-vādins* (proponents of exclusion), held that, through exclusion (*apoha*) and difference (*anya*) of individuals, an *ersatz* universal could be found that would provide an empirically adequate account of the generality of a general expression.

Though no explicit semantic account of these two forms of negation, exclusion (*apoha*) and difference (*anya*), is given by the *apoha-vādins*, it is still tempting to think that they might have had such an account in mind, at least implicitly. This conjecture[1] is plausible for several reasons. To begin with, Indian grammarians had devised a generative grammar and semantics of classical Sanskrit. In addition, Indian grammarians themselves did distinguish two forms of negation, *prasajya-pratiṣedha* and *paryudāsa*, dubbed "verbally bound negation" and "nominally bound negation" by Matilal, though these forms were never given an explicit analysis (Matilal 1971, 162–165). Indeed, what precisely these negations consist in is still not clear (Gillon 1987). Finally, Dharmakīrti, one of the early developers of the apoha theory, himself had given an explicit semantic analysis of the Sanskrit word *eva* (*only*), which plays a crucial role in the statement of the truth conditions of the classical Indian syllogism (Kajiyama 1973; Gillon and Hayes 1982; Gillon 1999).

The aim of this paper is to show that the two most obvious candidates from contemporary logic that one might use to explicate the *apoha-vādin*'s notions of exclusion (*apoha*) and difference (*anya*), namely, internal and external negation, do not provide the *apoha-vādins* with the *ersatz* universals they were looking for. Below, I shall first state what a semantic theory is and rehearse its principal features, availing myself of the semantics of monadic predicate logic (without quantifiers). I shall then set out various semantics for monadic predicate logic using internal and external negation in various combinations. I shall conclude by showing that their combination does not permit the definition of an *ersatz* universal that appeals only to individuals and thereby satisfies the scruples of those, such as Buddhists, who found the positing of universals to be metaphysically repugnant.

WHAT IS A SEMANTIC THEORY?

The idea that provides the basis of contemporary semantic theory dates back to Pāṇini, the great Sanskrit grammarian of ancient India (c. fourth century B.C.E.). His idea can be summarized as follows: each Sanskrit sentence can be analyzed into minimal constituents and the sense of each minimal constituent contributes to the sense of the entire sentence. Thus, a grammar of a natural language should not only generate all and only its acceptable constituents, it should also make clear how meanings, once associated with the simplest constituents, determine the meanings of the complex constituents they constitute. The meanings associated with minimal constituents comprise, among other things, the constituents of situations, namely, actions (*kriyā*) and their participants (*kāraka*).[2] Here, however, is where the grammar fell short, for neither Pāṇini nor his successors had a clear way to give a mathematical treatment of situations and their components. The remedy to this obstacle did not appear until the second quarter of the twentieth century, when began to emerge the one discipline whose business it is to make precise how values associated with constituents determine the value of the constituents they make up, namely, the subdiscipline of logic known as model theory. Its founder, Alfred Tarski (1901–1983), recognized the pertinence of such investigations to the study of how complex constituents in a natural language acquire their meaning from the constituents that constitute them, though he himself doubted that a satisfactory formal account of this property of natural language constituents could be worked out (Tarski 1935, 1944).

Natural language semantics, then, addresses two central questions: What values are to be associated with the basic constituents of a language? And, how do the values of simpler constituents contribute to the value of the complex constituents the simpler ones make up? The utility of model theory is to enlighten us on how to proceed with answering the latter question. If a satisfactory answer is obtained, then it will explain not only how changes in our understanding of complex constituents changes with changes in their constituents but also how it is that humans are able to understand completely novel complex constituents. After all, one's understanding of complex constituents cannot be accounted for by appeal to memorization of a language's constituents, as explicitly noted by Patañjali twenty-three hundred years ago,[3] any more than an appeal to memorization can explain how it is that humans knowing elementary arithmetic can understand previously unencountered numerals. Rather, one is able to understand novel complex constituents since they are combinations of simple ones that are not novel and that are antecedently understood.

MODEL THEORY

For the sake of illustration, let us consider a portion of classical predicate logic: monadic predicate logic without quantifiers. The formulae are obtained from two disjoint sets, the set of individual constants (*CN*) and the set of one-place predicates (PD_1). Atomic formulae (*AF*) are obtained by prefixing a one-place predicate to an individual constant. Thus, if *P* is a one-place predicate and *c* is an individual constant, then *Pc* is an atomic formula.

Definition 1: *Atomic Formulae of Monadic Predicate Logic*

$\alpha \in \mathrm{AF}$, the atomic formulae of Monadic Predicate Logic, iff $\alpha = \Pi c$, where $\Pi \in \mathrm{PD}_1$ and $c \in \mathrm{CN}$.

The set of formulae is the smallest set that includes the atomic formulae and those formulae obtained from prefixing a formula with a unary connective or enclosing a pair of formulae between parentheses and placing between the formulae a binary connective. There is but one unary connective $\neg$ and there are four binary connectives: $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$.

Definition 2: *The Formulae of Monadic Predicate Logic (MPL)*

FM, the set of formulae of monadic predicate logic, is defined as follows:

(1)	If AF $\subseteq$ FM;
(2.1)	if $\alpha \in$ F M, then $\neg \alpha \in$ FM;
(2.2.1)	if $\alpha, \beta \in$ F M, then $(\alpha \wedge \beta) \in$ FM;
(2.2.2)	if $\alpha, \beta \in$ F M, then $(\alpha \vee \beta) \in$ FM;
(2.2.3)	if $\alpha, \beta \in$ F M, then $(\alpha \rightarrow \beta) \in$ FM;
(2.2.4)	if $\alpha, \beta \in$ F M, then $(\alpha \leftrightarrow \beta) \in$ FM;
(3)	nothing else is in FM.

A model for this notation comprises a nonempty universe *U* and an interpretation function that assigns to each member of *CN* a member of *U* and that assigns to each member of PD_1 a subset of *U*. Values assigned to symbols will be called their denotations.

Definition 3: *Model for MPL*

Let CN and PD_1 be the constants and the monadic predicates, respectively. Then, $\langle U, i\rangle$ is a model for monadic predicate logic iff U is a nonempty set and i is an interpretation function of CN into U and of PD_1 into the subsets of U.

An atomic formula is said to be true in a model if and only if the individual assigned to the constant is a member of the set assigned to the predicate.

Thus, let *i* be the interpretation function of the model and let *Pc* be the formula, then *Pc* is true in the model iff $i(c) \in i(P)$. Composite formulae are true by dint of the usual definition.

Definition 4: *A Classical Valuation for the Formulae of MPL*

Let $\langle U, i\rangle$ be a model for MPL.

(1) ATOMIC FORMULAE:
Let Π be a member of PD1 and let c be a member CN. Then,

(1.1)	$v_i(\Pi c) = T$	iff	$i(c) \in i(\Pi)$.

(2) COMPOSITE FORMULAE:
Then, for each α and for each β in FM,

(2.1)	$vi\ (\neg\alpha) = T$	iff	$v_i\ (\alpha) = F$;
(2.2.1)	$vi\ (\alpha \wedge \beta) = T$	iff	$v_i\ (\alpha) = T$ and $v_i\ (\beta) = T$;
(2.2.2)	$v_i\ (\alpha \vee \beta) = T$	iff	either $v_i\ (\alpha) = T$ or $v_i\ (\beta) = T$;

(2.2.3) $v_i(\alpha \rightarrow \beta) = T$ iff either $v_i(\alpha) = F$ or $v_i(\beta) = T$;
(2.2.4) $v_i(\alpha \leftrightarrow \beta) = T$ iff $v_i(\alpha) = v_i(\beta)$.

SIMPLE COPULAR SANSKRIT CLAUSES

Of course, we are not interested in the semantics of monadic predicate logic, rather we are interested in the semantics of natural language, in particular, the semantics of classical Sanskrit. Now the syntax and semantics of a natural language is vastly more complicated than that of monadic predicate logic, even if we confine our attention to simple copular clauses. For the sake of simplicity, let us confine our attention to simple copular clauses consisting of a common noun (e.g., *kāka*, *gau*, *puruṣa*, etc.) and a proper noun (e.g., *Devadatta*, *Yajñadatta*, *Kṛṣṇa*, etc.). In such a case, the model theory developed earlier for monadic predicate logic without quantifiers can be adopted virtually without change to such elementary Sanskrit clauses. An atomic (simple) copular clause, as just stated, comprises a proper noun and a common noun (both in the nominative case). The formation of composite simple copular clauses is straightforward but involves complications regarding the placement of the connectors, which need not detain us here.

A model for this language comprises a nonempty universe and an interpretation function that assigns to each proper noun an individual in the universe and to each common noun a subset of *U*. An atomic simple copular Sanskrit clause is true if and only if the individual in the universe denoted by the proper noun (i.e., the individual assigned to the proper noun by the interpretation function) is a member of the set denoted by the common noun (i.e., the set assigned to the common noun by the interpretation function).

Consider the sentence:

(1) Devadattaḥ puruṣaḥ.
Devadatta is a man.

This sentence is true if and only if the thing denoted by the proper noun *Devadattaḥ* (*Devadatta*) is a member of the denotation of the common count word *puruṣaḥ* (man).

THE PROBLEM OF LEXICAL SEMANTICS

While model theory is enlightening with respect to the fundamental question of how the meanings of simpler constituents contribute to the meaning of the complex constituents the simpler ones make up, model theory tells us nothing about how values are assigned to the simplest constituents. Indeed, model theory is not concerned with this question, for it is interested in the properties of formulae as the interpretations of their minimal constituents are allowed to vary freely, holding the interpretation of the logical symbols constant.

Although, in natural language, it is clear that the assignment of meanings to words is arbitrary, nonetheless, once fixed, the assignment does not vary freely from occasion of use to occasion of use, for without these invariable meanings, communication would be impossible. But what imparts this invariability and how is what imparts it learned?

Let us confine our attention to proper nouns and common nouns. If we set aside the problem of the indeterminacy of demonstration, noted by Wittgenstein, which afflicts both the learning of the meaning of proper nouns and the learning of the meaning of common nouns, we notice another contrast between the learning of proper nouns and the learning of common nouns. Once one associates the individual denoted by the proper noun with the proper noun, nothing more need be said. In contrast, even after one associates an individual of which the common noun is true with the common noun, one still has to figure out how to associate another individual of which the common noun is true with the common noun. In other words, once one knows to whom or to what a proper noun applies, one need know nothing further to apply it again, since there is nothing else to which it applies; however, even after one knows that a common noun applies to a particular individual, one is no further ahead ipso facto with respect to the problem of knowing to what else it applies. The contrast resides not so much in the fact that a single individual is associated with a proper noun, while more than one individual is associated with a common noun; rather, it lies in the fact that what is associated with a proper noun is bounded, whereas what is associated with a common noun is not. Thus, when one learns what is associated with a proper noun, one learns that it applies to one individual and no other. Indeed, a proper noun cannot apply to anything else without losing its meaning. But when one learns what is associated with a common noun, even if one knows what all the things it applies to are, nothing in its meaning rules out its applying to something that comes into existence subsequently, provided the individual that comes into existence is of the right sort. (Of course, we are putting aside the problem of vagueness.)

How, then, does one manage to know to what a common noun applies and to what it does not? Model theory provides no help here. Indeed, the only answers to this problem are answers put forth long before the advent of model theory. One answer is to say that they are, in a sense, just like proper nouns: associated with each common noun is but one thing, a universal, by dint of which association, one knows, for any new individual, whether it applies or not. But, what is a universal and how does one grasp it? Many of those who have felt that these questions have no satisfactory answer have been driven to seek answers elsewhere.

APOHA MODEL THEORY

Buddhists maintained that *anya-apoha* (exclusion of what is different) provides an answer to the question of how one knows the meaning of a general expression, without having recourse to universals. The idea is that one can use two kinds of negation: *anya* (what is different) and *apoha* (exclusion). Unfortunately, these thinkers never specified what precisely these two negations are. An obvious and natural suggestion is the negations found in contemporary three-valued logic, internal negation and external negation. Let T, F, and N be the three values of three-valued logic. The two unary operators are defined as follows:

$$
\begin{array}{ll}
\text{I:} & \mathrm{T} \mapsto \mathrm{F} \\
 & \mathrm{F} \mapsto \mathrm{T} \\
 & \mathrm{N} \mapsto \mathrm{N} \\
\text{E:} & \mathrm{T} \mapsto \mathrm{F} \\
 & \mathrm{F} \mapsto \mathrm{T} \\
 & \mathrm{N} \mapsto \mathrm{T}
\end{array}
$$

These operations were devised to interpret two unary connectives in a three-valued propositional logic. As such, they have no role to play in the problem before us. Rather, what we require is their set-theoretic counterparts. To arrive at this, we must alter the model theory for monadic predicate logic from a classical model theory to one that I shall call apoha model theory. An apoha model for monadic predicate logic, like a standard model, comprises a nonempty universe *U* and an interpretation function that assigns to each constant an individual from the universe and that assigns to each monadic predicate a pair of disjoint subsets of the universe.

Definition 5: *Apoha Model for MPL*

Let CN and PD_1 be the constants and the monadic predicates. Then, $\langle U, i\rangle$ is a model for monadic predicate logic iff U is a nonempty set and i is an interpretation function from CN into U and from PD_1 into $\{ \langle X, Y\rangle: X, Y$ are disjoint subsets of $U\}$.
The idea is that with each predicate must be associated not only the set of things of which the predicate is true but also the set of things of which it is false. In other words, each predicate is associated with an ordered pair. Let us call the ordered pair the predicate's denotation, while we shall call the first set its positive denotation and the second its negative denotation.

There are features of natural language that seem to warrant this increase in the complexity of a general term's denotation. These are the cases of presuppositional failure and category mistakes. Consider the sentence "four is blue" by way of an illustration of a category mistake. "Blue" is a predicate that applies to physical objects that can have color. It is true of things that have color and are blue and it is false of things that have color and are not blue. Many think that it is neither true nor false of things that have no color. If one interprets natural language negation as internal negation and one assigns to the word "blue" a positive denotation of the set of colored blue things and a negative denotation of the set of colored nonblue things, then neither the sentence "four is blue" nor the sentence "four is not blue" is true.

Now, the set-theoretical counterparts of the two kinds of propositional negation given earlier can now be defined as follows:

I: $\quad (X, Y) \mapsto (Y, X)$
E: $\quad (X, Y) \mapsto (-X, X)$

(where - is complementation with respect to U). The first function simply swaps the positive and negative denotation associated with a monadic predicate, while the second function maps the positive denotation to its set-theoretic complement and maps the negative denotation to the set whose complement has become the positive denotation.

Now the question arises: how do these two kinds of negation permit universals to be dispensed with in lexical semantics in favor of just individuals? To answer this question, let us distinguish between an expression's denotation, which is dictated by our semantic theory, and an expression's extension, that of which an expression is true and can be ascertained empirically by asking a speaker of the expression's language whether or not the expression applies to some entity. In light of this distinction, we can rephrase our question as follows: how do these two kinds of negation permit universals to be dispensed with in lexical semantics of common nouns in favor of just individuals so that, given a common noun's denotation, one can arrive at its extension?

The answer of the non-*apoha-vādin* is that the denotation of a common noun is a universal and its extension is just the set of individuals in which the universal inheres. The answer of the *apoha-vādin* is that the denotation of a common noun is an individual (or what is set-theoretically equivalent, the singleton set containing that individual) and its extension is obtained by the application of the two negation operations.

As we shall now see, the answer of the *apoha-vādin* does not furnish a suitable extension for any common noun whose extension contains at least two individuals. There are exactly four possible combinations of the operations *I* and *E*: namely, *II*, *EE*, *IE*, and *EI*. Four propositions can be proved, one corresponding to each of the combinations. They are the following:

PROPOSITION 1:
$II(X, Y) = (X, Y)$,

PROPOSITION 2:
$EE(X, Y) = (X, -X)$,

PROPOSITION 3:
$IE(X, Y) = (X, -X)$,

PROPOSITION 4:
$EI(X, Y) = (-Y, Y)$

(where *X* ranges over the positive denotations, *Y* over negative denotation, and "-"is set theoretic complementation). As the first three propositions show for the combinations of *II*, *EE*, and *IE*, the positive denotation remains fixed under the application of the combined negations. Thus, under these forms of negation, the positive denotation and the positive extension are the same. Thus, for example, suppose the common noun *puruṣa* (man), whose extension is *M* and one of whose members, among others, is *m*, is assigned the denotation of *m* (that is, $X = \{m\}$), then the result of applying any of the first three combinations of negation only yields the set $\{m\}$, and not the set *M*, as its extension, which is, by hypothesis, different from *M*. In other words, given {*m*} as the positive denotation, one obtains {*m*} as the extension—not, of course, what is desired.

The only remaining possibility is the combination of *EI*. But here the fourth proposition shows us that only the negative denotation is relevant to the determination of positive and negative extension, the positive denotation is utterly irrelevant. Thus, to arrive at the extension of *M*, one must know to start with -*M*, the set of all nonmen. But this is just the problem of finding the denotation of a general term all over again. *Anya-apoha*, understood as a combination of internal and external negation, does not provide any semantic alternative to universals. Perhaps another way of spelling out *anya-apoha* will. But that remains to be seen.

Notes

1. I attributed this view to Mark Siderits (1991, 93–100), perhaps mistakenly, when I wrote my review (Gillon 1992) of his important and insightful book on the philosophy of language in classical Indian philosophy.
2. For details, see Gillon 2007.
3. MB (ed. Kielhorn) v. 1, 5–6; translated in Staal 1969, 501–502.