EDWARD L. KEENAN

FACING THE TRUTH: SOME ADVANTAGES OF DIRECT INTERPRETATION*

1. THE PROBLEM

Within any model theoretic framework for natural language semantics, we associate with each category C in the language a set of *logically possible denotations* defined in terms of the semantic primitives of the model. We shall refer to this set as the *type* for C (relative to the primitives) and denote it T_C. The problem we are concerned with here is how to specify the interpretations of syntactically simple expressions in a category C where, on the one hand, the expressions are not logical constants, and on the other hand, are not interpretable freely in T_C. Two examples below illustrate the problem.

First, consider that proper nouns (*John*, *Mary*, etc.) and expressions like *every man, no student, exactly two students*, etc. share a category (NP) in English. Proper nouns are not logical constants and cannot be freely interpreted in T_{NP}. If they could, then the argument in (1) below would not be valid, since *John* might denote the same as *no student* or *exactly two students*, and the argument is clearly not valid when *John* is everywhere replaced by any of those NP's.

(1) John is a vegetarian. All vegetarians are socialists. Therefore, John is a socialist.

How are we to guarantee then that proper noun denotations are to be restricted so that they can never denote the same as *no student*, etc.?

Second, consider (cf. Montague, 1973) that (2) below is logically true, and remains so when the adjective *male* is replaced by ones like *tall* and *skillful*.

(2) Every male thief is a [illegible]

But replacing *male* by AP's like *fake* and *apparent* yields sentences which are not L-true. None of the AP's are logical constants, and if all were freely interpreted in T_{AP}, (2) would be L-true for no choice of AP (Montague's option in fact). How then may we constrain the interpretation of AP's like *male, tall*, and *skillful* so that sentences like (2) are L-true for those choices?

Linguistics and Philosophy **6** (1983) 335–371. 0165–0157/83/0063–0335$03.70

Some examples of decreasing Dets are of course the overt negations of increasing ones, e.g. *not every, not one, not more than ten*, etc. and include as well ones like *at most ten, fewer than ten, no*, etc. They also include non-logical Dets like *no student's, fewer than ten students'*, etc. As can be seen below, they do in fact trigger negative polarity items in the VP:

$$\text{(12)} \qquad \left\{\begin{array}{l} \textit{not one} \text{ student} \\ \textit{at most ten} \text{ students} \\ \textit{no student's} \text{ mother} \end{array}\right\} \text{saw } \textit{any} \text{ deer in the forest}$$

However, if the Dets in (12) are replaced by increasing ones preserving *any* in the VP, the result is ungrammatical (try e.g. *one, more than ten, some student's*, etc.). Further it is generally the case that Dets which are neither increasing nor decreasing fail to trigger negative polarity items: e.g. *exactly three, between five and ten, all but three, more male than female*, etc.

Once again then by examination of the types we associate with subcategories, we have been able to find significant and non-obvious correlations between surface form and semantic structure. Let us now compare T_{Det} with the other sorts of types we have needed in a semantics for natural language.

6.2. *Similarities Between Dets and Other Categories*

Dets are different from all other categories so far considered in that they are semantically functions from an algebra into its power set. So there is no other category we know of whose denotations can be conservative.

Notwithstanding T_{Det} does show certain affinities with restricting adjectives. Thus, while T_{Det} is a ca algebra, as is $T_{\mathrm{AP}_{+,r}}$, neither are ca-free algebras, so Dets and the major AP's differ from Argument categories in the same way. Similarly Dets are basically not structure-preserving functions in the same sense in which AP's are not: only a few trivial elements in T_{Det} are homomorphisms (see KS for the characterization of just which elements in T_{Det} are homomorphisms). Third, and most obvious of course, Dets and restricting AP's are functions with the same domain. And in one non-obvious respect these functions behave similarly on their arguments. Namely, the possible values of a Det or AP at an argument p increase directly with the number of individuals which have p. For a restricting AP function f we have that $f(p) \leqslant p$, so there are 2^k possible values for such functions at p, where k is the number of individuals with p.

For Det functions, on the other hand, we may show that there are 2^{2^k}

possible values of such functions at p (as above). In fact the value of a Det function f at p is determined by the properties $q \leqslant p$ which are in $f(p)$. So we may think of elements of T_{Det} as given by functions which associate with each property p a set of properties q, all $\leqslant p$. In this loose sense then we may think of a Det semantically as a kind of higher order restricting AP.

6.3. *The Major Difference Between Dets and Other Categories: Its Logical Nature*

Almost certainly the most striking difference between Det and other categories is the high percentage of logical constants among the Dets. When we try to think of expressions which are (intuitively still) logical constants in the categories NP, CN, n-place Predicate, or AP, only a handful of cases come to mind. For example, in an extensional logic *exist* is a logical constant among the P_1's, and *be* (the identity function on individuals) is one among the P_2's. Perhaps *individual* (or *existent*, etc.) is among the CN's, etc. But among the Dets most of the expressions we think of initially are logical in nature: *every, some, at east k, all but three*, etc. We would like to know whether this apparent listribution is an accident or not. If we think harder, will we find many nore logical expressions among the P_1's? A little trial and error suggests hat this is no in the following sense. Any logical element we find among he P_1's (they may be syntactically complex) appears to have the same lenotation as *exist* (e.g. *sings or doesn't sing*, etc.) or its complement *not xist* (*is singing and isn't singing*, etc.). Are these apparent facts in any vay predictable given the sets of possible denotations we associate with hese categories? We show below that basically the answer is yes.

To show this, we must first have some clear idea what logical onstants are, and this is a matter of some discussion in the philosophi-al and logical literature. However, one point on which all theorists vould agree is the following:

HESIS. If an expression e in L is a logical constant then all inter-retations of the language relative to a given universe agree on e, i.e. ssign it the same denotation.

he logical constants in any category then will be among the expressions hich are constantly interpreted in the sense of the Thesis above. mong the Dets, for example *every, at least two, between five and ten,* c., all meet the condition though ones like *John's* do not, since given a universe E, many functions in T_{Det} may interpret *John's* depending on which individual interprets *John* and depending on which objects he 'has'. Among the P_1's *exist* and *not exist* are clearly constantly interpreted but many others, e.g. *sleep*, *swim*, etc. are not, since given an E, and thus the individuals of the universe, it is arbitrary which ones are *sleeping*, *swimming*, etc.

We may now, for any category C, look at the elements of the type for C which are the denotations of the constantly interpreted expressions in C. And we may reasonably ask just what proportion of T_C is occupied by the logical elements. But is there any way of identifying these elements except on a case by case basis? Is there not some property they have in common regardless of what category C is chosen? The answer is yes. Consider first the special case of Dets.

Intuitively a 'logical' element of T_{Det} should not be able to discriminate among properties which have the same boolean structure. To make that notion more precise, consider than an isomorphism on an algebra is by definition a function which preserves all of the structures of the algebra. We may think then of two elements p, q in T_{CN} as having the same boolean structure, if there is an isomorphism from T_{CN} onto itself which maps one to the other. (An isomorphism from an algebra onto itself is standardly called an *automorphism*.) So p and q have the same boolean structure iff there is an automorphism i on T_{CN} such that $i(p) = q$. And to say that an element f in T_{Det} cannot discriminate among properties with the same boolean structure is to say that in deciding whether to put p in the set of properties it associates with q, f cannot tell whether it is looking at the pair (p, q) or an automorphic image of such a pair $(i(p), i(q))$. Thus,

DEFINITION. $f \in T_{\mathrm{Det}}$ is *automorphism invariant* (AI) iff for all elements $p, q \in T_{\mathrm{CN}}$ and all automorphisms i on T_{CN}, $p \in f(q)$ iff $i(p) \in f(i(q))$.

It is not hard to show that the denotations of *every*, *at least k*, etc. lie in $T_{\mathrm{Det}_{\mathrm{AI}}}$, the set of automorphism invariant elements of T_{Det}. And it is easy to see that this fails for Dets like *John's*. To see this we note without proof that in a finite universe, two properties are identifiable under an automorphism iff the number of individuals with one is identical to the number with the other. Now consider a finite universe in which there are exactly two cars and two boats, so there is an automorphism i such that $i(\mathrm{car}) = \mathrm{boat}$; John has exactly one car and no boats. Then *John's boat* denotes the empty set of properties and *John's car* the individual with the car property which John has. So car $\in$ John's car but i(car) = boat $\notin$ John's (i(car)) = John's (boat), therefore John's is not AI.

Finally we note the following theorem (to be generalized) which holds for all forms of model theoretic semantics for natural language:

THEOREM 12′. For all expressions $d \in$ Det, if d is constantly interpreted as per the Thesis above, then for any universe E and any interpretation m of the language relative to E, $m(d)$ is automorphism invariant.

Thus the set of AI functions in T_{Det} provides a subset of T_{Det} in which the logical constants of Det always take their denotations.

One may query just how big this set is. Again the boolean nature of T_{Det} allows a rather easy answer to this question. First we may easily show that T_{DetAI} is complete and atomic subalgebra of T_{Det} and thus has cardinality 2^k, where k is the cardinality of the set of atoms. In a finite universe with n individuals the number of these atoms may be computed to be $(n+1)(n+2)/2$. Thus,

THEOREM 13. In a finite universe with n individuals there are $2^{(n+1)(n+2)/2}$ automorphism invariant functions in T_{Det}.

The absolute figure here is of little interest but two properties it has should be noted. First, it is a fairly small portion of T_{Det} itself. The reader may compute that in a world of two individuals there are, as noted, 512 elements in T_{Det}, only $2^6 = 64$ of which are AI.

The second point to notice is that the cardinality of $T_{\mathrm{Det}_{\mathrm{AI}}}$ increases with the number of individuals. In this respect we will see that Dets differ from many other categories where typically the number of AI elements in the type has an upper bound which is constant and independent of how many individuals there are. For this claim to make sense, of course, we must define what we mean by a P_1 denotation being AI, etc., as above we gave the definition only for the case of Dets. We sketch the more general definition below. It is in some respects simpler than the special case given above, though it may appear at first to be less intuitive (again see KS for details).

Given a universe E, let us consider all the automorphisms i from E^* ($= T_{\mathrm{CN}}$) onto E^*. Each of those automorphisms naturally extends to a function i^* which is an automorphism on every type. That is, for every category C, i^* restricted to T_C is an automorphism from T_C into T_C. For every C, we shall refer to these automorphisms on T_C as the *basic automorphisms* on T_C. We may then define:

DEFINITION. For all categories C and all $x \in T_C$, x is *automorphism invariant* (AI) iff for every basic automorphism i^*, $i^*(x) = x$.

So the AI elements in any type then are just the elements which a basic automorphism must map onto themselves. We show (easily) in KS that this definition coincides with the earlier one on T_{Det}. Further we now have the appropriately general version of Theorem 12:

THEOREM 12. If e in L is constantly interpreted then for all universes E and all interpretations m of L, $m(e)$ is automorphism invariant.

So it now makes sense to talk of the AI elements in the types for the AP's, n-place predicates, etc. And when we check what proportion of these types is taken by AI elements, we obtain some interesting results. For example,

THEOREM 14.

(a) No element in the type for proper nouns is automorphism invariant.
(b) Exactly two elements in the type for the absolute AP's are AI.
(c) Exactly two elements in the types for the P_1's and the CN's are AI.
(d) At most four elements in the type for the P_2's are AI.

Universal 6 below is basically a corollary of Theorem 14.

(a) No language has logically constant proper nouns.
(b) The number of syntactically simple logical constants among CN's, the n-place predicates, and absolute AP's is small and bounded

Universal 6a is reasonable since by Theorem 14a no element in T_{PN} is AI, whence by Theorem 12 Universal 6a follows. Theorems 14b,c tell us that there are at most two non-synonymous logical constants among the absolute AP's and the P_1's. So the number of syntactically simple such constants is limited by whatever constraints limit the number of lexical synonyms in a language and can be expected to be small. Analogous claims hold regarding Theorem 14d.

Matters are quite otherwise for Det, however. We saw that in a world of only two individuals there are 64 AI elements in T_{Det}, so even a small world can discriminate among many logical constants among the Dets. And that number increases steadily as the size of the universe increases. For universes of different sizes we can find distinct Dets which have the same extension in the smaller but distinct extensions in the larger one. To take an easy example, in a world of just seven individuals the Dets *at least eight* and *at least nine* both denote the zero element of T_{Det}, the function sending each property p to the empty set. No property q can be in at least eight (nine) p's since there are at most seven p's. But if we enrich the world by one individual, then *at least eight* will not denote the zero Det, though *at least nine* still will, so they have distinct denotations in a world of eight individuals.

We do then have some reason to expect large numbers, even an infinite number, of logical constants among the Dets, since for any number we pick we can always find an E such that T_{Det} provides distinct denotations for that many distinct constants. Thus the presence of many logical constants among the Dets is a semantic possibility in a way in which it isn't for predicates, absolute AP's, common nouns, and proper nouns.

The semantic notion of automorphism invariance turns out to be interesting in the analysis of English Dets in yet another way. Namely, it allows us, together with other notions, to represent on the one hand the sense in which certain rather simple Dets like *several* and *a few* are 'logical' compared with ones like *John's*, and yet on the other hand they are less 'determinate' in meaning than ones like *every*, *at least three*, etc.

To take the latter point first, we clearly want to put *several* in a subcategory of Det which is +AI and +increasing, thus guaranteeing that it will always be interpreted by an increasing automorphism invariant function. We doubtless want to put other conditions on its interpretation as well, e.g. that it is booleanly $\leqslant$ *at least three*. But these conditions do not uniquely determine the interpretation of *several* in general. That is, for sufficiently large E there will be many functions from E^* into E^{**} satisfying these conditions. So *several* is not a logical constant. Moreover, its interpretation is not determined by the interpretations of other expressions in the language, since *several* is syntactically simple. So in this sense *several* is vague. Within admittedly quite narrow bounds, there is more than one function it might denote, and we have no way of telling which.

On the other hand, *several* is 'logical' compared with e.g. *John's*, since it shares the characteristic property of logical constants – namely it always denotes an AI element of T_{Det}. Note that it is only the AI requirement on *several* which distinguishes it from expressions like *John's*. For example, *John's* itself is increasing. *John's three* is booleanly $\leq$ at least three, etc.

Thus we can see that explicitly defining the property which the logical constants among the Dets have in common – something we are not motivated to do on a translation approach – has enabled us to represent, at least partially, the meaning differences between certain Dets, including ones which are not themselves logical constants. (We may say that Dets like *several* are logical, but not constant).

7. Conclusion, and a Last Universal

For the categories we have considered it has generally been the case that expressions in that category took their denotations in a set which possessed a boolean structure, enabling us to interpret them as conjunctions, disjunctions, and negations of expressions in the category. In addition, this boolean structure was used in a variety of more specific ways in defining particular classes of types, e.g. in defining homomorphisms, restricting functions, conservative functions, ca-free generators, etc. The fact that the types for basically all productive categories in English have a boolean structure strongly suggests that the boolean operations, the meanings of *and*, *or*, and *not* if you like, are not semantic operations specific to any particular category. E.g. the meaning of *and*, etc. is not that of a function which only applies to sentence meanings, as one might infer by default from the semantics given for standard first order logic. And what this suggests is that the boolean operations, rather than representing properties of objects in the world (denotations), represent properties of the mind, ways we conceive of objects in the world. This assumption allows us to account not only for the syntactic ubiquity of *and*, *or*, and *not* as illustrated in this paper; it also allows us to account for the naturalness with which we interpret apparent conjunctions, disjunctions, etc. of expressions which we would not normally assign a category to and interpret. Thus we easily interpret such cumbersome expressions as *John bought and Mary cooked the turkey*, even though we do not normally assign an interpretation to expressions like John bought. Let us conclude this paper then with a 'Universal' which is not specifically linguistic, though linguistically instantiated by the boolean generalizations about human language we have made in this paper.

UNIVERSAL 7. The boolean operations, expressed by *and*, *or*, and *not* in English, are operations of the mind.

COROLLARY. The truths of boolean algebra are 'Laws of Thought'.

The 'corollary' above is intended to express the indebtedness of this work to the original (in both senses) work of George Boole (1854).

UCLA

Notes

* Research for this paper was supported by NSF grant BNS 79-14141.

[1] The smallest subset of a set satisfying certain conditions is by definition the intersection of all the subsets which satisfy the conditions.

References

Barwise, J. and R. Cooper: 1981, 'Generalized Quantifiers and Natural Languages', *Stanford Working Papers in Semantics*, Vol. 1, pp. 1–80.

Ben-Chorin, S.: 1982, 'Sentence Meanings as a Formal Basis for Intensional Logic', in *Proceedings of the first West Coast Conference on Formal Linguistics*, Stanford.

Boole, G.: 1854, *An Investigation into the Laws of Thought*, Cambridge.

Chomsky, N.: 1957, *Syntactic Structures*, Mouton, The Hague.

Fauconnier, G.: 1979, 'Implication Reversal in Natural Language', in *Formal Semantics and Pragmatics for Natural Language*, F. Guenthner and S. J. Schmidt (eds.), Reidel, Dordrecht.

Keenan, E. L.: 1981, 'A Boolean Approach to Semantics', in *Formal Methods in the Study of Language*, J. A. G. Groenendijk *et al.* (eds.), Mathematisch Centrum, Amsterdam.

Keenan, E. L.: 1982, 'Boolean Algebra for Linguists', UCLA Working Papers in Semantics, Vol. 1, Dept. of Linguistics, UCLA.

Keenan, E. L. and L. Faltz (to appear), *Logical Types for Natural Language* (completely revised edition: Reidel, *Dordrecht*.

Keenan, E. L. and L. Faltz: 1978, *Logical Types for Natural Language*, UCLA Occasional Papers in Linguistics, No. 3, Dept. of Linguistics, UCLA.

Keenan, E. L. and J. Stavi: (to appear), 'A Semantic Characterization of Natural Language Determiners', in *Linguistics and Philosophy*.

Ladusaw, W.: 1980, *Polarity Sensitivity as Inherent Scope Relations*, Garland Press, New York (also in IULC).

Montague, R.: 1973, 'The Proper Treatment of Quantification in Ordinary English', in J. Hintikka *et al.* (eds.), *Approaches to Natural Language*, Reidel, Dordrecht.

Partee, B. and R. Mats: (to appear), 'Generalized Conjunction and Type Ambiguity', in *Meaning, Use and Interpretation*, A. von Stechow *et al.* (eds.), de Gruyter.

Strawson, P. F.: 1959, *Individuals*: *An Essay in descriptive Metaphysics*, Methuen, London.