

TWO THEORIES FOR COMPUTING THE LOGICAL FORM OF MASS EXPRESSIONS

Francis Jeffrey Pelletier
Lenhart K. Schubert

Dept. Computing Science
University of Alberta
Edmonton, Alberta T6G 2E1
Canada

ABSTRACT

There are various difficulties in accommodating the traditional mass/count distinction into a grammar for English which has a goal the production of "logical form" semantic translations of the initial English sentences. The present paper surveys some of these difficulties. One puzzle is whether the distinction is a syntactic one or a semantic one, i.e., whether it is a well-formedness constraint or whether it is a description of the semantic translations produced. Another puzzle is whether it should be applied to simple words (as they occur in the lexicon) or whether it should apply only to longer units (such as entire NPs). Of the wide variety of possible theories, only two seem to produce the required results (having to do with plausible inferences and intuitively satisfying semantic representations). These two theories are developed and compared.

According to Montague (Thomason 1974), Gazdar (Gazdar *et al* 1984) and a rapidly growing number of linguists, philosophers, and AI researchers, the logical form underlying sentences of a natural language are systematically--and simply--determined by the syntactic form of those sentences. This view is in contrast with a tacit assumption often made in AI, that computation of logical translations requires throngs of more or less arbitrary rules operating upon syntactic forms.¹

The following are a few grammar rules in approximately the style of Gazdar's Generalized Phrase Structure Grammar (GPSG). They differ from Gazdar's primarily in that they are designed to produce more or less "conventional" logical translations, rather than the intensional ones of Montague and Gazdar (for details see Schubert & Pelletier 1982). Each rule consists of a rule number, a phrase structure rule, and a semantic (logical translation) rule.

1. $S \rightarrow NP VP, VP(NP')$
2. $VP \rightarrow [V +be] PRED, PRED'$
3. $PRED \rightarrow N, N' \quad N_s = \{water, wine, food, furniture, \dots\}$

Parsing and translating in accordance with such rules is a fairly straightforward matter. Since the syntactic rules are context free, standard context-free parsing methods can be employed, except that allowance must be made for the propagation of features, with due regard for concord.

¹The work reported herein was partially supported by NSERC grants A5525 (FJP) and A8818 (LKS). We also wish to thank Matthew Dryer, David Justice, Bernard Linsky, and other members of the Univ. Alberta Logical Grammar Study Group for discussions on these topics.

Applying the rules of translation is even simpler. In essence, all that is needed is a mechanism for arranging logical expressions into larger expressions in conformity with the semantic rules. (For examples of parsers see Thompson 1981, Schubert & Pelletier 1982, Gawron *et al* 1982, Rosenschein & Shieber 1982).

The topic of mass terms and predicates has a substantial literature within both linguistics and philosophical logic, with much of the recent research deriving inspiration from Montague Grammar (e.g., see Pelletier 1979, ter Meulen 1980, Bunt 1981, Chierchia 1982). There are three views on the mass/count distinction, namely that the distinction is (a) syntactic, (b) semantic, and (c) pragmatic. Orthogonal to these views we have the further possibilities (i) that the mass/count distinction is lexical, and (ii) that it is determined by the context in which the expression occurs. We shall present arguments in the full paper to eliminate position (c), leaving us with four possible kinds of theories. (1) a syntactic expression (lexical) approach, (2) a syntactic occurrence approach, (3) a semantic expression approach, and (4) a semantic occurrence approach. This raises the question of what is the difference between syntactic approaches generally and semantic approaches generally. A syntactic approach treats +mass and +count as syntactic classifications or features, that is as features to be used by the syntactic rules in determining whether some longer stretch of words is well-formed. Central to the semantic approach is the claim that +count and +mass are *not* syntactic features or categories, but rather are a description of the semantic representation of the expression. In this approach, no syntactic rules refer to +count or +mass (since these are not syntactic objects). Rather, in sentences like *Mary put apple in the salad* vs. *Mary put an apple in the salad*, the semantic approaches allow us to say that it was a mass or count semantic representation of *apple* only after inspecting the kind of thing that *apple* is true of in the sentences.

There are reasons for rejecting options (2) and (3), thus leaving us with only a syntactic expression approach and a semantic occurrence approach. (The reasons are given in Pelletier & Schubert 1985). These are the two theories of mass expressions that are to be discussed in the paper. They seem to us to be the most plausible candidates for an adequate theory of the logical form of sentences involving mass expressions. The fragment of English that the two theories of mass expressions are concerned with is roughly those sentences with a copular verb and either a mass or count expression as predicate, and whose subjects are either bare noun phrases or quantified noun phrases. A sentence is a noun phrase and a verb phrase. A verb phrase is a copula followed by a

PRED.

which in turn is either a bare noun (as in *Claret is wine* or *This puddle is man* --the latter said after an application of the universal grinder)² or an *a* followed by a noun (as in *John is a man* or *Claret is a wine*) or is an entire noun phrase (as in *John is the man most likely to succeed* or *Claret is my favourite red wine*). A noun phrase is either a bare noun (as in *Claret is a dry red wine* or *Dogs are barking outside*) or else is a quantified term (as in *All men are mortal* or *Sm red wine is tasty* --we include as determiners *this, all, some, sm, much, little, each, every*, and the numeral quantifiers). Nouns may themselves be either an adjective-phrase noun combination, or just a noun. We consider here two cases of adjective modification: intersective and non-intersective. For the former we have in mind such adjectives as *red*, while for the latter we think of such adjectives as *fake*.

The rules which give alternatives, such as 3p vs. 3s, are those rules which are different for the two theories of mass terms. The p-rules are for the semantic occurrence approach while the s-rules are for the syntactic expression approach. The ontological underpinnings of these theories are that "reality" contains two sorts of items: (1) "ordinary objects" such as rings, sofas, puddles (and including here what many theorists have called "quantities of matter"), (2) "kinds", that is, "varieties", "substances", etc. We have in mind here such items as wine, claret, red wine, and the like, and also servings of such items. We wish to make no special metaphysical claims about the relationships that might hold between "ordinary objects" and "kinds"--instead we content ourselves with describing how such an ontology leads to a simple and natural description of various of the facts concerning mass (and possibly plural) expressions. Linguistically, that is semantically, we take there to be three distinct types of predicates: (a) those which apply only to "kinds", e.g., *is a substance, is scarce, is a kind of wine, is abundant*, (b) those which apply only to "objects", e.g., *is a quantity of gold, is a puddle*, and (c) those which can apply to both "kinds" and "objects". In this last group we have in mind mass predicates such as *is wine, is furniture, is food, and is computer software*.

Both of these theories take it that *is wine* is true of the (abstract) kind claret in addition to an individual quantity such as the contents of this glass. Moreover, they take *is wine* to be true of an object such as a drop or puddle of wine, occupying the same region as some quantity of wine. (*This ring is gold* or *This hamburger is food* are clearer examples of the application of mass predicates to objects.) Generally speaking, the theories view the kinds of M as forming an upper semilattice of kinds with M at the top. This is a "formal" semilattice in that the union of any two elements of it is a member of the semilattice, and we view *is wine* as being true of any of these formal kinds. So a sentence like *Cheap wine is wine* will be true, since *cheap wine* names an element of the semilattice. Predicates like *is a wine* are true of conventionally recognized kinds (*Claret is a wine* is true) but not of every "formal" kind since, e.g., *Cheap wine is*

² The universal grinder (Pelletier 1975) takes objects corresponding to any count noun, grinds them up and spews the result from the other end. Put a table into it and after a few minutes there is sm table on the floor. (We regularly represent the unstressed *some* by *sm*.)

a wine is not true. (*Sauterne mixed with claret is a wine* is also not true, showing that *is a wine* is not true of unions of elements of the semilattice). These predicates are not only true of the conventional kinds but also of conventional servings such as the bottle of wine on the table or the 250ml in this glass. Note that these can again be abstract entities: but rather than potentially being abstract conventional kinds of wine, they can be abstract conventional kinds of servings of wine. Finally such predicates are true of individual quantities--as when we say we have ordered four wines, all of the same kind and size. When a bare mass noun phrase (or indeed other bare noun phrases, although we shall not dwell on them here) is used as a subject (or object, but again we shall not consider that here), it is taken to name the kind. So in *Cheap wine is wine*, the subject *cheap wine* names a kind; and since the sentence is true it must name a "formal kind" so that *is wine* can be predicated of it. But since *Cheap wine is a wine* is not true, the formal kind cannot be a conventionally recognized kind (nor, for that matter, a conventional serving nor an individual quantity). Both theories hold that mass CN's should be translated into the semantics as predicates. Strictly this is not required: for, all we have given direct evidence for is that mass VP's be translated as predicates with a mixed object/kind extension. It could be the case that mass CN's are quite different, yet in the formation of a mass VP the entire VP gets assigned a mixed, predicate denotation. Still, it would be simple, and in keeping with much philosophical and linguistic analysis, to assume coincidence of CN and "is CN" denotations (at least when tense is ignored, as here).

With just this much of the theory sketched, we can overcome various of the difficulties that plagued other theories. For example, it is most unclear that any other theory can adequately translate sentences like

Tap water is water
This puddle is water
Consider also sentences like

All wine is wine
wherein the subject *all wine* seems to quantify over both kinds of wine and quantities of wine, entailing both *White wine is wine* and *The litre of wine in this bottle is wine*, for example. It seems to us that no other theory allows this comprehensiveness. An even clearer example of such comprehensive denotation is (a), from which both of (b) and (c) follow, given that rice is edible and this sandwich is edible. (Note also the comprehensive denotation of *edible*). No other theory we know of can account for the validity of these two arguments.

- a. Everything edible is food
- b. Rice is food
- c. This sandwich is food

Both of these theories will want to be able, in the semantics, to form predicates which are true of kinds, or of servings, or of individuals, given a predicate which has comprehensive extension. So, for example, from the predicate *water'* which is assumed to be true of quantities, servings, and kinds, we shall want to be able to form (*k water'*) which is true of conventional kinds of water, to form (*p water'*) which is true of conventional portions (and kinds of portions) of water, and to form (*q water'*)

which is true of quantities of water. Conversely, if we have a predicate which is true of individuals and kinds, we shall want to form a predicate true of all the entities that mass predicates are true of--quantities of stuff, kinds of stuff, and objects coincident with quantities of stuff. For example, if *man'* is a predicate true of objects and kinds, then (*s man'*) is the mass predicate formed therefrom. Also, we shall want to be able to form the name of a kind from a predicate: (μ *water'*) is the name of the kind water and (μ (*cheap'(wine')*)) is the name of the kind cheap wine.

The rules for the relevant portion of our two theories are (λ is our symbol for lambda abstraction):

1. $S \rightarrow NP VP, VP(NP')$
2. $VP \rightarrow [V +be] PRED, PRED'$
- 3p. $PRED \rightarrow N, N'$
- 3s. $PRED \rightarrow [N +MASS], N'$
- 4p. $PRED \rightarrow [DET +a] N, (\lambda x)[(k N')(x) \vee (p N')(x)]$
- 4s. $PRED \rightarrow [DET +a] [N +COUNT], N'$
5. $PRED \rightarrow NP, (\lambda x)(x=NP')$
6. $PRED \rightarrow ADJP, ADJP'$
- 7p. $NP \rightarrow N, (\mu N')$
- 7s. $NP \rightarrow [N +MASS], (\mu N')$
8. $NP \rightarrow DET N, DET(N')$
9. $[N +ADJP] \rightarrow [ADJP +INTERSECT] N, (\lambda x)[ADJP'(x) \& N'(x)]$
10. $[N +ADJP] \rightarrow [ADJP -INTERSECT] N, ADJP(N')$

The S-theory distinguishes in the lexicon mass from count nouns. And it has what might be called "lexical extension" rules to give us the "stretched" meaning of nouns that we have earlier talked about. For example, it has

$[N +COUNT] \rightarrow sofa, man, substance, \dots$
 $[N +MASS] \rightarrow wine, water, \dots$

$[N +COUNT] \rightarrow [N +MASS], (k N')$
 $[N +COUNT] \rightarrow [N +MASS], (p N')$
 $[N +MASS] \rightarrow [N +COUNT], (s N')$

Now, both of these theories can give the correct semantic representation to a wide range of sentences involving mass terms, given certain meaning postulates. (The two theories do it slightly differently, as might be expected since they have somewhat different semantic understandings of the lexical nouns. For example, the s-theory takes *man* to be true of individual men and of kinds of men, while the p-theory takes it also to be true of the stuff of which men are made. In the p-theory, when a sentence uses *a* --as in *a man* --then the semantic operators convert this "basic" meaning into one that is true of individual men and of kinds of men. The s-theory rather has a lexical-extension rule which will convert the lexical count noun *man* into one which is a mass noun and is true of the stuff of which men are made. They will also take a different tack on what quantified terms designate, although that has been hidden in rule 8 above by assigning the same logical form to both theories. Nonetheless, the meaning postulates of the two theories will differ for these.) In addition to the sorts of examples stated above, both these theories can generate and give the correct logical form to such sentences as

Wine is wine (two readings, both analytic)

Wine is a wine (false)
 All wine is wine (analytic)
 Claret is a wine (true)
 Cheap wine is a wine (false)
 *All wine is a wine (semantically anomalous)
 Water is dripping from the faucet (entails: sm water is dripping from the faucet)
 Water is a liquid (entails: water is liquid)
 Both theories make the following six inferences valid

1. Claret is a wine, wine is a liquid, so claret is a liquid
2. Claret is a wine, wine is a liquid, so claret is liquid
3. Claret is a wine, wine is liquid, so claret is a liquid
4. Claret is a wine, wine is liquid, so claret is liquid
5. Claret is wine, wine is a liquid, so claret is liquid
6. Claret is wine, wine is liquid, so claret is liquid
 And they both make these two inferences invalid

7. Claret is wine, wine is a liquid, so claret is a liquid

8. Claret is wine, wine is liquid, so claret is a liquid
 We know of no other theories which can do all these things. Yet the two theories are radically different: one has a mass/count distinction in the syntax and the other doesn't, and they have different extensions assigned to the lexical items. So the question naturally arises--which is better? What can be said against the two theories? There is not space in a paper of this size to go into this in detail, so we shall content ourselves with just hurling the main charge that each one directs against the other.

Briefly, the p-theory charges the s-theory with pretending to use syntactic features +mass and +count but allowing them to do no syntactic work. For every sentence which has a mass term in a given location, there is another sentence which has a count term in that position. No constructions are ruled out; the only use of the +mass/+count features is in directing the semantic translation process. And that suggests that the features should all along have been semantic. The s-theory charges the p-theory with being unable to give coherent meaning postulates because of its commitment to a comprehensive extension to the lexical terms. For example, suppose one wanted to give as a meaning (or factual) postulate that *A lamb has fur*. The s-theory can do this without difficulty: *lamb'* is true of individual lambs and the meaning postulate says of each of them that they have fur. But the p-theory cannot easily do this: *lamb'* is true of stuff, so the predicate must be converted to one which is true of individuals. But there is no provision in the p-theory for doing this--the closest that it could come is with a predicate that is true of both conventional kinds and "conventional portions" (i.e., ordinary lambs).

Given the above rules (augmented with additional features such as number and person agreement features in rule 1) we are able to extend the capabilities of our parsers (Schubert & Pelletier 1982) so that they deliver logical form translations of sentences involving mass expressions. These translations have the desired semantic properties and, with an extension of the inference mechanisms to allow for predicate modification and μ -abstraction, allow the above valid arguments to be duplicated. So, which theory is to be preferred? That is a topic for further research. The time for studies of mass

expressions with only casual reference to the syntax and semantics of language is past. Only systematic attempts to account for large classes of mass expressions within formal syntactic-semantic-pragmatic frameworks can hope to resolve the remaining issues.

WORKS CITED

- Bunt, H.C. (1981) *The Formal Semantics of Mass Terms* Dissertation, University of Amsterdam.
- Chierchia, G. (1982a) "Bare Plurals, Mass Nouns and Nominalization" in D. Flickinger, M. Macken & N. Wiegand (eds) *Proceedings of the First West Coast Conference on Formal Linguistics* 243-255.
- Gawron, J., J. King, J. Lamping, E. Loebner, A. Paulson, G. Pullum, I. Sag, & T. Wasow (1982) "The GPSG Linguistics System" *Proc. 20th Annual Meeting of the Association for Computational Linguistics* 74-81.
- Gazdar, G., E. Klein, G. Pullum, I. Sag (1984) *English Syntax* (forthcoming).
- Pelletier, F.J. (1975) "Non-Singular Reference: Some Preliminaries" *Philosophia* 5. Reprinted in Pelletier (1979), 1-14. Page references to the reprint.
- Pelletier, F.J. (ed.) (1979) *Mass Terms: Some Philosophical Problems* (Reidel: Dordrecht).
- Pelletier, F.J. & L.K. Schubert (1985) "Mass Expressions" to appear in D. Gabbay & F. Guentner *Handbook of Philosophical Logic, Vol. 4* (Reidel: Dordrecht).
- Rosenschein, S. & S. Shieber (1982) "Translating English into Logical Form" *Proc. 20th Annual Meeting of the Association for Computational Linguistics*.
- Schubert, L.K. & F.J. Pelletier (1982) "From English to Logic: Context-Free Computation of 'Conventional' Logical Translation" *American Journal of Computational Linguistics* 8, 26-44.
- ter Meulen, A. (1980) *Substances, Quantities and Individuals*. Ph.D. Dissertation, Stanford University. Available through Indiana University Linguistics Club.
- Thomason, R. (1974) *Formal Philosophy: Writings of Richard Montague*. (Yale UP: New Haven).
- Thompson, H. (1981) "Chart Parsing and Rule Schemata in PSG" *Proc. 19th Annual Meeting of the Association for Computational Linguistics* 167-172.