File Change Semantics and the Familiarity Theory of Definiteness*

Irene Heim, Stanford

1. Introduction

What is the difference in meaning between definite noun phrases and indefinite ones? Traditional grammarians, in particular Christophersen and Jespersen, worked on this question and came up with an answer that nowadays finds little favor with semanticists trained in twentieth century logic. It amounts to the following, in a nutshell:

(1) A definite is used to refer to something that is already *familiar* at the current stage of the conversation. An indefinite is used to introduce a ***new*** referent.

This has been labeled the "familiarity theory of definiteness"[1].

When confronted with (1), the logically minded semanticist will notice immediately that it presumes something very objectionable: that definites and indefinites are referring expressions. For only if there is a referent at all can there be any question of its familiarity or noveltly. Advocates of (1) cannot happily admit that there are non-referring uses of definites or indefinites (or both), because that would be tantamount to admitting that their theory leaves the definite-indefinite-contrast in a significant subset of NP uses unaccounted for.

But the existence of nonreferential uses of definite and indefinite NPs can hardly be denied, and I will take it for granted without repeating the familiar arguments[2]. Just think of examples like (2) and (3).

(2) Every cat ate its food.

(3) John didn't see a cat.

(2) has a reading where "its", a personal pronoun, i.e. a type of definite NP, functions as a so-called "bound variable pronoun" and doesn't refer to any particular cat. Under the preferred reading of (3), with negation taking widest scope, the indefinite "a cat" fails to refer.

---

* The ideas contained in this article are elaborated more fully in my Ph.D. thesis (Heim 1982). All the people whose help I acknowledge there should also be mentioned here, in particular Angelika Kratzer and my thesis advisor Barbara Partee.

[1] The label is due to Hawkins (1978).

[2] See in particular Russell (1919, Ch. 16), Quine (1960), Kaplan (1972), and Geach (1962).

So the cards appear to be stacked against the familiarity theory of definiteness. Nevertheless, I will try to revive and defend it, or a theory very much like it. The version I will defend is just different enough from (1) to avoid the problematic presumption of referentiality. Otherwise it agrees with (1) – and accordingly deviates from standard assumptions in logical semantics – in fundamental respects: It involves familiarity and novelty as a central pair of notions, and it takes neither definites nor indefinites to be quantifiers.

What is the point of rehabilitating a problem-ridden traditional approach when much more sophisticated alternatives have become available through the work of logical semanticists from Russell to the present? – I would like to argue that a familiarity theory of definiteness, if construed along the lines of this article, enables us to make better predictions than competing theories about the behavior of definites and indefinites in natural languages, in particular about their participation in anaphora relationships. I return to this point in section 7 below, but first I must lay out the theory I am proposing.

2. Karttunen's "Discourse Referents"

Mine is not the first attempt to rehabilitate the familiarity theory of definiteness by dissociating it from the problematic presumption that definites and indefinites are referring expressions. In the late 1960s, Karttunen wrote some papers[3] directed at the same goal. In order to avoid untenable claims about reference, Karttunen reformulates the familiarity theory by using a new notion, that of "discourse reference", in place of "reference". So instead of principle (1), he has a requirement that a definite NP has to pick out an already familiar *discourse* referent, whereas an indefinite NP always introduces a new *discourse* referent. Since discourse reference is distinct from reference, and since, in particular, an NP may have a discourse referent even when it has no referent, this reformulation makes the familiarity theory immune to the objections encountered by its traditional version (1).

Let met illustrate with two examples how the distinction between discourse reference and genuine reference can be exploited in evading dilemmata that the traditional familiarity theorist must find fatal. Consider the text under (4).

(4) John came, and so did Mary. *One of them* brought a cake.

The underlined NP "one of them" is indefinite, therefore (1) would seem to predict that it must refer to an as yet "unfamiliar" person, i.e. a person not already introduced in the previous discourse. Now the first sentence of (4) mentions both John and Mary, hence both of them are familiar when "one of them" gets uttered and should consequently be excluded as potential referents for "one of them". But that is counterintuitive, since (4) is naturally read as saying that one of John and Mary, not some third person, brought a

[3] Karttunen 1968a, 1968b, 1976.

cake. "One of them" – if we are to admit that it refers to anything at all – clearly can refer to John or Mary here, in apparent violation of the familiarity theory. – But now suppose we have replaced (1) by Karttunen's version in terms of discourse referents. Then the prediction about "one of them" will be that, since it is indefinite, its discourse referent must be new and must be distinct from the discourse referents of "John" and "Mary" in particular. There is no prediction about the referents of these three NPs, and we may consistently hold any assumption we please about those. In particular, we may assume that NPs with distinct discourse reference sometimes happen to coincide in reference, and that (4), being a case of this kind, involves three discourse referents, but only two referents.

Next, consider (5).

(5)(a) Everybody found a cat and kept *it*. (b) *It* ran away.

The relevant facts here are that the "it" in (5a), but not the "it" in (5b), can be interpreted as anaphoric to "a cat", (the intended reading being one with "everybody" taking wider scope than "a cat"). Since the first "it" and its antecedent "a cat" both fail to refer, the traditional version of the familiarity theory cannot really be applied to them at all. Talking in terms of discourse referents, however, we can describe what is going on in (5) as follows: The indefinite "a cat" introduces a discourse referent. The first "it" picks up that same discourse referent, which – having just been introduced – is familiar, as required. At the end of (5a), this discourse referent ceases to exist and is no longer available when the second "it" comes along. Therefore that second "it" must find the familiar discourse referent it requires elsewhere, or the text is inacceptable. – Note that this way of talking about (5) implies that discourse referents behave in ways which it wouldn't make any sense to attribute to real referents: not only are there discourse referents for NPs that have no referents, but moreover, discourse referents may suddenly go out of existence, depending on certain properties of the utterance. In this case, the relevant generalization is that if a discourse referent gets first introduced inside the scope of a quantifier (here: "everybody"), then its lifespan cannot extend beyond the scope of that quantifier.

But what are discourse referents? We have seen that for this new concept to be useful we must dissociate it from certain properties inherent in the notion of a referent. But a merely negative characterization is of course not enough if we don't want to be reduced to vacuity. Karttunen (in the papers cited) formulates a number of generalizations about discourse referents, i.e. about the conditions under which they get introduced and the factors that determine their lifespan, such as for instance the generalization about quantifier scope limiting the lifespan of discourse referents that I just alluded to above. Taken together, these generalizations combine with Karttunen's version of the familiarity theory into a theory that yields empirical predictions and in the context of which "discourse reference" is a non-vacuous theoretical concept. In this sense, the question what discourse referents are has a satisfactory answer implicit in Karttunen's work, although there is no explicit definition.

Still, it has remained puzzling in many ways just what discourse referents are and where they fit into semantic theory. It seems appropriate to say that we are describing some aspect of the meaning of a word or construction of English when we talk about its capacity for introducing, picking up, or influencing the lifespan of, discourse referents. But is that an entirely separate aspect of meaning, or is it dependent upon other aspects of meaning, such as the referential and truth-conditional aspect? – Questions of this sort I hope to shed light on by suggesting that Karttunen's discourse referents be identified with what I will call "file cards", i.e. elements of a so-called "file", a theoretical construct which mediates in a way to be made precise between language and the world.

3. Conversation and File-keeping

A listener's task of understanding what is being said in the course of a conversation bears relevant similarities to a file clerk's task. Speaking metaphorically, let me say that to understand an utterance is to keep a file which, at every time in the course of the utterance, contains the information that has so far been conveyed by the utterance.[4] Suppose, for instance, someone is listening to an utterance of the following three-sentence-text.

(6)(a) A woman was bitten by a dog. (b) She hit it. (c) It jumped over a fence.

Before the utterance starts, the listener has an empty file, i.e. a collection of zero file cards. Call that empty file "F_0". As soon as (6a) has been uttered, the listener puts two cards into the file, gives each card a number – say "1" and "2", and writes the following entries on them: on card 1, he writes "is a woman" and "was bitten by 2", and on card 2, "is a dog" and "bit 1". He now has a two card file, call it "F_1", which looks like this:

F_1:

1	2
– is a woman – was bitten by 2	– is a dog – bit 1

[4] The file metaphor was first suggested to me by Angelika Kratzer, in response to an earlier attempt of mine to modify Grice's and Stalnaker's notion of "common ground" (cf. especially Stalnaker 1979) in such a way as to impose on common grounds an essentially file-like structure. I subsequently found uses of the file metaphor for more or less similar purposes elsewhere in the literature, e.g. in Karttunen (1976). With respect to their role in a model of semantics, my files are closely related not only to Stalnaker's "common grounds", but particularly to the "discourse representation structures" of Kamp (1981).

Next, (6b) gets uttered, which prompts the listener to update card 1 by adding the entry "hit 2", and to update card 2 by adding "was hit by 1". He now has F_2, still a two card file, but a different one:

F_2:

1	2
– is a woman – was bitten by 2 – hit 2	– is a dog – bit 1 – was hit by 1

Now comes the utterance of (6c). The listener takes a new card, numbers it "3", writes on it "is a fence" and "was jumped over by 2", and also updates card 2 by adding on it "jumped over 3". This leaves him with F_3, a three card file:

F_3:

1	2	3
– is a woman – was bitten by 2 – hit 2	– is a dog – bit 1 – was hit by 1 – jumped over 3	– is a fence – was jumped over by 2

With this illustration in mind, let us repeat our initial question: How do definites differ from indefinites? We may now answer: They differ in the way they influence the development of the file; the listener treats them differently, apparently following an instruction like (7) in his file keeping.

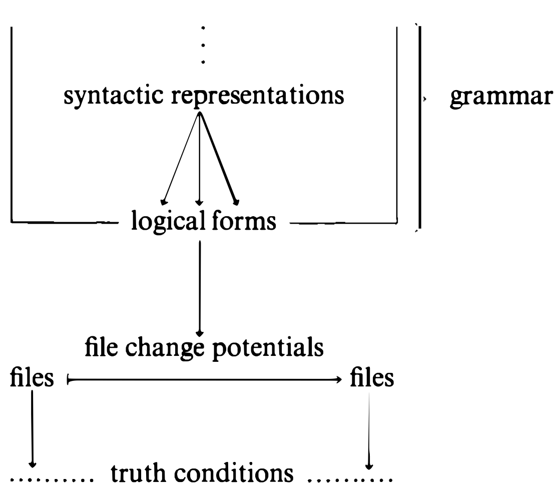
(7) For every indefinite, start a new card. For every definite, update an old card.

For instance, cards 1 and 2 were newly introduced in response to the indefinites "a woman" and "a dog" that occurred in (6a). Only definites, namely "she" and "it", occurred in (6b), therefore F_2 only contained the same cards that were already in F_1, albeit updated. (6c) had both an indefinite ("a fence") and a definite ("it") in it, hence it prompted both introduction of a new card (card 3) and updating of an old one (card 2). All of this conformed to (7).

Instruction (7) is reminiscent of principle (1) above and can in fact be seen as incorporating a version of the familiarity theory of definiteness: Like (1), (7) links definiteness to familiarity (= "oldness") and indefiniteness to novelty. The only difference of (7) from (1) is that not referents are supposed to be old or new, but rather file cards. By substituting file cards for referents in the formulation of the familiarity theory of definiteness, I have made basically the same move as Karttunen, who substituted discourse referents for referents, and like in Karttunen's case, this move enables me to avoid the presumption of referentiality which caused such problems for the traditional version (1) of the familiarity theory. Examples like (4) and (5) are easily accommodated, once we think of file cards instead of referents, since it is quite conceivable for there to be a file card that fails to describe a referent, or for two different file cards to happen to describe the same thing, or for file cards to be introduced into and be removed from the file, depending on what is getting uttered. In short, just the properties we have found it necessary to attribute to Karttunen's discourse referents are properties that fit right into the file card metaphor. This is why I would like to suggest that Karttunen's talk about "discourse referents" be rephrased by substituting "file card" for "discourse referent": once we realize that discourse referents are essentially like file cards, their identity criteria and their relation to referents come to look much less mysterious.

In this section, I have introduced the file metaphor and have applied it informally to examples. Now it remains to give a more precise account of the theoretically relevant properties of files and of the role they play in the semantic interpretation of natural language. Roughly, the model of semantics that I am going to present will embody the following assumptions. The grammar of a language generates sentences with representations on various levels of analysis, among them a level of "logical form". Each logical form is assigned a "file change potential", i.e. a function from files into files. Given an utterance with a certain logical form, this function will determine how you get from the file that obtains prior to the utterance to the file that comes to obtain as a result of the utterance. The system moreover includes an assignment of truth conditions to files. Note that logical forms themselves are not assigned truth conditions, only files are. Only in an indirect way, i.e. via the files they affect, will logical forms be associated with truth conditions. The diagram under (8) shows how this model of semantic interpretation is organized. I will elaborate on its various components in the next few sections, starting with the association of files with their truth conditions.

(8)



4. Files and the World

A file can be evaluated as to whether it corresponds to the actual facts or misrepresents them. Take e.g. the file F_1 of our example above. If it so happens that among all the women and dogs that there are there is not a woman-and-dog-pair such that the dog bit the woman, then F_1 obviously misrepresents the facts. I will speak of a "false" file in such a case, and correspondingly will call a file "true" if it fits the facts.

What does it take for a file to be true? To establish the truth of a file, we have to, so to speak, line up the sequence of cards in the file with a sequence of actual individuals, such that each individual fits the description on the corresponding card. Or, as I will put it, we have to find a sequence of individuals that *satisfies* the file. For file F_1, for instance, we have to find a two-membered sequence, i.e. a pair, that consists of a 1st member a_1 and a 2nd member a_2 such that a_1 fits card 1, and a_2 fits card 2 of F_1. Any such pair will satisfy F_1, i.e. we have:

$\langle a_1, a_2 \rangle$ satisfies F_1 iff a_1 is a woman, a_2 is a dog, and a_2 bit a_1.

Depending on how many cards a file contains, it will take pairs, triples, quadruples, or what not to satisfy it, therefore I speak generally of "sequences". Technically, a sequence is a function from some subset of N (the natural numbers) into A (the domain of all individuals). The pair $\langle a_1, a_2 \rangle$, for instance, is the function which maps 1 to a_1 and 2 to a_2. (Notice that I also admit sequences whose domains are not initial segments of N. E.g. a function that assigns an individual each to the numbers 2 and 7, but is not defined for any other numbers, also qualifies as a sequence. This would be the sort of sequence to satisfy a file whose only two cards are numbered "2" and "7".) A degenerate sort of sequence is the one whose domain is the empty set ϕ and which is therefore ϕ itself. ϕ is the only sequence that satisfies file F_0, the file of zero cards in our example above.

I will often want to refer to the set of all sequences that satisfy a given file, therefore I introduce a piece of notation, "Sat(F)" (read: "the satisfaction set of F").

(9) $\mathrm{Sat}(F) =_{\mathrm{def}} \{a_N\colon a_N \text{ satisfies } F\}$.

(Here and elsewhere, "a_N", "b_M", and the like range over sequences, where the subscripts "N", "M", etc. stand for each sequence's domain.) I also need a short way of referring to all the card-numbers that are used in a given file, so I use the notation "Dom(F)" (read: "the domain of F").

(10) $\mathrm{Dom}(F) =_{\mathrm{def}} \{n \in N\colon F \text{ contains a card with number } n\}$.

As I said before, a file is to count as true if some satisfying sequence for it can be found. Definition (11) expresses this.

(11) F is true iff $\mathrm{Sat}(F) \neq \phi$ (and false otherwise).

In the remainder of this article, I will often describe a file solely in terms of its domain and satisfaction set. It should be clear that that does not suffice to pick out a unique file. There are always many distinct files that happen to have the same domain and satisfaction set. To give an extreme example, any two false files which happen to employ the same set of card numbers are indistinguishable if you look only at domains and satisfaction sets, the satisfaction sets being empty for all false files. Yet, two such files may have completely different entries on their cards. So by specifying only the domains and satisfaction sets, I am leaving the files I am talking about grossly underspecified. Nevertheless, for certain selected purposes, such as those of the present article, it is possible to abstract away from all the ways in which files with identical domains and satisfaction sets may differ, and to still formulate the relevant principles.

5. Semantic Categories and Logical Forms

I will now turn to the upper part of diagram (8) and highlight some of the assumptions about logical forms that I need to rely on. Following standard practice, I assume that logical forms differ from surface structures and other syntactic levels of representation in that they are disambiguated in two respects: scope, and anaphoric relations. Scope is marked configurationally, with an element c-commanding its scope, and anaphoric relatedness is marked by coindexing, with two anaphorically related elements bearing identical numerical subscripts. The relation between sentences and their logical forms, generally a one-to-many relation, is defined by a set of transformational rules that derive logical forms from syntactic representations and by wellformedness constraints on the output of those rules.[5]

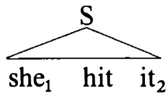
Both the rules that derive logical forms and the rules that interpret them by assigning them file change potentials appear to discriminate between elements of different semantic categories, such as variables, operators, and the like. Here I will not go into such questions as how many and which semantic categories there are, and to what extent the syntactic category of an element determines its semantic category. I just assume that there are at least the following semantic categories and they include at least the kinds of things listed, whether as a matter of stipulation or of principle.

variables: pronouns, empty NPs, indices on NPs with predicate heads (see below for illustration of the latter);
predicates: verbs, nouns;
operators: "every", negation.

[5] These assumptions about logical form are taken over from Chomsky's work and other work in the framework of the "Revised Extended Standard Theory", see in particular May (1977) and Reinhart (1976).

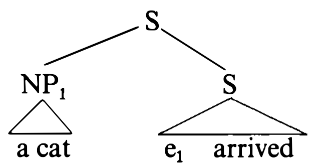
As for the rules and constraints that define the relation between the syntactic representation of a sentence and its logical forms, I will be very informal and incomplete here. Consider the structures in (12), each of which represents one of the logical forms that the English sentence below it can have.

(12)(a)



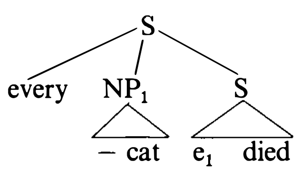
"She hit it."

(b)



"A cat arrived."

(c)



"Every cat died."

("e" marks an empty NP-position; the blank before "cat" in (12c) indicates an empty determiner-position.) These three examples may serve to illustrate a few general assumptions about logical forms:
- Every NP in logical form carries a numerical index.
- Only variables occur in the argument positions of predicates.[6]
- NPs that are not variables, i.e. those headed by predicates, are adjoined to their scopes and coindexed with the argument position they originate from.[7]
- Operators are adjoined as sisters to their argument(s). (Most operators are 2-place operators, in particular the quantifiers; some may be 1-place, e.g. negation.)

There is more to be said about how these assumptions follow from the way in which rules of logical-form-construction, wellformedness constraints on logical forms, and limitations on semantic interpretability interact with each other.

[6] This is similar to the "predication condition" of May (1977).

[7] May (1977) makes this assumption only for quantifying NPs, whereas I extend it to all predicate-headed NPs, quantifying or not.

Note the contrast between structures (12b) and (12c), which is due to an assumption whose significance I will have more to say about, viz. the assumption that the articles "a" and "the" are not operators, whereas certain other determiners, e.g. "every", are. What then is the semantic category of articles? None at all. They are treated as though they weren't there at all when it comes to semantic interpretation.

What I have so far said about semantic categorization applies only to lexical items and other basic units, but fails to specify a semantic category for the complex building blocks of logical forms, such as S-constituents and predicate-headed NP-constituents. With some simplification, we may assume that all complex constituents that are of any semantic category at all are *propositions*. These subdivide into atomic propositions, which consist of a predicate and its arguments, and molecular propositions, which are made up of other propositions and may or may not involve an operator. One kind of atomic proposition is dominated by S and made up of a verb and its subject and complements (if any), where the verb is the predicate and the variables in the subject and complement positions are its arguments. In (12), $[_S \text{ she}_1 \text{ hit it}_2]$, $[_S \text{ e}_1 \text{ arrived}]$, and $[_S \text{ e}_1 \text{ died}]$ exemplify this kind of atomic proposition. The other option for an atomic proposition is to have a noun as the predicate, in which case the dominating node is NP. (12) contains the examples $[_{NP_1} \text{ a cat}]$ and $[_{NP_1} \text{ \_ cat}]$. Nominal predicates always have one of their arguments realized as a mere numerical index which appears on the dominating NP-node. "Cat", for instance, is a 1-place predicate, and its argument in the examples just cited is the index 1. This is why I included "indices on NPs with predicate heads" in the above list of variables.

6. Logical Forms and their File Change Potentials

We can now proceed to the heart of the system diagrammed in (8), the assignment of file change potentials to logical forms. It will be useful to have another piece of notation, the symbol "+", which stands for the file change operation. Suppose we have a logical form p that determines a file change from F to F′. We express this by writing:

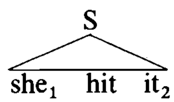
$$F + p = F'$$

(read: "the result of updating F on account of p is F′"). The task of assigning file change potentials to logical forms can now be seen as amounting to the task of defining "F + p" for files F and logical forms p of arbitrary composition and complexity. Actually, I will limit my efforts to a more modest task than that: Instead of committing myself to a full specification of the formal properties of files and the changes they undergo, I will characterize only one aspect of file change, namely how the satisfaction set is affected. As I noted earlier, this means that I am leaving a lot about the files I am talking about wide open. What I will define, thus, is not actually "F + p", but rather "Sat(F + p)".

A standard way of assigning interpretations to a language with expressions of unlimited complexity is by means of a recursive definition. Following this format, I will begin by specifying the file changes induced by atomic propositions and then characterize the file changes that molecular propositions bring about in terms of the file change potentials of their parts.

Consider (12a), repeated below, one of the logical forms of the simple sentence "She hit it."

(12a)



In the informal introduction of the file metaphor in section 3, I had the file clerk react to this sentence by changing a certain file F_1 into a certain file F_2. Recall what F_1 and F_2 were supposed to look like. Using the terminology I have since made available, they can now be described as follows:

$\mathrm{Dom}(F_1) = \mathrm{Dom}(F_2) = \{1, 2\}$
$\mathrm{Sat}(F_1) \;\; = \{\langle a_1, a_2\rangle : a_1$ is a woman, a_2 is a dog, and a_2 bit $a_1\}$
$\mathrm{Sat}(F_2) \;\; = \{\langle a_1, a_2\rangle : a_1$ is a woman, a_2 is a dog, a_2 bit a_1, and a_1 hit $a_2\}$
It is apparent that the transition from $\mathrm{Sat}(F_1)$ to $\mathrm{Sat}(F_2)$ consists in eliminating from $\mathrm{Sat}(F_1)$ all those pairs which fail to satisfy the sentence being processed, i.e. those pairs which fail to stand in the relation that the predicate of the sentence denotes, in this case the relation of hitting. Put formally:

$\mathrm{Sat}(F_2) = \{\langle a_1, a_2\rangle : \langle a_1, a_2\rangle \in \mathrm{Sat}(F_1)$ and $\langle a_1, a_2\rangle \in \mathrm{Ext}(\text{"hit"})\}$

(I write "Ext" for "the extension of".) The general rule under which this transition falls may be given as follows (subject to later revision):

(13) Let F be a file, and let p be an atomic proposition that consists of an n-place predicate R and an n-tuple of variables whose indices are $i_1, \ldots,$ and i_n respectively. Then:
$\mathrm{Sat}(F + p) = \{a_N : a_N \in \mathrm{Sat}(F)$ and $\langle a_{i_1}, \ldots, a_{i_n}\rangle \in \mathrm{Ext}(R)\}$.

Applied to the file F_1 and the logical form (12a), (13) gives us:

$\mathrm{Sat}(F_1 + (12a)) = \mathrm{Sat}(F_2)$,
as intended.

We just focussed on a particular logical form that grammar provides for the sentence "She hit it", namely (12a). But there are infinitely many others, since the choice of indices is supposed to be free. So (12a) represents really only one of many readings that the sentence may be uttered with, and we have yet to talk about the others. We also have to say something to explain the puzzling fact that despite the infinity of distinct logical forms assigned to each sentence by the grammar, most real-life utterances can be immediately understood in an unambiguous way. To appreciate the problem, put yourself once more into the imaginary file clerk's shoes. You have so far constructed the file F_1, and now you hear the speaker say: "She hit it". How do you guess that the intended reading is "She_1 hit it_2", and not, say, one of the following?

(14)(a) She_1 hit it_1.
(b) She_3 hit it_7.
(c) She_2 hit it_1.

(14a) is pretty easy to exclude: there is a well-known constraint, called "Disjoint Reference", which we may think of as a wellformedness condition on some level of representation in the grammar (logical form or one of the levels it is derived from). By this constraint, coindexings like the one in (14a) are ruled out, so (14a) doesn't count as a wellformed logical form and thus doesn't represent an available reading for any utterance of the sentence "She hit it" whatsoever.

With (14b), it's a rather different matter. No known constraint on indexing applies here, and it would quite clearly be wrong-headed to expect that anything would rule (14b) an ill-formed logical form. It can't be ill-formed, because we can imagine utterances of "She hit it" where (14b) would be precisely the logical form that represents the intended reading. Suppose, for instance, the preceding conversation had taken its course in such a way that you, the file clerk, had come up with a file F_4, which, unlike F_1, is characterized by the domain $\text{Dom}(F_4) = \{3, 7\}$ and the satisfaction set $\text{Sat}(F_4) = \{\langle a_3, a_7\rangle$: a_3 is a woman, a_7 is a dog, and a_7 bit $a_3\}$. If at this point you were confronted with an utterance of "She hit it", (14b) rather than (12a) would be the reading to construe it with. – What this shows is that in order to disambiguate the uttered sentence as (12a) as opposed to (14b), the file clerk must take into account what his current file looks like. What is at work here is thus not a constraint on logical forms considered in isolation, but rather a principle that constrains the choice of logical form *relative to a given file*. I want to propose that a principle of this sort, and in fact just the right principle to help us rule out (14b) when given F_1, is suggested to us by the familiarity theory of definiteness. The principle, which I call the "Novelty/Familiarity Condition", is this:

(15) Let F be a file, p an atomic proposition. Then p is appropriate with respect to F only if, for every noun phrase NP_i with index i that p contains:
if NP_i is definite, then $i \in \text{Dom}(F)$,
and if NP_i is indefinite, then $i \notin \text{Dom}(F)$.

With respect to the file F_1, for instance, (14b) is inappropriate because it contains two definite NPs, "she_3" and "it_7", whose indices fail to be in $\text{Dom}(F_1)$. (12a), on the other hand, with the definites "she_1" and "it_2", meets

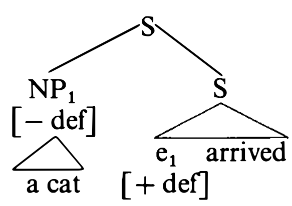
(15)'s requirement for appropriateness w.r.t. F_1. Note that for (15) to be applicable in the intended way, we must generally assume that NPs in logical form are marked for the feature [± definite].

(15) is presumably only one among other conditions on when a logical form is appropriate w.r.t. a file. Much of what has been discussed under the name of „presupposition" seems to be a matter of conditions of this sort.[8] From the point of view of the task of assigning file change potentials to logical forms, we may take appropriateness conditions as delimiting the range of pairs $\langle F, p \rangle$ for which the file change operation $F + p$ is at all defined. Unless p is appropriate w.r.t. F, there is no file change result $F + p$ determined. – As you will come to see shortly (once I have discussed indefinites), (15) interacts with the rules for file change in such a way that files will in effect always develop in accordance with instruction (7), which I formulated in section 3 as a first informal way of incorporating the familiarity theory of definiteness into a file-based semantics.

Returning to the file clerk's problem of eliminating all but the intended one among the infinity of logical forms for a given sentence, the Novelty/Familiarity Condition (15) will certainly help to rule out a lot of unwanted options, but it will still let through some. (14c) above is a case in point: Given the file F_1, the indexing "she$_2$"/"it$_1$" violates (15) no more than "she$_1$"/"it$_2$" (and (14c) is of course not ill-formed as a logical from either). In order to predict the inappropriateness of (14c) w.r.t. F_1, we need some account of gender, which I will not provide here. Another problem whose solution I must leave for another occasion[9] is the fact that different kinds of definites, e.g. personal pronouns in comparison with definite descriptions, differ in their appropriateness conditions in a way that the Novelty/Familiarity Condition, which is sensitive only to the distinction between definites and indefinites, is incapable of predicting.

Let us now turn to an example with an indefinite, such as the sentence "A cat arrived", one of whose logical forms is (12b), repeated from above, this time with the definiteness features filled in.

(12b)



To determine the file change that (12b) induces, we will have to consider two questions: First, since (12b) is a molecular proposition, we have to ask our-

[8] Heim (in preparation) argues that this view of what presuppositions are throws light on the behavior of presuppositions with respect to the so-called "projection problem".

[9] See Heim (1982).

selves how its overall effect on the file may be calculated on the basis of the file changes that each of its two parts would induce. Second, which rules of file change pertain to each of those parts?

The answer to the first question is as simple as it could be: We compute the file change of (12b) as a whole by subjecting the file first to the change that the left constituent dictates, and subsequently to the file change that the right constituent dictates. The general rule for this successive left-to-right mode of file change is this:

(16) Let F be a file, and let p be a molecular proposition whose immediate constituents are the propositions q and r (in that order). Then:
$\mathrm{Sat}(F + p) = \mathrm{Sat}((F + q) + r)$.

Applied to (12b), this means that we get from a given file F with satisfaction set Sat(F) to Sat(F + (12b)) by first calculating $\mathrm{Sat}(F + [_{NP_1}\ \text{a cat}])$ and then, from that, $\mathrm{Sat}((F + [_{NP}\ \text{a cat}]) + [_{S} e_1\ \text{arrived}])$.

$[_{NP_1}\ \text{a cat}]$ is an atomic proposition. Before we try to determine Sat (F + + $[_{NP_1}\ \text{a cat}]$), we have to make sure that it is even welldefined, i.e. that $[_{NP_1}\ \text{a cat}]$ is appropriate w.r.t. F in the sense of the Novelty/Familiarity Condition (15). Since $[_{NP_1}\ \text{a cat}]$ contains (in this case, exhaustively contains) an indefinite with index 1, (15) requires that $1 \notin \mathrm{Dom}(F)$. Let's assume F meets that requirement. Then $\mathrm{Sat}(F + [_{NP_1}\ \text{a cat}])$ is defined and should, by rule (13) above, equal the set:

(17) $\{a_N\colon a_N \in \mathrm{Sat}(F) \quad \text{and} \quad a_1 \in \mathrm{Ext}(\text{“cat”})\}$

That doesn't seem right, however. The problem is that if, as we are assuming, $1 \notin \mathrm{Dom}(F)$, then no element $a_N \in \mathrm{Sat}(F)$ will have a first member a_1 at all, let alone one that is in the extension of "cat". So the set described in (17) would of necessity be empty. This is not consistent with our intuition that "A cat arrived" is a contingent statement and should, at least sometimes, lead to a non-empty, i.e. true, file. We have to fix up rule (13) accordingly. The revised version under (18) is more adequately equipped to handle the example under consideration, while it still works just like (13) in cases of the sort that made us first design (13).

(18) Let F be a file, and let p be an atomic proposition that consists of an n-place predicate R and an n-tuple of variables whose indices are $i_1, \ldots, i_n$ respectively. Then:
$\mathrm{Sat}(F + p) = \{a_N \cup b_M \in A^{N \cup M}\colon a_N \in \mathrm{Sat}(F),\ M = \{i_1, \ldots, i_n\},$
and $\langle b_{i_1}, \ldots, b_{i_n} \rangle \in \mathrm{Ext}(R)\}$.

In contrast with (13), (18) allows for cases where F + p has a larger domain than F, i.e. where the sequences in Sat(F + p) have to be longer than those in Sat(F). Put informally, (18) says that every sequence in Sat(F + p) has to include as subsequences a sequence a_N satisfying F and a sequence b_M

satisfying the proposition p. Whenever you can find an a_N satisfying F and a b_M satisfying p, where a_N and b_M agree on the intersection of their domains, link them together and the result, $a_N \cup b_M$, will be a member of Sat(F + p). (That a_N and b_M have to agree on the common part of their domains is expressed in (18) by requiring "$a_N \cup b_M \in A^{N \cup M}$". $A^{N \cup M}$ denotes the set of functions from $N \cup M$ into A, and the union of two sequences is itself a sequence (i.e. a function) just in case they coincide on their common domain.) (18) reduces to (13) whenever $\{i_1, \ldots, i_n\}$ happens to be a subset of Dom(F).

Returning to our example, assume for concreteness that we start out with the empty file F_0, i.e. the one which has $\mathrm{Dom}(F_0) = \phi$ and $\mathrm{Sat}(F_0) = \{\phi\}$. F_0 is of course among those files w.r.t. which $[_{NP_1}$ a cat$]$ is appropriate in the sense of (15). What, then, does (18) tell us about the file-change result $F_0 + [_{NP_1}$ a cat$]$? We calculate:

(19) $\mathrm{Sat}(F_0 + [_{NP_1}\text{ a cat}]) = \{\langle b_1 \rangle : b_1 \in \mathrm{Ext}(\text{“cat”})\}$.

It is now easy to compute $\mathrm{Sat}(F_0 + (12\mathrm{b}))$ by applying (16) and, once more, (18).

(20) $\mathrm{Sat}(F_0 + (12\mathrm{b})) =$
$= \mathrm{Sat}((F_0 + [_{NP_1}\text{ a cat}]) + [_S e_1 \text{ arrived}]) =$
$= \{\langle b_1 \rangle : b_1 \in \mathrm{Ext}(\text{“cat”}) \text{ and } b_1 \in \mathrm{Ext}(\text{“arrived”})\}$.

This result is in line with our earlier, metaphorical, characterization of file change: Starting from a zero-card file, the sentence "A cat arrived" has brought us to a one-card file which is satisfied by any one-membered sequence whose one member is a cat and arrived.

Before I conclude this section, let me substantiate a remark that I made at the end of section 3. There I said that, although logical forms are not directly mapped onto truth conditions in a semantics that is organized along the lines of diagram (8), they still receive truth conditions indirectly, via the files they affect. I had in mind the following truth criterion for logical forms:

(21) Let F be a true file and p a logical form. Then p is true w.r.t. F if F + p is true, false w.r.t. F if F + p is false, and truth-value-less w.r.t. F if F + p is undefined.

(21) makes reference to the notion of truth that I defined for files in (11) above, and it basically equates the truth conditions of what is being said with the truth conditions of the resulting file. However, the applicability of this truth criterion is limited to cases where we can assume the truth of the file we start out with. If we have a false file to begin with, then we will always end up with another false file, however "true", in an intuitive sense, the utterance under consideration may be.

7. The Non-quantificational Analysis of Indefinites

I am only half way through with my recursive set of rules for assigning file change potentials to logical forms. But this is a good point to take a break and have a critical look at the present analysis of indefinite NPs and how it compares with the widely accepted Russellian analysis. Russell[10] argued that intuitively correct truth-conditions for sentences with indefinites result when the indefinite article is treated as an existential quantifier and sentences of the form (22) are assigned logical analyses of roughly the form (23).

(22) $[_S X [_{NP} a\, Y] Z]$

(23) $\exists x(Y(x) \mathbin{\&} (X x Z))$

"A cat arrived", for instance, would be logically analyzed as: "$\exists x(\text{cat}(x) \mathbin{\&}$ x arrived)". This "quantificational analysis of indefinites", as I will call it, is nowadays accepted in one variant or another by the vast majority of philosophers and linguists.

This paper contains what I will call, by contrast, a "non-quantificational analysis of indefinites". The logical analysis of an indefinite, as presented above, is just a proposition with a variable free in it. E.g. "a cat" corresponds to something like "cat(x)". When an indefinite occurs in a sentence, as in schema (22), the logical analysis of that sentence is again a proposition with a variable free in it:

(24) $Y(x) \mathbin{\&} (X x Z)$

The free variable in the indefinite remains free in the sentence as a whole. An existential quantifier is not part of the indefinite or of the sentence that contains it, neither is a quantifier of some other force than existential.[11] This section is intended to bring to bear some linguistic evidence on the choice between a quantificational and a non-quantificational analysis of indefinites. But first let me clarify to what extent the two analyses agree in their empirical predictions.

Despite the absence of an existential quantifier in the logical forms of sentences with indefinites, my theory predicts what are, in effect, existential truth-conditions for such sentences. Consider again "A cat arrived" with the logical form (12b). By the truth criterion (21) for logical forms, we know that (12b) is true w.r.t. a true file F if and only if F + (12b) is true. For F + (12b) to be true, in turn, means two things: First, (12b) must be appropriate w.r.t. F, in particular, Dom(F) must not contain 1, for F + (12b) to be defined.

[10] Russell (1919, Ch. 6).

[11] When I say (here and elsewhere in this article) that the indefinite is not a quantifier, I am of course not using "quantifier" in the sense of Barwise and Cooper (1981). In their sense of "quantifier", anything that denotes a function from predicate-meanings to proposition-meanings is a quantifier, and every kind of NP, even proper names and pronouns, can therefore be construed as quantifiers.

Second, Sat(F + (12b)) must be non-empty. Rules (16) and (18) determine that Sat(F + (12b)) = $\{a_N \cup b_{\{1\}} \in A^{N \cup \{1\}}: a_N \in \text{Sat}(F), b_1$ is a cat, and b_1 arrived$\}$. Given that Sat(F) is non-empty (since F is true), this set is non-empty just in case there is at least one cat that arrived. What we have just shown is that (12b) is true w.r.t. F if and only if at least one cat arrived. Since the proof did not depend on any particular properties of F other than that it be true and that (12b) be appropriate w.r.t. it, we may suppress relativization to F and simply say that (12b) is true if and only if at least one cat arrived. Moreover, since an analogous proof would have gone through for any other wellformed logical form of the sentence that (12b) represents, we can say that we have shown that the sentence "A cat arrived" is true if and only if at least one cat arrived. This prediction coincides of course with the familiar existential truth-condition that a quantificational analysis would have predicted as well.

At first sight, one might have thought it impossible that an existential truth-condition can be predicted while assuming a quantifier-free logical form like (12b) or (24). But there was of course no magic involved in the proof I just gave. The truth-condition came out existential because the notion of truth of a file has, so to speak, existential quantification built into it: truth of a file was defined as *there being at least one* satisfying sequence. So my disagreement with the quantificational analysis of indefinites is not a disagreement about whether or not we understand statements with indefinites in them as existentially quantified. It is rather a disagreement as to what is to be held responsible for the existential force of such statements: the indefinite article itself, or rather the way in which files generally relate to the facts that verify them? If we are to find any empirical evidence that will discriminate between these two points of view, it won't help to simply examine our intuitions about what sentences like "A cat arrived" mean. We will have to resort to considerations based on relatively indirect evidence like the following.

It is well-known of certain undebatable cases of quantifying NPs in natural language that they are subject to tighter restrictions on anaphora than certain other NPs. I have in mind contrasts like this one:

(25) Every soldier is armed. He will shoot.
(26) He is armed. He will shoot.

The two "he"'s can be anaphorically related in (26), but no anaphoric relation is possible between "every soldier" and "he" in (25). Why should this be so? – An explanation suggests itself if we assume that "every" is a quantifier, pronouns are variables, and (25) and (26) have logical analyses of essentially these forms:

(25′) $\forall x_i(\text{soldier}(x_i) \rightarrow \text{armed}(x_i))\ \&\ (x_j \text{ will shoot})$

(26′) $\text{armed}(x_i)\ \&\ (x_j \text{ will shoot})$

Is $i = j$ or $i \neq j$? Let us try to get away with the simplest possible assumption, i.e. that both texts permit readings with any arbitrary choice of i and j, and in particular readings with $i = j$ as well as readings with $i \neq j$. Now look at the satisfaction conditions that formulas like (25′) and (26′) receive under standard interpretations of predicate calculus. If (26′) has two different variables $x_i \neq x_j$, then a sequence satisfying it will have to contain an armed person and a (possibly distinct) person that will shoot. If the variables are the same in (26′), then a satisfying sequence has to include a person that is both armed and will shoot. So in the case of $i = j$, we have a substantially different satisfaction condition than in the case of $i \neq j$.

Now take (25′) and compare the satisfaction conditions that we get with $i = j$ to those we get with $i \neq j$. It turns out that it makes no difference: A sequence that satisfies (25′) must contain a person that will shoot, and provided it does, will satisfy (25′) only if every soldier is armed. This same satisfaction condition applies regardless of whether x_i and x_j are different variables or the same. This seems to be what is behind our judgment that (25) has no reading where "every soldier" and "he" are "anaphorically related": Even if we make a point of coindexing "every soldier" with "he", i.e. of picking identical variables in the logical analysis of (25), the coindexing is of necessity a semantically "vacuous" coindexing.[12]

What we have just observed about (25′) falls under a general law, so to speak a design feature of quantificational logic:

(27) If x_i is bound by a quantifier whose scope does not include x_j, then coindexing between x_i and x_j can only be vacuous.

(27) makes explicit what it is about the logical analysis of (25) that makes it different from the logical analysis of (26) in such a way that (25) will permit only vacuous coindexing where (26) permits the sort of non-vacuous coindexing that we perceive as an anaphoric reading. The crucial point is that "every soldier" was analyzed as a quantifying NP, whereas there was no quantifier assumed to occur in the corresponding position in (26).

What does all this have to do with the choice between a quantificational and a non-quantificational analysis of indefinites? Well, since (27) makes reference to quantifiers, we might try to exploit it as a diagnostic test for quantifyingness: If indefinites turn out to bear non-vacuous coindexing relations to variables outside their scope, then that ought to show they are not

[12] The relevant notion of "vacuity" could be defined as follows:

Def.: Let p be a formula, x a variable, and A the set of all occurrences of x in p. Suppose B and C are two disjoint subsets of A, with $A = B \cup C$. Then the members of B are *vacuously coindexed* with the members of C iff for some variable $y = x$, p and p′ have identical satisfaction conditions; where p′ results from p by substituting y for every occurrence of x that is in C.

Note that the "law" under (27) in the text is not a definition of vacuity, but rather a theorem that follows from the definition above, given the standard interpretation of quantifiers. This is why one could not simply choose to replace (27) by a stipulation that permits certain quantifiers to be coindexed non-vacuously with variables beyond their scope – unless one were to use logic as an uninterpreted formalism altogether.

quantifiers. Unfortunately, this test is not as foolproof in application as one might hope. But let's try it first.

Consider (28).

(28) A soldier will accompany us. He will shoot.

Presumably, (28) would be analyzed as (28′) under a quantificational treatment of indefinites, but as (28″) under a non-quantificational treatment.

(28′) $\exists x_i(\text{soldier}(x_i) \,\&\, (x_i \text{ will accompany us})) \,\&\, (x_j \text{ will shoot})$

(28″) $(\text{soldier}(x_i) \,\&\, (x_i \text{ will accompany us})) \,\&\, (x_j \text{ will shoot})$

By (27), the coindexing $i = j$ in (28′) is bound to be vacuous, while (28″) contains no obstacle to non-vacuous coindexing. Our intuitive judgment is that anaphora is possible in (28), just like in (26), and unlike in (25). We can straightforwardly predict the anaphoric reading by assuming a logical form along the lines of (28″), with $i = j$, a non-vacuous coindexing. (28′), on the other hand, would seem to preclude an anaphoric reading. This is prima facie evidence in favor of the non-quantificational analysis of indefinites.

There are various ways in which the conclusion just drawn can be, and has been, challenged. First, one might call into question a tacit assumption I have been making about the scope-options for quantifying NPs. With both (25) and (28), I took it for granted that a quantifying NP that occurred in the first sentence of each text could take scope at most over that sentence, not over the entire bisentential text. Had I permitted the quantifying NP "every soldier" in (25) and the putatively quantifying NP "a soldier" in (28) to take wider scope than the sentence, then the variable x_j in (25′) and (28′) could have come under the scope of $\forall$ or $\exists$, in which case $i = j$ would have been a non-vacuous coindexing. (Cf. (27).) This suggests that the quantificational analysis of indefinites could be saved if one were to maintain that indefinites, unlike certain other quantifying NPs, can take scope across several sentences.[13]

A second way of undermining my use of (28) as evidence against a quantificational analysis of indefinites goes like this: What we customarily describe as "anaphoric relations" may not be one and the same kind of logical relation in all cases, and in particular, need not always be non-vacuous identity of variables. So even if the logical analysis of (28) is (28′) (with either $i = j$ or $i \neq j$, it doesn't matter), we may still use (28) with the intention that x_j refer to whatever individual is responsible for the truth of "$\exists x_i(\text{soldier}(x_i) \,\&\, (x_i \text{ will accompany us}))$". Viewed in this way, the so-called "anaphoric" use of the pronoun in (28) has really a lot more in common with deictic pronoun uses than with bound-variable anaphora: The pronoun is here taken to refer to a contextually salient individual, just like deictic pronouns do, except that in this case the crucial factor in making the intended referent salient is the fact that it verifies a piece of immediately preceding discourse.[14]

[13] This is basically what Geach (1962) suggests.

Both of these objections deserve serious consideration before we can be sure that the ability of indefinites to serve as antecedents for anaphoric pronouns beyond their scope is indeed a symptom of the non-quantificational nature of indefinites. I will have to be brief here.[15] My answer to both objections is that the alternative accounts they give of the anaphoric relation between "a soldier" and "he" in (28) do not carry over to certain other examples of an analogous nature. Consider (29).

(29) Every time a soldier accompanies us he shoots.

Under a quantificational analysis of indefinites, (29) ought to get the following logical analysis:

(29′) $\forall t(\exists x_i(\text{soldier}(x_i) \;\&\; (x_i \text{ accompanies us at } t)) \rightarrow (x_j \text{ shoots at } t))$

Unlike in the case of (28), the truth-conditions of (29) are clearly inconsistent with an alternative analysis under which "a soldier" takes wide enough scope to include x_j.[16] This shows that if a quantificational analysis of indefinites is to be reconciled with their behavior w.r.t. anaphora, it will not suffice to appeal to their unconstrained scope options.

But (29) also doesn't lend itself to an account in terms of the sort of quasi-deictic use of "he" that had some plausibility for examples like (28). The problem is that the "he" in (29) fails to refer, and that deixis without reference is a contradiction in terms by all available explications of the concept.

So (29), more compellingly than (28), shows that indefinites enter into anaphoric relations where this is not to be expected from the point of view of a quantificational analysis. I have yet to show that the non-quantificational alternative that I am developing in this article covers examples like (29) in a natural way. This leads us to the topic of quantification.

8. File Change Rules for Quantified Sentences

Before I give the file change potentials for operator-headed logical forms, in particular universally quantified and negated ones, I should say something about "closed" propositions (i.e. propositions without free variables) in general. Take a simple sentence with a 0-place predicate:

(30) It is raining.

In the context of the file metaphor, one doesn't quite know how to deal with (30): As an informative sentence, it ought to call for an updating of the file

[14] This line is taken by Kripke (1977), Lewis (1979), and elsewhere.

[15] For more careful argumentation, see Heim (1982, Ch. I), where I also address a third way of undermining the use of (28) as evidence against the quantificational analysis of indefinites, advocated by Evans (1977) and Cooper (1979), among others.

[16] Unless one assumes, moreover, that the wide-scope taking indefinite switches its quantificational force from existential to universal. That assumption has been pursued in Engli (1979) and Smaby (1979), whose proposals are discussed in depth in Heim (1982, Ch. I).

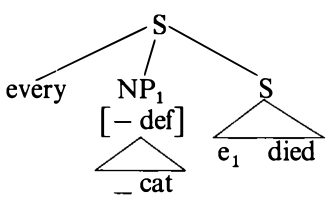
somehow; but what exactly is the file clerk supposed to do? The information that it is raining does not belong on any particular file card, it seems, since each file card is a description of an individual, but (30) is not about any individual. Should the file clerk perhaps write on some arbitrary card: "is such that it is raining"? Or should he write that on all cards? And what if the file so far doesn't contain any cards yet? – Fortunately, we can leave these questions unanswered here. Recall that we have already resigned ourselves to characterizing file change only as far as the domain and satisfaction set are concerned. So we need not specify anything else about the file change potential of (30) than its impact on domain and satisfaction set. And that is already taken care of by rule (18) above. We only need to assume that the extension of a 0-place predicate is empty if the corresponding state of affairs fails to obtain, and is the unit set of the empty sequence if it does obtain. E.g. we have Ext("rain") = $\{\phi\}$ if it rains, Ext("rain") = ϕ otherwise. This way we can apply (18) to give us:

$$\text{Sat}(F + (30)) = \begin{cases} \text{Sat}(F), \text{ if Ext(“rain”)} = \{\phi\}, \\ \phi, \text{ otherwise.} \end{cases}$$

This amounts to the correct truth conditions for such sentences. The reason why I dwelled on this point is that quantified and negated propositions are similarly puzzling if we are so ambitious as to want to say what exactly the file clerk does in response to them. Under the modest aspect of domain and satisfaction set change, however, they pose no problem.

An example of a universally quantified logical form is (12c), repeated from above.

(12c)



Note that I have here marked the determinerless NP [_ cat] as indefinite. I assume that NPs which have had their determiners moved out generally qualify as [– definite].

Unlike in the case of operator-free molecular propositions, the file change induced by (12c) cannot be broken down into a simple, so to speak "linear", succession of smaller steps that correspond to each of the sub-propositions. The presence of an operator makes considerably higher demands on the file clerk's memory and computational abilities. We may think of the evaluation of (12c) as proceeding in three steps as follows:

Step 1: Tentatively update the original file F by incorporating $[_{NP_1}$ _ cat$]$ into it. This gets you to $F' = F + [_{NP_1}$ _ cat$]$ with the following satisfaction set, as determined by rule (18):

$\mathrm{Sat}(F') = \{a_N \cup b_{\{1\}} \in A^{N \cup \{1\}}\colon a_N \in \mathrm{Sat}(F) \text{ and } b_1 \text{ is a cat}\}$.

The change from F to F′ is only "tentative" insofar as the file clerk retains F in his memory and is prepared to make his next actions depend not only on F′, but also on F.

Step 2: Tentatively update F′ by incorporating $[_S e_1 \text{ died}]$ into it. This results in F″, determined by rule (18) as follows:

$\mathrm{Sat}(F'') = \{a_N \cup b_{\{1\}} \in A^{N \cup \{1\}}\colon a_N \in \mathrm{Sat}(F), \text{ and } b_1 \text{ is a cat, and } b_1 \text{ died}\}$.

Again, F′ is retained in memory, which now contains both F and F′.

Step 3: For each sequence a_N in Sat(F), do the following: Determine whether all "continuations" of a_N that are in Sat(F′) are also in Sat(F″). (By a "continuation" of a_N I mean a sequence that includes a_N as a subsequence.) If yes, carry a_N along into the satisfaction set of the new file F + (12c); if no, eliminate a_N. – After you have done this for each $a_N \in \mathrm{Sat}(F)$, you will thus have:

$$\begin{aligned}\mathrm{Sat}(F + (12\text{c})) = \{&a_N \in \mathrm{Sat}(F)\colon \text{for every } b_M \supseteq a_N,\\ &\text{if } b_M \in \mathrm{Sat}(F') \text{ then } b_M \in \mathrm{Sat}(F'')\}.\end{aligned}$$

You may now clear the memory of F, F′, and F″.

Step 3 is obviously the one which takes into account the specific force of the operator involved, here universal quantification. The preceding two steps serve only to set up the auxiliary files on which the calculation in step 3 is based. These two steps are the same for all two-place operators.

Let us figure out the result of this three-step procedure for a concrete choice of initial file, the empty file F_0. Starting from F_0, the outcomes of steps 1 and 2 will look like this:

$\mathrm{Sat}(F_0 + [_{NP_1} \_ \text{ cat}]) = \{b_{\{1\}}\colon b_1 \text{ is a cat}\}$.
$\mathrm{Sat}((F_0 + [_{NP_1} \_ \text{ cat}]) + [_S e_1 \text{ died}]) = \{b_{\{1\}}\colon b_1 \text{ is a cat and } b_1 \text{ died}\}$.

The result of step 3 is then the following:

$$\mathrm{Sat}(F_0 + (12\text{c})) = \begin{cases}\mathrm{Sat}(F_0), \text{ if every cat died;}\\ \phi, \text{ otherwise.}\end{cases}$$

In view of the truth criterion (21) above, this implies that (12c) is true w.r.t. F_0 just in case every cat died, an intuitively adequate prediction. However, we still have to show that equally adequate predictions are generated with choices other than F_0 for the initial file.

At first glance, there seems to be a problem with initial files F that already contain a card number 1. For instance, if we assumed $\mathrm{Dom}(F) = \{1\}$ and $\mathrm{Sat}(F) = \{a_{\{1\}}\colon a_1 \text{ is a pet}\}$, then each sequence in Sat(F) could have at most one continuation in $\mathrm{Sat}(F + [_{NP_1} \_ \text{ cat}])$, namely the trivial continuation, which is itself. The result of step 3 would then be this:

$\mathrm{Sat}(F + (12\text{c})) = \{a_{\{1\}}\colon a_1 \text{ is a pet, and if } a_1 \text{ is a cat, then } a_1 \text{ died}\}$.

This conflicts with the intuitive truth conditions of (12c) and in particular with its universal force.
However, we have no reason to worry about this result, because it only arises if we neglect the constraints which the Novelty/Familiarity Condition imposes on the choice of F. Recall that the Novelty/Familiarity Condition (= (15) above) has to be met each time an atomic proposition is incorporated into the file, or else there won't be a file change result defined at all. Applied to the evaluation of (12c), this means in particular that step 1 cannot be carried out unless $[_{NP_1}$ _ cat$]$ is appropriate w.r.t. the initial file F. According to (15), F is therefore not permitted to contain the number 1 in its domain, "_ cat_1" being indefinite. In particular, the choice of F which in the example above seemed to lead to inadequate predictions about the truth conditions of (12c) was inconsistent with the Novelty/Familiarity Condition, and we should have realized that neither F + (12c) nor, consequently, the truth of (12c) w.r.t. F is at all defined for such choices of F.

Turning to examples of greater complexity than (12c), we find that the three step procedure that I have proposed applies analogously, and that it interacts with the Novelty/ Familiarity Condition in such a way as to predict the contrast between definites and indefinites when they appear inside a universally quantifying NP. Compare (31) and (32).

(31) Every man who likes a donkey buys it.
(32) Every man who likes it buys it.

(31) expresses a generalization about man-donkey-pairs; it is as though the universal quantifier "every" was here binding the donkey-variable along with the man-variable. (32), by contrast, is read as generalizing over all men that like a fixed object. The variable corresponding to the "it" in "every man who likes it" may refer to a contextually supplied object, or may be anaphoric to an antecedent in the larger text in which (32) appears. Either way, it is not understood as bound to "every" in the way that "a donkey" in (31) is. Let me briefly show how this contrast is derived from the assumptions I have introduced.

(31) is represented on the logical form level roughly as follows.

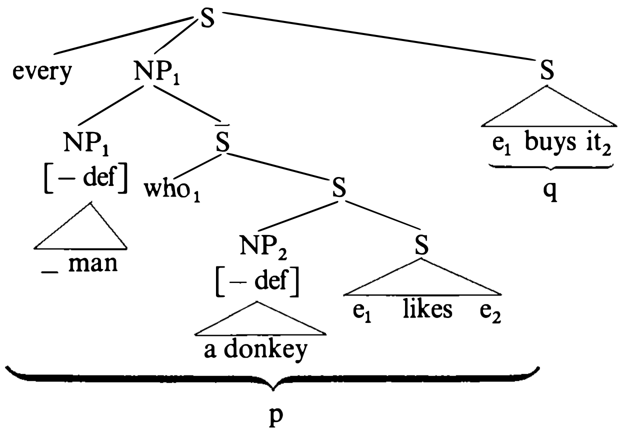
Starting from an initial file F, steps 1 and 2, in analogy to the specifications given above, provide us with auxiliary files $F' = F + p$ and $F'' = F' + q$. These have the following satisfaction sets, according to rules (16) and (18).

$\text{Sat}(F') = \{a_N \cup b_{\{1,2\}}: a_N \in \text{Sat}(F)$, b_1 is a man, b_2 is a donkey, and b_1 likes $b_2\}$.

$\text{Sat}(F'') = \{a_N \cup b_{\{1,2\}}: a_N \in \text{Sat}(F)$, b_1 is a man, b_2 is a donkey, b_1 likes b_2, and b_1 buys $b_2\}$.

Concerning F, we must assume that Dom(F) contains neither 1 nor 2, because otherwise the Novelty/Familiarity Condition would not let F′ be

(31′)

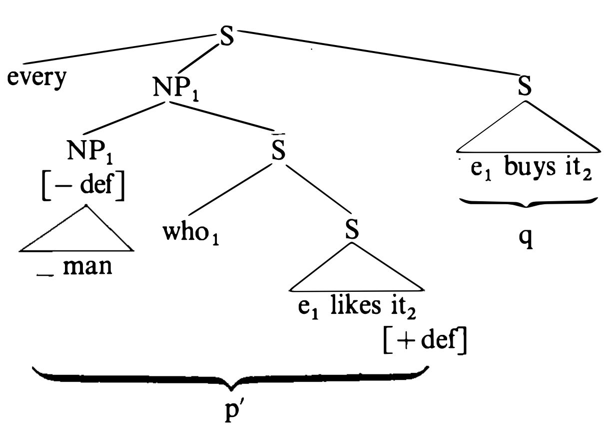


defined. We now proceed to step 3, in which we consider one by one the members a_N of Sat(F). For each such a_N, we form every continuation of a_N that is in Sat(F′) and determine whether it is also in Sat(F″). To satisfy F′, a continuation of a_N has to contain two members, number 1 and number 2, which are a man and a donkey he likes, respectively. Every man/donkey-pair of this sort will figure in some continuation of a_N, because a_N itself does not contain any members number 1 and number 2. Therefore the requirement that every continuation of a_N that satisfies F′ must also satisfy F″ amounts to the requirement that every man-donkey pair in which the man likes the donkey is also such that the man buys the donkey. The result of step 3 is therefore:

$$\mathrm{Sat}(F + (31')) = \begin{cases} \mathrm{Sat}(F), \text{ if every man who likes a donkey buys it,} \\ \phi, \text{ otherwise} \end{cases}$$

The logical form of (32) differs from that of (31) in that it has the definite "it" instead of the indefinite "a donkey":

(32′)



This time, steps 1 and 2 will produce the auxiliary files $F' = F + p'$ and $F'' = F' + q$ (where F stands again for the initial file):

$\mathrm{Sat}(F') = \{a_N \cup b_{\{1,2\}} \in A^{N \cup \{1,2\}}: a_N \in \mathrm{Sat}(F)$, b_1 is a man, and b_1 likes $b_2\}$.
$\mathrm{Sat}(F'') = \{a_N \cup b_{\{1,2\}}: a_N \in \mathrm{Sat}(F)$, b_1 is a man, b_1 likes b_2, and b_1 buys $b_2\}$.

Unlike in the previous example, the Novelty/Familiarity Condition this time requires that Dom(F) doesn't contain 1, but does contain 2. This has important consequences for how step 3 applies. In step 3, we look at each $a_N \in \mathrm{Sat}(F)$ and form all continuations of a_N that satisfy F′. Because $2 \in \mathrm{Dom}(F)$, a_N includes a member a_2, and every continuation of a_N has that same a_2 as its member number 2 as well. Therefore, not every pair of a man and an individual he likes will necessarily be part of a continuation of a_N, but rather, only those pairs where the individual the man likes is none other than a_2. The predicted result of step 3 is a file with the following satisfaction set:

$\mathrm{Sat}(F + (32')) = \{a_N \in \mathrm{Sat}(F)$: for every b_1, if b_1 is a man and b_1 likes a_2, then b_1 buys $a_2\}$.

The difference between this and $\mathrm{Sat}(F + (31'))$ above reflects the intuition that (31) involves universal quantification over pairs, whereas (32) quantifies over men which like a "fixed" individual.

It remains to write up explicitly the file change rule which dictates the three step procedure I have described. We want this rule to be general enough to work not only for examples like (12c), (31), and (32), but also for examples like (33):[17]

(33) Every man who owns a donkey sells it to a merchant.

(33) contains an indefinite ("a merchant") in the right-hand argument of the quantifier, and this creates complications for step 3 as I have specified it so far. The problem is that in a case like this, F″ will contain more cards than F′, and it will therefore be impossible in principle for any sequence that satisfies F′ to also satisfy F″. The following formulation of the file change rule for universally quantified propositions is designed to deal with this additional complication. This is why it doesn't simply require that every continuation of a given a_N that satisfies F′ also satisfy F″, but rather that a further continuation of the continuation satisfy F″.

(34) Let F be a file, and let p be a molecular proposition whose immediate constituents are a universal quantifier and the propositions q and r (in that order). Then:

$\mathrm{Sat}(F + p) = \{a_N \in \mathrm{Sat}(F)$: for every $b_M \supseteq a_N$ such that $b_M \in \mathrm{Sat}(F + q)$, there is a $c_L \supseteq b_M$ such that $c_L \in \mathrm{Sat}((F + q) + r)\}$.

[17] The example is from Kamp (1981), whose treatment of quantification (designed to go with his version of the non-quantificational analysis of indefinites) made me aware that I had overlooked cases like (33) in a earlier version of my theory.

I leave it to the reader to verify that (34) applies satisfactorily to example (33).

I complete this section by formulating the file change rule for negated propositions, trusting that the reader can come up with his or her own illustrations.

(35) Let F be a file, and let p be a molecular proposition whose immediate constituents are a negation operator and the proposition q. Then:
$\mathrm{Sat}(F + p) = \{a_N \in \mathrm{Sat}(F)\text{: there is no } b_M \supseteq a_N \text{ such that } b_M \in \mathrm{Sat}(F + q)\}$.

References

Barwise, J. and R. Cooper (1981): Generalized quantifiers and natural language, *Linguistics and Philosophy.*

Christophersen, P. (1939): *The Articles: A Study of their Theory and Use in English*, Copenhagen.

Cooper, R. (1979): The interpretation of pronouns, in: F. Heny and H. Schnelle (eds.), *Syntax and Semantics 10, Selections from the Third Groningen Round Table*, New York.

Egli, U. (1979): The Stoic concept of anaphora, in: R. Bäuerle, U. Egli, and A. v. Stechow (eds.), *Semantics from Different Points of View*, Berlin.

Evans, G. (1977): Pronouns, quantifiers, and relative clauses, *Canadian Journal of Philosophy 7.*

Geach, P. (1962): *Reference and Generality*, Ithaca N.Y.

Hawkins, J. (1978): *Definiteness and Indefiniteness*, Croom Helm Linguistics Series, R. Lass (ed.), London.

Heim, I. (1982): The semantics of definite and indefinite noun phrases, Ph. D. dissertation, University of Massachusetts, Amherst.

Heim, I. (in preparation): On the projection problem for presuppositions (working title).

Jespersen, O. (1949): *A Modern English Grammar, Vol. VII*, completed by N. Haislund, London 1961.

Kamp, H. (1981): A theory of truth and semantic representation, in: Groenendijk et al. (eds.), *Formal Methods in the Study of Language*, Amsterdam.

Kaplan, D. (1972): What is Russell's theory of definite descriptions?, in: D. Davidson and G. Harman (eds.), *The Logic of Grammar*, New York 1975.

Karttunen, L. (1968 a): What makes definite noun phrases definite?, RAND corporation report P 3871, unpublished.

Karttunen, L. (1968 b): What do referential indices refer to?, RAND corporation report P 3854, unpublished.

Karttunen, L. (1976): Discourse referents, in: J. McCawley (ed.), *Syntax and Semantics 7*, New York.

Kripke, S. (1977): Speaker's reference and semantic reference, in: P. French, T. Uehling, and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*, Minneapolis 1979.

Lewis, D. (1979): Score-keeping in a language game, in: R. Bäuerle, U. Egli, and A. v. Stechow, *Semantics from Different Points of View*, Berlin.

May, R. (1977): The grammar of quantification, Ph. D. dissertation, M.I.T.

Quine, W. v.O. (1960): *Word and Object*, Cambridge, Mass.

Reinhart, T. (1976): The syntactic domain of anaphora, Ph.D. dissertation, M.I.T.

Russell, B. (1919): *Introduction to Mathematical Philosophy*, London.

Smaby, R. (1979): Ambiguous coreference with quantifiers, in: F. Günthner and S.J. Schmidt (eds.), *Formal Semantics and Pragmatics of Natural Languages*, Dordrecht.

Stalnaker, R. (1979): Assertion, in: P. Cole (ed.), *Syntax and Semantics 9, Pragmatics*, New York.