ANGELIKA KRATZER

AN INVESTIGATION OF THE LUMPS OF THOUGHT

CONTENTS

0. WHAT THIS PAPER IS ABOUT

This paper is about situation semantics and about the meaning of counterfactuals. It argues that there is a close connection between the laws of counterfactual reasoning and a relation between propositions that I want to call 'lumping'. Capturing this relation seems to require a component of semantic interpretation which recognizes parts of possible worlds (situations) as primitives and implies a new approach to genericity and negation.

* Research for this paper was supported by the National Science Foundation under grant BNS 87-19999. Different parts of the paper were presented in talks at the University of Massachusetts at Amherst in February and August, 1985, at Yale University in January, 1986, at the University of Texas at Austin in April, 1987, at the sixth Amsterdam colloquium in April, 1987 (in absentia), at MIT in April, 1988, and at CSLI in Stanford in September, 1988. 1 would like to thank the audiences of these talks for their comments and suggestions. A number of improvements in an intermediate version of the paper are the result of discussions with Emmon Bach and Barbara Partee while co-teaching a course on events. I am grateful for Edmund Gettier's efforts to find counterexamples. In particular, I would like to thank Arnim von Stechow for presenting the paper in Amsterdam and giving me helpful comments. Elisabeth Selkirk, Barbara Partee, Irene Heim, Fred Landman, Steve Berman, Nirit Kadmon, Karina Wilkinson, Alice ter Meulen, Max Cresswell, Jon Barwise, John Bigelow, Gennaro Chierchia, Peter Staudacher, and Roger Schwarzschild all read some version of the paper and gave me much appreciated feedback. Two anonymous reviewers for *Linguistics and Philosophy* helped to improve the final version. Elisabeth Selkirk was always generous in providing me with semantic intuitions. Her concern for focus inspired the proposed semantics for negation. The paper owes an obvious debt to the work of David Lewis.

Linguistics and Philosophy **12**: 607–653, 1989.

1. What lumps of thought are

Imagine the following situation: One evening in 1905, Paula painted a still life with apples and bananas. She spent most of the evening painting and left the easel only to make herself a cup of tea, eat a piece of bread, discard a banana or look for an apple displaying a particular shade of red. Against the background of this situation, consider the following two dialogues that might have taken place the following day:

Dialogue with a pedant

Pedant: What did you do yesterday evening?

Paula: The *only* thing I did yesterday evening was paint this still life over there.

Pedant: This cannot be true. You must have done something else like eat, drink, look out of the window.

Paula: Yes, strictly speaking, I did other things besides paint this still life. I made myself a cup of tea, ate a piece of bread, discarded a banana, and went to the kitchen to look for an apple.

Dialogue with a lunatic

Lunatic: What did you do yesterday evening?

Paula: The *only* thing I did yesterday evening was paint this still life over there.

Lunatic: This is not true. You *also* painted these apples and you *also* painted these bananas. Hence painting this still life was not the only thing you did yesterday evening.

In both dialogues, Paula exaggerated in claiming that painting a still life was the only thing she had done that evening. She had done other things, and the pedant correctly noticed this. Being a captive of his unfortunate character, he could not help insisting on the truth, and this is really all we can blame him for.

The lunatic case is very different. I don't think that Paula has to accept this person's criticism. She didn't paint apples and bananas *apart* from painting a still life. Painting apples and painting bananas was part of her painting a still life, like my arms and legs are part of me. Wherever I go, my arms and legs will come along. Is it true, then, that I can never be alone? I think not. Somehow, when I talk about myself, my parts have no independent existence, their presence doesn't count. Likewise, on that memorable evening, a very special relationship between three propositions was established: It was true that Paula painted a still life. And it was also true that she painted apples and that she painted bananas. But once we consider the first proposition a fact of this world (at the time in question), we are not entitled any longer to consider the latter two propositions as *separate* facts. If you count the entities in this room and you count me as one of them, you'd better forget about my ears. And if you count the facts of our world and you count Paula's painting a still life as one of them, you'd better overlook her painting apples. Quite generally, whenever we start counting, we have to make sure that the entities in our domain are truly distinct. Consider the following example (inspired by Carlson (1977, p. 346ff.) which illustrates the lunatic's fallacy in a different domain.

Noah's ark

How many kinds of animals did Noah take into the ark? He took a pair of dogs. That's one kind. He also took a pair of cats. That's another kind. Hence he took at least two kinds of animals. He also took a pair of doves. Now we have three kinds. He also took mammals. That's certainly a kind we haven't had before. That makes four kinds of animals. And he took birds, which gives us five kinds....

We like to think about the facts of a world in terms of the set of propositions which are true in it. And we are used to construe propositions as sets of possible worlds. The proposition that Paula painted a still life is the set of possible worlds in which Paula painted a still life. And the proposition that Paula painted apples is the set of possible worlds in which she painted apples (at the time under consideration). Both of these propositions happen to be true in our world. Possible worlds semantics captures this property in that the corresponding sets of possible worlds contain our world as a member. As far as our world is concerned, the two propositions are even more closely related, though: they are not *distinct* facts of our world. There is an aspect of the actual world that makes the proposition that Paula painted a still life true. And that very same aspect of our world also makes the proposition that she painted apples true. It will be useful to have a technical term for the relationship we are after. Let us say that the proposition that Paula painted a still life *lumps* the proposition that she painted apples in the actual world.[1] Note that whatever aspect of the actual world makes the proposition that Paula

[1] A proposition lumps another proposition in a world w in virtue of certain part-whole relationships holding between situations of w. The two propositions don't stand in a part-whole relationship themselves.

painted apples true is presumably not sufficient to make the proposition that she painted a still life true. The proposition that Paula painted apples, then, does not lump the proposition that she painted a still life in the actual world. If Paula's still life had contained only apples and no bananas, the case would be different. Whatever aspect of the actual world would make the proposition that Paula painted apples true would also make the proposition that Paula painted a still life true. Hence the two propositions would lump each other in the actual world.

Like many interesting semantic relationships, the lumping relation is affected by vagueness. Consider the following example: My neighbor's house burnt down. His kitchen burnt down as part of it. The proposition that his house burnt down, then, lumps the proposition that his kitchen burnt down in the actual world. My neighbor's barn was destroyed by the same fire. Was the barn part of the house? Does the proposition that his house burnt down lump the proposition that his barn burnt down in the actual world? We may or may not be able to settle on an answer to this question. We don't have to. There will be cases where the lumping relationship clearly holds. There will be other cases which are not so clear. If the lumping relationship plays a role in the semantics of certain constructions, we expect that its vagueness will contribute to the vagueness of these constructions in a systematic and detectable way.

We have seen that traditional possible worlds semantics construes propositions as sets of possible worlds. On this approach, it is not obvious how we can formally capture the lumping relationship. It seems, then, that we may be missing something in possible worlds semantics. We may be missing something, but it may not be very important. Or is it? I am going to argue in this paper that once we pay close attention to the lumping relationship, we will gain some new insights into the labyrinth of counterfactual reasoning, an area which has puzzled semanticists for a long time.

2. How lumps of thought can be characterized in terms of situations

As a result of our previous considerations, we are faced with the following task: We have to characterize the special relationship holding (in our world on some evening in 1905) between the proposition that Paula painted a still life and the propositions that she painted apples and that she painted bananas. Obviously, this relationship is not logical implication. Paula's painting a still life doesn't logically imply her painting apples and bananas. Material implication isn't a better candidate. In the actual world, at the time considered, Paula's painting a still life materially implies her making herself a cup of tea, for example. (Assuming that our scenario is true, both of these propositions were true at the time). But the proposition that Paula is making herself a cup of tea is not part of the 'lump of facts' whose properties we are trying to capture.

I suggested above that the proposition that Paula painted a still life and the proposition that she painted apples are not distinct facts of our world, since whatever aspect of our world makes the first proposition true will also make the second proposition true. On the other hand, *not* every aspect of our world in which "Paula painted a still life" is true, is also an aspect of our world in which any of the following is true: "Paula made herself a cup of tea", "Paula ate a piece of bread", "Paula discarded a banana", "Paula went to the kitchen". It seems, then, that we might be able to characterize the lumping relationship as soon as we grant that 'aspects' or 'parts' of possible worlds can make propositions true. What is an 'aspect' or a 'part' of a possible world? It seems that an 'aspect' or a 'part' of a possible world is nothing else but a *possible situation*. What I want to suggest, then, is that the lumping relationship be characterized in terms of the notion 'truth in a possible situation'. If propositions are sets of possible *situations* rather than simply sets of possible *worlds* we will be able to actually define the lumping relationship rather than taking it as a primitive notion.

Assume that we are given a set of possible worlds. Strictly speaking, we will mainly consider worlds without much of a history, slices of worlds, worlds at a time. Yet I will continue talking about 'worlds'. (Time is not a concern in this paper. Let us put it aside whenever we can.) Along with the worlds, we are given their parts. The parts of a world are its situations. Since worlds are parts of themselves, they are also situations. They are maximal situations, situations that are not part of other situations. Given all of this, consider the following definition of the *lumping relationship*

Lumping

A proposition p lumps a proposition q in a world w if and only if (i) and (ii) both hold:

(i) p is true in w

(ii) Whenever a situation s is part of w and p is true in s, then q is true in s as well

The above definition assumes that propositions can be true not only in whole worlds, but also in parts thereof, in situations. This assumption is not very common in possible worlds semantics (in spite of Kripke (1965)), but it is popular elsewhere. Recent advocates are Barwise and Perry (1983), Veltman (1985), and Landman (1986). The idea seems simple enough, but is not easy to execute. There is the danger of losing classical two-valued logic, and there are insecurities concerning negation and quantification. While being indebted to all of my predecessors, my proposals will differ from theirs in significant detail. The motivation for these deviations will come from a close look at natural language semantics, and – quite surprisingly – from an in-depth investigation of counterfactual reasoning.

3. A SEMANTICS BASED ON SITUATIONS

3.1. *The Metaphysics of Situations*[2]

What are situations? I suggested above that situations may help us define the lumping relation. In this section, I will in turn use intuitions about the lumping relation to sharpen our understanding of the nature of situations. Section 4 will then use facts about counterfactual reasoning to further clarify the intuitions relied on here.

Situations cannot be identified with space-time chunks. The following example shows why: As a matter of fact, I am hungry and tired right now. Let us consider that slice of our world history which comprises just this present moment. Is every part of this slice in which I am hungry a part in which I am tired? And likewise, is every part in which I am tired a part in which I am hungry? If situations were simply space-time chunks, the answer would probably be 'yes'. The minimal space-time chunk in which I am hungry now would be the space-time chunk presently occupied by me. But this would also be the minimal space-time chunk in which I am tired now. We would have to conclude, then, that the proposition that I am hungry now and the proposition that I am tired now lump each other in the actual world. They would be one and the same fact. But they are not.

For the proposition that I am hungry now to be true in a situation, the situation has to contain me. It also has to contain something that supports the truth of the proposition that I am *hungry* now. For the proposition that I am tired now to be true in a situation, the situation must again contain me. And it must contain something that supports the truth of the

[2] The sort of metaphysics I will be concerned with in this section is 'natural language metaphysics' in the sense of Bach (1986).

proposition that I am *tired* now. If neither proposition lumps the other in the actual world, then a situation must be able to contain me and whatever it is about me that makes it true that I am hungry now, *without* containing whatever it is about me that makes it true that I am tired now (and the other way round). This seems an almost absurd requirement. How can a situation contain me (or my present slice) without also containing whatever it is that makes it true that I have certain properties like being hungry or tired?

What we seem to need is a way of distinguishing between an individual per se and those aspects of the world that make it true that the individual has properties like the ones just mentioned. A distinction of precisely this kind is made in theories of universals. We speak of the 'residue' of a particular when we want to talk about that part of the particular that gives it its particularity. The residue is what we get when we abstract away from all the universals that the particular instantiates. This leads us into the heart of the debate on the status of universals and their relationship to the particulars that instantiate them. In what follows, I am not going to engage in a debate that has been going on for centuries. Instead, I will briefly and quite tentatively consider a proposal made by David Armstrong (Armstrong, 1978, volume 1, see also Lewis, 1986, Section 1.5). Armstrong's distinction between thick and thin particulars seems to give us the distinction we are looking for.

We started out with the assumption that situations are parts of worlds. But what, then, are worlds? For Armstrong, the actual world is a world of states of affairs. A state of affairs is a particular's having a 'property' ('property' in the sense of a monadic universal), or two or more particulars standing in some 'relation' ('relation' in the sense of a polyadic universal). We may consider particulars with all their 'properties'. This gives us the notion of a 'thick' particular. Alternatively, we may have a conception of a 'thin' particular. A thin particular is a particular with all its 'properties' stripped off (the 'residue' in more traditional terminology). When we say that a state of affairs is a particular's having a 'property' or two or more particulars standing in some 'relation', the notion of a thin particular is involved. Thick particulars are themselves states of affairs (but not every state of affairs is a thick particular, of course). Note that we really should not confuse universals on the one hand with properties and relations on the other. Universals are parts of states of affairs. Properties and relations are denotations of predicates. Some properties and relations may directly correspond to universals, others will not. I have been using scare quotes whenever I used the words "property" or "relation" and really meant 'universal'.

Suppose the actual world is made up of states of affairs in the sense of Armstrong. If there are other worlds they should be made up of possible states of affairs. A possible state of affairs will be a possible particular's having a 'property', or two or more possible particulars standing in some 'relation'. If possible situations are parts of possible worlds, they will be made up of possible states of affairs, too.

Let us now return to the problem we started out with: If I am hungry and tired at this very moment, does this mean that my being hungry now and my being tired now are propositions which are true in the same situations? There may or may not be a universal corresponding to the property of being hungry. Suppose there is. Then there will be a state of affairs (and hence a situation) consisting of the relevant thin slice of me and that universal. Suppose further that there is a universal corresponding to the property of being tired. Again there will be a state of affairs and a situation consisting of the relevant thin slice of me and that universal. The two situations both contain the same thin particular. But they contain different universals. Hence they are different. Hence the proposition that I am hungry and the proposition that I am tired are true in different situations of the slice of our world under consideration. If there aren't universals corresponding to the properties of being hungry and being tired a different story has to be told. Maybe my being hungry and my being tired will each be factored into several states of affairs. Maybe . . . It does not matter how the story goes. My being hungry and my being tired will now have a chance to come out as distinct facts. What is important is that a situation doesn't have to contain thick particulars, particulars with all their 'properties'. Thin particulars will often do.

3.2. *Some Ingredients for Situation Semantics*

This section gives an overview of some of the crucial features of the situation semantics that I want to propose. We will only pay attention to those features that will play a role in illuminating the lumping relation, and eventually in shedding light on counterfactual reasoning. We will also introduce some convenient abbreviations to be used in the rest of the paper. Here is a list of the basic ingredients that we are going to need:

S — a set, the *set of possible situations* (including the set of thick particulars)

A — a set, the set of *possible thin particulars*

$\leqslant$ — a *partial ordering* on $S \cup A$ such that at least the following conditions are satisfied:

(i) For no $s \in S$ is there an $a \in A$ such that $s \leqslant a$

(ii) For all $s \in S \cup A$ there is a unique $s' \in S$ such that $s \leqslant s'$ and for all $s'' \in S$: if $s' \leqslant s''$, then $s'' = s'$

$\mathbb{P}(S)$ the power set of S, the *set of propositions*

W a subset of S, the set of maximal elements with respect to $\leqslant$. W is the set of *possible worlds*. For all $s \in S$, let w_s be the maximal element s is related to by $\leqslant$.

Intuitively, $\leqslant$ is the 'part of'-relation. No possible situation is part of a thin particular. Every possible thin particular or situation is related to a unique maximal element, its *world*. We will need this assumption later when we discuss accidental interpretations for certain quantifiers and negation (see Sections 3.5, 5, and 6). If individuals are each related to a unique world, it follows that you and me, for example, cannot exist in other possible worlds. Other individuals very much like us may represent us there (our *counterparts*), but we are not there ourselves. Lewis (1986) presents a detailed defense of these views concerning counterparts. See also Lewis (1968) and (1973).

3.3 *The Logical Properties and Relations*

The situation semantics outlined above gives us the possibility of defining classical and non-classical versions of the basic semantic properties and relations. We might consider a non-classical notion of validity in terms of truth in all possible situations, for example, or else stick to the classical notion in terms of truth in all possible worlds. The following definitions capture the classical notions. While it would be worthwhile to explore the non-classical ones as well, I will not pursue the matter here.

Truth
A proposition $p \in \mathbb{P}(S)$ is *true* in a situation $s \in S$ if and only if $s \in p$.

Validity
A proposition $p \in \mathbb{P}(S)$ is *valid* if and only if p is true in all $w \in W$.

Consistency
A set of propositions $\mathbb{A} \subseteq \mathbb{P}(S)$ is *consistent* if and only if there is a $w \in W$ such that all members of $\mathbb{A}$ are true in w.

Compatibility
A proposition $p \in \mathbb{P}(S)$ is *compatible* with a set of propositions $\mathbb{A} \subseteq \mathbb{P}(S)$ if and only if $\mathbb{A} \cup \{p\}$ is consistent.

Logical Consequence
A proposition $p \in \mathbb{P}(S)$ *follows* from a set of propositions

$\mathbb{A} \subseteq \mathbb{P}(S)$ if and only if p is true in all those $w \in W$ in which all members of $\mathbb{A}$ are true.

Logical Equivalence

Two propositions p and $q \in \mathbb{P}(S)$ are logically equivalent iff $p \cap W = q \cap W$

The notions of 'validity', 'consistency', 'compatibility', 'logical consequence', and 'logical equivalence' depend only on the possible worlds part of propositions. This will insure that our semantics will be a classical one.

Our last definition will be the 'official' definition of the lumping relationship which now looks as follows:

Lumping

For all propositions p and $q \in \mathbb{P}(S)$ and all $w \in W$: p lumps q in w if and only if the following conditions hold:

(i) $w \in p$

(ii) For all $s \in S$, if $s \leq w$ and $s \in p$, then $s \in q$.

3.4 *Persistence*

This section addresses the question whether all propositions expressible by utterances of natural language sentences are persistent. This is a question that naturally arises within any semantic framework based on partial objects like situations. It is also a question that has received different answers from different scholars in the field.

Suppose a proposition is true in a situation s. Will this proposition be true in all situations of which s is a part? If so, it is *persistent* in the terminology of Barwise and Perry or *T-stable* in the terminology of Veltman and Landman. (Recall that we are neglecting matters of time. We are *not* talking about persistence through time. The situations we are considering all have the same temporal location). Are all propositions expressible by utterances of sentences in natural languages persistent? If you say something which is true in a situation, will it stay true once we consider bigger and bigger situations containing the situation we started out with? If you establish that a proposition is true in the limited situation accessible to your senses, can you conclude that this proposition is also true in the actual *world*? You can indeed, if propositions of the sort human beings can know are persistent. Persistence, then, seems to be a very desirable property of propositions. Yet there are problems.

Consider the proposition p that is true in a situation s if and only if whenever there is a tree in s, then this very tree is laden with wonderful apples in s. Now look at this orchard over there. All its trees are laden with wonderful apples. Obviously p is true in the limited part of our world which comprises just my orchard and nothing else. Let us now review some situations of which this orchard is a part: Amherst with all its trees, Hampshire County, Massachusetts, the United States, the Planet Earth. In all of those situations p is false. The proposition p, then, is an example of a proposition which is not persistent. Are propositions like p ever expressed by utterances of sentences in natural languages? Some scholars have thought so.

To see why, let us return to my orchard. A man from Boston wants to buy it. He wonders whether all its trees are apple trees. I inform him: "Yes, and every tree is laden with wonderful apples". In uttering

(1) Every tree is laden with wonderful apples

I didn't make a claim about every tree in the world. It is clear from the context of use that I only claimed that every tree in my orchard is laden with wonderful apples. How can we account for the limitations of my assertion? We might say that in uttering (1), I did indeed express the proposition p, but I only claimed p to be true in a very limited part of our world. This is the view expressed by Barwise and Perry concerning parallel examples with definite descriptions. For them, taking p to be the proposition expressed by my utterance of (1) would be the appropriate way of accounting for implicit quantifier restrictions. As a consequence, they are committed to the view that not all propositions expressible by utterances of sentences of natural languages are persistent.

There is an alternative, however. We might suppose that the limitations observed when I uttered (1), are part of the very proposition expressed. On this account, quantifiers like 'every', 'most', 'all', and so forth depend for their interpretation on a restricting property which may be provided in part by the context of use.[3] Which proposition is expressed by an utterance of sentence (1) varies, then, with the restricting property assumed in the utterance situation. Depending on the situation, it may be the proposition that every tree in the whole world is laden with wonderful apples, or that every tree in my orchard is laden with wonderful apples, or that every tree in your orchard is laden with wonderful apples, or that every tree as far away as Mindelheim is laden with wonderful apples, and so forth. On this picture, we are not committed to the view that my actual utterance of (1) expressed a non-persistent proposition. The proposition expressed (on this particular occasion) would now be the same proposition

[3] That quantifiers can be restricted by the utterance context was clearly seen by George Boole (1854) who coined the term 'universe of discourse'.

as the one expressed if I had uttered the following sentence:

(2) Every tree in my orchard is laden with wonderful apples

Section 3.5 below will explore different options for assigning persistent propositions to universally quantified sentences like (1) or (2).

The above treatment of implicit quantifier restrictions will allow us to tentatively hold on to the principle that all propositions expressible by utterances of sentences of natural languages are persistent. That is, we will assume that all these propositions obey what I want to call the 'persistence constraint'. Constraints like the persistence constraint are interesting since they narrow down the range of possible utterance meanings in a radical way. The persistence constraint has also quite specific empirical consequences. As we will see shortly, it will force us to posit sentence meanings and logical forms which have independent justifications in different areas of semantics (see especially the discussion of negation in Section 6). Within the framework presented above, the formal definition of persistence looks as follows:

Persistence

A proposition $p \in \mathbb{P}(S)$ is persistent if and only if for all s and $s' \in S$ the following holds: Whenever $s \leqslant s'$ and $s \in p$, then $s' \in p$.

In the remainder of this paper, I am going to neglect most matters of context-dependency. To facilitate exposition, let us pretend that sentence meanings are simply propositions rather than functions assigning propositions to utterance contexts as assumed, for example in Stalnaker (1972), Cresswell (1973), Kaplan (1977), Kratzer (1978), or Lewis (1980).

3.5. *Sentence Meanings*

This section gives examples for sentence meanings within the semantic framework developed so far. The emphasis will be on the interpretation of quantifiers and logical connectives since these words will be important in our discussion of counterfactual reasoning later on.

In what follows let α and β be variables for sentences. For any sentence α let $[\![\alpha]\!]^g$ be the proposition expressed by α given a variable assignment g. Until further notice, we will assume that our language contains only individual variables. A variable assignment will then be a function assigning a member of A (the domain of individuals) to each such variable. Following, for example, Chomsky (1981), semantic interpretation will take place at a level of Logical Form. At this level, quantifiers have been raised out of their surface positions to form restricted quantifier structures of the sort discussed in Cushing (1976) and McCawley (1981) among others. Let us first consider atomic sentences:

(D1) *Atomic sentences*
For any variable assignment g:
$[\![x \textit{ is sleeping}]\!]^g$ is true in a situation $s \in S$ if and only if $g(x) \leq s$ and $g(x)$ is sleeping in s.
$[\![\textit{Paula is sleeping}]\!]^g$ is true in a situation $s \in S$ if and only if there is an individual $a \in A$ such that a is a counterpart of Paula in w_S, $a \leq s$, and a is sleeping in s.

Our variables range over thin particulars, elements of the set A. Likewise, our names denote thin particulars. For it to be true that Paula is sleeping in a situation we only require her *thin* self (or her counterpart's thin self) to be part of that situation. But that much we have to require in this particular case. Verbs behave differently as to such 'physical presence' requirements. That I am talking to you can only be true in a situation which has the thin residues of both of us (or our counterparts) as parts. That I am longing for a piece of bread can be true in a situation which doesn't contain the tiniest ('thinnest') bread crumb.

Within a compositional semantics, we don't usually specify the meanings of atomic sentences as a whole. We break down these sentences into their constituents and specify for each of these parts the contribution it makes to the truth-conditions of the whole sentence. This procedure is not of particular interest in the cases considered here, so I leave it at the illustrations given above.

Let us now look at the truth conditions for some complex sentences. Paying careful attention to the persistence constraint, we don't have to give special treatment to conjunction, disjunction, and existential quantification. The definitions familiar from possible worlds semantics seem to suit situation semantics as well.

(D2) *Conjunction*
For any variable assignment g:
$[\![\alpha \textit{ and } \beta]\!]^g$ is true in a situation $s \in S$ if and only if $[\![\alpha]\!]^g$ and $[\![\beta]\!]^g$ are both true in s.

(D3) *Disjunction*
For any variable assignment g:
$[\![\alpha \textit{ or } \beta]\!]^g$ is true in a situation $s \in S$ if and only if $[\![\alpha]\!]^g$ or $[\![\beta]\!]^g$ is true in s.

(D4) *Existential Quantification*
For any variable assignment g:
$[\![(\textit{There is an } x\text{: } \alpha)\beta]\!]^g$ is true in a situation $s \in S$ if and only if there is a variable assignment g' which is just like g except possibly for the value it assigns to x (call such an assignment an "x-alternative of g") such that $[\![\alpha]\!]^{g'}$ is true in s and $[\![\beta]\!]^{g'}$ is true in s.

As defined above, conjunction, disjunction, and existential quantification all preserve persistence. That is, given that $[\![\alpha]\!]^g$ and $[\![\beta]\!]^g$ are persistent (for some variable assignment g), so are $[\![\alpha \textit{ and } \beta]\!]^g$, $[\![\alpha \textit{ or } \beta]\!]^g$, and $[\![(\textit{There is an } x\text{: } \alpha)\beta]\!]^g$.

Let us now turn to universal quantification. Universal quantification does require special attention in a semantics based on situations. If we want to pursue the hypothesis that all propositions expressible by utterances of natural language sentences are persistent, we cannot adopt the familiar truth conditions as given in (D5).

(D5) *Non-Persistent Universal Quantification*
For any variable assignment g:
$[\![(\textit{For all } x\text{: } \alpha)\beta]\!]^g$ is true in a situation $s \in S$ if and only if for all x-alternatives g' of g the following holds: If $[\![\alpha]\!]^{g'}$ is true in s, then $[\![\beta]\!]^{g'}$ is true in s

If we interpreted our old sentence (1) (via its logical form (1′))

(1) Every tree is laden with wonderful apples
(1′) (*For all* x: x is a tree) x *is laden with wonderful apples*

as definition (D5) tells us to, this sentence would express the proposition p we encountered in Section 3.3. We have seen that p is not persistent. If we are right in holding on to the persistence constraint, definition (D5) doesn't give the correct truth-conditions for sentence (1).

While ruling out definitions like (D5), the persistence constraint still allows conceivable definitions of the sort given in (D6), (D7) or (D8), for example.

(D6) *Radical Universal Quantification*
For any variable assignment g:
$[\![(\textit{For all } x\text{: } \alpha)\beta]\!]^g$ is true in a situation $s \in S$ if and only if $s \in W$ and for all x-alternatives g' of g the following holds: If $[\![\alpha]\!]^{g'}$ is true in s, then $[\![\beta]\!]^{g'}$ is true in s.

(D7) *Generic Universal Quantification*
For any variable assignment g:

$[\![(For\ all\ x{:}\ \alpha)\beta]\!]^g$ is true in a situation $s \in S$ if and only if for all situations $s' \in S$ such that $s \leq s'$ and all x-alternatives g' of g the following holds: If $[\![\alpha]\!]^{g'}$ is true in s', then there is an $s'' \in S$ such that $s' \leq s''$ and $[\![\beta]\!]^{g'}$ is true in s''.

(D8) *Accidental Universal Quantification*
For any variable assignment g:
$[\![(For\ all\ x{:}\ \alpha)\beta]\!]^g$ is true in a situation $s \in S$ if and only if for all x-alternatives g' of g the following holds: Whenever $[\![\alpha]\!]^{g'}$ is true in w_S, then $[\![\alpha]\!]^{g'}$ and $[\![\beta]\!]^{g'}$ are true in s.

The four definitions (D5) to (D8) assign four different propositions to a given universally quantified sentence. The four propositions will be true in the same possible *worlds*, but not in the same possible *situations*. They will be *logically equivalent*, but *not identical*. We might say that these propositions differ as to how the truths of a world are 'distributed' over the situations of that world. We have seen that definition (D5) doesn't assign persistent propositions to universally quantified sentences. Definition (D6) assigns propositions which are persistent, but can only be true in *worlds*. Such propositions are very strong lumpers. If they are true, they lump every other true proposition in the world under consideration. So definition (D6) doesn't give the correct truth-conditions for sentences like (1) either. It may be true that every tree is laden with wonderful apples, and it may likewise be true that I bought three cords of wood. Yet the latter fact is certainly not part of the former.

Definition (D7) gives the sort of truth-conditions which resemble some definitions in Kripke's semantics for intuitionistic logic (they are also reminiscent of definitions familiar from model-theoretic forcing). I called this kind of universal quantification 'generic' for reasons that will become clear in Section 5. On this account, universal quantification preserves persistence as desired, but it creates propositions with very poor lumping properties. Those propositions are true in all or none of the situations of a world. But then there is no world in which they could lump a proposition which is true in that world without being true in all of its situations. This is undesirable in our case as shown by the following example: Suppose that being a friend of exaggerations as much as of apple trees, I foolishly claim that the only thing which is the case in our world at this time of the year is that every tree in my orchard is laden with wonderful apples. And here comes the lunatic again. Pointing to one of the trees he counters: "That's not true, it is also the case that *this* tree here is laden with wonderful apples". This is the kind of remark we have come to expect from him, and the characteristic oddity of his reasoning is an indication that the proposition expressed by (3) (as uttered by him)

(3) This tree here is laden with wonderful apples

is lumped by the proposition expressed by (2) (as uttered by me) in our world.

(2) Every tree in my orchard is laden with wonderful apples
(2′) *(For all x: x is a tree and x is in my orchard) x is laden with wonderful apples*

Definition (D7) does not account for this. Given definition (D7), the proposition expressed by sentence (2) can be true in situations of our world which don't contain the lunatic's tree as a part. Hence in our world, the proposition which definition (D7) assigns to sentence (2), though being true, does not lump the proposition expressed by sentence (3).

Definition (D8) is what we are looking for. (As before, don't yet pay attention to the name I chose for it). On this proposal, the proposition expressed by (2) is true in a situation s only if s is big enough to contain all the individuals which, in the world of s, are trees in my (counterpart's) orchard. Quite generally, on the accidental interpretation, the proposition expressed by an utterance of a universally quantified sentence can be true in a situation s only if all the individuals which satisfy its restrictive clause in the world of s satisfy its restrictive and its matrix clause in s. This means that given interpretation (D8), and given that the proposition expressed by sentence (2) is true in our world we predict that in our world, this proposition will lump the proposition expressed by sentence (3).

Truth-conditions for other quantifiers are given in a similar way. Take the case of 'exactly two'. 'Exactly two' is another quantifier that presents us with a potential persistence problem (unlike 'at least two'). Here is a proposal for a persistent interpretation:

(D9) *Exactly Two (Accidental Interpretation)*
For any variable assignment g:
$[\![(\textit{exactly two } x\text{: } \alpha)\beta]\!]^g$ is true in a situation s if and only if there are x-alternatives g' and g'' ($g' \neq g''$) of g such that $[\![\alpha]\!]^{g'}$, $[\![\beta]\!]^{g'}$, $[\![\alpha]\!]^{g''}$, and $[\![\beta]\!]^{g''}$ are true in s, and whenever there is an x-alternative g''' of g such that $[\![\alpha]\!]^{g'''}$ and $[\![\beta]\!]^{g'''}$ are true in w_S, then $g''' = g'$ or $g''' = g''$.

Under interpretation (D9), a sentence like (4) with logical form (4′)

(4) Exactly two trees in my yard have bird nests in them
(4′) *(Exactly two x: x is a tree in my yard) x has bird nests in it*

expresses a proposition which can be true in a situation *s* only if whenever an individual satisfies the restrictive and matrix clauses in the world of *s* it also satisfies the two clauses in *s*. This ensures that the proposition assigned to (4) (via 4′) is persistent and has the desired lumping properties: in our world, the fact that this tree here (I am pointing at a tree in my yard) has bird nests in it is part of the fact that exactly two trees in my yard have bird nests in them. Definition (D9) in interaction with our definition of 'lumping' correctly captures this relationship.

We have discussed the meanings of conjunction, disjunction and of various sorts of quantifiers. What is still missing is an account of negation. In a semantics based on situations, truth conditions for negation are a difficult matter which need careful justification. I will not be able to give this justification before having shown how the lumping relation enters into counterfactual reasoning. Let us stop here, then, and briefly summarize what we have achieved so far.

3.6. *A First Conclusion: A Semantics Based on Situations*

In traditional frameworks, truth conditions derive their empirical justification from their predictive power concerning truth in a world, logical consequence, and so forth. In developing the situation semantics presented above, an additional criterion of adequacy was imposed: We wanted to predict the correct lumping properties of the propositions involved. Imposing this criterion made us reject the 'radical' as well as the 'generic' truth conditions for universal quantification (the rejection of the generic truth conditions is only temporary. We will find a use for these truth conditions in Section 5). Contrary to views articulated by Barwise and Perry, Veltman and Landman, our semantics obeys the persistence constraint throughout. I argued in Section 3.4 that once we recognize the importance of quantifier restrictions contributed by the utterance situation, we are not forced anymore to assume that propositions expressed by sentences involving universal quantifiers or definite descriptions are not persistent. Veltman's and Landman's reasons for rejecting persistence have mainly to do with epistemic 'may'. I have argued in Kratzer (1977) and related work that modals require for their interpretation a 'conversational background' to be provided by the context of use. This means that like contextually provided quantifier restrictions, conversational backgrounds play a role in determining the very proposition expressed.

I have been using a couple of lunatic stories as heuristic devices helping me to point to intuitions regarding the lumping relation. These stories are only of very limited use however. More often than not, phrases like "the only thing I did", "the only thing that is the case" or "the only thing which is going on" sound very unnatural or impose special conditions, and quite generally, they don't combine very easily with negative sentences. The sections to follow will show that the investigation of counterfactual reasoning may be a much more interesting testing ground.

4. Counterfactual Reasoning

4.1. *Some Facts About Counterfactuals*

Counterfactuals come in different varieties. Here are two examples:

(5) If Mr. Brown read a newspaper, he would read the Morning Union

(6) If Mr. Brown read a newspaper, he might read the Morning Union

(5) expresses a 'would'-counterfactual, (6) expresses a 'might'-counterfactual. There are other sorts of counterfactuals, but for our purposes, the two types mentioned will suffice (I am using the term 'counterfactual' whenever I want to talk about the proposition expressed by a given counterfactual sentence. The antecedent of the counterfactual is the proposition expressed by the 'if'-clause of the sentence, the consequent is the proposition expressed by the 'then'-clause, leaving out the modal).

Finding the truth-conditions for counterfactuals has been one of the most hotly debated questions in semantics and more generally in the philosophy of science. Most scholars working in the field agree that the truth of a counterfactual in a world depends, in some way or other, on what is the case in that world (maybe at a particular time). What makes the semantics of counterfactual sentences so difficult is that not all facts have equal weight: some are important, others are altogether irrelevant.

There are two approaches to this problem. Philosophers like Nelson Goodman (1947) actually took it upon themselves to try to say exactly what the facts are which have to be taken into account in the evaluation of a counterfactual sentence. The idea was that after adding these facts to the antecedent as additional premises, the consequent of the counterfactual should follow logically from the resulting set (Goodman did not consider propositions, but this is not important here). Goodman eventually reached the conclusion that the additional premises don't seem to be specifiable in a non-circular way.

An alternative view was advanced by Robert Stalnaker (1968) and David Lewis (1973) who carefully avoided any precise characterization of the additional premises relevant for a particular piece of counterfactual reasoning. Stalnaker and Lewis both stress the vagueness of counterfactuals. Their analyses crucially rely on the inherently vague concept of similarity. As an example, let us look at Lewis' analysis of counterfactuals (Stalnaker's analysis differs from Lewis' analysis in ways which are not essential for our present purposes). Consider the following sentence.

(7) If I were looking into a mirror, I would see a face with brown eyes

The counterfactual expressed by (7) is true in a world w if and only if there is a world w' such that in w', I am looking into a mirror and see a face with brown eyes, and w' is closer to the actual world than any world in which I am looking into a mirror and don't see a face with brown eyes. 'Might'-counterfactuals are interpreted as 'duals' of the corresponding 'would' counterfactuals. The 'might'-counterfactual corresponding to (7), for example, is true in a world w if and only if the counterfactual expressed by (7′) is false in w.

(7′) If I were looking into a mirror, I would not see a face with brown eyes

A semantics for counterfactuals along these lines may look innocent, but using such an analysis and some obvious properties of the similarity relation, it has been possible to formally characterize an interesting body of valid counterfactual reasoning. Is this a success? I think it is. I also think, however, that it is not yet a complete success.

Consider the following scenario:[4] Last year, a zebra escaped from the Hamburg zoo. The escape was made possible by a forgetful keeper who forgot to close the door of a compound containing zebras, giraffes, and gazelles. A zebra felt like escaping and took off. The other animals preferred to stay in captivity. Suppose now counterfactually that some other animal had escaped instead. Would it be another zebra? Not necessarily. I think it might have been a giraffe or a gazelle. Yet if the similarity theory of counterfactuals were correct, we would expect that, everything else being equal, similarity with the animal that actually escaped should play a role in evaluating this particular piece of counterfactual reasoning. Given that all animals in the compound under consideration had an equal chance of escaping, the most similar worlds to our world in which a different animal escaped are likely to be worlds in which another zebra

[4] The zebra example incorporates some very helpful suggestions from an anonymous reviewer.

escaped. That is, on the similarity approach, the counterfactual expressed by (8) should be false in our world.

(8) If a different animal had escaped instead, it might have been a gazelle

However, I don't think that I would make a false claim if I uttered (8), given the circumstances described above. The fact that overall similarity with the animal that actually escaped seems to be irrelevant in this case, suggests that the similarity involved in counterfactual reasoning is not our everyday notion of similarity. It must then be a very special sort of similarity. In fact, this has been the usual reaction to examples of this kind (Lewis, 1979). Lewis and Stalnaker characterize the special sort of similarity relevant for counterfactuals in very general terms. There is nothing in their approach that would explain why it is that in our example, overall similarity with the actual zebra is not a concern. Note that it is not that the similarity theory says anything false about examples of this kind. It just doesn't say enough. It stays vague where our intuitions are relatively sharp. I think we should aim for a theory of counterfactuals that is able to make more concrete predictions with respect to particular examples.

4.2. *Truth-Conditions for Counterfactuals*

There is a very intuitive and appealing way of thinking about the truth-conditions for counterfactuals. It is an analysis, that in my heart of hearts, I have always believed to be correct (see Kratzer, 1978, 1981). Taken at face value, however, this analysis turns out to be so obviously wrong, that it doesn't seem to merit any serious attention. The analysis is this:

'Would'-counterfactuals
A 'would'-counterfactual is true in a world *w* if and only if every way of adding propositions which are true in *w* to the antecedent while preserving consistency reaches a point where the resulting set of propositions logically implies the consequent.

'Might'-counterfactuals
A 'might'-counterfactual is true in a world *w* if and only if not every way of adding propositions which are true in *w* to the antecedent while preserving consistency reaches a point where adding the consequent would result in an inconsistent set.

In a context where situations play such a prominent role, it may be surprising that we specified the conditions under which counterfactuals are true for *worlds*. Shouldn't we say something general about their truth in *situations*? I think we should indeed, but all we should say is that counterfactuals can only be true in situations which are worlds. This seems to be the most general way of insuring that we will not have any problems with the persistence constraint. (Insuring persistence in this way will give us a welcome consequence in connection with our final analysis of counterfactuals presented below. We will take up this point at the end of Section 4.4.)

For a given world w, both our definitions rely on the set of propositions which are true in w. In Section 3.2, we specified the set of propositions as the power set of the set of possible situations S. Do we really want to consider all true propositions in $\mathbb{P}(S)$ as relevant for the truth of a counterfactual? $\mathbb{P}(S)$ is likely to contain propositions that could not possibly form the contents of human thoughts. It seems plausible to assume that only humanly graspable propositions matter for the truth of a counterfactual. Constraining the range of propositions relevant for the truth of a counterfactual in this way may be plausible, but it doesn't seem to save our analysis. The analysis seems to be wrong anyway, plain wrong as shown in the following section.

4.3. *Where the Analysis Goes Wrong*

The first example

Suppose that in the actual world, Paula is buying a pound of apples. Besides that, nothing special is going on. The Atlantic Ocean isn't drying up, for example, nor is the moon falling down. The propositions expressed by the following sentences are then all true in our world:[5]

(9)a. Paula is buying a pound of apples
b. The Atlantic Ocean isn't drying up
c. The moon isn't falling down
d. Paula is buying a pound of apples or the Atlantic Ocean is drying up
e. Paula is buying a pound of apples or the moon is falling down

[5] The 'or' in (9d) and (9e) is to be understood as the truth-functional inclusive 'or' familiar from propositional logic. English 'or' can also have a generic reading (see Section 5).

Given these facts, our analysis of 'might'-counterfactuals predicts that the propositions expressed by (10a) and (10b) should both be true in our world:

(10)a. If Paula weren't buying a pound of apples, the Atlantic Ocean might be drying up
b. If Paula weren't buying a pound of apples, the moon might be falling down

That Paula isn't buying a pound of apples is logically compatible with the proposition expressed by (9d). The consequent of the counterfactual expressed by (10a) follows from the antecedent in conjunction with this proposition. We can now add any propositions whatsoever to this set, the consequent of the counterfactual will always follow from the resulting set. This means that there is a way of adding true propositions to the antecedent while preserving consistency such that there will never come a point where the addition of the consequent would yield an inconsistent set of propositions. Very similar considerations apply to the counterfactual expressed by (10b). The proposition expressed by (9e) is compatible with its antecedent. Its consequent follows from the antecedent in conjunction with this proposition. Again, the consequent will then follow from every superset of this set of propositions. Hence both counterfactuals are wrongly expected to be true in our world.

The second example[6]

Consider the following situation: Paula and Otto are the only persons in this room. They are both painters. They have a good friend, Clara, who actually is a sculptor. If some such story is true in our world, the propositions expressed by the following sentences will all be true in it:

(11)a. Paula is in this room
b. Otto is in this room
c. There are exactly two persons in this room
d. All persons in this room are painters

Given these facts, our analysis predicts that the counterfactual expressed by the following sentence should be true in our world:

(12) If Clara were also in this room, she might be a painter

That Clara is in this room is logically compatible with the proposition

[6] The example is inspired by Goodman (1947): 'All the coins in my pocket are silver'.

expressed by (11d). And the consequent of the counterfactual expressed by (12) follows from the antecedent in conjunction with this proposition. We expect, then, that the counterfactual expressed by (12) should be true in our world. This is not a welcome result, since the mere act of entering a room all filled with painters has no effect on your becoming a painter yourself.

4.4. *We Forgot About Lumps*

In constructing these two counterexamples to our apparently very plausible analysis, we made, I think, a capital mistake. We were trapped. We forgot that propositions never come alone. In trying to consistently add true propositions to the antecedents of our counterfactuals, we did not remember that whenever we add a proposition, it will bring along all the propositions that are lumped by it in the world under consideration. With this perspective in mind, let us examine our examples all over again.

The first example reexamined
Consider all the sentences involved:

(9)a. Paula is buying a pound of apples
b. The Atlantic Ocean isn't drying up
c. The moon isn't falling down
d. Paula is buying a pound of apples or the Atlantic Ocean is drying up
e. Paula is buying a pound of apples or the moon is falling down

(10)a. If Paula weren't buying a pound of apples, the Atlantic Ocean might be drying up
b. If Paula weren't buying a pound of apples, the moon might be falling down.

We argued that, according to our analysis, the counterfactual expressed by (10a) had to be true in our world, since we can consistently add the proposition expressed by (9d) to its antecedent, and the consequent follows from the resulting set and all its supersets. But once we add (9d) to the antecedent, it will bring along the proposition expressed by (9a). In our world, in which the Atlantic Ocean isn't drying up, every situation in which Paula is buying a pound of apples or the Atlantic Ocean is drying up is a situation in which Paula is buying a pound of apples. Hence the proposition expressed by (9a) is lumped by the proposition expressed by (9d) in our world. (9a), however is not compatible with the antecedent of the counterfactual expressed by (10a). Seen in this way, we cannot consistently add the proposition expressed by (9d) to the antecedent of our counterfactual. Exactly the same type of argumentation applies to the counterfactual expressed by (10b). Our analysis, then – if reformulated so as to conform to the lumping requirement – doesn't imply anymore that implausible counterfactuals like the ones expressed by (10a) and (10b) are true in our world.

The second example reexamined
Recall all the sentences involved:

(11)a. Paula is in this room
b. Otto is in this room
c. Exactly two persons are in this room
d. All persons in this room are painters

(12) If Clara were also in this room, she might be a painter.

We believed that our analysis predicts that the counterfactual expressed by (12) should be true in our world, since we can consistently add the proposition expressed by (11d) to its antecedent, and the consequent follows from the resulting set (and all its supersets). But as soon as we add this proposition, it will bring along other propositions. Given our scenario and the semantics for accidental universal quantification, the proposition expressed by (11d) can only be true in a situation s of our world if Paula and Otto are in this room in s. But then s will always be a situation in which the propositions expressed by (11a) and (11b) are true. Hence the proposition expressed by (11d) lumps the propositions expressed by (11a) and (11b) in our world. The proposition expressed by (11d) also lumps the proposition expressed by (11c) in our world, given our scenario and the accidental interpretations of 'all' and 'exactly two'. Every situation of our world that contains all persons in this room will be a situation that contains exactly two persons in this room. The propositions expressed by (11a), (11b), and (11c), then, are all lumped by the proposition expressed by (11d) in our world. But adding all of these propositions to the antecedent of the counterfactual expressed by (12), yields an inconsistent set of propositions.

Taking the idea of lumping seriously, enabled us to discard two representative counterexamples to our analysis. Let us now see how this analysis handles the zebra example introduced above.

The zebra example
Recall the story: A zebra escaped from the Hamburg zoo (call it "John"). The escape was caused by a negligent keeper who forgot to close the door of the compound housing zebras, giraffes, and gazelles. We supposed

counterfactually that some other animal had escaped instead and, in ruminating about what sort of animal it might have been, we wondered why similarity with the original zebra didn't play a role here. On the present approach, we have an explanation for this: if the actual properties of the zebra mattered, then this would be because the propositions expressed by the following sentences would have to be taken into account:

(13)a. A zebra escaped
b. A striped animal escaped
c. A black and white animal escaped
d. A male animal escaped

. . . .

Given lumping, none of these propositions can be consistently added to the antecedent of a conditional sentence of the form: "If the animal which escaped had not been John . . .". In our world (at the time considered), every situation in which a zebra escaped, is a situation in which John escaped. But it is also true that every situation in which a striped animal escaped is again a situation in which John escaped and so forth for all the properties of John. Hence in our world, the proposition that John escaped is lumped by the propositions expressed by sentences (13a) to (13d) above.

We have seen that in drawing conclusions from our counterfactual assumption, similarity with the actual zebra doesn't play a role. Other sorts of similarities with the actual world do matter, though. If a different animal had escaped instead of John, there would still be the forgetful keeper (call him "Carl") who left the door open. There would still be the night house for kiwis and owls. There would still be the cages for lions and tigers. And there would still be the monkey rock. Our analysis makes us expect this. Take the proposition 'Carl left the door to the compound housing zebras, giraffes, and gazelles open'. This proposition doesn't lump the proposition 'John escaped from the Hamburg zoo' in our world. Nor does it seem to lump any other dangerous proposition. And the propositions 'there is a night house for kiwis and owls in the Hamburg zoo', 'there are cages for lions and tigers in the Hamburg zoo', or 'there is a monkey rock in the Hamburg zoo' are likewise quite innocent lumpers. We can add them safely to the antecedent of our counterfactual. No inconsistency will be produced.

The above examples suggest that our original analysis of counterfactuals might be tenable after all if it is enriched by lumping as illustrated above. Note that on this analysis, counterfactuals themselves can never be added consistently to the antecedent of a counterfactual, except when the antecedent happens to be true. Here is why. Pushed by the persistence constraint, we assumed above that counterfactuals can only be true in situations that are worlds. This means that a counterfactual that is true in a world lumps all the true propositions of that world. Hence using a true counterfactual as an additional premise for the evaluation of a counterfactual with a false antecedent will always produce an inconsistency. The added counterfactual will lump the negation of the antecedent of the counterfactual that is being evaluated.

4.5. *An Interesting Asymmetry*

In this section, we will look at yet another example supporting the analysis of counterfactuals proposed above. Assume that among all the people in the room, Otto and you are the only ones who are bored. In a situation like this, the following counterfactual is likely to be true (suppose that there aren't, say, any regulations requiring that, at all times, exactly two bored persons have to be in the room . . . etc.):

(14) If you and Otto weren't in the room, nobody in the room would be bored

The following counterfactual, however, would probably *not* be true (suppose that Otto and you aren't *necessarily* bored . . . etc.):

(15) If nobody in the room were bored, Otto and you wouldn't be in the room

Our analysis predicts this asymmetry.[7] Let us look at the relevant facts:

(16)a. Otto and you are in the room
b. Otto and you are bored
c. Exactly two people in the room are bored
d. Paula is in the room
e. Clara is in the room
f. Rainer is in the room
g. Exactly three people in the room are neither Otto nor you
h. Paula is not bored
i. Clara is not bored
j. Rainer is not bored

The antecedent of (14) is incompatible with the proposition expressed by

[7] An anonymous reviewer pointed out that replacing 'bored' with 'fat' or 'tall' in (15) reverses the judgements. Note that 'bored' is a stage-level predicate, while 'tall' and 'fat' are individual-level predicates in the sense of Carlson (1977). Individual-level predicates seem to behave like non-accidental generalizations with respect to counterfactual reasoning (see Section 5; see also Kratzer, 1988).

(16a). Hence this proposition cannot be added consistently. The proposition expressed by (16c) lumps the proposition expressed by (16a), hence the proposition expressed by (16c) cannot be added consistently either. The antecedent of the counterfactual expressed by (14) and the propositions expressed by the remaining sentences collectively imply the consequent. This is not the whole story, of course. The truth of (14) depends on *everything* that is the case in the world under consideration (at the time considered). Other true propositions and sets of true propositions may give rise to new inconsistencies which we cannot even think of. This is all true. The hope with the present example is that we have indeed picked out the relevant premises, that is, that the neglected propositions will either not cause any more clashes or will at least be irrelevant in not adding any new aspects to our story. If everything goes well, the proposition expressed by (14) will come out true. That is, every way of adding true propositions to the antecedent while preserving consistency will eventually reach a point where the resulting set logically implies the consequent. Consider now the counterfactual expressed by (15). Its antecedent is incompatible with the proposition expressed by (16c). So this proposition cannot be added consistently. The same antecedent is compatible with the proposition expressed by (16a). And this proposition doesn't lump any dangerous proposition like (16c). What this means is that there is a way of adding propositions to the antecedent of the counterfactual expressed by (15) while preserving consistency such that the resulting set logically implies the negation of the consequent. But then the counterfactual expressed by (15) cannot be true. Instead, we predict the 'might'-counterfactuals expressed by the following two sentences to be true:

(17)a. If nobody in the room were bored, Otto and you might (still) be bored (without being in the room)

b. If nobody in the room were bored, Otto and you might (still) be in the room (without being bored)

The counterfactual expressed by (17a) is true since the proposition expressed by (16b) is compatible with its antecedent. The counterfactual expressed by (17b) is true since the proposition expressed by (16a) is compatible with its antecedent. In neither case are there any dangerous propositions which enter the picture through lumping.

4.6. *The Formal Definitions*

The preceding sections gave an intuitive idea of how a relatively simple analysis of counterfactuals plus a lumping mechanism can account for some interesting pieces of counterfactual reasoning. In this section, I want to present the formal definitions. This section is mainly intended for those readers who want to see a connection between the present proposal and the premise semantics for modals and counterfactuals developed in Veltman (1976), Kratzer (1977), Kratzer (1981), and Veltman (1985). (See also Lewis (1981) for a comparison of premise semantics with other approaches to counterfactuals.) The presentation in the remainder of the paper will then be informal again and most of it should be accessible without recourse to the official definitions presented below.

The first two definitions state what it means for a set of propositions to be closed under lumping or logical consequence. In view of future applications, we will adopt a weak notion of closure under lumping and a strong notion of closure under logical consequence.

> *Closure under lumping*
> A set of propositions $\mathbb{A}$ is (weakly) closed under lumping in a world w if and only if the following condition is satisfied for all $p \in \mathbb{A}$ and all $q \in \mathbb{P}(S)$: if p lumps q in w, then $q \in \mathbb{A}$.
>
> *Closure under logical consequence*
> A set of propositions $\mathbb{A}$ is (strongly) closed under logical consequence if and only if the following condition is satisfied for all $\mathbb{B} \subseteq \mathbb{A}$ and all $q \in \mathbb{P}(S)$: If $\mathbb{B}$ logically implies q, then $q \in \mathbb{A}$.

We have seen above that we might want to impose some very general constraints on the set of propositions that are relevant for the truth of counterfactuals in a world. (As a consequence, the two closure definitions above have to be relativized.) One such constraint is that only propositions that are true in a world w are relevant for the truth of a counterfactual in w. We also conjectured that probably only humanly graspable propositions are considered. Another conceivable constraint might be that only persistent propositions are admitted. Note that we are assuming here that for any w, the set of propositions relevant for the truth of a counterfactual in w can indeed be characterized by a handful of very general properties like 'true in w', 'humanly graspable', or 'persistent'. In particular, we are still excluding the possibility that individual properties of the counterfactuals considered or the context of use may affect the nature of this set. This is a very strong assumption that we will have to modify later. Any such modification should be carefully controlled, however, so let us stick with the strong assumption for the time being and gradually introduce the modifications as we go along. For any world w, then, a set $\mathbb{F}_w$ is tentatively defined as follows:

The set of propositions relevant for the truth of counterfactuals
For any $w \in W$, $\mathbb{F}_W = \{p \in \mathbb{P}(S)$: $w \in p$, p is graspable by humans, p is persistent . . . (further conditions to be discussed later)$\}$

The next definition defines a set that the truth conditions for counterfactuals will crucially rely on.

The crucial set
For any world $w \in W$ and proposition $p \in \mathbb{P}(S)$ let $\mathbb{F}_{w,p}$ be the set of all subsets $\mathbb{A}$ of $\mathbb{F}_w \cup \{p\}$ such that the following conditions are satisfied:
(i) $\mathbb{A}$ is consistent
(ii) $p \in \mathbb{A}$
(iii) $\mathbb{A}$ is (weakly) closed under lumping in w
(iv) $\mathbb{A} - \{p\}$ is (strongly) closed under logical consequence

The last two definitions give the truth conditions for 'would' and 'might' counterfactuals.

'Would'-counterfactuals
A 'would'-counterfactual with antecedent p and consequent q is true in a world w if and only if for every set in $\mathbb{F}_{w,p}$ there is a superset in $\mathbb{F}_{w,p}$ which logically implies q.

'Might'-counterfactuals
A 'might'-counterfactual with antecedent p and consequent q is true in a world w if and only if there is a set in $\mathbb{F}_{w,p}$ such that q is compatible with all its supersets in $\mathbb{F}_{w,p}$.

The above truth conditions for counterfactuals are the same as the ones in Kratzer (1981) except for the conditions concerning closure under lumping and logical consequence. Closure under logical consequence was simply superfluous as long as we didn't have closure under lumping. The King Ludwig example in Section 5.2 illustrates the need for closure under logical consequence in addition to closure under lumping. The same example also shows that we want a strong notion of closure under logical consequence, whereas a weak notion of closure under lumping seems to be sufficient.

5. Non-accidental generalizations. The nature of genericity

5.1. *What Generic Propositions Are*

All of the examples discussed so far illustrate the lumping properties of the propositions expressed by quantified sentences and disjunctions. We will see in what follows that these constructions are in fact ambiguous.They also have a meaning where the propositions expressed are extremely weak in lumping capacity-

Suppose that all of the following is true in our world:

(a) In New Zealand, either the Queen or the Governor-General opens Parliament. Last year, the Queen opened Parliament.
(b) In all Chinese restaurants, fortune cookies are served with the check.
(c) A king rules this country.

Consider now the counterfactuals expressed by sentences (18a) to (18c):

(18)a. If the Queen hadn't opened Parliament, the Governor-General would have done so.
b. If this (we are sitting in an Italian restaurant) were a Chinese restaurant, fortune cookies would be served with the check.
c. If this man (I am pointing at a picture of the current king) weren't the king, someone else would be.

I think all three sentences are true in the circumstances described above. Their truth seems to be mainly supported by the truth of the propositions expressed by sentences (19a) to (19c) respectively:

(19)a. In New Zealand, either the Queen or the Governor-General opens Parliament.
b. In all Chinese restaurants, fortune cookies are served with the check.
c. A king rules this country.

(19a) is a disjunction, (19b) is a universally quantified sentence and (19c) involves an existential quantifier. If we interpret these sentences as proposed in Section 3.5, we are in trouble. It will forever remain a mystery why the propositions expressed by them can support the truth of the corresponding counterfactuals. Take the New Zealand case. We are contemplating the 'last year'-slice of our world history. Every part of this slice in which the proposition expressed by (19a) (on the interpretation of disjunction given above) is true is a part in which the Queen opened the New Zealand Parliament. The proposition expressed by (19a), then, lumps a proposition which is incompatible with the antecedent of the counterfactual expressed by (18a). But this means that the proposition expressed by (19a) can never support the truth of the counterfactual expressed by (18a).

Let us now turn to Chinese restaurants. Nowadays, every situation of our world in which the proposition expressed by (19b) (on the accidental interpretation for universal quantification given above) is true is a situation in which it is true that fortune cookies are served in Panda Garden, Amherst Chinese, Sze's etc., for all the Chinese restaurants in the universe of discourse. And furthermore, every situation in which the proposition expressed by (19b) (under the accidental interpretation) is true is a situation in which there are exactly so many Chinese restaurants. But then the proposition expressed by (19b) lumps a set of propositions with which the antecedent of the counterfactual expressed by (18b) is not compatible.

Consider last the proposition expressed by (19c) (on the interpretation of existential quantification given above). At this very moment, every situation of our world in which this proposition is true is a situation in which the man who is our current king is the king. But the proposition that he is the king is incompatible with the antecedent of the counterfactual expressed by (18c).

There is a difference between sentences (19a) to (19c) on the one hand, and the cases of disjunction and quantification we have discussed before. On their most natural readings, sentences (19a) to (19c) express *non-accidental generalizations*. That all Chinese restaurants serve fortune cookies with the check is a non-accidental generalization of our world. That all people in this room are painters is just an accidental fact. Likewise, that a king rules this country is a non-accidental fact of our world, but not that a zebra escaped from the Hamburg zoo. And it is a non-accidental generalization that the Queen or the Governor-General opens the New Zealand Parliament, but it is a mere accidental fact that Paula is buying a pound of apples or the moon is falling down. Which truths of our world are accidental and which are not? I wish I knew (but see the remarks below). The problem that concerns us here is a different one. It has been thought to be almost as hard, however. The problem is what the ambiguity between generic and non-generic (or accidental) readings observed with disjunctions, quantifiers, and a great many other constructions consists in. This question is a question for semantics to answer. A possible answer was already implicit in the preceding discussion: There is an interpretation for disjunction and quantifiers which endows the propositions expressed by these constructions with strong lumping capacities. This is the *accidental* interpretation. But there is another interpretation for disjunction and quantifiers which produces propositions with very weak lumping properties. This is the *generic* interpretation. Our sentences (19a) to (19c) have to receive a generic interpretation if they are to support the counterfactuals (18a) to (18c). But what do these generic interpretations look like? We know already one candidate: The 'generic' truth conditions for universal quantification (Definition (D7) in Section 3.5). We have seen that this interpretation gives us propositions which don't lump very well. But this property seems to be precisely the property associated with genericity. If definition (D7) is used for the interpretation of (19b), the resulting proposition will not lump any facts contradicting the antecedent of (18b). Generic propositions are propositions which are true in all or none of the situations of a world. The generic interpretation of disjunction and existential quantification will now look as follows.[8]

(D9) *Disjunction* (*generic interpretation*)
For any variable assignment g:
$[\![\alpha \textit{ or } \beta]\!]^g$ is true in a situation $s \in S$ if and only if there is a situation $s' \in S$ such that $s \leq s'$ and $[\![\alpha]\!]^g$ is true in s' or $[\![\beta]\!]^g$ is true in s'.

(D10) *Existential quantification* (*generic interpretation*)
For any variable assignment g:
$[\![(\textit{There is an } x\text{: } \alpha)\beta]\!]^g$ is true in a situation $s \in S$ if and only if there is an x-alternative g' of g and a situation $s' \in S$ such that $s \leq s'$ and $[\![\alpha]\!]^{g'}$ and $[\![\beta]\!]^{g'}$ are both true in s'.

For a given sentence, its generic interpretation and its accidental interpretation will always agree on the possible *worlds* part of the propositions assigned (recall that propositions are sets of situations and some of these situations are worlds). Accidental generalizations and their non-accidental counterparts are logically equivalent, they only differ in the way their truth is distributed over the situations of a given world. As a consequence, they differ in lumping ability, a property which in turn affects their ability to support the truth of counterfactuals.

Consider the sentence "All current superpowers are referred to by abbreviations the first letter of which is *U*" (the example is almost Dahl's example in Dahl (1975).[9]) This sentence is ambiguous. It may express an accidental or a non-accidental generalization. On the present account, we want to say that the accidental and the non-accidental interpretation both yield propositions which are true in our world (assuming that the US and the USSR are the only current superpowers). There is nothing in our approach, however, that would force us to say that the propositions expressed by the sentences "it is a law that all current superpowers are referred to by abbreviations the first letter of which is *U*" and "it is a mere accident that all current superpowers are referred to by abbreviations the first letter of which is *U*" are also both true in our world. Nor do we

[8] Generic disjunction cannot be expected to solve all well-known problems with 'or'. See Landman (1986).

[9] Dahl's example was brought to my attention by an anonymous reviewer.

have to assume that in the actual world, the accidental and the non-accidental generalization are both members of the set of propositions that are relevant for the truth of counterfactuals. The non-accidental proposition will only qualify if certain standards having to do with lawhood are satisfied. One way of thinking about lawlike propositions is in terms of propositions occupying certain privileged roles within "some integrated system of truths that combines simplicity with strength in the best way possible" (Lewis 1986a, p. 122, see also Lewis 1973 reporting proposals by Ramsey). The generalization expressed by Dahl's sentence doesn't seem to satisfy the standards for lawhood in our world. The set of propositions relevant for the truth of counterfactuals in the actual world, then, will only contain the accidental generalization. But then we correctly predict that counterfactuals of the sort "If Andorra were a current superpower, it would be referred to by an abbreviation the first letter of which is *U*" are false in our world.

The distinction between accidental and non-accidental generalizations is certainly not always a sharp one. This does not mean, however, that there are no objective standards involved. People agree to a great deal as to which facts are and which ones aren't accidental once we discard all doubts concerning matters of truth. The remaining unclear cases partly account for the vagueness of counterfactuals. That counterfactuals are vague is one of their characteristic properties which any analysis of counterfactuals has to account for (see e.g., Lewis (1973); Kratzer (1981)). On the present account, the accidental/non-accidental distinction is assumed to be a major source of vagueness (see Section 5.2 for concrete examples).

Generic propositions are true in all or none of the situations of a world, and this is what makes them such poor lumpers. The only propositions they are able to lump are other generic propositions. They will never bring along any accidental facts. While generic propositions are weak lumpers, they are strong lumpees. Whenever a proposition is true in a world, it will lump all the non-accidental generalizations of that world. We are now in the position to explain why in counterfactual reasoning, non-accidental generalizations have priority over accidental facts.

> *The lobster example*
> Watch this lobster crawling on the bottom of the ocean. It is all blackish green. Suppose now counterfactually that it were in boiling water. What color would it be? Might it still be blackish green?

The proposition that it is all blackish green is logically compatible with the counterfactual assumption. But given lumping, this proposition will bring along the non-accidental generalizations of our world. There is a non-accidental generalization stating that all lobsters which are in boiling water are red (suppose they change color instantly). We now have an inconsistent set. The non-accidental generalization, however, can be added to the antecedent without causing any trouble. The lobster would be red.

5.2. *Goodman's Puzzle*

The following example is a slightly simplified version of Goodman's famous match example (Goodman 1947). Consider the following scenario: King Ludwig of Bavaria likes to spend his weekends at Leoni Castle. Whenever the Royal Bavarian flag is up and the lights are on, the King is in the Castle. At the moment, the lights are on, the flag is down, and the King is away. Suppose now counterfactually that the flag were up. Well, then the King would be in the Castle and the lights would still be on. But why wouldn't the lights be out and the King still be away? This is Goodman's puzzle. Here are all the propositions involved:

(20)a. Whenever the flag is up and the lights are on, the king is in the castle.
b. The flag is down.
c. The lights are on.
d. The king is away.

Our counterfactual assumption is expressed by (21):

(21) The flag is up.

There seem to be two ways of consistently adding propositions to the proposition expressed by (21):

Possibility 1:	*Possibility* 2:
(21)	(21)
20(a)	20(a)
20(c)	20(d)

(20a) expresses a non-accidental generalization of the world under consideration. So (20a) will be interpreted generically. We haven't yet proposed interpretation rules for sentences like (20a). Yet we know what it means that (20a) will receive a generic interpretation. It will express a proposition which is true in all or none of the situations of a world. Hence it will be lumped by every true proposition. This means that we *have to* add the proposition expressed by (20a) to the antecedent of our counterfactual as long as we add any proposition at all.

Possibility 1 implies that the king is in the castle and the lights are on. Possibility 2 implies that the king is away and the lights are out. It looks as if both possibilities could be realized. But wait. Let's take a closer look at possibility 2. The propositions expressed by 20(a) and 20(d) jointly imply the proposition expressed by (22) (never mind that the flag may be neither up nor down):

(22) The flag is down or the lights are out

Assuming strong closure under logical consequence, we have to add (22) to the list illustrating possibility 2. But (22) lumps (20b) in the world under consideration. Now we have an inconsistency. Possibility 2 is discarded. Nothing dangerous seems to happen with possibility 1. King Ludwig would be in the castle.

Let us change the scenario just a little bit. Everything stays the same except now you and I are passing by the castle. Seized by some childish inclination, I say to you: "Suppose I hoisted the flag" . . . the consequences could be dramatic. Would my hoisting the flag bring the King back into the castle? No. The counterfactual expressed by (23) is false.

(23) If I hoisted the flag, the king would appear in the castle.

The counterfactual discussed before was mainly supported by the non-accidental generalization expressed by (20a). This generalization is just a generalization about the habits of the King and his staff. It has been true so far that whenever the flag was up and the lights were on, the King was in the Castle. In addition, the King made it public that the flag and the lights would be signs of his presence. The King's staff is absolutely reliable. Not a single violation so far. Yet I could destroy the regularity with a single action. Our second scenario suggests that this may indeed be my intention.

What was treated like a non-accidental generalization before has now been demoted to a simple accidental one. The relevant facts are as follows:

(24)a. Every time when the flag is up and the lights are on is a time when the king is in the castle.
b. The flag is down.
c. The lights are on.
d. The King is away.

(24a) now expresses an accidental generalization. Even though we are not really prepared to talk about matters of time, we can nevertheless sketch why the proposition expressed by (24a) cannot be added consistently to the antecedent of the counterfactual expressed by (23). On the accidental interpretation, the proposition expressed by (24a) will only be true in a situation which is big enough to contain e.g., all the occasions at which the flag was or will be up. That is, the proposition expressed by (24a) will lump propositions like the ones expressed by the following sentences:

(25) On occasion 1, the flag is up
On occasion 2, the flag is up
On occasion 3, the flag is up
. . . .
On occasion 115, the flag is up
The flag is up on exactly 115 occasions

The antecedent of the counterfactual expressed by (23) would add yet another occasion to this collection which would lead to an inconsistency. This means that among the propositions mentioned, only the ones expressed by (24c) and (24d) can be consistently added to the antecedent of our counterfactual. But then the flag would be up, the lights would be on, and the king would still be away. Non-accidental generalizations concerning natural phenomena are much more robust. Let us take one of the popular switch examples (Pollock 1984, p. 119): We have an open switch and a light wired in series with a battery. Whenever the switch is closed and the circuit is intact, then shortly thereafter the light will come on. At the moment, the light is out and the circuit is intact. Suppose now counterfactually that I closed the switch. Would the light come on (with the circuit still being intact) or would the circuit be defective (with the light still being out)? The example has exactly the same structure as our first King Ludwig example. It can be given exactly the same analysis. Yet I don't think that we can as easily demote the non-accidental generalization to a merely accidental one. The regularities in the life of a king can be violated much more easily than the laws of nature.

5.3. *The Linguistic Representation of Genericity*

So far, we have been dealing with genericity as an inherent property of quantifiers and connectives like 'or'. This may be the right approach in view of lexical ambiguities like 'every' versus 'any' (see Vendler, 1962). 'Every' would be associated with the accidental interpretation of the universal quantifier, 'any' would receive the generic interpretation. As a consequence, we predict that "any doctor will tell you what to do" can, but "every doctor will tell you what to do" cannot support a counterfactual like "if I were a doctor I would tell you what to do".[10] On the other hand, there is a very systematic connection between accidental and generic interpretations suggesting that generic interpretations should be derived from accidental interpretations with the help of a non-overt generic operator. Suppose we are given only accidental interpretations. Genericity will now arise through the effect of a one-place sentence operator G which is interpreted in the following way:

(D11) *Generic Operator*
For any variable assignment g:
$[\![G(\alpha)]\!]^g$ is true in a situation $s \in S$ if and only if there is a situation $s' \in S$ such that $s \leqslant s'$ and $[\![\alpha]\!]^g$ is true in s'.

The operator G acts like a possibility operator in modal logic. The reader can easily verify that if placed appropriately, G interacts with the accidental interpretations for universal and existential quantification and disjunction in the desired way. Hidden generic operators with various properties have been stipulated by Carlson (1977), Farkas and Sugioka (1983), Heim (1982) and Wilkinson (1986). Further research will have to investigate whether the conclusions reached there are compatible with the present proposal.

6. Negation

6.1. *In Search of an Accidental Interpretation*

In many ways, negation confronts us with the same sorts of considerations as universal quantification. Try to formulate the truth-conditions for negated sentences in the way familiar from classical logic. The result will be a definition like (D12).

(D12) *Non-persistent negation*
For any variable assignment g:
$[\![\textit{not}\ \alpha]\!]^g$ is true in a situation $s \in S$ if and only if $[\![\alpha]\!]^g$ is not true in s.

If you believe in the persistence constraint, you will have to reject (12). Let us look at an example. Suppose the Hampshire Gazette has been delivered today and is lying on the kitchen table. The proposition 'a paper is lying on a table' is then true in our world right now. It is also true in

[10] This example is due to Vendler. Dick Oehrle is credited with similar examples. Barbara Partee p.c.

some of its parts, but there are other parts in which it is not true. (Note that I have never used the word "false" so far. Saying that a proposition is 'false' in a situation immediately suggests that its negation is true in that situation. This is usually not what I mean when I talk about propositions being 'not true', however.) Consider that part of our world which consists of the sink and nothing else. Clearly, the proposition 'a paper is lying on a table' is not true over there. Definition (D12) tells us that in this situation, the proposition expressed by sentence (26)

(26) There isn't a paper lying on a table.

should be true. This proposition is not true in our world, however, since in it, the Hampshire Gazette is lying on the kitchen table. Hence according to definition (D12), sentence (26) expresses a proposition which can be true in a part of a world without being true in the world itself. This means that it expresses a non-persistent proposition. Definition (D12), then, will be ruled out by the persistence constraint. Next, let us try a Kripke style definition. This will give us the generic interpretation given in (D13).

(D13) *Generic negation*
For any variable assignment g:
$[\![not\ \alpha]\!]^g$ is true in a situation $s \in S$ if and only if for all $s' \in S$ such that $s \leqslant s'$: $[\![\alpha]\!]^g$ is not true in s'.

(D13) gives rise to propositions which are true in all or none of the situations of a world. We have seen that this is the characteristic property of generic propositions. Negated sentences *can* certainly have generic interpretations as when I claim "Cats don't bark". But they don't *have to* be generic. Consider the most natural interpretations of sentences like "I am not asleep", "I am not hungry", "I am not with you". What we are looking for, then, is an accidental interpretation for negation. Recall the basic characteristics of such an interpretation for universal quantification: In order to obtain strong lumping properties, we made sure that the propositions assigned could only be true in situations which were big enough to satisfy a certain condition. The condition was supplied by the restrictive clause of the quantifier construction. Unfortunately, we do not have such a restrictive clause in the case of negative sentences. Well, I am not so sure whether that's correct. Maybe negative sentences come with restrictive clauses after all. And here is why I think so: It has often been noted that there is a connection between the interpretation of negation and focus (see, e.g., Jackendoff, 1972). What is important for us is that on the analysis I have in mind, the semantic interpretation of focus constructions involves a syntactic procedure isolating the unfocused part from the rest of the sentence. In one way or other, some such procedure is at the heart of a variety of approaches to focus (Jackendoff, 1972; von Stechow, 1981; Cresswell and von Stechow, 1982; Jacobs, 1982, 1983; and Kadmon and Roberts, 1986; but see also Rooth, 1985). In what follows, I won't be able to discuss the details of these highly relevant proposals. My main concern here is to suggest that we should quite generally conceive of negation as an operator which is intimately connected to focus. And that Logical Form representations very much like the ones assumed by most scholars working on focus constructions may be just the sort of representations we need in order to endow negative sentences with interpretations exhibiting the desired lumping properties.

6.2. *Negation and Restrictive Clauses*

Recall that we are trying to obtain accidental interpretations for negative sentences. We have seen that the main feature of such an interpretation is a condition on situations provided by the restrictive clause of something like a quantifier construction. At the end of the preceding section, I expressed some hope that we might be able to obtain the appropriate logical form representations once we recognize the intimate relationship between negation and focus. In English, focused constituents can be signalized phonetically by the presence of a pitch accent and syntactically by means of cleft constructions. Usually, these tools are not sufficient to determine an unambiguous focus assignment to a sentence (see Jackendoff, 1972; Selkirk, 1984 for extensive discussion). For ease of exposition I am going to neglect these problems here. In what follows I am going to assume that we are given a focus assignment for the sentences of our language. I will use capital letters to indicate what the intended focused constituents are (regardless of what the actual means for focusing may have been). Consider now representations like (27) and (28).

(27) Paula isn't registered in PARIS.
(28) PAULA isn't registered in Paris.

Preserving the spirit of previous analyses of focus while emphasizing the similarity with restricted quantifier structures, we are led to the following logical forms for (27) and (28):

(27′) (*Not*: *x is a place and Paula is registered in x*)
Paula is registered in Paris.
(28′) (*Not*: *x is a person and x is registered in Paris*)
Paula is registered in Paris.

In (27′) and (28′), the restrictive clause corresponds to the unfocused part of the sentence. In addition to the material contained in the original sentences, the restrictive clause contains some sortal for the variable to be provided by the context of use. I propose the following truth-conditions for structures of this sort:

(D14) *Accidental Negation*
For all variable assignments g:
$[\![(Not\colon \alpha)\beta]\!]^g$ is true in a situation $s \in S$ if and only if the following two conditions hold:
(i) For all x-alternatives g' of g: Whenever $[\![\alpha]\!]^{g'}$ is true in w_s, then $[\![\alpha]\!]^{g'}$ is true in s.
(ii) $[\![\beta]\!]^g$ is not true in s.

On definition (D14), the proposition expressed by (27) can be true in a situation s only if whenever there is a place such that Paula is registered at this place in the world of s, then Paula is registered at this place in s. And the proposition expressed by (28) can be true in a situation s only if whenever there is a person such that this person is registered in Paris in the world of s, then the same person is registered in Paris in s.

Definition (D14) treats negation syntactically very much like an ordinary quantifier. It implies that there is no such thing as a one place negation operator. Every negation operator has a restrictive clause which results from the original clause by replacing the focused phrase by an appropriate variable. In richly typed languages like Cresswell's lambda categorial language (Cresswell, 1973) or Montague's intensional logic (Montague, 1974), we are given variables for each syntactic category, and we are given variable assignments assigning appropriate entities to these variables. If we have a variable of category S, for example, then every admissible variable assignment assigns a proposition to it. Consider now the extreme case where a whole sentence is focused. We will then have logical forms of the following kind:

(29) $(Not\colon x_S)\beta$

The proposition expressed by sentences of this form will only be true in a situation s when all the propositions which are true in the world of s are true in s. But this means that the propositions expressed by these sentences can only be true in worlds. It has frequently been observed that sentences corresponding to our representations (27) and (28)

(27) Paula isn't registered in PARIS.
(28) PAULA isn't registered in Paris.

differ with respect to the presuppositions associated with them. Sentence (27) presupposes that Paula is registered at some place which is not Paris. And sentence (28) presupposes that some person who is not Paula is registered in Paris. These presuppositions are typical for certain quantifier constructions. Quantified sentences often presuppose that their restrictive clauses are satisfied by something. Take sentence (30).

(30) All ghosts in this house bothered us last night
(For all x: is a ghost and x is in this house) x bothered us last night.

(30) presupposes that there are ghosts in this house. The 'quantifier' approach to negation, then, allows us to capture the presuppositions associated with negation and universal quantification in a unified way.

If negation is quite generally connected to focus as suggested here, the logical form of negated sentences is not as simple a matter as hitherto assumed in the logical tradition. Sentence negation and verb phrase negation will be special cases out of a much broader range of possibilities. Work on negation will have to pay close attention to work on focus and intonation, an area which is under active investigation in linguistics. Naturally, the above thoughts, while suggesting a possible road for further research, could not do justice to the full range of questions involved.

In this section, we have been looking at a promising candidate for an accidental interpretation of negation. What we haven't seen yet is whether this interpretation endows the propositions expressed by sentences involving negation with the correct lumping properties. The following section will examine this question.

6.3. *Negation and Counterfactual Reasoning*

You have probably heard of Clyde, a sweet young man who eventually married Bertha (Dretske, 1972). Suppose now counterfactually that he hadn't married BERTHA. Might he have married somebody else? If there weren't such a thing as lumping, life would look grim for Clyde. In the actual world, he didn't marry CATHERINE, for example. The proposition expressed by

(31) Clyde didn't marry CATHERINE
(Not: Clyde married x_N *and* x_N *is a woman) Clyde married Catherine.*

can be consistently added to the antecedent of our counterfactual and the proposition that Clyde didn't marry CATHERINE follows from the resulting set. Without lumping, it would be quite likely that each and every way of adding propositions which are facts of this world to the antecedent of our counterfactual while preserving consistency would eventually reach a point where the resulting set implies the proposition that Clyde didn't marry CATHERINE. Hence Clyde wouldn't marry CATHERINE. Clyde wouldn't have more luck with other women. Analogous reasoning would show that he wouldn't marry EVANGELINE, he wouldn't marry GUINEVERE, he wouldn't marry ISOLDE, he wouldn't marry KIRI, he wouldn't marry MIRIAM, he wouldn't marry OLGA, he wouldn't marry QUILLA, he wouldn't marry STEPHANIE, he wouldn't marry URSULA, he wouldn't marry WILHELMINA, he wouldn't marry YVONNE, . . . (Cresswell, 1981). But fortunately, there is lumping. Every situation of our world in which the proposition expressed by (31) is true is a situation in which it is also true that Clyde married Bertha (assuming that (31) is given an accidental interpretation and that Clyde married only Bertha). Hence the proposition expressed by (31) lumps the proposition that Clyde married Bertha in our world. As a consequence the proposition expressed by (31) cannot be added consistently to the antecedent of a counterfactual of the form "If Clyde hadn't married Bertha . . .". But then, it seems, Clyde might have married Catherine, he might have married Evangeline, he might have married Guinevere, and all the rest, provided only that there weren't other facts preventing this (supposing that no man marries his sister in our world, and this is a non-accidental generalization, Clyde wouldn't have married his sister, for example).

At this point, you may wonder what will happen with a proposition like the one expressed by the following sentence.

(32) Clyde didn't MARRY Catherine.
(*Not*: *Clyde* x_V *Catherine*) *Clyde married Catherine.*

The proposition expressed by (31) and the proposition expressed by (32) are logically equivalent. They are true in exactly the same possible worlds. In particular, they are both true in the actual world. But unlike the proposition expressed by (31), the proposition expressed by (32) does not lump the fact that Clyde married Bertha in the actual world. If we try to add the proposition expressed by (32) to the antecedent of a counterfactual like the one expressed by (33), no inconsistency is likely to arise via lumping.

(33) If Clyde hadn't married BERTHA, he might have married Catherine.

But then the counterfactual expressed by (33) might very well come out false (the exact outcome depends on the complete array of facts, of course). Similar points could be made with respect to all the other women in the list. But then Clyde would be likely to stay unmarried after all.

The example of Clyde and Bertha suggests that foregrounding and backgrounding of information may sometimes play a role in selecting or rejecting a proposition as relevant for the evaluation of a given counterfactual. In our case, we would have to assume that the proposition expressed by (31) is selected, while the logically equivalent proposition expressed by (32) is rejected. (If w is the actual world, then the proposition expressed by (31) but not the proposition expressed by (32) will be in the set $\mathbb{F}_w$ defined in Section 4.6.) It is actually not difficult to see why this should be so. The focus structure of the antecedent of (32) conveys that we are discussing alternatives as to who Clyde might have married. The focus structure of (31), again, conveys that we are discussing precisely those alternatives. This information is not only encoded by the restrictive clause of the logical form of (31). It is also present in the corresponding proposition. The proposition expressed by (31) can only be true in a situation s if s contains all the women who Clyde married in the world of s. In contrast, the focus structure of (32) conveys that we are discussing alternatives as to what Clyde might have done with respect to Catherine. Again, this information is retrievable from the restrictive clause of the logical form of (32), as well as from the proposition expressed. We may assume, then, that sometimes, the focus structure of a counterfactual sentence may lead to the rejection of a 'non-matching' proposition as relevant for its evaluation. The exact mechanism of this process will have to be explored in future work.

Let us turn to another topic. You probably remember those lines which might have been longer or shorter than they actually are (Lewis, 1973). Take this clothesline here. It is thirty feet long. Suppose now counterfactually that it were longer than it actually is. How long might it be? I think it might be 35 feet long, for example. On Lewis' account, that couldn't be. Given his interpretation of 'might'-counterfactuals (Lewis, 1973, p. 21), the proposition expressed by (34)

(34) If this clothesline were longer than it actually is, it might be 35 feet long

could only be true in a world if there is no world in which the line is longer than it actually is and is not 35 feet long which is closer to the actual world than any world in which the line is longer than it actually is and is 35 feet long. But there are plenty of such worlds: All the worlds in which the clothesline is longer than 30 feet but shorter than 35 feet. A proponent of the similarity theory would now have to argue that for some reason, trying to stay as close as possible to the actual length of the line isn't a consideration guiding the evaluation of counterfactuals of this sort. But why should this be so? On the present account, we can explain this rather puzzling fact. The explanation is similar to the one we gave in the case of the zebra. If closeness to the actual length of the line were to play a role in the evaluation of counterfactual sentences of the form "If this clothesline were longer than it actually is . . .", then this would mainly be due to the presence of a host of relevant negative facts like the ones expressed by sentences of the following kind (let us forget about units smaller than feet):

(35) (Not: x_N is a number and this clothesline is longer than x_N feet) this clothesline is longer than 31 feet
(Not: x_N is a number and this clothesline is longer than x_N feet) this clothesline is longer than 32 feet
(Not: x_N is a number and this clothesline is longer than x_N feet) this clothesline is longer than 33 feet
. . . .

The propositions expressed by the sentences in (35) will be true in all those situations of our world in which this clothesline is present and is longer than 1, 2, . . . , 29 feet . But all of those situations will be situations in which this clothesline is actually 30 feet long. Hence the propositions expressed by the sentences in (35) all lump the proposition that this clothesline is 30 feet long in our world. Adding any of those propositions to the antecedent of the counterfactual expressed by (34), then, will always lead to an inconsistency.

All the examples we have examined in this section were examples where negative propositions had to be eliminated. They were in the way and had to be knocked out by lumping. Or else they had to be considered as irrelevant for the evaluation of the counterfactual under consideration. Doesn't this suggest that, maybe, negative propositions are *never* relevant for the evaluation of a counterfactual? Why should we admit negative propositions as relevant only to eliminate them later through lumping? Are there ever any occasions when negative propositions are crucial for the truth of a counterfactual? I don't know about 'crucial', but I do know that there are pieces of counterfactual reasoning where negative propositions enter at least naturally. We had an example in Section 4.5. Here is another example: I am not wearing MY GLASSES right now. Suppose counterfactually that I tried to read the sign over there. I couldn't do it. My eyes are bad and I am not wearing GLASSES, and I am not wearing CONTACT LENSES, and I am not . . . . And there are laws stating under which circumstances a person like me can read signs which are that far away. It seems, then, that sometimes, we do want to use negative propositions in counterfactual reasoning. A proposition like the one that I am not wearing GLASSES will bring along other propositions: that I am wearing earrings, that I am not wearing SHOES, that I am wearing a sweater, that I am not wearing A SCARF, ... All these propositions will peacefully join the antecedent of our last counterfactual. If I tried to read the sign over there, I couldn't do it.

Conclusion

The above thoughts addressed a number of topics: the semantics of situations, counterfactual reasoning, quantification, negation, focus, genericity. Our discussion also left a number of questions open for further research. I have said very little about the linguistic representation of genericity, for example. I neglected all important issues of time. And I didn't examine counterfactuals with antecedents contradicting some non-accidental generalization. One way of thinking about these counterlegals is to take them as not only denying the *truth* but also the *non-accidental* status of their antecedents. Counterlegals may be the most explicit way of demoting a law to a mere accidental generalization.

If a theory of counterfactuals of the sort presented here is correct, then all the semantic complexities of natural languages could potentially add to the complexity of counterfactual reasoning: The truth-conditions of particular constructions lead to particular lumping properties of the propositions expressed. These properties in turn determine the role of those propositions in the evaluation of a piece of counterfactual reasoning. This means that the investigation of counterfactuals could give us invaluable insights into the semantics of natural languages. But it also means that it is not a topic which can be quickly settled once and forever.

September 1986
Revised July 1987, January 1988, September 1988

References

Armstrong, David M.: 1978, *Nominalism and Realism. Universals and Scientific Realism*, Cambridge University Press, Cambridge.

Bach, Emmon: 1986, 'Natural Language Metaphysics', in R. Barcan Marcus et al. (eds.), *Logic, Methodology and Philosophy of Science VII*, Elsevier, Amsterdam, pp. 573–95.

Barwise, Jon and John Perry: 1983, *Situations and Attitudes*, MIT Press, Cambridge, Massachusetts.
Boole, George: 1854, *An Investigation of the Laws of Thought*, reprint, Dover, New York, 1958.
Carlson, Gregory: 1977, *Reference to Kinds in English*, dissertation, University of Massachusetts.
Chomsky, Noam: 1981, *Lectures on Government and Binding*, Foris, Dordrecht.
Cresswell, Maxwell: 1973, *Logics and Languages*, Methuen, London.
Cresswell, Maxwell: 1981, 'Adverbs of Causation', in H.-J. Eikmeyer and H. Rieser (eds.), *Words, Worlds and Contexts*, de Gruyter, Berlin.
Cresswell, Maxwell and Arnim von Stechow: 1982, 'De Re Belief Generalized', *Linguistics and Philosophy* **5**, 503–35.
Cushing, Stephen: 1976, *The Formal Semantics of Quantification*,unpublished dissertation, University of California at Los Angeles.
Dahl, Östen: 1975, 'On Generics', in E. Keenan (ed.), *Formal Semantics of Natural Language*, Cambridge University Press, Cambridge.
Dretske, Fred: 1972, 'Contrastive Statements', *Philosophical Review* **81**, 411–37.
Farkas, Donka and Yoko Sugioka: 1983, 'Restrictive If/When Clauses', *Linguistics and Philosophy* **6**, 225–58.
Goodman, Nelson: 1947, 'The Problem of Counterfactual Conditionals', *Journal of Philosophy* **44**, 113–28.
Heim, Irene: 1982, *The Semantics of Definite and Indefinite Noun Phrases*, unpublished dissertation, University of Massachusetts.
Jackendoff, Ray: 1972, *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, Massachusetts.
Jacobs, Joachim: 1982, *Syntax und Semantik der Negation im Deutschen*, Fink, München.
Jacobs, Joachim: 1983, *Fokus und Skalen. Zur Syntax und Semantik der Gradpartikeln im Deutschen*, Niemeyer, Tübingen.
Kadmon, Nirit and Craige Roberts: 1986, 'Prosody and Scope: The Role of Discourse Structure', in A. M. Farley, P. Farley and K.-E. McCullough (eds.), *Papers from the Parasession on Pragmatics and Grammatical Theory*, Chicago Linguistic Society.
Kaplan, David: 1977, *Demonstratives*, unpublished manuscript.
Kratzer, Angelika: 1977, 'What "Must" and "Can" Must and Can Mean', *Linguistics and Philosophy* **1**, 337–55.
Kratzer, Angelika: 1978, *Semantik der Rede*, Scriptor, Königstein.
Kratzer, Angelika: 1981, 'Partition and Revision: The Semantics of Counterfactuals', *Journal of Philosophical Logic* **10**, 201–16.
Kratzer, Angelika: 1988, 'Stage-Level and Individual-Level Predicates', paper presented at the Tübingen Conference on Genericity, August 1988.
Kripke, Saul: 1965, 'Semantical Analysis of Intuitionistic Logic I'. in J. N. Crossley and M. A. E. Dummett (eds.), *Formal Systems and Recursive Functions*, North-Holland, Amsterdam.
Landman, Fred: 1986, *Towards a Theory of Information. The Status of Partial Objects in Semantics*, Foris, Dordrecht.
Lewis, David K.: 1968, 'Counterpart Theory and Quantified Modal Logic', *Journal of Philosophy* **65**, 113–26.
Lewis, David K.: 1973, *Counterfactuals*, Basil Blackwell, Oxford.
Lewis, David K.: 1979, 'Counterfactual Dependence and Time's Arrow', *Nous* **13**, 455–76.
Lewis, David K.: 1980, 'Index, Context, and Content', in Stig Kanger and Sven Öhmann (eds.), *Philosophy and Grammar*, Reidel, Dordrecht.
Lewis, David K.: 1981, 'Ordering Semantics and Premise Semantics for Counterfactuals', *Journal of Philosophical Logic* **10**, 217–34.

Lewis, David K.: 1986, *On the Plurality of Worlds*, Basil Blackwell, Oxford.

Lewis, David K.: 1986a, *Philosophical Papers*, Vol. II, Oxford University Press, Oxford.

McCawley, James: 1981, *Everything that Linguists have always Wanted to Know about Logic but Were Ashamed to Ask*, The University of Chicago Press, Chicago.

Montague, Richard: 1984, *Formal Philosophy: Selected papers of Richard Montague*, Yale University Press, New Haven.

Pollock, John L.: 1984, *The Foundations of Philosphical Semantics*, Princeton University Press, Princeton, New Jersey.

Rooth, Mats: 1985, *Association with Focus*, unpublished dissertation, University of Massachusetts.

Selkirk, Elisabeth O.: 1984, *Phonology and Syntax, The Relation between Sound and Structure*, MIT Press, Cambridge, Massachusetts.

Stalnaker, Robert: 1968, 'A Theory of Conditionals', in N. Rescher (ed.), *Studies in Logical Theory*, Basil Blackwell, Oxford.

Stalnaker, Robert: 1972, 'Pragmatics', in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, Reidel. Dordrecht.

von Stechow, Arnim: 1981, 'Topic, Focus and Local Relevance', in W. Klein and W. Levelt (eds.), *Crossing the Boundaries in Linguistics*, Reidel, Dordrecht.

Veltman, Frank: 1976, 'Prejudices, Presuppositions and the Theory of Counterfactuals', in J. Groenendijk and M. Stokhof (eds.), *Amsterdam Papers in Formal Grammar*, Vol. 1.

Veltman, Frank: 1985, *Logic for Conditionals*, unpublished dissertation, University of Amsterdam.

Vendler, Zeno: 1962, 'Each and Every, Any and All', *Mind* **71**, 145–60.

Wilkinson, Karina: 1986, 'Generic Indefinite NPs', unpublished manuscript, University of Massachusetts at Amherst.

Dept. of Linguistics
University of Massachusetts
Amherst, MA 01003
U.S.A.