DAVID MAKINSON

HOW TO GIVE IT UP: A SURVEY OF SOME FORMAL ASPECTS OF THE LOGIC OF THEORY CHANGE

ABSTRACT. The paper surveys some recent work on formal aspects of the logic of theory change. It begins with a general discussion of the intuitive processes of contraction and revision of a theory, and of differing strategies for their formal study. Specific work is then described, notably Gärdenfors' postulates for contraction and revision, maxichoice contraction and revision functions and the condition of orderliness, partial meet contraction and revision functions and the condition of relationality, and finally the operations of safe contraction and revision. Verifications and proofs are omitted, with references given to the literature, but definitions and principal results are presented with rigour, along with discussion of their significance.

1. INTRODUCTION

1.1. *Consequence Operations*

Logic has always been concerned with the business of inference – the passage from a given stock of propositions to others which they support or imply. And the study of deductive logic, where the implication envisaged is one of necessity, gives rise to the very useful concept of a *consequence operation*. Introduced by Tarski in 1928, this concept is helpful in providing a framework in which to do "universal logic", that is, to discuss problems that arise not only for classical logic, but also for many systems that in one way or another extend, restrict, or modify classical logic. A consequence operation is any operation Cn that takes sets of propositions to sets of propositions, such that three conditions are satisfied: $A \subseteq Cn(A)$ for any set A of propositions, known as *inclusion*; $Cn(A) = Cn(Cn(A))$, or *iteration*; and $Cn(A) \subseteq Cn(B)$ whenever $A \subseteq B$, or *monotony*. Mathematicians will recognize this as simply the concept of a closure operation, in the sense of Kuratowski, on sets of propositions. Observe that this notion is extremely general; there is no reference to the properties of specific logical operators such as $\neg, \wedge, \vee, \supset, \forall, \exists, \ldots$ and so it is applicable to just about any deductive logic, classical or nonclassical. It is not, however, useful for inductive logic, or any kind of logic dealing with degrees of support, for the third condition, of monotony, fails for such inferences. The addition of

Synthese **62** (1985) 347–363. 0039–7857/85.10

premises to a set A, forming a larger set B, notoriously does not always increase or even preserve the *degree* of support or credibility of a conclusion, and in some cases may reduce the support so drastically as to transform an acceptable inference into a quite unacceptable one. But our concern here is with *deductive* relationships. We shall make use of the concept of a consequence operation Cn, saying also A *implies* x and writing $A \vdash x$ to mean $x \in Cn(A)$, for any given consequence operation Cn. We shall assume that Cn includes classical consequence, and that it satisfies a couple of additional conditions: first, that it is compact, i.e., $y \in Cn(A')$ for some finite subset A' of A whenever $y \in Cn(A)$; and second, that it satisfies the rule of "introduction of disjunction in the premises", i.e., whenever $y \in Cn(A \cup \{x_1\})$ and $y \in Cn(A \cup \{x_2\})$ then $y \in Cn(A \cup \{x_1 \vee x_2\})$. As is customary, we shall use the term *theory* to indicate any set X of propositions that is closed under Cn, i.e., such that $X = Cn(X)$, or equivalently, such that $X = Cn(Y)$ for some set Y. When $X = Cn(Y)$ we say that Y is a *base* for X, whether or not Y is finite, and even in the limiting case that Y is X itself.

Now given a consequence operation Cn and a theory $Cn(A)$, it is clear what is meant by *adding* a proposition x to $Cn(A)$. It consists in adding in x set-theoretically, and then closing under Cn. That is, we can define $Cn(A) + x$ to be $Cn(Cn(A) \cup \{x\})$. And it makes *no difference* whether the addition is done to the entire theory $Cn(A)$, as above, or to its base A. For, as is easily verified from the defining conditions of a consequence operation, $Cn(Cn(A) \cup \{x\}) = Cn(A \cup \{x\})$ for any proposition x and consequence operation Cn.

1.2. *The Intuitive Process of Contraction*

Now nobody likes giving things up, but sometimes we have to. If A is a body of propositions that implies x, we may find ourselves, for one reason or another, wishing to reject, or entertain the rejection, of x and thereby of whatever in A implies it. In other words, we may wish to get rid of x from the theory $Cn(A)$, or as we shall say, *contract* the theory to exclude x. When A is a code of norms of some kind – say a legal code, a moral code, or a body of regulations governing some activity – this process is the one familiar to legal theorists as the *derogation* of x from the code.

But this process of contraction is a tricky one, for two main reasons. *First*, it is not in general fully determined by A and x; without further specification it does not give a unique result. This contrasts with the addition of x to A; for a given A and x, and a given consequence operation Cn, $A + x = Cn(A \cup \{x\})$ *is* fully determined. But in general there will be *many* subsets of A that fail to imply x, and worse still, there will in general be many *maximal* such subsets (maximal, that is, under set inclusion). In order to give a unique result, contraction has to choose among these sets. *Second*, the result of the operation will depend upon the *formulation* of the theory as well as upon its *logical force*; contraction from a theory $Cn(A)$ taken as a whole does not in general give the same or even an equivalent result as contracting from its base A. There are in general maximal subsets of $Cn(A)$ that fail to imply x, that are not determined by, and do not even include, any maximal subsets of A that fail to imply x. Similarly, contracting from one base of a theory $Cn(A) = Cn(A')$ is not in general equivalent to contracting from another, even when each, taken separately, is uniquely determined. To take a very trivial, but very distressing example, consider two logically independent propositions a and b, and put $A = \{a, b\}$ whilst $A' = \{a \wedge b\}$. Clearly $Cn(A) = Cn(A')$, but $A \dot{-} a = \{b\}$ whilst $A' \dot{-} a = \phi$, where we write $\dot{-}$ for contraction that picks out a maximal subset (in this case unique) of A that fails to imply x. And if $Cn(A) \dot{-} a$ is constrained only to be a maximal subset of $Cn(A)$ that does not imply a, then it might be significantly different from both of them; it could for example contain $a \equiv b$ but neither a nor b alone. This is a very nasty and complexifying situation. Contraction of x out of A depends upon the formulation of A as much as upon its logical power.

Now one's first reaction to this might be to shrug one's shoulders and say that there is nothing much that formal logic can do to illuminate the business, beyond a few trivial remarks. One might take it for granted that the result of contracting x out of A should always be a maximal subset of A that fails to imply x – an assumption which, as we shall see, may *perhaps* need some revision, and certainly needs careful consideration. One might also surmise that the criteria for choosing among the various subsets are largely nonformal in character – matters of simplicity, convenience, degrees of "entrenchment", and various other epistemic and pragmatic considerations; and in the author's view this may often be the case. But it also turns out that the process of contraction, and the composite process of revision, can nevertheless be illuminated a great deal by a careful *formal* analysis. Such an analysis helps to identify and delineate various options, clarify their consequences, advantages and limitations, and reveal unsuspected regularities and irregularities.

1.3. *General Remarks on the Strategies of Postulation and Explicit Construction*

When tackling a problem like this – the logical or mathematical understanding of an intuitive concept or process – there are two general strategies that tend to present themselves: postulation on the one hand, and explicit construction on the other. On the former approach, we seek to formulate a number of postulates, preferably of a more or less equational nature, that seem plausible for the process, and then investigate their consequences and interrelations. On the latter approach, one seeks to formulate explicit definitions or constructions of the central concepts, and then investigate how far the concepts thus constructed satisfy various conditions, including in particular those which on the former approach may have been suggested as postulates.

Bertrand Russell, when working on the foundations of mathematics, had some hard words to say about the method of postulation, describing it as having the advantages of theft over honest toil. But it has become clear with the passage of time and the accumulation of examples, that *both* strategies have an essential role to play: the former in articulating our intuitions about the processes under study, in so far as we *have* any, and the latter in providing specific and well-defined structures that to some extent may satisfy, and to some extent may correct or be corrected by, these intuitions. Each of these two approaches is needed to guide and help correct the other. Moreover, it sometimes turns out that the two approaches grow rather close to each other. The postulates may involve some conditions of a nonequation nature, with a more complex logical structure, that begin to look like parts of a construction. On the other hand, an enterprise that sets out to provide an explicit construction may turn out not to yield a unique object, nor even a class of mutually isomorphic objects, but some other circumscribed class of objects, so that the "construction" is really a complex condition. This can happen particularly when there are choice functions in the affair. Now, in so far as the process of contraction is concerned, Peter Gärdenfors has suggested a number of conditions that, on reflection, he believes it should satisfy, and has isolated some of these as postulates for the others. And Carlos Alchourrón, Gärdenfors, and the author have considered some explicit constructions which, whilst not yielding a unique process, do provide us with circumscribed classes of operations of a clear and well-defined nature. In what follows, we shall outline some of that work.

2. POSTULATION

2.1. *Gärdenfors' Postulates for Contraction*

Gärdenfors' postulates for contraction are quite simple. They are, in their most convenient versions, and in an order of presentation that differs from his (cf. [5, 6]) as follows:

($\dot{-}$1) $A \dot{-} x$ is a theory whenever A is a theory (closure);
($\dot{-}$2) $A \dot{-} x \subseteq A$ (inclusion);
($\dot{-}$3) If $x \notin Cn(A)$ then $A \dot{-} x = A$ (vacuity);
($\dot{-}$4) If $x \notin Cn(\phi)$ then $x \notin Cn(A \dot{-} x)$ (success);
($\dot{-}$5) If $Cn(x) = Cn(y)$ then $A \dot{-} x = A \dot{-} y$ (preservation);

and the rather less "obvious" condition,

($\dot{-}$6) $A \subseteq Cn((A \dot{-} x) \cup \{x\})$ whenever A is a theory (recovery).

Most of these postulates are fairly straightforward. *Inclusion*, for example, tells us that when we contract A to get rid of x, we always get either a proper subset of A or else A itself. *Vacuity* tells us that in the limiting case that $x \notin Cn(A)$, where x is *already* excluded from the consequences of A, the contraction is vacuous and leaves us with A itself. *Success* tells us that a contraction of A to exclude x does in fact get rid of x in all cases except where x is a logical truth and so impossible to get rid of. *Preservation* tells us that getting rid of x is the same as getting rid of any other proposition to which x is equivalent. Contraction of A to exclude x is thus taken to depend on the logical force alone of the excluded proposition x, and not on its formulation; although as we have seen in our general discussion, contraction cannot in general be invariant under different formulations of A. The postulate of *recovery* tells us that when we contract a theory to get rid of x, and then "change our minds" and add x back again to the result of the contraction, we recover or recuperate all of the initial theory.

Clearly $A \dot{-} x$ will have to be "fairly big" as a subset of A in order to satisfy this; just how big will be made clear later. The postulate of *closure* ensures that contraction on a theory gives a theory; it is included for convenience so that we can formulate principles about $A \dot{-} x$ rather than about $Cn(A \dot{-} x)$.

2.2. *Revision via the Levi Identity, and More Postulates*

The concept of contraction leads us to the concept of *minimal change of belief*, or briefly, *revision*. We shall use the notation $A \dot{+} x$ (in words, *A plus x*), to indicate a result of revising A, in a minimal way, so as to get x in whilst keeping the set consistent. Formally, we define $A \dot{+} x$ to be $Cn((A \dot{-} \neg x) \cup \{x\})$ where $\dot{-}$ is a contraction function. This is called the *Levi identity*, as it was suggested by Isaac Levi in a paper [7] of 1977. It is immediate from the definition of $A \dot{+} x$ that:

($\dot{+}$1) $A \dot{+} x$ is always a theory;
($\dot{+}$2) $x \in A \dot{+} x$.

Moreover, the remaining postulates for contraction have immediate consequences for revision, namely:

($\dot{+}$3) If $\neg x \notin Cn(A)$ then $A \dot{+} x = Cn(A \cup \{x\})$;
($\dot{+}$4) If $\neg x \notin Cn(\phi)$ then $A \dot{+} x$ is consistent;
($\dot{+}$5) If $Cn(x) = Cn(y)$ then $A \dot{+} x = A \dot{+} y$;
($\dot{+}$6) $(A \dot{+} x) \cap A = A \dot{-} \neg x$ whenever A is a theory.

The last of these consequences is particularly interesting, and we shall call it the *Gärdenfors identity*. It states that just as revision can be defined from contraction, so too, for a theory A, contraction can be characterized in terms of revision. In principle, it would thus be *possible* to take revision as the basic notion, with contraction defined from it, and with the present postulates for contraction *derived* from the basic properties of revision. In the author's view, however, to do this would be to analyse the conceptually simpler in terms of the conceptually more complex, and so while formally possible, it is heuristically unwise. For the notion of contraction does seem conceptually more basic: we naturally *think* of revisions, or minimal changes of belief, as obtained by a composite process of first eliminating and then adding. In the context of a legal code, for example, we can think of amendments as obtained by a composite process of explicit, or more often implicit, derogation of old material followed by addition of new norms; whereas we do not think of a derogation as based upon an (even implicit) prior amendment, even when it can be so specified.

The conditions $(\dot{+}1)$–$(\dot{+}6)$ are, then, derivable from Gärdenfors' postulates for contraction, via the Levi identity defining revision in terms of contraction. But Gärdenfors has suggested that revision *also* satisfies some "supplementary conditions" that are *not* derivable from the above, namely:

$(\dot{+}7)$ $A \dot{+} (x \wedge y) \subseteq Cn((A \dot{+} x) \cup \{y\})$ for any theory A;

and its conditional converse:

$(\dot{+}8)$ $Cn((A \dot{+} x) \cup \{y\}) \subseteq A \dot{+} (x \wedge y)$ for any theory A, provided that $\neg y \notin A \dot{+} x$.

Given the presence of the conditions $\dot{\pm}(1)$–(6), these two supplementary conditions on $\dot{+}$ can be shown to be equivalent to various conditions on $\dot{-}$. Some rather complex such conditions are given in [6], but a particularly elegant and simple pair, equivalent respectively to $(\dot{+}7)$ and $(\dot{+}8)$, are the following:

$(\dot{-}7)$ $(A \dot{-} x) \cap (A \dot{-} y) \subseteq A \dot{-} (x \wedge y)$ for any theory A,

and its conditional converse,

$(\dot{-}8)$ $A \dot{-} (x \wedge y) \subseteq (A \dot{-} x)$ for any theory A, provided that $x \notin A \dot{-} (x \wedge y)$.

Gärdenfors adopts $(\dot{+}7)$ and $(\dot{+}8)$ as "supplementary postulates" for revision. As in this presentation we are resolutely taking contraction as primitive with revision defined via the Levi identity, we prefer to choose the simpler $(\dot{-}7)$ and $(\dot{-}8)$ as our "supplementary postulates", with the conditions $(\dot{+}7)$ and $(\dot{+}8)$ as properties following from them.

We shall not here go far into Gärdenfors' reasons for including these supplementary postulates. They lie in part in reflections on the intuitive processes that convince him that these conditions are reasonable, and in part on the *need* that he has for them in order to carry out certain *applications* of the logic of theory change to an epistemic account of counterfactual conditionals. We shall however make two remarks. First, that these supplementary postulates are satisfied by *some* of the explicit constructions that we shall consider shortly, the former being rather more easily satisfied than the latter.

Second, that the supplementary postulates are quite powerful, and Gärdenfors manages to derive from them various other principles for contraction and revision. One of the "nicest" among them is the following: If $x \in A \dot{+} y$ and $y \in A \dot{+} x$ then $A \dot{+} x = A \dot{+} y$, for any theory A. This is reminiscent of a principle concerning the selection of neighbouring "possible worlds" that was postulated by Robert Stalnaker in his seminal 1968 paper [8] on counterfactual conditionals. For this reason, we shall call it the *Stalnaker property*.

3. AN EXPLICIT CONSTRUCTION: MAXICHOICE CONTRACTION AND REVISION FUNCTIONS

3.1. *The Maxichoice Operations*

We turn now to the strategy of explicit construction, and our first and principal construction simply takes seriously the idea that $A \dot{-} x$ should be a *maximal* subset of A that does not imply x.

We define $A \perp x$, in words, A *less* x, to be the class of all maximal subsets B of A such that $x \notin Cn(B)$. It is easily verified that $A \perp x$ is nonempty just when $x \notin Cn(\phi)$. Now let A be a fixed set of propositions, and let γ be any choice function on the family of all classes $A \perp x$, that is, $\gamma(A \perp x) \in A \perp x$ whenever the latter is nonempty. We then define $A \dot{-} x$ by the rule:

$$A \dot{-} x = \gamma(A \perp x) \text{ when } x \notin Cn(\phi)$$
$$A \dot{-} x = A \text{ otherwise.}$$

Thus $A \dot{-} x$ will be some maximal subset of A that fails to imply x, if such exists, and otherwise will be A itself. We call such contraction functions *maxichoice contraction functions*.

Two kinds of questions immediately arise. One is, are there any interesting ways in which such choice functions may be *built up* out of some kind of ordering of the elements of A itself? This is of particular interest in the context of legal or administrative codes, where we can often discern some kind of hierarchy among the norms. Such a hierarchy will sometimes be vague, and sometimes well delineated. It may be determined in part by the formulations of the regulations themselves, involving, for example, cross-reference from one to another or clauses of the kind "notwithstanding . . .". It may be determined in part by considerations of a more extrinsic nature, such as conventions of priority associated with the powers and competences of the issuing bodies, dates of promulgation and amendment, and the degrees of specificity or generality of the regulations made. This kind of question has been studied by Carlos Alchourrón and the author [1]. There it was shown how partial orderings of the underlying code A can be lifted to various kinds of ordering (not in general partial orderings) of its power set 2^A, and how these can be used to constrain and eventually render unique the choice of a set from $A \perp x$. However, we shall here focus on a second kind of question: when $A \dot{-} x$ is a maxichoice operation, and revision is defined from contraction via the Levi identity, *what properties do they have*? In particular, do they satisfy the Gärdenfors' postulates?

3.2. *Checking Out the Gärdenfors' Postulates*

It is easily verified that all six of Gärdenfors' postulates for contraction (as we have formulated them) hold true of maxichoice contraction functions. This of course also means that the six corresponding conditions for revision, derivable from their contraction counterparts, are likewise true of maxichoice revision. We shall not provide verification of these points here, although they are short and in some instances (recovery and the Gärdenfors identity) interesting. Verifications can be found in [2].

On the other hand, the two supplementary postulates ($\dot{-}7$) and ($\dot{-}8$) do not hold in general for maxichoice revision. To obtain them, the maxichoice functions must be further constrained.

We say that a maxichoice contraction function $\dot{-}$ is *orderly over A* iff there is some partial ordering (i.e., reflexive, transitive, antisymmetric) $\leq$ of 2^A, such that $B \leq A \dot{-} x$ for all propositions x and all $B \in A \perp x$. That is, roughly, $\dot{-}$ is orderly over A iff there is a partial ordering of the power set of A that always selects $A \dot{-} x$ as a *best* element of $A \perp x$. Note that the relation $\leq$ is required to be *the same* for all propositions x, rather than vary with them – otherwise all maxichoice contraction functions would be trivially orderly over A. There is considerable "tolerance" in this definition, in the sense that we can weaken the condition on $\leq$ to mere antisymmetry, or strengthen it by adding connectivity, without affecting the logical force of the concept of orderliness; that is, such definitions can be shown to be equivalent.

Now the great thing about orderly contraction functions is that for them we have that if $A \dot{-} y \in A \perp z$ and $A \dot{-} z \in A \perp y$, then $A \dot{-} y = A \dot{-} z$. For if $A \dot{-} y \in A \perp z$ and $A \dot{-} z \in A \perp y$ then for some partial ordering $\leq$, we have both $A \dot{-} y \leq A \dot{-} z$ and $A \dot{-} z \leq A \dot{-} y$, so by antisymmetry, $A \dot{-} y = A \dot{-} z$. This lemma provides an avenue for proving the identity of contractions $A \dot{-} y$ and $A \dot{-} z$ of different propositions y and z from a given set A. Armed with this lemma, and some others like it, it is not too difficult to show that every maxichoice revision function that is orderly over a theory A satisfies both of the supplementary postulates as applied to A, and also has a number of other interesting properties. Among these is of course the Stalnaker property, for as we mentioned earlier, that is itself derivable from Gärdenfors' supplementary postulates. But there are also others which were not considered by Gärdenfors. One particularly powerful one is a *decomposition condition* for contraction: when a maxichoice contraction function $\dot{-}$ is orderly over a theory A, then for all propositions x, y, $A \dot{-} (x \wedge y) = A \dot{-} x$ or $A \dot{-} (x \wedge y) = A \dot{-} y$. Finally, it can be shown that for *maxichoice* contraction and revision functions, all of the above properties – the decomposition property, the Stalnaker property, the supplementary postulates, and the lemma that we mentioned – and also various other properties we have not mentioned here, are severally *equivalent* to the orderliness of that maxichoice function, over a given set A. The details and verifications are given in [2].

3.3. *Some "Unpleasant" Features of the Maxichoice Operations*

That was the good news. Now for the bad news. It is that in certain kinds of application, maxichoice contraction and revision give us sets that must be considered as *too large* to serve as reasonable formal representations of their intuitive counterparts. To explain the situation, let us recall that a theory A is said to be *complete* iff for every proposition y of its language, either $y \in A$ or $\neg y \in A$. Now we would perhaps expect maxichoice revision to *preserve* completeness, in the sense that $A \dot{+} x$ will be complete whenever A is. And so it does. But it does more, and worse, than that: it *creates* completeness. It can be shown that if A is any theory, complete or not, then if $\neg x \in A$ then $A \dot{+} x$ will be complete. The set $A \dot{+} x$ will thus in general be much too large to serve as even an idealization of any intuitive process of rational revision.

Underlying the above observation is one about maxichoice contraction. If A is any theory with $x \in A$, then for every proposition w, either $(x \vee w) \in A \dot{-} x$ or $(x \vee \neg w) \in A \dot{-} x$. This too seems counterintuitive, for w may presumably be chosen to be a proposition that has "nothing to do" with x or the rest of A, so that $x \vee w$ and $x \vee \neg w$ are in A "only because" x is there. Withdrawal of x would then leave no apparent reason for retaining either one of them. In general, neither $x \vee w$ nor $x \vee \neg w$ should be retained in the process of eliminating x from A unless there is "some reason" in A for their continued presence.

There are three remarks that may help clarify the interpretation of these formal results. The first is to suggest that one thing that the result for contraction reveals is that when A is *really a theory*, that is, closed under Cn, and x is an element of A, then no proposition w (in the language of the theory) really has anything to do with any part of the theory A. For both of $x \vee w$ and $x \vee \neg w$ *will* be in A; and as w occurs in each of *them*, w does have, indirectly, quite a lot to do with A. The "reason" for the continued presence of at least one of $x \vee w$ and $x \vee \neg w$ in $A \dot{-} x$ is the presence in A of *both* of the two.

Nevertheless, the formal result on contraction remains at least surprising, and the formal result on revision, which follows from it, downright anomalous for any representation of the intuitive process of revision. The author would suggest that the reason for this is that in real life, when we perform a contraction or derogation, *we never do it to the theory itself* (in the sense of a set of propositions closed under consequence) but rather on some finite or recursive or at least recursively enumerable *base* for the theory. And as such bases are in general far from being closed under Cn – indeed they are usually closer to the opposite extreme, of being irredundant – the application of the maxichoice operations to them does not give rise to the above formal results. In other words, contrary to casual impressions, the *intuitive* processes of contraction and revision are always applied to more or less clearly identified finite or otherwise manageable *bases* for theories, which will in general be either irredundant or reasonably close to irredundant. And when the maxichoice operations are applied to such bases, as they should be, they do not yield inflated sets. The fault lies not in our operations, but in the items to which they were applied.

The third remark concerns the application of the maxichoice operations to *complete* theories. Although the formal results hold for complete theories, they are not in that case particularly anomalous. It is not very surprising or counterintuitive to find that maxichoice revision *preserves* completeness. Of course, if the reflections that we have just made are on the mark, then in real life we never revise theories, and *a fortiori* complete theories, themselves; we revise only a finite or otherwise manageable base of a complete theory, and in general that will not yield a complete theory. But the application of the maxichoice operations directly to complete theories remains interesting for *another* reason. So applied, they can help clarify, not so much human efforts at theory revision, as transcendental notions of "similarity" between "possible worlds".

To see how this may be done, consider for example the classic paper [8] of Robert Stalnaker on the logic of counterfactual conditionals. There Stalnaker invites us to consider functions $f(\alpha, x)$ which take a pair consisting of a "possible world" α and a proposition x to a world β that is most like α except that it renders x true. Various conditions are imposed by Stalnaker and others on such functions, but these conditions have seemed to be justified more by their ability to validate and invalidate favoured arrays of logical principles about counterfactuals, than by a clear intuition. Now it has often been observed that a possible world may be identified with the set of all propositions that it renders true. But although this may help us understand better the notion of a possible world, it does not take us far towards understanding the nature and behaviour of the similarity functions $f(\alpha, x)$. To do that, it is convenient to regard them as *composite* functions, that can be broken down into an initial step of contraction or elimination, followed by one of addition. A similarity function $f(\alpha, x)$ may be regarded as a maxichoice revision function $A \dot{+} x$ applied to the set A of all propositions true in α, and may thus be seen as a *composite* function $Cn((A \dot{-} \neg x) \cup \{x\})$ formed by using an underlying maxichoice contraction function $\dot{-}$. In this way, the properties of maxichoice contraction functions yield properties of similarity functions on possible worlds. If, as in the work of David Lewis and others, $f(\alpha, x)$ is taken as giving not a single world, but rather a *set* of worlds closest to α in which x is true, we can *still* decompose the process, by first forming a *family* $\mathcal{B} \subseteq A \perp \neg x$ of maximal subsets of A that do not imply $\neg x$, and then forming the family of all sets $Cn(B \cup \{x\})$ where $B \in \mathcal{B}$.

Why, one might ask, was this breakdown of similarity functions on possible worlds so slow to appear, despite heuristic hints in its direction back in Stalnaker's paper of 1968? Because, we suspect, it involves breaking out of the ring of all possible worlds: $A \dot{-} \neg x$ will not itself be a possible world (more accurately speaking, will not be the set of all propositions true in some possible world) even when A is. Such has been the hypnotic power of the notion of a possible world since 1963 that it has come to confine logicians as well as to serve them. To understand operations on possible worlds, we sometimes need to look at *more* than possible worlds.

4. OTHER EXPLICIT CONSTRUCTIONS: PARTIAL MEET CONTRACTION AND HIERARCHICAL CONTRACTION

4.1. *Partial Meet Contraction*

The "unpleasant" features of maxichoice contraction and revision bring out the interest, even if (as we have argued) not the urgency, of looking around for other formal processes that might be used as idealized representations of the intuitive ones. One idea which is tempting is of course to take $A \dot{-} x$ to be $\bigcap(A \perp x)$ – the *intersection* of *all* the maximal subsets of A that fail to imply x, whenever $A \perp x$ is not empty. In the limiting case that $A \perp x$ is empty – which happens just when $x \in Cn(\phi)$ – $A \dot{-} x$ would be set at A. But this notion of *meet contraction* will not do: it can be shown that it gives a set that is far *too small*, whether the operation is applied to bases or to theories (even to irredundant bases or complete theories). In particular, as shown in [2], when A is a theory and $x \in A$, $\bigcap(A \perp x) = Cn(\neg x) \cap A$. That is, if we eliminate x from a theory A in this way, we are left with only the propositions of A that are already consequences of $\neg x$. Nevertheless the operation of meet contraction is a useful one to be able to refer to, for it does appear to provide a definite *lower bound* on any reasonable contraction operation: for any contraction operation worthy of the name we should have $\bigcap(A \perp x) \subseteq A \dot{-} x \subseteq A$ for every theory A and proposition $x \notin Cn(\phi)$. Moreover, it can be shown that any operation satisfying this lower bound condition (and which puts $A \dot{-} x = A$ as usual in the limiting case that $x \in Cn(\phi)$), satisfies recovery.

Following this lead, Alchourrón, Gärdenfors, and the author [3] have been working on an idea that generalizes both maxichoice and meet contraction. As with maxichoice contraction, a choice function γ is used, to give however not an arbitrary *element* of $A \perp x$, but rather a nonempty *subclass* $\gamma(A \perp x) \subseteq A \perp x$ when the latter is itself nonempty, and set at $\{A\}$ in the limiting case that $A \perp x$ is empty. The *partial meet contraction* $A \dot{-} x$ is then defined to be $\bigcap \gamma(A \perp x)$, the intersection of all the sets in $\gamma(A \perp x)$. Thus $A \dot{-} x$ is taken to be the meet of *some* class of maximal subsets of A that fail to imply x. The maxichoice operations thus form a limiting case of partial meet contraction, where $\gamma(A \perp x)$ is a singleton; and full meet contraction can be seen as the special case where $\gamma(A \perp x) = A \perp x$. *Partial meet revision* is then introduced by the usual Levi identity $A \dot{+} x = Cn((A \dot{-} \neg x) \cup \{x\}) = Cn(\bigcap \gamma(A \perp \neg x) \cup \{x\})$.

This idea is also of interest in that it is closely related to the construction that we mentioned at the end of section 3 for decomposing the "possible world" selections of David Lewis. For if $\mathscr{B}$ is a nonempty subset of $A \perp x$, it is easily verified that $Cn(\bigcap \mathscr{B} \cup \{x\}) = \bigcap\{Cn(B \cup \{x\}); B \in \mathscr{B}\}$.

Like the maxichoice operations, the operations of partial meet contraction and revision can be shown to satisfy Gärdenfors' postulates $\dot{\pm}(1)$–(6). Conversely, it can also be shown that every operation over a theory A satisfying the Gärdenfors' postulates for contraction *is itself* a partial meet contraction operation. This result serves as a *representation theorem* for the class of operations satisfying the Gärdenfors' postulates.

As with the maxichoice operations, the supplementary postulates do not hold unless we impose further restraints. An idea that comes naturally to mind is to restrain the choice function γ by some condition akin to orderliness that both makes formal sense and has plausibility in the more general context of partial meet operations. One such condition is that of "relationality". We call γ *relational* over A iff there is some relation $\leq$ over 2^A that marks off $\gamma(A \perp x)$ in the sense that for all $x \notin Cn(\phi)$,

$$\gamma(A \perp x) = \{B \in A \perp x : B' \leq B \text{ for all } B' \in A \perp x\}.$$

We call γ *transitively relational* over A iff there is a transitive relation $\leq$ that marks off $\gamma(A \perp x)$ by the same equation. It is easy to show that if γ is relational over a theory A, then the first supplementary postulate, numbered ($\dot{-}$7) in our exposition, holds; and if γ is transitively relational over a theory A, then both ($\dot{-}$7) and ($\dot{-}$8) hold. It is also possible to prove a converse, that serves as another representation theorem. Let A be any theory and $\dot{-}$ an operation satisfying the Gärdenfors' postulates ($\dot{-}$1)–($\dot{-}$6) and the supplementary postulates ($\dot{-}$7) and ($\dot{-}$8). Then $\dot{-}$ is a partial meet contraction function determined by a choice function γ that is transitively relational over A. This, and a number of other results in a rather complex web of interrelationships between conditions around relationality, is proven in [3].

Since transitively relational partial meet contraction functions satisfy Gärdenfors' supplementary postulates as well as his basic ones, they also have all the properties that are derivable from those postulates, and notably the Stalnaker property noted in section 2.2. But a word of warning is in order here. We noted in section 3.2 that for *maxichoice* contraction and revision functions, the condition of orderliness is in fact *equivalent* to each of the supplementary postulates, to the Stalnaker property, and to the "decomposition property" $A \dot{-} (x \wedge y) = A \dot{-} x$ or $A \dot{-} (x \wedge y) = A \dot{-} y$. Care is needed in attempting to generalize these results to partial meet functions. In particular, decomposition does not always hold for partial meet contraction functions, even when ($\dot{-}$7) and ($\dot{-}$8) are both satisfied, and that even in the finite case. And in the context of partial meet contraction functions, the supplementary postulates ($\dot{-}$7) and ($\dot{-}$8) are not in general equivalent to each other, indeed neither implies the other, again even in the finite case. These negative results are also noted, with an outline of proof, in [3].

4.2. *Safe Contraction and Revision*

Another approach to a formal account of contraction, which *prima facie* is rather different from the preceding ones, makes use of the notion of a "safe" element of A. Let $<$ be an ordering of A itself (rather than of its power set as with the constructions considered so far): we need only suppose that $<$ is irreflexive and transitive. Now let x be any proposition that we may wish to eliminate from among the consequences of A. We say that an element a of A is *safe* with respect to x (modulo $<$) iff every minimal subset B of A that implies x either does not contain a or else contains at least one element $b < a$. In other words, a is safe iff no inclusion-minimal subset of A that implies

x contains a as a $<$-minimal element. Intuitively speaking, a is safe with respect to x iff it can never be "blamed" for the implication of x; that is, if we can never find a minimal subset of A implying x that contains a and does not contain any element "worse" than a. The notion of safety is a refinement of the cruder notion of "normality" defined in [1]. We write A/x for the set of all elements of A that are safe with respect to x, modulo $<$, and we define $A \dot{-} x$ to be $Cn(A/x) \cap A$, that is, the set of all elements of A that are implied by A/x. When A is itself a theory, this of course reduces to $Cn(A/x)$.

Like the maxichoice and the partial meet operations, these operations of *safe contraction and revision*, as we shall call them, satisfy all of the Gärdenfors' postulates $\dot{\pm}(1)$–(6); success and recovery are the interesting ones to verify. Consequently, using the representation theorem of the previous section, the safe contraction operations form a *subclass* of the partial meet contraction operations. Once more, the supplementary postulates appear to fail unless special conditions are imposed. We can regain the postulate $(\dot{-}7)$ by imposing an appropriate condition on $<$. We say that $<$ *continues up* $\vdash$ iff (in addition to the background conditions of transitivity and irreflexivity) we have for all a, b, c in A, if $a < b$ and $b \vdash c$ then $a < c$. Similarly, we say that $<$ *continues down* $\vdash$ iff $a \vdash b$ and $b < c$ implies $a < c$. Then it can be shown that if $<$ continues up or continues down $\vdash$ the supplementary postulate $(\dot{-}7)$ holds. There are also interesting conditions on $<$ that enable us to recuperate for safe contraction the other supplementary postulate $(\dot{-}8)$. Proofs of these and other results on safe contraction will appear in [4].

Having surveyed all this formal work, we end with a general remark. Even if it turns out that none of the formal operations that we have been considering, or others like them, provide entirely satisfactory representations or idealizations of the intuitive processes, they at least serve a purpose. By studying them we can begin to get a grasp of the options available, the consequences and limitations of each, with the regularities and irregularities that may be encountered, and thereby reach a better understanding of the fine points of the intuitive processes themselves.

NOTE

* This paper was prepared while the author was on leave from Unesco; neither institution bears any responsibility for content. The author expresses his gratitude to

Professors Peter Gärdenfors and Carlos Alchourrón for permission to include in section 4.1 concepts and results developed in mutual correspondence. The material of sections 4.1 and 4.2 will be presented in detail, with proofs, in two forthcoming papers written by the three authors as indicated in the bibliography. A draft of this survey has benefited from comments of Alchourrón, Gärdenfors, and Haim Gaifman, and was presented at the American University of Beirut, the Hebrew University of Jerusalem, and the CNRS Toulouse le Mirail.

REFERENCES

[1] Alchourrón, Carlos E. and David Makinson: 1981, 'Hierarchies of regulations and their logic', in R. Hilpinen (ed.), *New Studies in Deontic Logic*, D. Reidel, Dordrecht, pp. 125–148.
[2] Alchourrón, Carlos E. and David Makinson: 1982, 'On the logic of theory change: contraction functions and their associated revision functions', *Theoria* **48**, 14–37.
[3] Alchourrón, Carlos E., Peter Gärdenfors and David Makinson: 'On the logic of theory change: partial meet contraction and revision functions', forthcoming.
[4] Alchourrón, Carlos E. and David Makinson: 'On the logic of theory change: safe contraction and revision functions', forthcoming.
[5] Gärdenfors, Peter: 1979, 'Conditionals and changes of belief', in I. Niiniluoto and R. Tuomela (eds.), *The Logic and Epistemology of Scientific Change*, Acta Philosophica Fennica 30, 1978, issues 2–4, North-Holland, Amsterdam.
[6] Gärdenfors, Peter: 1982, 'Rules for rational change of belief', in T. Pauli (ed.), *320311: Philosophical Essays Dedicated to Lennart Åqvist on his Fiftieth Birthday*, University of Uppsala, Philosophical Studies n°. 34, Uppsala.

Philosophy Department
American University of Beirut
Beirut
Lebanon