# Contextual Modals

Horacio Arló Costa[1] and William Taysom[2]

[1] Carnegie Mellon University,
hcosta@andrew.cmu.edu
[2] Institute for Human and Machine Cognition,
wtaysom@ihmc.us

**Abstract.** In a series of recent articles Angelika Kratzer has argued that the standard account of modality along Kripkean lines is inadequate in order to represent context-dependent modals. In particular she argued that the standard account is unable to deliver a non-trivial account of modality capable of overcoming inconsistencies of the underlying conversational background. She also emphasized the difficulties of characterizing context-dependent conditionals. As a response to these inadequacies she offered a two-dimensional account of contextual modals. *Two* conversational backgrounds are essentially used in this characterization of contextual modality.

We show in this paper that Kratzer's *double relative* models (with finite domains) are elementary equivalent to well known neighborhood models of normal modalities originally proposed by D. Scott [S] and R. Montague [M]. We also argue that neighborhood models can be also used to represent some (non-normal) graded modalities that are difficult to represent in her framework (like 'it is likely that' or 'it is highly probable that', etc). Finally we show that an extension of the neighborhood semantics of conditionals is able to capture some of her proposals concerning dyadic modals. DR models with infinite domains can be shown to be pointwise equivalent to neighborhood models, but they are not guaranteed to have relational counterparts. So DR models surpass the representational power of relational (Kripkean) models. Neighborhood representations are, nevertheless, always possible, making clear as well that the central feature of double relative modals is that they are capable of encoding two central aspects of context: its propositional content, and its dynamic properties (which in Kratzer's models are represented via an *ordering source*).

## 1 Introduction

Phrases like *what the law provides*, *what we know*, or *what we presuppose* are usually employed to introduce what semanticists and logicians call *conversational backgrounds.* And such types of contexts are usually encoded as world-dependent sets of propositions. In other words conversational backgrounds (or context sets) are usually represented as functions which assign to every member $w$ of a universal set of worlds $W$ a subset $f(w)$ of the power set of $W$.

Then modals like, 'must', 'might', etc. can be analyzed in terms of this given background as follows: 'It must be the case that $P$' is true at $w$ with respect to $f(w)$ if and only if the proposition expressed by $P$ follows from everything one knows at $w$. By the same token: 'It might be the case that $P$' is true at $w$ with respect to $f(w)$ if and only if the proposition expressed by $P$ is compatible with everything one knows at $w$.

This is the simple analysis of contextual modals along Kripkean lines that Kratzer proposes in various writings ('the standard model' of modalities according to her terminology). The surface description of these standard models for modalities might not sound nevertheless familiar to readers used to relational semantics. But it is easy to show that models of these type can be straightforwardly presented as neighborhood models obeying certain constraints, and that, in turn, these neighborhood models are always pointwise equivalent to relational models.[1] Parenthetical references below point to [Ch].

**Definition 1 (Ch Def. 7.1).** *A model $M = \langle W, f, P\rangle$ is a* neighborhood model *iff:*

1. *$W$ is a set.*
2. *$f$ is a function which maps each world to a set of propositions (i.e. $f : W \rightarrow \wp(\wp(W))$ ).*
3. *$P : N \rightarrow \wp(W)$.*

**Definition 2 (Ch Def. 7.2).** *Let $w$ be a world in a neighborhood model $M = \langle W, f, P\rangle$. Then:*

1. $\models^M_w \Box \mathrm{A}$ *iff* $||\mathrm{A}||^{\mathrm{M}} \in \mathrm{f(w)}$.
2. $\models^M_w \Diamond \mathrm{A}$ *iff* $W - ||\mathrm{A}||^{\mathrm{M}} \notin \mathrm{f(w)}$.

This is the basic presentation of modalities in terms of neighborhood models. Kratzer's models can be presented as the following particular case:

**Definition 3.** *Let $w$ be a world in a Kratzer model $M = \langle W, f, P\rangle$. Then:*

1. $\models^M_w \Box \mathrm{A}$ *iff* $\bigcap f(w) \subseteq ||\mathrm{A}||^{\mathrm{M}}$.
2. $\models^M_w \Diamond \mathrm{A}$ *iff* $||\mathrm{A}||^{\mathrm{M}} \cap \bigcap \mathrm{f(w)} \neq \emptyset$.

In this particular case the value of the function $f$ at $w$ encodes the propositional content of the context at $w$. The propositions in $f(w)$ are the ones that are presupposed, or known, or the propositions that encode what the law provides, etc. There is, of course, the question of how structured this body of information should be. Should it be closed under logical consequence? Should it be closed under conjunctions? Kratzer is right, nevertheless in proceeding gradually from

[1] A neighborhood model $M$ is pointwise equivalent to a relational model $N$, when for any sentence $A$ and any world $w$, $A$ is true at $w$ in $M$ iff $A$ is true at $w$ in $N$. One can show, as suggested above, that for every relational model, there is a pointwise equivalent neighborhood model obeying some constraints, and vice versa.

the less structured situation (where $f$ is unconstrained) to the adoption of constraints that fit models capable of handle modalities along Kripkean lines. The following definitions will help to clarify which are these assumptions.

A neighborhood function $f$ is *supplemented* if and only of if $p \in f(w)$ and $q \in f(w)$, when $p \cap q \in f(w)$. In addition a neighborhood model is *augmented* if and only of if it is supplemented and for every world $w \in W$, $\bigcap f(w) \subseteq f(w)$.

A Kratzer model can consist of a frame containing a function $f$ encoding finite information ($f(w)$ might have finite cardinality). The augmentation of such a frame is obtained by taking the intersection of $f(w)$ and closing under supplementation. We can see the augmentation ($f!$) of a function $f$ as the representation of the logical consequences of the information encoded in $f(w)$ for every $w$ in the domain. It is also easy to see that the augmentation of the frame of a Kratzer model validates exactly the same modal sentences that the Kratzer model itself. Moreover it is well known that augmented neighborhood models are pointwise equivalent to standard relational models.[2]

So, Kratzer's 'standard account' is indeed relatable to the standard relational account of modalities. But as Kratzer correctly observes, the standard account cannot tolerate a weakly inconsistent context (i.e. a case where $\bigcap f(w)$ is empty). In this case every proposition is both necessary and impossible. Kratzer also argues (correctly, we think) that the standard account cannot handle correctly *graded* modalities and that its treatment of dyadic modality is also poor.

## 2 Double Relative Models

Kratzer's solution to the former problem is the adoption of models dependent on two conversational backgrounds rather than one:

**Definition 4.** *A model $\langle W, f, g, P\rangle$ is* doubly relative *(a DR model) iff:*

1. *$W$ is a set.*
2. *$f$ and $g$ are conversational backgrounds.*
3. *$P : N \rightarrow \wp(W)$.*

The background $f$ is called the *modal base* for the model; $g$ is called the *ordering source*. The role of the modal base is slightly different than in the Kratzer models presented above. Its central function is to determine for each world, which worlds are accessible from it. A conversational background $e$ is *empty* just in case $e_w = \emptyset$ at every world $w$. We conventionally use $e$ to denote the empty conversational background. When the DR model $\langle W, f, g, P\rangle$ is understood and we have fixed a world $w$, we let, for every world $u$, $\gamma(u) = \{p \in g_w | u \in p\}$.

**Definition 5.** *The ordering source $g$ is used to fix a* typicality ordering $\leq_{g(w)}$ *such that for all $w, u, v \in W$, $v \leq_{g(w)} u$ iff $\gamma(u) \subseteq \gamma(v)$.*

[2] See [Ch], page 221.

So, for each world the second conversational background induces a (partial) ordering of the set of worlds accessible from that world (where this accessible set is determined by the first background). Truth conditions in this model function in a simple way, only complicated by technical concerns abut infinity. The basic idea is that a proposition is necessary if and only of it is true in all accessible worlds which come close to the ideal established by the ordering source. The formal definition is as follows;

**Definition 6.** *Let $w$ be a world in a DR model $M = \langle W, f, g, P\rangle$:*

1. $\models^M_w \Box$A *iff for all $u \in \bigcap f_w$, there is a $v \in \bigcap f_w$ such that $v \leq_{g(w)} u$ and for all $z \in \bigcap f_w$ if $z \leq_{g(w)} v$, then $z \in ||A||^M$.*
2. $\models^M_w \Diamond$A *iff* $\models^M_w \neg\Box\neg$A.

The definition for $\models^M_w \Box$A is a bit difficult to parse. The idea is that $\models^M_w \Box$A just in case for every world $u$ if we look at a series of worlds each progressively more typical than $u$, there comes a point in the sequence $v$ such that for any world $z$ more typical than $v$, A is true at $z$. Or in other words, $\models^M_w \Box$A when no matter what world $u$ you choose, you can find a world more typical than $v$ such that A is true in every world more typical than $v$. The only difference between what has just been expressed and the full formal definition is that in the definition the domain of quantification for $u$, $v$, and $z$ is restricted to $\bigcap f_w$. Also observe that when the typicality ordering has most typical worlds in $\bigcap f_w$, the truth conditions for a sentence A can be simplified: $\models^M_w \Box$A just in case $||A||^M$ is a superset of the set of most typical worlds in $\bigcap f_w$. When this happens, we say that a doubly relative model is reduced.

We introduce some terminology and show that any partial ordering can be a typicality ordering.

**Definition 7.** *Let $\leq$ be a partial ordering on $U$ (i.e. $\leq \subseteq U \times U$ that is reflexive, antisymmetric, and transitive). If $V \subseteq U$, then the ordering of $\leq$* restricted to *$V$ is the ordering $I(\leq, V) = \{\langle u, v\rangle \in \leq | u, v \in V\}$.*

**Definition 8.** *A point $u$ is* minimal *in an ordering $\leq$ iff for every $v \leq u$, $v = u$. An ordering $\leq$ of $U$ has a* bottom *iff for every $u \in U$, there is a minimal point $m \in U$ such that $m \leq u$. If $\leq$ does not have a bottom, $\leq$ is* bottomless. *If $\leq$ has a bottom, then $bot(\leq)$ denotes the set of minimal points in $\leq$. If $\leq$ is bottomless, $bot(\leq)$ is undefined.*

A minimal point is one which has has no points less than it. Suppose $\leq$ is an ordering on $U$. If every element in $U$ is greater than or equal to a minimal point of $U$, then $bot(\leq)$ is the set of minimal points of $U$. On the other hand, if there is some element $u$ of $U$ such that there is no minimal point $m$ of $U$ with $m \leq u$, then $bot(\leq)$ is undefined.

**Theorem 1.** *If $\leq$ is a partial ordering of $W$, then there is a DR model $\langle W, f, g, P\rangle$ such that $\leq = \leq_{g(w)}$.*

This theorem shows that in general, $\leq_{g(w)}$ can be any partial ordering of $W$.

### 2.1 Relating DR Models and Augmented Models

For some DR models, the doubly relative condition DRC can be simplified substantially. We call the simplified condition the reduced doubly relative condition (abbreviated RDRC). If a DR model can be characterized by RDRC, we say that it is a reduced doubly relative model (or RDR model for short). We show that the class of RDR models is pointwise equivalent to the class of augmented models. Furthermore, a DR model is pointwise equivalent to an augmented model just in case the DR model is reduced.

**Definition 9.** *A DR model $M = \langle W, f, g, P\rangle$ is* reduced *(alternatively, a* reduced doubly relative model *or RDR model) just in case $bot(I(\leq_{g(w)}, \bigcap f_w))$ is defined.*

Given a sentence A, we call the condition $bot(I(\leq_{g(w)}, \bigcap f_w)) \subseteq ||A||^M$ the *reduced doubly relative condition* (RDRC for short). Intuitively, The reduced doubly relative condition just says after restricting $\leq_{g(w)}$ to $\bigcap f_w$, the resulting restricted ordering has a bottom. So a DR model is not reduced, if the restriction of $\leq_{g(w)}$ to $\bigcap f_w$ is bottomless.

**Theorem 2.** *The class of RDR models is pointwise equivalent to the class of augmented minimals.*

The final theorem of this section shows that the class RDR models characterizes those DR models which are pointwise equivalent to augmented minimals.

**Theorem 3.** *A DR model is reduced iff it is pointwise equivalent to an augmented minimal $N = \langle W, N, P\rangle$ such that for all $w \in W$, $\bigcap N_w = ||\mathrm{B}||^{\mathrm{N}}$ for some sentence* B.

### 2.2 A concrete example

Kratzer offers a concrete example, which we can analyze here in order to illustrate some of the main constructions. The example uses a set $\{\mathrm{M}, \mathrm{G}, \neg\mathrm{G}\}$ containing background information about laws passed in a hypothetical country. 'M' stands for 'Murder is a crime' and 'G' for Owners of goats are liable for damage caused by them'. The idea is that two courts in different parts of the country have passed contradictory laws - the intuition is that different geographical conditions justify different jurisprudence.

With this example in mind, Kratzer believes that the following sentences should be true in $M$ at a given world $w$:

1. $\Box$M – "Murder is necessarily a crime."
2. $\Diamond$G – "Owners of goats are possibly liable for damage caused by their goats."
3. $\neg\Box$G – "Owners of goats are not necessarily liable for damage caused by their goats."

But these conditions, though feasible, are not the only ones that fit with the scenario of several courts and conflicting judgments. The way Kratzer has it, any conflict between the courts mutes the precedent. Conflicting rulings essentially cancel each other out. Yet the judicial system might work differently. Suppose that each ruling of a court is binding individually but that the normative force of independent rulings is not combined. So each ruling serves to establish precedent, but the precedent of two or more rulings cannot be conjoined. Such cases need to be decided individually. Then, the following sentences are true in $M$ at a given world $w$:

1. $\Box(\mathrm{G})$
2. $\Box(\neg\mathrm{G})$
3. $\neg\Box(\mathrm{G} \wedge \neg\mathrm{G})$

No DR model gives us these truth conditions. A class of minimal models, on the other hand, does.

In what follows, we show how Kratzer uses a DR model for her truth conditions. Then we show that there is no DR model for the modified scenario. We close by showing that there are minimal models which capture both Kratzer's intuition and the alternative.

### 2.3 How a Doubly Relative Model handles Kratzer's Intuition

Let $M = \langle W, f, g, P\rangle$ be a DR model such that the modal base $f$ is empty and the ordering source $g_w = \{||\mathrm{M}||^{\mathrm{M}}, ||\mathrm{G}||^{\mathrm{M}}, ||\neg\mathrm{G}||^{\mathrm{M}}\}$ for some world $w$. Let $W = \{a, b, c, d\}$. Let $||\mathrm{M}||^{\mathrm{M}} = \{\mathrm{a}, \mathrm{b}\}$. And let $||\mathrm{G}||^{\mathrm{M}} = \{\mathrm{a}, \mathrm{c}\}$. We show that the three sentences mentioned above are all true in $M$ at $w$ (i.e. $\models^M_w \Box\mathrm{M}$, $\models^M_w \Diamond\mathrm{G}$, and $\models^M_w \neg\Box\mathrm{G}$).

By definition, a conversational background $f$ is empty just in case $f_w = \emptyset$ for every world $w$. Recall that $f_w$ is a set of propositions and that a proposition is just a set of worlds. So $\bigcap f_w$ is the set of worlds such that each $u \in \bigcap f_w$ is in every proposition $p \in f_w$. Now, since $f_w$ is empty, there are no propositions in $f_w$. Therefore, every world is in every proposition in $f_w$ vacuously. So, $\bigcap f_w = W$. Thus, $I(\leq_{g(w)}, \bigcap f_w) = \leq_{g(w)}$. We show that $R$ is defined. We determine the value of $R$, and then use RDRC to show the three sentences are true.

Now, consider the typicality ordering $\leq_{g(w)}$ induced by the ordering source $g_w$. We know that $v \leq_{g(w)} u$ iff $\gamma(u) \subseteq \gamma(v)$ where $\gamma(x) = \{p \in g_w | x \in p\}$ for any world $x$. The following table shows the value of $\gamma(x)$ for each $x \in W$:

| $x$ | $\gamma(x)$ |
|---|---|
| $a$ | $\{||\mathrm{M}||^{\mathrm{M}}, ||\mathrm{G}||^{\mathrm{M}}\}$ |
| $b$ | $\{||\mathrm{M}||^{\mathrm{M}}, ||\neg\mathrm{G}||^{\mathrm{M}}\}$ |
| $c$ | $\{||\mathrm{G}||^{\mathrm{M}}\}$ |
| $d$ | $\{||\neg\mathrm{G}||^{\mathrm{M}}\}$ |

Hence, $\leq_{g(w)}$ contains $\langle a, b\rangle$ and $\langle b, c\rangle$ in addition to the reflective pairs $\langle x, x\rangle$. The following diagram shows, graphically, the ordering:

It is clear from the diagram that $bot(\leq_{g(w)}) = \{a, b\}$. Therefore, if the proposition expressed by a sentence is a super set of $\{a, b\}$ it is necessary in $M$ at $w$. Clearly, $||\mathrm{M}||^{\mathrm{M}}$ is a super set of $\{a, b\}$. Therefore, $\models_w^M \Box\mathrm{M}$.

Since $\Diamond$ is the dual of $\Box$, $\models_w^M \Diamond\mathrm{G}$ iff $\models_w^M \neg\Box\neg\mathrm{G}$ iff it is not the case that $\models_w^M \Box\neg\mathrm{G}$ iff $||\neg\mathrm{G}||^{\mathrm{M}}$ is not a super set of $\{a, b\}$. So since $||\neg\mathrm{G}||^{\mathrm{M}} = \{\mathrm{b}, \mathrm{d}\}$, $\models_w^M \Diamond\mathrm{G}$.

Likewise, since $||\mathrm{G}||^{\mathrm{M}} = \{\mathrm{a}, \mathrm{c}\}$ is not a superset of $\{a, b\}$, it is not the case that $\models_w^M \Box\mathrm{G}$. Therefore, $\models_w^M \neg\Box\mathrm{G}$. Thus, we see that the DR model has the truth condition that Kratzer intends.

### 2.4 Why no Doubly Relative Model handles the Alternative Intuition

Suppose there is a doubly relative model that handles the alternative intuition. Then there is a pointwise equivalent filter $M = \langle W, N, P\rangle$ that also handles the alternative intuition. Thus for every world $w$, $\models_w^M \Box\mathrm{G}$, $\models_w^M \Box\neg\mathrm{G}$), and $\models_w^M \neg\Box(\mathrm{G} \wedge \neg\mathrm{G})$. So by definition, $||\mathrm{G}||^{\mathrm{M}} \in \mathrm{N_w}$ and $||\neg\mathrm{G}||^{\mathrm{M}} \in \mathrm{N_w}$. Then since $M$ is closed under finite intersections, $||\mathrm{G}||^{\mathrm{M}} \cap ||\neg\mathrm{G}||^{\mathrm{M}} \in \mathrm{N_w}$. Therefore, $\models_w^M \Box\mathrm{G} \wedge \neg\mathrm{G}$ counter to the intuition.

### 2.5 Minimal Models Capture both Intuitions

Previously, we showed that every RDR model is propositionally equivalent to an augmented minimal model. We apply the construction used in proving the theorem to find the minimal model which captures Kratzer's intuition.

Let $M = \langle W, f, g, P\rangle$ be the DR model we showed to capture Kratzer's intuition. Then according to the theorem, we let $N = \langle W, N, P\rangle$ be an augmented minimal such that $\bigcap N_w = bot(\leq_{g(w)})$. By definition, $\models_w^N \Box\mathrm{A}$ iff $||\mathrm{A}||^{\mathrm{N}} \in \mathrm{N_w}$ iff $\bigcap N_w \subseteq ||\mathrm{A}||^{\mathrm{N}}$ iff $bot(\leq_{g(w)}) \subseteq ||\mathrm{A}||^{\mathrm{N}}$ iff $\models_w^M \mathrm{A}$. So, clearly, a sentence is true in $M$ at $w$ iff it is true in $N$ at $w$.

With Kratzer's intuition out of the way, we return to the alternative intuition. The alternative intuition is that residents of New Zealand are obligated by the consequences of each ruling individually, but that combinations of rulings do not establish obligations. In terms of neighborhoods, closure under logical consequence corresponds to the idea of supplementation.

Let $N = \langle W, N, P\rangle$ be a minimal model such that for all $w$, $N_w$ is supplementation of $\{||\mathrm{M}||^{\mathrm{M}}, ||\mathrm{G}||^{\mathrm{M}}, ||\neg\mathrm{G}||^{\mathrm{M}}\}$. Then since $||\mathrm{G}||^{\mathrm{N}} \in \mathrm{N_w}$, $||\mathrm{G}||^{\mathrm{M}} \in \mathrm{N_w}$. Therefore, $\models_w^N \Box\mathrm{G}$. Likewise, $\models_w^N \neg\mathrm{G}$. But since the empty set is not a superset of any set, it is not $N_w$. Therefore, $\models_w^N \neg\Box(\mathrm{G} \wedge \neg\mathrm{G})$.

### 2.6 Relating DR Models and Filters

A neighborhood model which is supplemented and in addition is closed under finite intersections and contains the unit is called a *filter*.

**Theorem 4.** *The class of DR models is pointwise equivalent to the class of filters.*

*Proof outline.* This is the main theorem of the paper. Presenting its detailed proof would take too much space here (we would need new technical notions, which only have instrumental value for the completion of the proof). So, we will only present an outline with the main ideas.

In the first part half of the proof, we show that for any DR model we can construct a pointwise equivalent filter. We accomplish the construction by defining the neighborhood of a minimal model to validate just those sentences which are valid in the DR model. Thus pointwise equivalence follows trivially. The rest of this part is dedicated to showing that the minimal model is indeed a filter.

The second part requires much more work. We take advantage of the fact that every filter has a nesting.[3] Using one such nesting, we construct the ordering source for the DR model. (We just let the modal base be empty.) The construction is accomplished by slicing the intersection of the defined nesting out of the nesting itself and letting the ordering source be the sliced nesting[4] together with the intersection, which was removed.

We prove four properties about the constructed ordering source. These give us enough properties to show that the DR model validates a sentence iff the filter does. We split the proof into two cases. First, we show modal sentences validated by the filter are validated by the DR model. Then we show the converse. This concludes the proof•

The theorem shows that the $DR$ models go beyond relational models. In fact there are examples of filters which are not augmented and that therefore have no relational (Kripkean) counterparts. The following example is adapted from David Lewis [L]. Suppose that two particles are separated by a distance of just over an inch. Let the set $W$ of possible worlds be a continuum expressing possible distances between the two particles. So, $16 \in W$ is the possible world in which the particles are 16 inches apart. Relative to $W$, propositions (sets of worlds) allow is to specify ranges for the distance separating the particles. For instance, the assertion that, "the two particles are separated by a distance of less than 3 inches" is represented with the proposition, $\{x \in W | x < 3\}$.

Let $w$ be a world in which our original assertion – that the two particles are separated by a distance of just over an inch – is true. Let $g_w$ be a nest of propositions compatible with this assertion so that a proposition $p \in g_w$ just in case $p$ is an interval $p = (a, b)$ such that $a = 1$ and $1 < b$. Now, choosing to call this set $g_w$ is no coincidence: $g_w$ as viewed as an ordering source, produces a doubly relative model for the world $w$ in which the particles are just over an inch apart. (Let the ordering source $f$ be empty.) To see this, just consider the typicality ordering $\leq_{g(w)}$ induced by $g_w$. Let $u \in W$. If $u < 1$, then $u$ is in none of the propositions of $g_w$ (i.e. $\gamma(u) = \emptyset$) – we know that the particles are more

[3] A set **C** of sets is a *nest* if and only if for every pair of sets $U, V$ in **C**, either $U \subseteq V or V \subseteq U$.

[4] This sliced nest is the nest obtained from the original nest by subtracting the intersection of the original nest from each element of the original nest.

than an inch apart, so any world where they are less than an inch a part isn't typical at all. If $1 < u$, on the other hand, then $\gamma(u) = \{(1, x)|u \leq x\}$. Therefore, if $v \in W$ such that $1 < v < u$, then $\gamma(u) \subseteq \gamma(v)$. Hence, $v \leq_{g(w)} u$. Furthermore, there is no most typical world under $\leq_{g(w)}$, since for all $u > 1$, there is a $v > 1$ such that $v < u$. The typicality ordering is bottomless, hence the DR model is not reduced.

Notice that for use in natural language, it doesn't really matter whether or not particles or points or that space is continuous. The very fact that we can consider an example of this sort shows that people can talk about points and continuous spaces. Notice further, that a filter equivalent to this model will not be augmented since there is no most typical world under the typicality ordering. Therefore there is no Kripke model pointwise equivalent to it.

## 3 Going beyond Double Relative Models

Even when Double Relative models can go beyond relational semantics when the underlying domain is infinite, they are guaranteed to have Kripkean counterparts when the domain is finite. In order to see this consider;

$B_w = \{||A||^M \colon M, w \models \Box A\}$

It is a consequence of Kratzer's definition that if W is finite then $B_w$ is guaranteed to be consistent in the sense that $\bigcap B_w$ is non-empty. But many of the graded modals that she or other people have considered do not have this feature. For example, the modality 'highly probable' lacks this feature. Or the modality 'legally obligatory' also lacks this feature when the underlying legal context encodes jurisprudence containing contradictory norms. Kratzer seems to confine her attention to a particular class of neighborhood models, namely those conforming filters. This seems to be limitative from a representational point of view.

Say that you are reasoning about an assembly line implementing certain standard for quality control. The background information is determined by high probability judgments about the line. Given the standards for quality control enforced in the line we have that it is highly probable for each piece fabricated in the line that it is non-defective (a piece is defective when its length is less than a threshold *min* and when its length is more than a threshold *max*). At the same time it is highly probable that some piece is defective. It seems that one should be able to draw sound inferences from this background. That one can conclude for example for lengths $t'$ between *min* and *max* that: 'If the length is $t'$ then the piece is non-defective'; or 'It is not (seriously) possible that there will not be defective pieces' ('It is expected that there will be defective pieces'); or 'It is likely that there will be some defective pieces', etc.

This background cannot be represented directly as a modal base. Apparently in a situation of this kind Kratzer needs to proceed indirectly by assuming a space of points (in the propositional representation) each one of which encodes a situation where some number of pieces are defective, and where there is information about the length of each piece. Then an ordering is constructed for this space. The points where certain small number $n$ of pieces (for a threshold value $n$) are defective and the rest sane will be in the bottom of the ordering (maximally probable). Points where all pieces are defective will be maximally improbable, and so on. Any proposition that is a superset of the bottom of the ordering will be considered highly probable. This gives us: 'it is highly probable that there is some defective piece'; but it does not give us for each piece p, 'it is highly probable that p is defective'. We can have at most: 'it is possible that piece p is defective'.

A situation of this sort can be straightforwardly represented via neighborhood models. For each $w$ we can directly use $f(w)$ as in the standard case. The propositions in $f(w)$ are the propositions receiving high probability. What is a possibility in this model? We can extrapolate from the standard case. In this case $A$ is possible if $||A||^M$ is compatible with $\bigcap f(w)$. In this new type of scenario we can define that $A$ is possible if $|A|$ intersects each proposition in $f(w)$. Then we can add rationality conditions to the neighborhood representation that fit the modality in question. For example in this case one would like to add supplementation and the constraint establishing that neighborhoods contain the unit. But it is clearly unreasonable to require closure under finite intersections.

So, it seems that one can do with neighborhoods all that one can do with Double Relative Models and more. But it is also true that some of the basic insights in neighborhood semantics can be supplemented with some of the ideas that inspired Double Relative models. Kratzer's analysis of contextual conditionals offers a clear illustration of this possibility.

## 4 Conditionals

Kratzer's basic idea about conditionals is that the function of the antecedent in evaluating an if-clause at a world $w$ is to restrict the set of worlds which are accessible from $w$. This is done formally by adding the proposition expressed by the antecedent to the modal base.

This proposal (based in David Lewis's ideas about the semantics of counterfactuals) has various well-known problems representing conditionals of various kinds. One of the them is the validation of an axiom establishing that $A \wedge B$ entails $A > B$, for all propositions $A, B$. It is easy to find counterexamples to this axiom for various types of conditionals, both of ontic and epistemic type. Stochastic counterexamples are the simplest that one can exhibit. Suppose that at time $t$ and with respect to some determinate chance set-up (officiating as background) a coin is tossed ($T$) and that it lands heads ($H$). Does one want to infer from that that $T > H$ (i.e. 'If the coin is tossed at time $t$ in this chance set-up, it lands heads')? This does not seem reasonable. Other types of counter-examples abound.

There are remedies for this kind of problem. We will propose a possible solution, and we will show that the solution in question can be implemented in an extension of the neighborhood semantics for conditionals proposed in [Ch]. Our intention is to show that the central semantic ideas defended by Kratzer can all be accommodated in a slight extension of the classical presentation of neighborhood semantics. Neighborhood semantics has the adequate resources and the representational flexibility needed to implement a two dimensional analysis of modalities (both monadic and dyadic) that is both sensitive to context and capable of tolerating weakly inconsistent background information.

### 4.1 Neighborhood models for conditionals

Neighborhood models of conditionals are offered in chapter 10 of [Ch]. In comparison with other kind of semantic account of conditionals they remain relatively unexplored. Perhaps due to the fact that these models for conditionals were proposed independently of models for unary modalities, and with different intended applications, the two accounts of modality remain also unrelated. One of the virtues of Kratzer's models is that she presents an integrated account of both modalities.

**Definition 10 (Ch Def. 10).** *A model $M = \langle W, f, P\rangle$ is a* neighborhood conditional model *iff:*

1. *$W$ is a set.*
2. *$f$ is a function which maps pairs of worlds and propositions to a set of propositions (i.e. $f : W \times \wp(W) \rightarrow \wp(\wp(W))$ ).*
3. *$P : N \rightarrow \wp(W)$.*

The truth conditions for conditionals proceed along familiar lines:

$\models^M_w A > B$ iff $||\mathrm{B}||^{\mathrm{M}} \in \mathrm{f}(\mathrm{w}, ||\mathrm{A}||^{\mathrm{M}})$.

Models where the conditionals neighborhoods form filters have been recently considered in order to represent non-monotonic notions of consequence - see, among other papers [Ben]. Nevertheless in an integrated representation one would like to establish some lawful connection between the content of $f(w)$ and the conditional neighborhood $f(w, ||\mathrm{A}||^{\mathrm{M}})$. Intuitively the latter encodes the result of supposing that $A$ is the case with respect to the conversational background given by $f(w)$. And one can capture this intuition by having a fixed partial ordering which regulates suppositional background changes. Obviously the more unstructured is the content of $f(w)$ the more difficult is to establish the desired relation. We will consider here only cases already analyzed by Kratzer, where the conversational background $f(w)$, for every world $w$, forms at least a filter, and we will restrict our analysis to finite domains. The extension to the infinite case is technically more complicated but it preserves the central ideas of the model. The extension for backgrounds that do not manage to form filters requires modifications that are beyond the scope of this note. It should be noted in passing though that the use of a partial, rather than a total ordering, extends some of the most common analysis that one can find in the literature on conditional logic, while preserving most of the basic intuitions advanced by Kratzer.

Given a fixed partial ordering $\leq$ we can define some useful notions. $C(\leq)$ is a *downward cone* for $\leq$ if and only if for every $z$ in $C(\leq)$ we have that for every $x$, such that $x \leq z$, that $x \in C(\leq)$. We can restrict here our attention to backgrounds $f(w)$ obtained by taking supersets of a given downward cone of $\leq$ for $f(w)$.

A chain of worlds will be called minimal if and only if it has a minimal point as endpoint. The intuitive idea is that the background $f$ is given by an augmented neighborhood whose intersection is determined by taking the union of a set of minimal chains of worlds. At least this is so if the domain is finite. Otherwise the notion of downward cone might also include infinite descending chains of worlds.

So, what is supposing that A is the case with respect to $f(w)$? When $C(\leq)$ is compatible with the proposition expressed by $A$ we can just take $C(\leq) \cap ||\mathrm{A}||^{\mathrm{M}}$. Otherwise we can implement the idea that supposing $A$ requires opening our mind with respect to $A$ first and then inputing $A$ in this suppositional scenario. This requires contracting both $A$ and its negation from $C(\leq)$ – this idea is first presented by Isaac Levi in [Le]. This requires, in turn, an explanation of how can we contract $A$ from a view that affirms $A$ by appealing to the fixed ordering $\leq$. This can be done in two steps. First we identify a halo of $\neg A$ points with respect to $C(\leq)$. This can be done via the following operation;

$\mathrm{h}(\neg A) = \{w \in ||\neg\mathrm{A}||^{\mathrm{M}}$: for some $z \in C(\leq)$, $z \leq w$ and for all $y$ such that $w > y$, $y \in ||\mathrm{A}||^{\mathrm{M}}$ $\}$

As a second step we can define the contraction in question as follows:

$||\mathrm{C}(\leq) \div \mathrm{A}||^{\mathrm{M}} = \{w \in ||\mathrm{A}||^{\mathrm{M}}$: $w \leq z$ for $z \in h(\neg A)$ $\}$

With the help of these preliminary definitions we can introduce now a conditional neighborhood model as follows:

**Definition 11.** *A model* $M = \langle W, s, P\rangle$ *is a* conditional neighborhood model *iff:*

1. *$W$ is a set.*
2. *$s$ is a function which maps pairs of world and propositions to propositions (i.e. $f : W \times \wp(W) \rightarrow \wp(W)$ ).*
3. *$\leq$ is a partial ordering on $W$.*
4. *$P : N \rightarrow \wp(W)$.*

The selection function $s$ is defined in terms of the contraction operation presented above;

$s(w, ||\mathrm{A}||^{\mathrm{M}}) = ||(\mathrm{C}(\leq) \div \mathrm{A}) \div \neg\mathrm{A}||^{\mathrm{M}} \cap ||\mathrm{A}||^{\mathrm{M}}$.

**Definition 12 (Ch Def. 7.2).** *Let $w$ be a world in a neighborhood model $M = \langle W, f, P\rangle$. Then:*

1. $\models^M_w A > B$ *iff* $s(w, ||\mathrm{A}||^{\mathrm{M}}) \subseteq ||\mathrm{B}||^{\mathrm{M}}$

Basically we have introduced one more selection function in a model that contains the double relative construction of Kratzer. $W$ officiates here as the modal base of Kratzer. The ordering has the role of her ordering source. And the selection function is constructed by performing a more sophisticated manipulation in the ordering than just taking a restriction with respect to the antecedent. Of course, when $C(\leq)$ entails the proposition expressed by $A$ the result of $s(w, ||\mathrm{A}||^{\mathrm{M}})$ need not yield $C(\leq)$ back. One might then select a strictly larger set of worlds as the representative of the supposition with $A$.

## 5 Conclusions and Related Work

Kratzer's account of modalities is based on two central contextual parameters given by the modal base and the ordering source. Formally these conversational backgrounds provide two main tools: an accessibility relation and a partial ordering defined over it. Kratzer manages to show that these two parameters can be used advantageously in order to actually go beyond the common approach to modality based on the use of relational semantics (what she calls the 'standard approach'. We argued above that neighborhood semantics can be used in order to represent most of what she is interested to represent, and that actually neighborhood representations can go beyond the expressive power of her approach when infinity matters (and when one wants to represent genuine conflict in the background).

But we also argued that some of the insights deployed in Double Relative models can be transposed into an enriched version of neighborhood semantics. The central idea is to enrich neighborhood models with a partial ordering. This ordering can play an explanatory role in determining the behavior of selection functions in modeling conditional reasoning.

But the use of an ordering both in hypothetical reasoning and in fixing the content of background information should be familiar from its use in neighboring disciplines. We are thinking about Economics as a paradigmatic example. It is true nevertheless that it is only recently that economists began to think about rationality without the help of a *total* ordering. And the use of partial orderings in encoding typicality or entrenchment is also relatively recent in philosophy and mathematical psychology. But it is clear that abandoning the assumption of completeness of an underlying ordering helps to obtain more realistic descriptions useful in all these disciplines.

Notice that in the previous section we followed Kratzer in distilling augmented neighborhoods from a given partial ordering. Our $C(\leq)$ is a generalization of Kratzer's use of the bottom of the partial ordering in order to determine the content of neighborhoods (used in the semantics of monadic modal operators). But an approach in terms of neighborhoods is more flexible and allows for more scattered representations of epistemic context. For example, we can follow the lead of Amrtya Sen in [Sen] and we can construct neighborhoods from partial orderings as follows:

$$Min(\leq) = \{P \in 2^W\text{: for every } x, y \text{ in } P,\ x = y \text{ and for no } z \in P,\ r \in W\ z > r\}$$

$Min(\leq)$ picks the propositions composed by (comparable and equi-preferred) worlds that fail to dominate any other world. If we take the supplementation of $Min(\leq)$ there is no guarantee that the corresponding neighborhood is even weakly consistent (its intersection can be empty). Nevertheless, we can define monadic and dyadic modalities with respect to this neighborhood by appealing to minor variants of the techniques used above. So, for example, the propositions included in the neighborhood obtained by taking the supplementation of $Min(\leq)$ might indicate what is known 'all things considered' - if the ordering is interpreted as a typicality ordering.

Moreover as Sen arguend in [Sen] (by appealing to ideas first developed by Stig Kanger in modal logic) it is also reasonable to make the partial ordering dependent on the menu of preferences as well as dependent on the chooser. As he argued in [Sen] much of the tools used in maximization (minimization) remain useful in this case, while the representation gains in realism.

## References


[AC] Arlo Costa, H., First order extensions of classical systems of modal logic:The role of the Barcan schemas, forthcoming in Studia Logica 71: 2002, 87-118.

[HAC] Arlo Costa, H., and Pacuit, E. First order classical modal logic, Technical Report No. CMU-PHIL-164, Carnegie Mellon University, 2004. An abbreviated version is forthcoming in TARK X.

[Ben] Ben-David, S., and Ben-Eliyahu, R., A Modal Logic for Subjective Default Reasoning". *Artificial Intelligence*, Volume 116, pp 217-236, 2000.

[Ch] Chellas, B. (1980). Modal logic. An introduction, Cambridge UP, Cambridge.

[K] Kratzer, A. (1991). Modality. In: von Stechow, A. and Wunderlich, D. ( eds.). Semantik. Ein internationales Handbuch der zeitgenossischen Forschung. S. 639-650. Walter de Gruyter, Berlin.

[K2] Kratzer, A. (1977). What 'must' and 'can' must and can mean. In: Linguistics and Philosophy, 1(3):337-356.

[Le] Levi, I.: 1996, *For the sake of the argument: Ramsey test conditionals, Inductive Inference, and Nonmonotonic reasoning*, Cambridge University Press, Cambridge.

[L] Lewis, D. (1973). Counterfactuals, Blackwell.

[M] Montague, R. (1968). Pragmatics. Contemporary Philosophy, 101-121. La Nuova Italia Editrice.

[Sen] Sen, A. (2003) Non-Binary Choice and Preference, in *Rationality and Freedom*, Harvard University Press, Cambridge, Mass.

[S] Scott, D. (1970). Advice on modal logic. Philosophical Problems in Logic, 143-173.