In: *Journal of Logic, Language and Information*, 1997.

Counterfactuals and updates as inverse modalities

Mark Ryan
School of Computer Science. University of Birmingham, Birmingham B15 2TT, UK. mdr@cs.bham.ac.uk, http://www.cs.bham.ac.uk/~mdr.

Pierre-Yves Schobbens
Institut d'Informatique, Facultés Universitaires de Namur, Rue Grandgagnage 21, 5000 Namur, Belgium pys@info.fundp.ac.be, http://www.info.fundp.ac.be/~pys.

Abstract. We point out a simple but hitherto ignored link between the theory of *updates*, the theory of *counterfactuals*, and classical modal logic: update is a classical existential modality, counterfactual is a classical universal modality, and the accessibility relations corresponding to these modalities are inverses. The Ramsey Rule (often thought esoteric) is simply an axiomatisation of this inverse relationship.

We use this fact to translate between rules for updates and rules for counterfactuals. Thus, Katsuno/Mendelzon's postulates U1–U8 are translated into counterfactual rules C1–C8 (table 7), and many of the familiar counterfactual rules are translated into rules for updates (table 8). Our conclusions are summarised in table 5.

From known properties of inverse modalities we deduce that not all rules for updates may be translated into rules for counterfactuals, and vice versa. We present a syntactic condition which is *sufficient* to guarantee that a translation from update to counterfactual (or vice versa) is possible.

1. Introduction

Background. An intuitive connection between theory change and counterfactuals was observed by F. P. Ramsey [19], who proposed what has become known as the Ramsey Rule:

> To find out whether the counterfactual 'if A were true, then B would be true' is satisfied in a state S, change the state S minimally to include A, and test whether B is satisfied in the resulting state.

(Actually, Ramsey proposed the rule only for non-counterfactual conditionals, but the term 'Ramsey Rule' is now taken to refer to counterfactuals too.)

It was initially hoped that the AGM theory of belief revision [7, 14] would provide the right notion of minimal change. However, the intuitively acceptable AGM postulates for belief revision are known to be incompatible with the Ramsey Rule [6, 7].

It turns out that the theory of updates proposed by Katsuno and Mendelzon [11] is compatible with the Ramsey Rule [9]. Updates, like revisions, are a formalisation of theory change; but whereas revisions are intended to model changing beliefs about a fixed world, updates are intended to model a changing world. The difference between the formalisations of updates and revisions can be seen in terms of postulates; for example, the AGM postulate

$$A * B = A \wedge B \quad \text{if } A \wedge B \text{ is consistent}$$

is accepted for revisions, but rejected for updates. The difference can also be seen in terms of operations on models; in revision, we measure the distance to the models of the old theory as a whole, while in update we measure the distance to them pointwise. The intuition behind this crucial difference (expanded upon in section 3.1) is that in updates we want to change total states of the world, whereas in revisions we change partial descriptions of it. Though they were discovered more recently, updates are conceptually simpler than revisions, because of this pointwise character.

Our contribution. We capitalise on the realisation that updates are the right notion of theory change for the Ramsey rule. We show that the standard treatments of updates (e.g. [11]) is a system of multimodal logic. Furthermore, this modal logic bears a particular relationship with the modal logic of counterfactuals [22, 13, 17]: it has the *inverse* accessibility relation. It is therefore appropriate to present a single modal logic, with positive and negative modalities; the positive modalities (used for updates) refer to the accessibility relation R, and the negative ones (used for counterfactuals) refer to the relation R^{-1}.

The Ramsey rule turns out to be an axiomatisation of the inverse relationship between the two sets of modalities.

We work out the *correspondence properties* of the accessibility relation for the standard rules for updates and counterfactuals (tables 4 and 6), and we provide a *systematic translation* of rules for updates into rules for counterfactuals, and conversely (theorem 14). The paper extends two previous papers [21, 20].

Structure. The paper is arranged as follows. Section 2 contains preliminaries about modal logic and (inverse) accessibility relations. In section 3, we show that updates and conditionals are systems of multimodal logic, and that they have inverse accessibility relations. In section 4, we show that this is equivalent to the Ramsey rule, and we translate the standard rules for update into conditional logic rules, and vice versa. Conclusions are in section 5.

2. Preliminaries

This section sets our notation and provides background technical details about the version of multi-modal logic we will use, and about inverse modalities. The modally-expert reader can skim it quite fast.

2.1. Multi-modal logic

We assume a propositional language L with atomic propositions $p, q, r, \ldots$ and connectives $\wedge, \vee, \neg, \rightarrow, \leftrightarrow, \Box, \Diamond, \boxminus, \diamondminus$. The unary connectives bind most tightly, then $\wedge, \vee$, and $\rightarrow, \leftrightarrow$ bind least tightly. The modal connectives take two arguments; if A, B are formulas then so are $\Box_A B$, $\Diamond_A B$, $\boxminus_A B$, $\diamondminus_A B$. (The formula $\Diamond_A B$ will later be read as the result of updating B by A; the formula $\boxminus_A B$ will be read as the counterfactual 'if A were the case, B would be the case'. The other two modal formulas are their duals.) The set L is the set of atomic formulas $p, q, r, \ldots$; the set $\mathbf{L}$ is the set of all formulas over L.

The semantics of multi-modal logic with inverses is given as follows. A *model* $M = \langle W, R, V \rangle$ of the multi-modal language L is a set W of *worlds*, an *accessibility relation* $R \subseteq \mathcal{P}(W) \times W \times W$ and a *valuation* $V : L \rightarrow \mathcal{P}(W)$.

The relation $\Vdash$ of *satisfaction* between a model $M = \langle W, R, V \rangle$, a world $x \in W$ and a formula A is defined inductively on A as follows.

$$
\begin{array}{rl}
x \Vdash_M p & \text{iff } x \in V(p) \\
x \Vdash_M \neg A & \text{iff } x \nVdash_M A \\
x \Vdash_M A \wedge B & \text{iff } x \Vdash_M A \text{ and } x \Vdash_M B \\
x \Vdash_M \Box_A B & \text{iff for each } y \in W,\ R_{|A|}(x, y) \text{ implies } y \Vdash_M B \\
x \Vdash_M \boxminus_A B & \text{iff for each } y \in W,\ R_{|A|}(y, x) \text{ implies } y \Vdash_M B
\end{array}
$$

The missing connectives $\vee$, $\rightarrow$, $\leftrightarrow$, $\Diamond$, $\diamondminus$ are defined by similar (standard) clauses. As can be seen, $\boxminus, \diamondminus$ are like $\Box, \Diamond$ except that they refer to the inverse accessibility relation.

In the context of a model M, $|A|$ is defined to be $\{x \in W \mid x \Vdash_M A\}$. The subscript on $\Vdash_M$ will usually be dropped in order to make the notation lighter.

The model M satisfies the formula A, written $M \Vdash A$, if $x \Vdash_M A$ for each $x \in W$. A *frame* $F = \langle W, R \rangle$ consists of a set of worlds and an accessibility relation. Such a frame F satisfies A, written $F \Vdash A$, if for each valuation V, we have $\langle W, R, V \rangle \Vdash A$. A formula A is *valid*, written $\models A$, if it is satisfied by every frame. A formula A is *satisfiable* in a model M if $|A| \neq \emptyset$. The formula A over L is *complete* *w.r.t.* $M = \langle W, R, V \rangle$ if there is precisely one $x \in W$ with $x \Vdash A$. If $A_1, A_2, \ldots, A_n, B$ are formulas, the rule

$$\frac{A_1 \quad A_2 \quad \ldots \quad A_n}{B}$$

holds in a frame F if: for all V, if $M \Vdash A_i$ for each i then $M \Vdash B$, where $M = \langle F, V \rangle$. Notice that this is weaker than asserting the axiom $A_1 \wedge \ldots \wedge A_n \rightarrow B$, but stronger than asserting 'if each A_i is valid, then B is valid'. The double-barred rule $\overset{A}{\underset{B}{=}}$ holds in F if for all $M = \langle F, V \rangle$, $M \Vdash A$ iff $M \Vdash B$.

This multi-modal logic is not the straightforward one discussed in [8] and [18], because the indices are formulas. However, our modalities are not quite the binary modalities found in [25], because they treat the two arguments A and B in quite different ways (and, for example, $\neg\Box_A B$ is equivalent to $\Diamond_A \neg B$, but not to $\Diamond_{\neg A} \neg B$ as in [25]). Nevertheless, our definitions also seem natural, and as will be seen, they are right for our application.

2.2. Inverse modalities

Inverse modalities have already been used in modal logics: in linear temporal logic, they are the *past* modalities. Table 1 summarises their intuitive meaning. These inverse modalities should not be confused with the dual modalities, nor with the inverse of the dual (which is of course dual of the inverse).

Not all modalities have intuitively interesting inverses. Temporal and dynamic modalities do, but epistemic and doxastic modalities do not. If one interprets $\Box B$ as 'I believe B', then $\boxminus B$ seems to say: 'in all situations where my beliefs admit the current situation, B is true'.

In later sections, we discover that the counterfactual modality has an interesting inverse: its inverse is (the dual of) update.

2.2.1. *Axiomatising inverse modalities*

How can we axiomatise the link between $\Box$ and $\boxminus$? There are two ways. The first is to add a pair of axioms (whose monomodal versions are already known from temporal logic [8, ex. 6.1]). The second way is the Ramsey Rule.

Theorem 1 (Folklore) *Assume $\Box, \Diamond$ are interpreted by the accessibility relation R and $\boxminus, \diamondminus$ are interpreted by S in a frame F.*

The following are equivalent:

1. *For each $T \subseteq W$, $R_T = S_T^{-1}$.*

Table 1. Modalities, their inverses and duals

modality $\Box B$	inverse $\boxminus B$	dual $\Diamond B$	inverse dual $\diamondminus B$
henceforth B	up to now, B	eventually B	once upon a time, B
tomorrow B	yesterday B	tomorrow B	yesterday B
any execution of program P in the current state results in a state satisfying B	if P has just been executed, the starting state satisfied B	there is an execution of program P in the current state which results in a state satisfying B	P could have just executed in a state satisfying B

2. *F satisfies the following axiom schemes (each of which is also given in its dual form):*

$$\begin{array}{lll} (1) & B \to \Box_A \diamondminus_A B & \Diamond_A \boxminus_A B \to B \\ (2) & B \to \boxminus_A \Diamond_A B & \diamondminus_A \Box_A B \to B. \end{array}$$

3. *F satisfies the following rule (which is also given in its dual form):*

$$\frac{B \to \boxminus_A C}{\Diamond_A B \to C}\ \mathrm{RR} \qquad\qquad \frac{\diamondminus_A C \to B}{C \to \Box_A B}\ \mathrm{RRd}.$$

(The proof of this and all other theorems is given in the Appendix.)

Observe that with the reading we will give to $\boxminus$ and $\Diamond$, the rule RR exactly expresses the Ramsey rule for counterfactuals. $\boxminus_A C$ is read as 'if A were true, then C would be true; $\Diamond_A B$ is read as the update of B by A; so the rule states that the counterfactual 'if A, C' is supported in a state B iff the state obtained by updating B with A supports C. Notice that formulas are used to denote both *pieces of information* and *(descriptions of) states*.

[Aside. Those knowledgeable about category theory may like to see the rule of Theorem 1(3) as an *adjunction between pre-orders*. The pre-order in question is $\mathcal{P}(W)$ (given a fixed model $M = \langle W, R, V\rangle$) ordered by inclusion. For any fixed A, $\boxminus_A$ and $\Diamond_A$ may be considered as the following monotonic operators on the preorder:

$$\Diamond_A S = \{x \in W \mid \exists y \in S\ R_{|A|}(x, y)\}$$
$$\boxminus_A S = \{y \in W \mid \forall x \in W\ (R_{|A|}(x, y) \Rightarrow x \in S)\}$$

and we have $S \subseteq \boxminus_A T$ iff $\Diamond_A S \subseteq T$.]

2.3. Other preliminaries

If $(X, \leq)$ is a pre-ordered set (i.e. $\leq$ is a reflexive and transitive ordering on X) and $Y \subseteq X$, then $\mathrm{Min}_{\leq}(Y)$ is the set of $\leq$-minimals in Y, i.e. $\mathrm{Min}_{\leq}(Y) = \{y \in Y \mid \forall x \in Y. x \not< y\}$.

If R is a relation in $W \times W$, then $R(a) = \{b \in W \mid R(a, b)\}$ and $R^{-1}(b) = \{a \in W \mid R(a, b)\}$.

3. Updates and counterfactuals

Our aim in this section is to show that updates and conditionals are systems of modal logic, having inverse accessibility relations to each other. In 3.1 we clarify the difference between 'update' and its cousin 'revision'. In 3.2 we show that updates are existential modalities. This enables us to look at correspondence properties, in the style of [23]. In 3.3 we recall the known result that counterfactuals are universal modalities, and see that the accessibility relation is the inverse of the one for updates. We recall some known correspondence theory, and use it to relate rules for update with rules for counterfactuals.

3.1. Updates vs revisions

This paper concerns the notion of 'update'. It is related to the notion of belief revision, and shares some properties. Indeed, there has been a historical confusion between the two notions. They differ in motivation, however, and they have some important technical differences. Before focussing on updates, we briefly recall the differences between update and revision.

The need for different notions was pointed out in [12]. Updates are intended to model a world which changes, while revisions are intended to model a static world, about which one's information changes. In [11] new postulates for updates were proposed. These postulates are similar to those for revisions [7] and are presented in Table 2. In the last column, we have indicated the name used in revision.

Katsuno/Mendelzon use $\diamond$ as an infix operator; $q \diamond p$ means q updated by p. It may appear surprising that the two arguments of $\diamond$ are the

Table 2. Update postulates according to Katsuno/Mendelzon [11], using their syntax. They use $\diamond$ as an infix operator; $q \diamond p$ means q updated by p.

name [11]	postulate	name [7]
U1	$q \diamond p \rightarrow p$	K*2
U2.1	$q \rightarrow p$ implies $q \rightarrow q \diamond p$	
U2.2	$q \rightarrow p$ implies $q \diamond p \rightarrow q$	K*4w
U3	$q \diamond p$ satisfiable, if p, q satisfiable	$\sim$ K*5
U4.1	$q \leftrightarrow r$ implies $q \diamond p \leftrightarrow r \diamond p$	
U4.2	$q \leftrightarrow r$ implies $p \diamond q \leftrightarrow p \diamond r$	K*6
U5	$(q \diamond r) \wedge p \rightarrow q \diamond (r \wedge p)$	K*7
U6	$q \diamond p \rightarrow r, q \diamond r \rightarrow p$ imply $q \diamond p \leftrightarrow q \diamond r$	
U7	q complete implies $(q \diamond p) \wedge (q \diamond r) \rightarrow q \diamond (p \vee r)$	
U8	$(q \vee r) \diamond p \leftrightarrow (q \diamond p) \vee (r \diamond p)$	

same syntactic type (namely, formulas), since one usually expects *states* to be updated by formulas. The explanation is that q is a formula denoting a set of states, and $q \diamond p$ is the formula denoting the resulting set of states after each of them has been updated by p.

Comparing the update postulates and the standard AGM postulates [7], one sees that updates and revisions have much in common, explaining the historical confusion. The fundamental difference between the two notions may be seen by looking at the *representation theorems* associated with each of them, rather than the postulates. [In the following two theorems, $|A|$ is the set of valuations making A true; thus, KM and AGM work on specific frames where valuations are in bijection with worlds.]

Theorem 2 ([7]) $*$ *is a revision operator iff*

$$|B * A| = \mathrm{Min}_{\leq_{|B|}}(|A|)$$

where $\leq_{|B|} \subseteq W \times W$ *is a preorder of closeness to* B *(*$x \leq_S z$ *means that* x *is at least as close to the set of worlds* S *as* z *is).*

Theorem 3 ([11]) $\diamond$ *is an update operator iff*

$$|B \diamond A| = \bigcup_{y \in |B|} \mathrm{Min}_{\leq_y}(|A|)$$

where $\leq_y \subseteq W \times W$ *is a preorder of closeness to* y *(*$x \leq_y z$ *means that* x *is at least as close to the world* y *as* z *is).*

Thus, while revisions measure closeness to $|B|$ as a whole, updates measure closeness to each element of $|B|$ individually. This is the reason for which, in [10], updates are referred to as *pointwise revisions*. The justification for the pointwise character can be explained in terms of the changing-world motivation behind updates. Suppose the world is in a state satisfying $p \vee q$. Suppose now a change is made, whose result guarantees $\neg p$. Belief revision would tell us that the resulting state must satisfy $\neg p \wedge q$, which is equivalent to the conjunction of these formulas; it says that whenever the conjunction is consistent, then it is the result of the revision. Updates, however, argue that if all we had was $p \vee q$, then the world either satisfied p or it satisfied q. Considering these cases independently, we see that after $\neg p$ is imposed the world satisfies $\neg p$ in the first case, or $\neg p \wedge q$ in the second. As we don't know which of the two cases it was, all we can be sure about now is that it satisfies the disjunction, namely $\neg p$.

3.2. Updates are existential modalities

Suppose we write $\Diamond_A B$ instead of $B \diamond A$. Looking at the representation theorem for updates given above, it is easy to see that it can be written

$$x \Vdash \Diamond_A B \quad \text{iff} \quad \text{there exists } y \text{ s.t. } R_{|A|}(x,y) \text{ and } y \Vdash B$$

in the multi-modal model $M = \langle W, R, V \rangle$, where W is the set of valuations and the relation R is given by

$$R_S(x,y) \Leftrightarrow x \in \mathrm{Min}_{\leq_y}(S).$$

This fact shows that updates are an existential modality, and justifies the decision to write $\Diamond_A B$ instead of $B \diamond A$. We can therefore think of Katsuno/Mendelzon's U1-U8 as multi-modal axioms and rules. They are presented as such in Table 3.

Katsuno/Mendelzon's theory of updates may be seen as a particular multi-modal logic, the one generated by the axioms and rules of Table 3.

To guarantee the classical properties of a modality, we should also have necessitation:

$$\frac{B}{\Box_A B}$$

Theorem 4 *Necessitation follows from the axioms and rules in Table 3.*

A corollary of our observation that update is an existential modality is that the Katsuno/Mendelzon theory represents one particular logic in a hierarchy, whose intuitive *base level* is weaker, being a minimal normal

Table 3. Update postulates rewritten as modal logic rules

name [11]	rewritten as
U1	$\Diamond_A B \to A$
U2.1	$\dfrac{B \to A}{B \to \Diamond_A B}$
U2.2	$\dfrac{B \to A}{\Diamond_A B \to B}$
U3	$\dfrac{\neg\Diamond_A B}{\neg B}$ if A satisfiable
U4.1	$\dfrac{B \leftrightarrow C}{\Diamond_A B \leftrightarrow \Diamond_A C}$
U4.2	$\dfrac{B \leftrightarrow C}{\Diamond_B A \leftrightarrow \Diamond_C A}$
U5	$\Diamond_A B \wedge C \to \Diamond_{A\wedge C} B$
U6	$\dfrac{\Diamond_A B \to C \quad \Diamond_C B \to A}{\Diamond_A B \leftrightarrow \Diamond_C B}$
U7	B complete implies $\Diamond_A B \wedge \Diamond_C B \to \Diamond_{A\vee C} B$
U8	$\Diamond_A (B \vee C) \leftrightarrow \Diamond_A B \vee \Diamond_A C$

modal logic. One could consider stronger or weaker logics, according to applications.

Some of the rules in Table 3, namely U4.1, U4.2 and U8, are automatically valid, simply by virtue of the modal semantics:

Theorem 5 *The rules U4.1, U4.2 and U8 hold in any frame.*

The other rules are valid if we suitably constrain the accessibility relation. As usual within the framework of modal logic, we can study the 'correspondence properties' on R imposed by an axiom or rule.

Theorem 6 *A rule in Table 3 holds in a frame $F = \langle W, R\rangle$ iff R has the corresponding property stated in Table 4.*

Table 4. Correspondence conditions for the update rules

name	property of R
R1	$R_S(x, y)$ implies $x \in S$
R2.1	$y \in S$ implies $R_S(y, y)$
R2.2	$y \in S$ and $R_S(x, y)$ imply $x = y$
R3	$S \neq \emptyset$ implies $\forall y \exists x . R_S(x, y)$
R5	$x \in S$ and $R_T(x, y)$ imply $R_{S \cap T}(x, y)$
R6	$R_S^{-1}(y) \subseteq T$ and $R_T^{-1}(y) \subseteq S$ imply $R_S^{-1}(y) = R_T^{-1}(y)$.
R7	$R_S \cap R_T \subseteq R_{S \cup T}$

Compare correspondence theorems for standard modal logic, e.g. [18, §5.2], [8, theorems 1.12, 1.13],[23]. The proofs (given in the Appendix) follow the usual pattern in correspondence theory. In the $\Leftarrow$ direction, we add to the frame $\langle W, R \rangle$ an arbitrary valuation V to form the model $M = \langle W, R, V \rangle$, and show that the constraint on R is enough to guarantee that the rule is satisfied in M. In the $\Rightarrow$ direction, we make a judicious choice of the valuation and an instance of the scheme, to show that the constraint on R must hold.

Notice that these conditions are second-order, and (unlike the case for mono-modal logic) none of them can be reduced to first order conditions. This is because the R is always indexed by a set.

Note that the 'pointwise character' of updates is crucial in their ability to be represented as an existential modality: it implies to the distribution of $\Diamond$ over $\vee$ (U8). Revisions are not existential modalities in this sense. (Revisions may however be analysed as modalities in a rather different sense, for example [4, 24].)

3.3. Counterfactuals are universal modalities

According to [22, 13, 17], the counterfactual 'if A was the case, then B would be the case' may be interpreted by: "In all closest worlds satisfying A, we find that B holds." It is well-known that counterfactuals have the properties of classical universal modalities [3, 13]. The counterfactual 'if A was the case, then B would be the case' holds at a world x if B holds in all y in $\mathrm{Min}_{\leq_x}|A|$. But this relation between x and y is simply the inverse of the relation $R_{|A|}$ given at the beginning of section 3.2. So counterfactuals and updates are *inverses* (at the level of accessibility relations), and *duals* (since updates are existential modalities, while counterfactuals are universal). In terms of our logic, the counterfactual sentence 'if A was the case, then B would be the case' can be written $\boxminus_A B$.

Slogan: Counterfactuals are the inverse dual of updates.
Updates are the inverse dual of counterfactuals.

One can perform the same analysis as we did for updates, namely, the correspondence theory for standard rules for counterfactuals. Much of this is known, but tends to be scattered in the literature, and often the semantic conditions are not put into one-to-one correspondence with the axiom schemes. We have collated results from many sources in Tables 5 and 6.

These tables show that several authors have re-discovered axiom schemes under different names, including ourselves! They also shows the correspondence between some update rules and counterfactual rules, obtained by noticing that they correspond to the same semantic condition on the relation. This is the topic of the next section.

Note that we are choosing to work with the Kripke accessibility relation R, while many authors cited in Table 5 use the selection function f. Of course, the relation between them is trivial, and which one is used is merely a matter of convention or personal preference. $R_{|A|}(x, y)$ expresses that x is closest among $|A|$ to y, while $f(A, y)$ selects exactly those x's. We therefore have $x \in f(A, y)$ iff $R_{|A|}(x, y)$, and may write $f(A, y) = R^{-1}_{|A|}(y)$.

As a convention, we will use the leftmost names from the table in the remainder of the paper. Note that our assumption that counterfactuals are *normal modalities* renders some postulates in the table (e.g. ID, RI in the first row) equivalent, though these are sometimes distinguished in the literature.

4. Inter-translating systems for counterfactuals and updates

4.1. Via the correspondence conditions

We have observed that rules for updates correspond to particular properties of the accessibility relation R, and similarly for counterfactuals, whose rules correspond to properties of the inverse relation R^{-1}. This gives us a criterion for identifying a particular rule for updates and a rule for counterfactuals.

Example 7 The update axiom scheme U1, $\Diamond_A B \to A$, corresponds to the semantic condition $R_S(x, y)$ implies $x \in S$ (theorem 6). The

Table 5. The correspondence between update rules and counterfactual rules. The table has three major columns (the minor subcolumns show how different authors give different names to the same thing). The first column shows the name of an update axiom scheme or rule, the third column shows the name of the corresponding counterfactual rule. The second column shows the name of the condition on the accessibility relation.

Update rule		semantic condition			Counterfactual rule							
[11]	here	here	[16]	[3]		[16, §12]	[3]	[2]	[1]	[7, §7.2]	[5]	[15]
U1		R1	CS1	id	C1	ID	ID	A0, B0	ID, RI	A4	G4	RCE
U2.1		R2.1	CS2.1	mp	C2.1	MP	MP	D1(b)	MP	A5	G7	A4
U2.2		R2.2	CS2.2		C2.2	CS		D1(a)	CS	A6	G6	CS
U3		R3			C3							
U5		R5			C5							
U6		R6	CS4		C6	CSO		B3	CE	A8		CSO
U7		R7			C7							
	U CS3		CS3		C CS3							
	UV		CS5			CV			AS		G9	CV
	U A3	R A3						A3	ASC			
	U SDA	R SDA		cm′		SDA, SA	CM′					S*, WA
	UA	RA		cc′		CA	CC′	A4	AD		G8	A7
	UN′	RN′		cn′			CN′					
	U MOD	R MOD				MOD						MOD
	UEM	REM				CEM						CEM
	U Tr	R Tr				Tr						
	U Contr	R Contr				Contr						
	Ut3	Rt3							t3			
	Ut4	Rt4							t4			
	UB2	RB2						B2				
	UB4	RB4						B4				
	UD′	RD′						D′				
	UD″	RD″						D″				
	UD0	RD0						D0				
	UA1	RA1										A1
	UA2	RA2										A2
	URT	RRT										RT, A6
	U Triv	R Triv										Triv

Table 6. Correspondence conditions for counterfactuals.

	Counterfactual rule		Condition on accessibility relation
C1	$\boxminus_A A$	R1	$R_S(x,y)$ implies $x \in S$
C2.1	$\boxminus_A B \to (A \to B)$	R2.1	$y \in S$ implies $R_S(y,y)$
C2.2	$A \wedge B \to \boxminus_A B$	R2.2	$y \in S$ and $R_S(x,y)$ imply $x = y$
C3	$\neg\boxminus_A -$ if A satisfiable	R3	$S \neq \emptyset$ implies $\forall y \exists x . R_S(x,y)$
C5	$\boxminus_{C \wedge A} B \to \boxminus_C (A \to B)$	R5	$x \in S$ and $R_T(x,y)$ imply $R_{S \cap T}(x,y)$
C6	$\boxminus_A C \wedge \boxminus_C A \to$ $(\boxminus_A D \leftrightarrow \boxminus_C D)$	R6	$R_S^{-1}(y) \subseteq T$ and $R_T^{-1}(y) \subseteq S$ imply $R_S^{-1}(y) = R_T^{-1}(y)$
C7	B complete implies $\boxminus_{A \vee C} \neg B \to \boxminus_A \neg B \vee \boxminus_C \neg B$	R7	$R_S \cap R_T \subseteq R_{S \cup T}$
CCS3	$\boxminus_A - \to \boxminus_C \neg A$	CS3	$R_S^{-1}(y) = \emptyset \to R_T^{-1}(y) \cap S = \emptyset$
CV	$\boxminus_A B \wedge \neg(\boxminus_A \neg C) \to \boxminus_{A \wedge C} B$	CS5	$R_T^{-1}(y) \cap S = \emptyset \vee R_{T \cap S}^{-1}(y) \subseteq R_T^{-1}(y)$
A3	$\boxminus_A B \wedge \boxminus_A C \to \boxminus_{A \wedge B} C$	RA3	$R_S^{-1}(y) \subseteq T \to R_{S \cap T}^{-1}(y) \subseteq R_S^{-1}(y)$
SDA	$\boxminus_{A \vee A'} B \to \boxminus_A B$	cm$'$	$S \subseteq T \to R_S \subseteq R_T$
CA	$\boxminus_A B \wedge \boxminus_{A'} B \to \boxminus_{A \vee A'} B$	cc$'$	$R_{S \cup T} \subseteq R_S \cup R_T$
CN$'$	$\boxminus_{-} A$	cn$'$	$R_\emptyset = \emptyset$
MOD	$\boxminus_{\neg A} A \to \boxminus_B A$	RMOD	if $R_{\neg S}^{-1}(y) \subseteq S, R_T^{-1}(y) \subseteq S$
CEM	$\boxminus_A B \vee \boxminus_A \neg B$	RCEM	R_S^{-1} is a (partial) function
Tr	$\boxminus_A B \wedge \boxminus_B C \to \boxminus_A C$	RTr	$R_S^{-1}(y) \subseteq T \to R_S^{-1}(y) \subseteq R_T^{-1}(y)$
Contr	$\boxminus_A B \to \boxminus_{\neg B} \neg A$	RContr	if $R_S^{-1}(y) \subseteq T, R_{W \setminus T}^{-1}(y) \subseteq S$
t3	$\boxminus_A B \to (\boxminus_A C \leftrightarrow \boxminus_{A \wedge B} C)$	Rt3	if $R_S^{-1}(y) \subseteq T, R_S^{-1}(y) = R_{S \cap T}^{-1}(y)$
t4	$\boxminus_A B \to (\boxminus_B C \leftrightarrow \boxminus_{A \vee B} C)$	Rt4	if $R_S^{-1}(y) \subseteq T, R_T^{-1}(y) = R_{S \cup T}^{-1}(y)$
B2	$\boxminus_A B \to \boxminus_{A \vee C}(B \vee C)$	RB2	if $R_S^{-1}(y) \subseteq T, R_{S \cup U}^{-1}(y) \subseteq T \cup U$
B4	$\dfrac{B \to A,\ A \wedge C \to B}{\boxminus_A C \to \boxminus_B C}$	RB4	$S \cap U \subseteq T \subseteq S$ and $R_S^{-1}(y) \subseteq U$ imply $R_T^{-1}(y) \subseteq U$
D$'$	$\boxminus_{A \vee B} \neg A \to$ $\boxminus_{A \vee C} \neg A \vee \boxminus_{C \vee B} \neg C$	RD$'$	if $R_{S \cup T}^{-1}(y) \cap S = \emptyset$ then $R_{S \cup U}^{-1}(y) \cap S = \emptyset$ or $R_{T \cup U}^{-1}(y) \cap U = \emptyset$
D$''$	$\boxminus_{A \vee B} A \vee \boxminus_{A \vee B} B$	RD$''$	RCEM and R1
D0	$\neg\boxminus_{\top} -$	RD0	$R_W^{-1}(y) \neq \emptyset$
A1	$\boxminus_{A \wedge \neg B}(A \to B) \to$ $(\boxminus_{\neg A} A \to \boxminus_{\neg B} B)$	RA1	if $R_{T \setminus S}^{-1}(y) \cap T \setminus S = \emptyset$ and $R_S^{-1}(y) \cap S = \emptyset$ then $R_T^{-1}(y) \cap T = \emptyset$
A2	$\boxminus_{A \wedge \neg B}(A \to B) \to \boxminus_A B$	RA2	$R_{S \setminus T}^{-1}(y) \cap S \setminus T = \emptyset \to R_S^{-1}(y) \subseteq T$
RT	$\boxminus_{A \wedge B} C \wedge \boxminus_A B \to \boxminus_A C$	RRT	$R_S^{-1}(y) \subseteq T \to R_S^{-1}(y) \subseteq R_{S \cap T}^{-1}(y)$
Triv	$\boxminus_A B \to \boxminus_{A \wedge \neg B}(A \to B)$	RTriv	if $R_S^{-1}(y) \cap T = \emptyset$ then $R_{S \cap T}^{-1}(y) \cap (S \cap T) = \emptyset$

counterfactual axiom scheme $\boxminus_A A$ known as ID corresponds to the same semantic condition:

$$\begin{aligned} F \Vdash \boxminus_A A &\Leftrightarrow \forall V, x, y\, (R^{-1}_{|A|}(x,y) \Rightarrow y \Vdash A) \\ &\Leftrightarrow \forall x, y\, (R_S^{-1}(x,y) \Rightarrow y \in S) \\ &\Leftrightarrow \forall x, y\, (R_S(x,y) \Rightarrow x \in S) \end{aligned}$$

Indeed, U1 and ID intuitively say the same thing.

Using the criterion, we can look for a counterfactual rule which corresponds to an update rule, or vice versa.

Example 8 The update rule U2.1

$$\frac{B \to A}{B \to \Diamond_A B}$$

corresponds to the semantic condition $y \in S$ implies $R_S(y,y)$. As it happens, this semantic condition is equivalent to the same condition on R^{-1}, so it corresponds to the counterfactual rule

$$\frac{B \to A}{B \to \diamondminus_A B}$$

This rule may be more simply stated as $\boxminus_A B \to (A \to B)$ (known as MP in the literature) [for proof, see appendix].

Example 9 U6 $\Leftrightarrow$ C6. Let's first work out the correspondence condition R6 for U6.

$$F \text{ satisfies } \frac{\Diamond_A B \to C \quad \Diamond_C B \to A}{\Diamond_A B \leftrightarrow \Diamond_C B} \qquad \text{U6} \qquad (1)$$

$$\Leftrightarrow \frac{\bigcup_{b \Vdash B} R^{-1}_{|A|}(b) \subseteq |C| \quad \bigcup_{b \Vdash B} R^{-1}_{|C|}(b) \subseteq |A|}{\bigcup_{b \Vdash B} R^{-1}_{|A|}(b) = \bigcup_{b \Vdash B} R^{-1}_{|C|}(b)}, \text{ all } V \qquad (2)$$

$$\Leftrightarrow \frac{R^{-1}_{|A|}(b) \subseteq |C| \quad R^{-1}_{|C|}(b) \subseteq |A|}{R^{-1}_{|A|}(b) = R^{-1}_{|C|}(b)}, \text{ all } V \qquad (3)$$

$$\Leftrightarrow \frac{R_S^{-1}(b) \subseteq T \quad R_T^{-1}(b) \subseteq S}{R_S^{-1}(b) = R_T^{-1}(b)} \qquad \text{R6} \qquad (4)$$

From 1 to 2: recall F satisfying the rule means that any $M = \langle F, V\rangle$ satisfying the top also satisfies the bottom; recall that M satisfies an implication if the worlds satisfying the antecedent are contained in those satisfying the consequent; and the worlds satisfying $\Diamond_A B$ are $\bigcup_{b \Vdash B} R^{-1}_{|A|}(b)$.

From 2 to 3: The direction $2 \Rightarrow 3$ is by taking the special case $|B| = \{b\}$. For the other direction, we use the fact that for any sets S_i, T, we have $\bigcup_{i\in I} S_i \subseteq T$ iff for each i, $S_i \in T$. From 3 to 4: The direction $4 \Rightarrow 3$ is immediate, since, given V, $|A|$, $|B|$ are just particular sets. For the reverse direction, if we are given S, T we pick atomic A, C and choose V such that $V(A) = S$, $V(C) = T$.

Now we must find a corresponding update rule. Let's 'guess' that it is C6 := $\boxminus_A C \wedge \boxminus_C A \rightarrow (\boxminus_A D \leftrightarrow \boxminus_C D)$, and verify that it corresponds to the same condition on R.

$$
\begin{aligned}
&x \Vdash \boxminus_A C \wedge \boxminus_C A \rightarrow (\boxminus_A D \leftrightarrow \boxminus_C D)\\
&\quad\Leftrightarrow\ (R^{-1}_{|A|}(x,y) \Rightarrow y \Vdash C) \wedge (R^{-1}_{|C|}(x,y) \Rightarrow y \Vdash A) \Rightarrow\\
&\qquad\quad ([R^{-1}_{|A|}(x,y) \Rightarrow y \Vdash D] \Leftrightarrow [R^{-1}_{|C|}(x,y) \Rightarrow y \Vdash D]) \qquad (2)\\
&\quad\Leftrightarrow\ R^{-1}_{|A|}(x) \subseteq |C| \wedge R^{-1}_{|C|}(x) \subseteq |A| \Rightarrow R^{-1}_{|A|}(x) = R^{-1}_{|C|}(x) \qquad (3)
\end{aligned}
$$

3 comes from 2 because D is arbitrary. So U6 and C6 both correspond to R6.

Experience can make the guessing easier! However, in the next section we describe a more deterministic way of doing the translation.

Theorem 10 *The update rules U1-U8 translate to the counterfactual rules given in table 7.*

Notice that this modal logic perspective gives us a *spectrum* of counterfactual logics (any selection of the rules defines a logic), and also a spectrum of update logics; and these spectra are put into one-one correspondence by the set of constraints they imply on R.

In the opposite direction, we have taken standard counterfactual rules in the literature, and worked out the corresponding update rules in Table 8.

4.2. Via the Ramsey rule

The proof of the equivalence between update and counterfactual rules can be performed

- either by going via the accessibility relation R, as in the examples above;

Table 7. Counterfactual rules corresponding to the update rules U1–U8

name	counterfactual rule	name [16, §12]
C1	$\boxminus_A A$	ID
C2.1	$\boxminus_A B \to (A \to B)$	MP
C2.2	$A \wedge B \to \boxminus_A B$	CS
C3	$\neg\boxminus_A \bot$ if A satisfiable	
C5	$\boxminus_{C\wedge A} B \to \boxminus_C (A \to B)$	
C6	$\boxminus_A C \wedge \boxminus_C A \to (\boxminus_A D \leftrightarrow \boxminus_C D)$	CSO
C7	B complete implies $\boxminus_{A\vee C}\neg B \to \boxminus_A \neg B \vee \boxminus_C \neg B$	

– or by working directly with the axioms of theorem 1(2); or, equivalently, the Ramsey Rule (or its dual) in theorem 1(3).

Here we do it working directly with the axioms and rules.

Example 11 U1 $\Leftrightarrow$ C1. Using the Ramsey rule, U1: $\Diamond_A B \to A$ translates immediately into $B \to \boxminus_A A$. But this can be further simplified, to $\boxminus_A A$. [Note: $B \to \boxminus_A A$ and $\boxminus_A A$ are not logically equivalent, but they are 'frame' equivalent, as was shown in example 7. This means they are equivalent axiom *schemes*.]

Example 12 U5 $\Leftrightarrow$ C5.

'$\Rightarrow$'

$$\begin{array}{lr}
\Diamond_C \boxminus_{C\wedge A} B \wedge A \to \Diamond_{C\wedge A} \boxminus_{C\wedge A} B & \text{U5} \\
\Diamond_C \boxminus_{C\wedge A} B \to (A \to \Diamond_{C\wedge A} \boxminus_{C\wedge A} B) & \text{equiv.} \\
\Diamond_C \boxminus_{C\wedge A} B \to (A \to B) & \text{by (1)} \\
\boxminus_C \Diamond_C \boxminus_{C\wedge A} B \to \boxminus_C (A \to B) & \text{Nec,K} \\
\boxminus_{C\wedge A} B \to \boxminus_C (A \to B) & \text{by (2)}
\end{array}$$

'$\Leftarrow$'

$$\begin{array}{lr}
\boxminus_{A\wedge C} \Diamond_{A\wedge C} B \to \boxminus_A (C \to \Diamond_{A\wedge C} B) & \text{C5} \\
B \to \boxminus_A (C \to \Diamond_{A\wedge C} B) & \text{by (2)} \\
\Diamond_A B \to (C \to \Diamond_{A\wedge C} B) & \text{by RR} \\
\Diamond_A B \wedge C \to \Diamond_{A\wedge C} B & \text{equiv.}
\end{array}$$

Table 8. Update rules corresponding to some counterfactual rules

Counterfactual rule	Update rule	
CCS3	U CS3	$\dfrac{\neg\Diamond_A B}{\neg(\Diamond_C B \wedge A)}$
A3	U A3	$\dfrac{\Diamond_A C \to B}{\Diamond_{A\wedge B} C \to \Diamond_A C}$
SDA	U SDA	$\Diamond_A B \to \Diamond_{A\vee A'} B$
CA	UA	$\Diamond_{A\vee A'} B \to \Diamond_A B \vee \Diamond_{A'} B$
CN′	UN′	$\neg\Diamond_{\_}\top$
MOD	U MOD	$\dfrac{\Diamond_{\neg A} B \to A}{\Diamond_C B \to A}$
Tr	U Tr	$\dfrac{\Diamond_A D \to B}{\Diamond_A D \to \Diamond_B D}$
Contr	U Contr	$\dfrac{\Diamond_A C \to \neg B}{\Diamond_B C \to \neg A}$
t3	Ut3	$\dfrac{\Diamond_A C \to B}{\Diamond_A C \leftrightarrow \Diamond_{A\wedge B} C}$
t4	Ut4	$\dfrac{\Diamond_A C \to B}{\Diamond_B C \leftrightarrow \Diamond A \vee BC}$
B2	UB2	$\Diamond_{A\vee B} C \to B \vee \Diamond_A C$
B4	UB4	$\dfrac{B \to A, A \wedge C \to B, \Diamond_A D \to C}{\Diamond_B D \to C}$
D0	UD0	$\dfrac{\neg\Diamond_\top A}{\neg A}$
A1	UA1	$\dfrac{\Diamond_{A\wedge\neg B} C \wedge A \to B, \Diamond_{\neg A} C \to A}{\Diamond_{\neg B} C \to B}$
A2	UA2	$\dfrac{\Diamond_{A\wedge\neg B} C \wedge A \to B}{\Diamond_A C \to B}$
RT	URT	$\dfrac{\Diamond_A C \to B}{\Diamond_A C \to \Diamond_{A\wedge B} C}$
Triv	UTriv	$\dfrac{\Diamond_A C \to B}{\Diamond_{A\wedge\neg B} C \wedge A \to B}$

Proving the equivalence of update rules and counterfactual rules by the Ramsey rule can be mechanised to some extent. We present syntactic criteria on counterfactual rules which allow them to be translated to update rules, and conversely, using the Ramsey Rule.

Definition 13 We say a formula is '$\boxminus$-simple' if it is formed from arbitrary non-modal formulas and the operators $\boxminus$ and $\wedge$, where all subscripts of $\boxminus$ are non-modal.

A rule of inference is $\boxminus$-simple if:

- its premises are implications whose antecedent is non-modal and whose consequent is $\boxminus$-simple; and
- its conclusion is an implication whose antecedent and consequent are $\boxminus$-simple.

Dually, $\Diamond$-simple formulas are formed from non-modal formulas and the operators $\Diamond$ and $\vee$, where all subscripts of $\Diamond$ are non-modal. A rule is $\Diamond$-simple if:

- its premises are implications whose antecedent is $\Diamond$-simple and whose consequent is non-modal; and
- its conclusion is an implication whose antecedent and consequent are $\Diamond$-simple.

$\Box$-simple, $\diamondminus$-simple formulas and rules are defined similarly.

Theorem 14 *Any $\boxminus$-simple counterfactual inference rule can be translated into a $\Diamond$-simple update inference rule (and conversely) using the Ramsey Rule.*

The proof is an algorithm performing the translation. For readability, we treat $\boxminus$-simple rules; the three other cases are similar. We first massage the conclusion for application of the Ramsey Rule, by introducing a fresh meta-variable, say X, and replacing the conclusion $S_1 \to S_2$ by the equivalent rule: $\frac{X \to S_1}{X \to S_2}$. Then we iterate the following steps for each premise and the conclusion, now all of the form $N \to S$, where N does not contain $\boxminus$:

- if S contains no $\boxminus$, we are done;
- if S has an outermost $\boxminus$, we use the Ramsey rule to replace $N \to \boxminus_M T$ by the equivalent $\Diamond_M N \to T$;
- if S has an outermost $\wedge$, we distribute $N \to S_1 \wedge S_2$ into $N \to S_1$, $N \to S_2$

We eventually obtain an equivalent $\Diamond$-simple rule, which can sometimes be further simplified.

Note that the conditions of Theorem 14 are quite liberal, since:

- an axiom is a rule with no premises
- any formula can be regarded as an implication with antecedent $\top$
- it allows nesting of modalities, though not in subscripts. This freedom is more than is needed, since classical rules contain no modal nesting.

Example 15 All of the counterfactual rules of Table 6 are $\boxminus$-simple, except CV, CEM, D$'$, D$''$, C3, C7.

Example 16 Let us see how this works in the case of counterfactual axiom Tr from Table 8.

$$\boxminus_A B \wedge \boxminus_B C \to \boxminus_A C \iff \frac{X \to \boxminus_A B \qquad X \to \boxminus_B C}{X \to \boxminus_A C}$$

$$\iff \frac{\Diamond_A X \to B \quad \Diamond_B X \to C}{\Diamond_A X \to C}$$

$$\iff \frac{\Diamond_A X \to B}{\Diamond_A X \to \Diamond_B X}$$

The rules C3 and C7 are simple if we ignore the side conditions; and this is safe if we factor them out of the proof. This is done for C3 in the proof of theorem 10 in the appendix.

For the remaining rules that are not simple, namely CV, CEM, D$'$, D$''$ one possibility would be to follow Katsuno/Mendelzon by allowing premises about whether formulas are complete or satisfiable (cf. U3, U7). Then the counterparts can be written

$$\frac{B \text{ compl.} \quad \Diamond_A B \wedge C \text{ sat.}}{\Diamond_{A \wedge C} B \to \Diamond_A B}\,\text{UV} \qquad \frac{B \text{ compl.}}{\Diamond_A B \text{ compl.}}\,\text{UEM}$$

$$\frac{D \text{ compl.} \quad \Diamond_{A \vee B} D \to \neg A \quad \Diamond_{A \vee C} D \wedge A \text{ sat.}}{\Diamond_{B \vee C} D \to \neg C}\,\text{UD}'$$

UD$''$ is simply UEM+U1.

5. Conclusions

The link between counterfactual and updates, often considered as esoteric, is only the usual link between a relation and its inverse; counterfactuals can be considered as a universal modality, and update as its inverse existential modality.

Update rules are thereby translated into counterfactual rules, and vice versa. We found that some, but not all, of the counterfactual counterparts to U1-U8 are known in the counterfactual literature; and, symmetrically, the translation from counterfactuals to updates gives us some known and some new rules in that field.

Theorem 14 has proved very powerful, but probably does not cover all the cases when translation is possible. We would like to have an 'iff' characterisation. It is somewhat fortuitous that it was possible to do all the translations we performed, since it is known in temporal logic that there are easy examples of axioms using the 'future' modalities for which there is no equivalent using 'past' modalities. One such is $\Box(A \wedge \Box A \rightarrow B) \vee \Box(B \wedge \Box B \rightarrow A)$, whose correspondence condition is forward-linearity: $R(x,y) \wedge R(x,z) \Rightarrow R(y,z) \vee (y = z) \vee R(z,y)$. This condition cannot be expressed using $\boxminus, \diamondminus$; to prove this, we show it is not preserved under p-morphisms of frames by exhibiting two appropriate frames.

5.1. Acknowledgements

We are grateful to Odinaldo Rodrigues, who co-authored an early draft of this paper [21]. We are also grateful to Johan van Benthem for his continuing interest and his invitation to contribute to the Special Issue.

Thanks also to the referees for ModelAge'96 Workshop, TARK'96, AAAI'96, ECAI'96, Natasha Alechina, Sofia Guerra, Andreas Herzig for their comments on various drafts of this paper. We also acknowledge Esprit for partial funding under the WG ModelAge (8319). The first author also acknowledges the Nuffield Foundation for partial support.

Appendix

A. Proofs of theorems

Theorem 1 *Assume $\Box, \Diamond$ are interpreted by the accessibility relation R and $\boxminus, \diamondminus$ are interpreted by S in a frame F.*

The following are equivalent:

1. *For each $T \subseteq W$, $R_T = S_T^{-1}$.*

2. *F satisfies the following axiom schemes (each of which is also given in its dual form):*

$$\begin{array}{lll} (1) & B \to \Box_A \diamondminus_A B & \Diamond_A \boxminus_A B \to B \\ (2) & B \to \boxminus_A \Diamond_A B & \diamondminus_A \Box B_A \to B. \end{array}$$

3. *F satisfies the following rule (which is also given in its dual form):*

$$\begin{array}{c} B \to \boxminus_A C \\ \hline\hline \Diamond_A B \to C \end{array}\ \text{RR} \qquad\qquad \begin{array}{c} \diamondminus_A C \to B \\ \hline\hline C \to \Box_A B \end{array}\ \text{RRd}.$$

Proof. Let $\langle W, (S, R)\rangle$ be a frame, where R is the accessibility relation for $\Box, \Diamond$, and S is the relation for $\boxminus, \diamondminus$.

($1 \Rightarrow 2$) is straightforward.

For ($2 \Rightarrow 1$), suppose $S_T(x, y)$; choose the valuation V s.t. $V(p) = \{x\}$ and $V(q) = T$; then $x \Vdash p$, so by axiom (2) $x \Vdash \boxminus_q \Diamond_q p$ so $y \Vdash \Diamond_q p$ so $\exists z, R_T(y, z) \wedge z \Vdash p$. But by V, $z = x$, so $R_T(y, x)$. So $S_T \subseteq R_T^{-1}$. The converse inclusion is similar, but uses axiom (1).

($2 \Rightarrow 3$.) The axiom scheme (1) implies downward direction of the rule:

$B \to \boxminus_A C$	by hyp
$\Diamond_A B \to \Diamond_A \boxminus_A C$	by K, MP
$\Diamond_A B \to C$	by (1) dual form

The proof that the axiom scheme (2) implies the upward direction of the rule is similar.

($3 \Rightarrow 2$.) The upward direction of the rule implies the axiom (1):

$\diamondminus_A B \to \diamondminus_A B$	
$B \to \Box_A \diamondminus_A B$	RRd

and the downward direction similarly implies (2).

Theorem 4 *Necessitation follows from the axioms and rules in Table 3.*

Proof.

$$\dfrac{\dfrac{\dfrac{B}{\neg B \leftrightarrow -}}{\Diamond_A \neg B \leftrightarrow \Diamond_A -}\ \text{U4.1} \qquad \dfrac{\dfrac{- \to -}{\Diamond_A - \to -}\ \text{U2.2}}{\Diamond_A - \leftrightarrow -}}{\dfrac{\Diamond_A \neg B \leftrightarrow -}{\neg \Diamond_A \neg B}}$$

Theorem 5 *The rules U4.1, U4.2 and U8 hold in any frame.*

Proof. Take any frame F, and any valuation V. Let $M = \langle F, V\rangle$.

U4.1. Suppose $M \Vdash B \leftrightarrow C$ and $x \Vdash \Diamond_A B$. Then there exists y with $R_{|A|}(x,y)$ and $y \Vdash B$. But also, $y \Vdash C$, so $x \Vdash \Diamond_A C$. The other half is similar.

U4.2. Suppose $M \Vdash B \leftrightarrow C$ and $x \Vdash \Diamond_B A$. Then there exists y with $R_{|B|}(x,y)$ and $y \Vdash A$. But also, $R_{|C|}(x,y)$, so $x \Vdash \Diamond_C A$. The other half is similar.

U8. $x \Vdash \Diamond_A(B \vee C)$ iff $\exists y\, R_{|A|}(x,y), y \Vdash B \vee C$ iff $\exists y_1\, R_{|A|}(x,y_1), y_1 \Vdash B$ or $\exists y_2\, R_{|A|}(x,y_2), y_2 \Vdash C$ iff $x \Vdash \Diamond_A B \vee \Diamond_A C$.

Theorem 6 *A rule in Table 3 holds in a frame $F = \langle W, R\rangle$ iff R has the corresponding property stated in Table 4.*

Proof.

U1 $\Leftrightarrow$ R1. '$\Leftarrow$'. Let V be any valuation, and let $M = \langle W, R, V\rangle$. Suppose $x \Vdash_M \Diamond_A B$. Let $S = |A|$, and take a y such that $R_S(x,y)$ and $y \Vdash B$. By R1, $x \in S$, so $x \Vdash_M A$.

'$\Rightarrow$'. Suppose $R_S(x,y)$ in the frame $\langle W, R\rangle$; pick V such that $V(p) = S$ and $V(q) = \{y\}$. Since $y \Vdash q$, we have $x \Vdash \Diamond_p q$; so $x \Vdash p$ by U1, and $x \in S$ by def. of V.

U2.1 $\Leftrightarrow$ R2.1. '$\Leftarrow$'. Suppose $M \Vdash B \to A$ and $x \Vdash B$. Then $x \Vdash A$. We have $x \Vdash \Diamond_A B$ iff $\exists y$ s.t. $y \Vdash B$, and $R_{|A|}(x,y)$, which is true if we set $y = x$.

'$\Rightarrow$'. Suppose $y \in S$; we prove that $R_S(y,y)$. Let V be such that $V(p) = S$ and $V(q) = \{y\}$. It follows that $M \Vdash q \to p$, and therefore, by U2.1, $M \Vdash \Diamond_p q$. Since $y \Vdash q$, we get $y \Vdash \Diamond_p q$. Therefore, $\exists z \Vdash q$ s.t. $R_S(y,z)$. But $V(q) = \{y\}$. Thus, z must be y, and therefore $R_S(y,y)$.

Another style of proof is also possible, and is more immediate once you see how it works. The equivalence between U1 and R1 may be seen as follows. Let $F = \langle W, R\rangle$ be a frame.

$$
\begin{aligned}
& F \Vdash \Diamond_A B \to A, \text{ all } A, B && \text{U1} \quad (1)\\
\Leftrightarrow\ & \textstyle\bigcup_{b \Vdash B} R^{-1}_{|A|}(b) \subseteq |A|, \text{ all } A, B, V && (2)\\
\Leftrightarrow\ & \textstyle\bigcup_{y \in T} R^{-1}_{S}(y) \subseteq S, \text{ all } S, T && (3)\\
\Leftrightarrow\ & R^{-1}_{S}(y) \subseteq S, \text{ all } S, y && (4)\\
\Leftrightarrow\ & R_S(x,y) \text{ implies } x \in S, \text{ all } S, x, y && (5)
\end{aligned}
$$

Notice the quantification is carefully stated: A and B are 'local' to each line, and line 2 works for all valuations V. This quantification on V allows us to pass to line 3: in the $\Rightarrow$ direction, whatever S, T are we can pick V such that $V(p) = S$ and $V(q) = T$, for example; while in the $\Leftarrow$ direction, it is immediate by setting $S = |A|$ and $T = |B|$. The passage from line 3 to 4 again works because S, T are locally quantified.

U2.2 $\Leftrightarrow$ R2.2. '$\Leftarrow$'. Suppose $M \Vdash B \to A$ and $x \Vdash \Diamond_A B$. Take y such that $R_{|A|}(x,y)$ and $y \Vdash B$. Since $M \Vdash B \to A$, $y \in |A|$; so by R2.2, $x = y$; so $x \Vdash B$.

'$\Rightarrow$'. Suppose $y \in S$ and $R_S(x,y)$. Let V be such that $V(p) = S$ and $V(q) = \{y\}$. It follows that $M \Vdash q \to p$, and therefore, by U2.2, $M \Vdash \Diamond_p q \to q$. Since $y \Vdash q$, $x \Vdash \Diamond_p q$, so $x \Vdash q$, so $x = y$ (by def. of V).

U3 $\Leftrightarrow$ R3. '$\Leftarrow$'. Take any valuation such that A is satisfiable in M. Assume $M \Vdash \neg\Diamond_A B$. Take any y; we will show $y \Vdash \neg B$. By R3, exists x such that $R_{|A|}(x,y)$. Since $M \Vdash \neg\Diamond_A B$, then $x \Vdash \neg\Diamond_A B$, so we have $y \Vdash \neg B$.

'$\Rightarrow$'. Suppose $S \neq \emptyset$ and y is given. Take V such that $V(p) = S$ and $V(q) = \{y\}$. So p is satisfiable, and $M \not\Vdash \neg q$, since $y \Vdash q$. Therefore, by U3, $M \not\Vdash \neg\Diamond_p q$, so exists x such that $x \Vdash \Diamond_p q$. Then there is a z with $R_S(x,z)$ and $z \Vdash q$. But by choice of V, $z = y$; so $R_S(x,y)$.

U5 $\Leftrightarrow$ R5. '$\Leftarrow$'. Suppose $x \Vdash \Diamond_A B \wedge C$. Since $x \Vdash \Diamond_A B$, $\exists y \Vdash B$, $R_{|A|}(x,y)$. By R5, $R_{|A|\cap|C|}(x,y)$ and thus $x \Vdash \Diamond_{A\wedge C} B$.

'$\Rightarrow$'. Suppose $x \in S$ and $R_T(x,y)$. Pick V such that $V(p) = S$, $V(q) = \{y\}$, and $V(r) = T$. Then $y \Vdash q$ and $R_T(x,y)$ imply that $x \Vdash \Diamond_r q$. Since $x \Vdash p$, by U5 $x \Vdash \Diamond_{r\wedge p} q$. Therefore, $\exists y'$ s.t $y' \Vdash q$ and $R_{S\cap T}(x,y')$. But y' must be y, since $V(q) = \{y\}$, so $R_{S\cap T}(x,y)$.

U6 $\Leftrightarrow$ R6. See example 9 in the text.

U7 $\Leftrightarrow$ R7

$$B \text{ complete implies } \Diamond_A B \wedge \Diamond_C B \to \Diamond_{A\vee C} B \qquad \text{U7} \quad (1)$$

$$\Leftrightarrow \; B \text{ complete implies } \bigcup_{b\Vdash B} R^{-1}_{|A|}(b) \cap \bigcup_{b\Vdash B} R^{-1}_{|C|}(b) \subseteq \bigcup_{b\Vdash B} R^{-1}_{|A|\cup|C|}(b) \quad (2)$$

$$\Leftrightarrow \; R^{-1}_{|A|}(b) \cap R^{-1}_{|C|}(b) \subseteq R^{-1}_{|A|\cup|C|}(b) \quad (3)$$

$$\Leftrightarrow \quad R_S^{-1}(y) \cap R_T^{-1}(y) \subseteq R_{S\cup T}^{-1}(y) \tag{4}$$

$$\Leftrightarrow \quad R_S \cap R_T \subseteq R_{S\cup T} \tag{5}$$

$2 \Rightarrow 3$ by setting $B = \{b\}$. $3 \Rightarrow 2$ by completeness of B! $3 \Rightarrow 4$ because we can choose arbitrary valuations; $4 \Rightarrow 3$ by setting S, T to be $|A|, |C|$.

Example 8 (Last part.) The counterfactual rule

$$\frac{B \to A}{B \to \diamondminus_A B}$$

may be more simply stated as $\boxminus_A B \to (A \to B)$ (known as MP in the literature).

Proof. From left to right:

$$\cfrac{\cfrac{\cfrac{A \wedge \neg B \to A}{A \wedge \neg B \to \diamondminus_A (A \wedge \neg B)}\text{ hyp.}}{A \wedge \neg B \to \diamondminus_A \neg B}\text{ prop. of diamond}}{\boxminus_A B \to (A \to B)}\text{ equiv.}$$

and right to left:

$$\cfrac{\cfrac{B \to A}{B \to A \wedge B} \qquad \cfrac{\boxminus_A \neg B \to (A \to \neg B)\ [\text{hyp.}]}{A \wedge B \to \diamondminus_A B}\text{ equiv.}}{B \to \diamondminus_A B}$$

Theorem 10 *The update rules U1-U8 translate to the counterfactual rules given in table 7.*

Proof. The proofs for U1, U2.1 and U6 are given in examples 7, 8, 9 respectively. The proof for U5 is given using the method of section 4.2 in example 12. The other proofs may be constructed similarly. For C3 and C7, we factor out the side condition. For example, for C3, assume that A is satisfiable; we show the equivalence between $\dfrac{\neg \diamond_A B}{\neg B}$ (U3) and

$\neg\boxminus_A -$ (C3).

$$\dfrac{\dfrac{\dfrac{\diamondminus_A\top \to \diamondminus_A\top}{\top \to \Box_A\diamondminus_A\top}\,\text{RRd}}{\neg\Diamond_A\boxminus_A -}\,\text{eq.}}{\neg\boxminus_A -}\,\text{U3} \qquad \dfrac{\dfrac{\Diamond_A B \to - \text{ hyp.}}{B \to \boxminus_A -}\,\text{RR} \qquad \boxminus_A - \to - \text{ C3}}{B \to -}\,.$$

References

1. J. Bell. The logic of nonmonotonicity. *Artificial Intelligence*, 41:365–374, 1989.
2. John P. Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic*, 22:76–84, 1981.
3. B. F. Chellas. Basic conditional logic. *Journal of Philosophical Logic*, 4:133–153, 1975.
4. M. de Rijke. Meeting some neighbors. In Jan van Eijck and Albert Visser, editors, *Logic and Information Flow*. MIT Press, 1994.
5. P. Gärdenfors. Conditionals and changes of belief. *Acta Philos. Fennica*, 30:381–404, 1978.
6. P. Gärdenfors. Belief revision and the Ramsey test for conditionals. *Philosophical Review*, 91:81–93, 1986.
7. Peter Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, Bradford Books, Cambridge, MA, 1988.
8. R. Goldblatt. *Logics of Time and Computation*. CSLI Lecture Notes, 1987.
9. G. Grahne. Updates and counterfactuals. In *Proc. Second International Conference on Principles of Knowledge Representation and Reasoning (KR '91)*, pages 269–276. Morgan Kaufmann, San Francisco, CA, 1991.
10. Hirofumi Katsuno and Alberto O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.
11. Hirofumi Katsuno and Alberto O. Mendelzon. On the difference between updating a knowledge base and revising it. In Peter Gärdenfors, editor, *Belief Revision*, number 29 in Cambridge Tracts in Theoretical Computer Science, pages 183–203. Cambridge University Press, 1992.
12. Arthur M. Keller and Marianne Winslett Wilkins. On the use of an extended relational model to handle changing incomplete information. *IEEE Transactions on Software Engineering*, 11(7):620–633, July 1985.
13. David K. Lewis. *Counterfactuals*. Harvard University Press, 1973.
14. D. Makinson. How to give it up: A survey of some formal aspects of the logic of theory change. *Synthèse*, 62:347–363, 1985.
15. Donald Nute. *Topics in Conditional Logic*. D. Reidel, Dordrecht, 1980.
16. Donald Nute. Conditional logic. In Dov M. Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic*, chapter Chapter II.8, pages 387–439. D. Reidel, Dordrecht, 1984.
17. John L. Pollock. A refined theory of counterfactuals. *Journal of Philosophical Logic*, 10:239–266, 1981.
18. S. Popkorn. *First Steps in Modal Logic*. Cambridge University Press, 1994.
19. F. P. Ramsey. Truth and probability. In R. B. Braithwaite, editor, *The Foundations of Probability and other Logical Essays*. Harcourt Brace, New York, 1931.
20. M. Ryan and P.-Y. Schobbens. Intertranslating counterfactuals and updates. In W. Wahlster, editor, *12th European Conference on Artificial Intelligence (ECAI)*, pages 100–104. John Wiley & Sons, Ltd., 1996.

21. M. Ryan, P.-Y. Schobbens, and O. Rodrigues. Counterfactuals and updates as inverse modalities. In Y. Shoham, editor, *TARK'96: Proc. Theoretical Aspects of Rationality and Knowledge*, pages 163–173. Morgan Kaufmann, 1996.
22. Robert C. Stalnaker. *A Theory of Conditionals*, volume 2 of *American Philosophical Quarterly Monograph Series (ed. Nicholas Rescher)*, pages 98–112. Blackwell, Oxford, 1968.
23. J. van Benthem. Correspondence theory. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 2. Dordrecht: D. Reidel, 1984.
24. J. van Benthem. Logic and the flow of information. In *Proc. 9th LMPS*. 1991.
25. Y. Venema. *Many-dimensional modal logic*. PhD thesis, Universiteit van Amsterdam, 1991.