A note on the rational closure of knowledge bases with both positive and negative knowledge

R.Booth and J.B.Paris
Department of Mathematics
University of Manchester
Manchester M13 9PL
UK

Abstract

The notion of the rational closure of a positive knowledge base K of conditional assertions $\theta_i \mid\!\sim \phi_i$ (standing for *if* θ_i *then normally* ϕ_i) was first introduced by Lehmann in [2] and developed by Lehmann and Magidor in [3]. Following those authors we would also argue that the rational closure is, in a strong sense, the minimal information, or simplest, rational consequence relation satisfying K. In practice however one might expect a knowledge base to consist not just of positive conditional assertions, $\theta_i \mid\!\sim \phi_i$, but also negative conditional assertions, $\theta_i \not\mid\!\sim \phi_i$ (standing for *not* $\{$*if* θ_i *then normally* $\phi_i\}$). Restricting ourselves to a finite language we show that the rational closure still exists for satisfiable knowledge bases containing both positive and negative conditional assertions and has similar properties to those exhibited in [3]. In particular an algorithm in [3] which constructs the rational closure can be adapted to this case and yields, in turn, completeness theorems for the conditional assertions entailed by such a mixed knowledge base.

Introduction, Notation and Preliminaries

The results presented in this paper can best be motivated, or justified, by first considering the following problem. Let us suppose that we have somehow learnt that an intelligent, rational, agent believes some finite set, K say, of so called conditional, or defeasible, assertions,

$$\textit{If } \theta_i \textit{ then normally (or usually) } \phi_i \qquad i = 1, \ldots, m$$

where θ_i, ϕ_i are sentences from some finite propositional language L. (To fix a little notation here let SL be the set of sentences of L, built up using the standard connectives $\wedge, \vee, \rightarrow, \neg$ from the propositional variables $p_1, p_2, \ldots, p_n$.) In this case what other assertions of this form can we, or should we, conclude that the agent believes? Much of the knowledge we possess and, more importantly communicate, is in the form of just such assertions and our ability to use them to correctly draw further conclusions is seen as a manifestation of our intelligence. In consequence this question has received considerable attention in the AI community with many different approaches and solutions being proposed (see, for example [4],[5],[6],[7],[8], [12]).

The key distinguishing feature of such assertions, as opposed to classical notions of logical consequence, is their possible *non-montonicity.* For it is now certainly possible for and agent to believe

$$\textit{If } \theta \textit{ then normally } \phi$$

whilst not believing

$$\textit{If } \theta \textit{ and } \psi \textit{ then } \phi.$$

[For example take θ to be 'the cake contains butter', ϕ to be 'the cake tastes good' and ψ to be 'the cake contains paraffin'.]

For our part however we find the approach of Lehmann, Magidor and Kraus (see [9],[3], and also the closely related work of Gabbay, Gärdenfors, Makinson, [1],[11],[10]) particularly attractive and the results we shall present here follow directly on from Lehmann et al's [3]. (Not unnaturaly therefore we shall be assuming that the reader has some familiarity with that paper.) The approach of Lehmann et al is to treat conditional assertions such as

$$\textit{If } \theta_i \textit{ then normally } \phi_i$$

as instances of a *consequence relation* on $SL \times SL$, usually denoted as $\mid\!\sim$, and to consider what closure properties $\mid\!\sim$ should have given that it corresponds to the beliefs (of this form) of an intelligent, rational agent (such as ourselves!). In this sense then we can think of the agent *as* such a consequence relation. As far as what properties $\mid\!\sim$ should satisfy one that we clearly do not want in its full generality is the rule of monotonicity, viz.,

$$\frac{\theta \mid\!\sim \phi}{\theta \wedge \psi \mid\!\sim \phi}$$

although by considering the intended interpretation of $\mid\!\sim$ it certainly does seem desirable that various weaker forms of this rule are satisfied. For example if in the monotonicity rule we had in addition that if θ then normally ψ, then $\theta \wedge \psi$ would, under normal circumstances, be synonomous with θ so, again under normal circumstances, ϕ should follow from $\theta \wedge \psi$ just if ϕ followed from θ alone. This formalises as the rule of *cautious monotonicity* (following [1]),

$$\frac{\theta \mid\!\sim \phi,\ \theta \mid\!\sim \psi}{\theta \wedge \psi \mid\!\sim \phi}.$$

In [9],[3] Kraus, Lehmann and Magidor discuss a number of possible rules and families of rules, based on considerations of the way the relation 'if ... then normally...' is used in natural language, to describe, via a purely logical analysis, a procession of ever more powerful consequence relations. At the same time they are able to use the justifications for these rules to criticise, through the failure of one or more of them, the main alternative approaches of circumscription [6], autoepistemic logics [7], [8] and default logic [5]. Of the various consequence relations which they name, we believe, for the reasons they give, that their notion of *rational consequence relation* best sums up the properties that $\mid\!\sim$ should possess given that it arises in this way. Precisely:

Definition 1 *A binary relation $\mid\!\sim$ on $SL \times SL$ is a rational consequence relation if it satisfies for all $\theta, \phi, \psi \in SL$*

- $\theta \mid\!\sim \theta$ *REF (Reflexivity)*
- $\dfrac{\theta \mid\!\sim \phi,\ \theta \equiv \psi}{\psi \mid\!\sim \phi}$ *LLE (Left Logical Equivalence)*
- $\dfrac{\theta \mid\!\sim \phi,\ \phi \models \psi}{\theta \mid\!\sim \psi}$ *RWE (Right Weakening)*

- $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \theta \mathrel{|\!\sim} \psi}{\theta \mathrel{|\!\sim} \phi \wedge \psi}$ *AND*
- $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \psi \mathrel{|\!\sim} \phi}{\theta \vee \psi \mathrel{|\!\sim} \phi}$ *OR*
- $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \theta \mathrel{|\!\sim} \psi}{\theta \wedge \phi \mathrel{|\!\sim} \psi}$ *CMO (Cautious Monotonicity)*
- $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \theta \not\mathrel{|\!\sim} \neg\psi}{\theta \wedge \psi \mathrel{|\!\sim} \phi}$ *RMO (Rational Monotonicity)*

A binary relation $\mathrel{|\!\sim}$ on $SL \times SL$ is called a preferential consequence relation if it satisfies all the above except, possibly, RMO.

Whilst the above conditions are intended to be thought of as closure conditions on a binary relation $\mathrel{|\!\sim}$ they might equally be thought of (with a possible question mark against RMO) as *axioms and rules of proof* for deriving *sequents* $\theta \mathrel{|\!\sim} \phi$. This viewpoint is closer to that taken by some other workers in the field (see for example [1],[10]). In particular, viewed in this way these axioms and rules of proof are often refered to as the *GM axioms and rules*, *GM* here standing for Gabbay-Makinson. For this paper it will be very convenient for us to jump back and forth between the two interpretations at will. We hope that it will always be clear as to whether $\mathrel{|\!\sim}$ is standing for a rational consequence relation or (essentially) a marker dividing the two sides of a sequent! For future reference let P stand for the GM axioms and rules *less* the RMO rule. Notice that in the case of P all the rules are limited to just 'positive' sequents so that the idea of proof and derivability are unambiguous.

Returning now to our initial problem, if we assume, as we henceforth do throughout this paper, that our agent is, at least as regards his beliefs in such conditional assertions, a rational consequence relation then on the basis of K alone we can conclude that our agent *must* believe *if θ then normally ϕ* just if $\theta \mathrel{|\!\sim} \phi$ holds of *all* rational consequence relations satisfying K. (We hope the reader will understand, and forgive, this third use of $\mathrel{|\!\sim}$, here as a variable standing for a binary, in this case, rational consequence, relation.) By lemma 2.25 amd theorem 3.12 of [3] Lehmann and Magidor give an elegant, and perhaps slightly surprising, characterisation of this set of conditional assertions:

Theorem 1 *[Lehmann and Magidor]*
The following are equivalent:
(i) $\theta \mathrel{|\!\sim} \phi$ holds in all rational consequence relations satisfying K.
(ii) $\theta \mathrel{|\!\sim} \phi$ holds in all preferential consequence relations satisfying K.
(iii) $\theta \mathrel{|\!\sim} \phi$ is derivable from K in P.

That (ii) and (iii) are equivalent is obvious, of course, but the surprise here is that (i) implies (ii). It says that in this context the condition RMO does not give us anything new in the way of consequences. Because of its relevance to some of the results which follow we shall say a few words about how this result is proved, but first we need a little more notation. We call sentences of L of the form $\pm p_1 \wedge \pm p_2 \wedge \ldots \wedge \pm p_n$ *atoms* (since they are the atoms from the corresponding Lindenbaum algebra). Since atoms determine the truth or falsity of all the propositional variables, and hence all sentences, they may equally be thought of as corresponding to worlds (although we will not assume that different worlds necessarily correspond to different atoms). Let At be the set of the 2^n atoms of L. By the disjunctive normal form theorem, for every $\theta \in SL$ there is a set $S_\theta \subseteq At$ such that $\bigvee S_\theta$ is logically equivalent to θ, $S_\theta \cap S_\phi = S_{\theta\wedge\phi}$, $S_\theta \cup S_\phi = S_{\theta\vee\phi}$, $At - S_\theta = S_{\neg\theta}$, and $S_\theta = At$ for θ a tautology, $S_\theta = \emptyset$ for θ a contradiction.

Now let $t_1, t_2, \ldots, t_k$ be subsets of At and define a relation $\mid\!\sim$ on $SL \times SL$ by

$$\theta \mid\!\sim \phi \textit{ iff either there is no } i \leq k \textit{ such that } S_\theta \cap t_i \neq \emptyset,$$
$$\textit{or there is such an } i \textit{ and for the least such, } S_\theta \cap t_i \subseteq S_\phi$$

The idea behind this definition is that an agent holds the 'worlds' in t_i to be more natural than, or preferable to, those in t_j just if $i < j \leq k$, whilst worlds not represented in any of the t_i the agent considers 'impossible'. The agent believes that *if* θ *then normally* ϕ if in all the most natural, or preferred, worlds in which θ is true ϕ is also true, i.e. if $\theta \mid\!\sim \phi$ holds. From this definition it is easy to check that $\mid\!\sim$ is a rational consequence relation on SL. What is rather less obvious is that the converse also holds, every rational consequence relation $\mid\!\sim$ on SL is of this form for some finite sequence $t_1, t_2, \ldots, t_k$ of subsets of At. (A proof of this result is given in [3], theorem 3.12, with, in their notation, the images under l of the sets of states of equal rank corresponding to our $t's$.) In such a case we shall say that $t_1, t_2, \ldots, t_k$ is a *model* of $\mid\!\sim$. It should be clear that there are many different models giving the same rational consequence relation. For example if we insert/remove copies of the empty set into/from the sequence $t_1, t_2, \ldots, t_k$, or add/subtract an atom to/from a particular t_i when it has already appeared in an earlier t_j, then none of this will affect the corresponding rational consequence relation. Indeed we could if we wanted have allowed our sequence of subsets t_i to be infinite without affecting the representation theorem. What we can easily show however is that every rational consequence relation has a *unique* model in which the sets t_i are disjoint and non-empty. We call this model its *normal model*. Clearly there are only finitely many such normal models (for L) and hence only finitely many rational consequence relations on L.

A great value of this representation theorem, from our point of view in this paper, is that it frequently allows us a simple route to showing that some $\theta \mid\!\sim \phi$ holds for all rational consequence relations satisfying K by arguing about a general model $t_1, t_2, \ldots, t_k$ satisfying K. Furthermore we can then in turn apply theorem 1 to show that there must be a *proof* of $\theta \mid\!\sim \phi$ from K in P. For example it is immediately clear from this representation that *Supra-Classicality (SC)*, i.e. *if* $\theta \models \phi$ *then* $\theta \mid\!\sim \phi$, is a derived rule of proof of P. We will, in fact, be repeatedly leaving it up to the reader to use this trick in the later sections of this paper to confirm that a derived rule is sound for rational consequence relations or that a certain formal proof exists! In such cases we will simply say that the result follows *by a semantic argument*, or words to that effect.

Returning now to theorem 1, it is through an analogous representation result for preferential consequence relations given in [9] that Lehmann and Magidor derive the key implication from (i) to (ii). Namely they show that a similar representation theorem to that for rational consequence relations holds for preferential consequence relations, it is just that instead of the (finitely many) t_i being linearly ordered they are now only partially ordered and instead of us talking about *the* least t_i we have to talk about *all minimal* t_i. The delightful observation is now that if we have such a model for a preferential consequence relation satisfying K but not $\theta \mid\!\sim \phi$, then we can simply complete this partial ordering to a *linear* ordering which still 'satisfies' K but not $\theta \mid\!\sim \phi$.

Theorem 1 tells us then what other beliefs we can infer our agent must hold given that we know only the agent believes K (and, of course, assuming as we do throughout, that our agent's beliefs in statements of this form correspond to a rational consequence relation). Indeed the set of such beliefs is actually the smallest *preferential consequence relation* satisfying K. Unfortunately this smallest preferential consequence relation will not in general satisfy RMO and so will not be a rational consequence relation. Whilst unavoidable, this is rather frustrating because it means that we know there are beliefs the agent holds which we cannot access through K alone.

Suppose however we had reason to believe that in imparting K to us the agent was, essentially, 'telling us all his, or her, knowledge on the subject in question', i.e. that there was nothing more that the expert knew which was not somehow grounded, or foretold, in K. For example, it is often assumed by the builders of expert systems that just such a set K can be acquired by observing and questioning the expert in the workplace over an extended period of time. In such a situation would we not be justified in assuming that whilst, in reality, K could not be the *totality* of the expert's relevant knowledge (since that surely is potentially infinite) it was, nevertheless, the case that the sum total of the expert's knowledge was in some sense the simplest, or minimal, rational consequence relation satisfying K?

The problem then is, what do we mean by 'simplest' or 'minimal' here? If we had been modelling the expert's knowledge as a preferential consequence relation here the answer would surely be the set theoretically smallest such preferential consequence relation, i.e. the intersection of all preferential consequence relations satisfying K, equivalently the set of consequences of K derivable in P – which in this case is itself a preferential consequence relation. However as we have already remarked there need be no set theoretically smallest *rational* consequence relation satisfying K, equivalently, the intersection of all rational consequence relations satisfying K need not itself be a rational consequence relation. So, what might we mean by 'simplest' here? In [3] Lehmann and Magidor suggest that their 'rational closure of K' is a contender, a view which we strongly support (perhaps even more than them!). For the purpose of this paper we shall take the definition of the rational closure of K to be as follows:

Definition 2 *Let K be a knowledge base. $\mid\!\sim_*$ is the rational closure of K if $\mid\!\sim_*$ is a rational consequence relation satisfying K and, for any other rational consequence relation $\mid\!\sim'$ satisfying K, the following two conditions hold:*

(RC1) *There exist $\theta, \phi \in SL$ such that $\theta \mid\!\sim' \phi$, $\theta \not\mid\!\sim_* \phi$ and, for any $\delta, \gamma \in SL$, if $\delta \vee \theta \mid\!\sim_* \neg\theta$ and $\delta \mid\!\sim_* \gamma$ then $\delta \mid\!\sim' \gamma$.*

(RC2) *For all $\lambda, \chi \in SL$, if $\lambda \mid\!\sim_* \chi$ and $\lambda \not\mid\!\sim' \chi$ then there exist $\eta, \psi \in SL$ such that $\eta \vee \lambda \mid\!\sim' \neg\lambda$, $\eta \mid\!\sim' \psi$ and $\eta \not\mid\!\sim_* \psi$.*

We should remark that this is not *exactly* the definition of rational closure given by Lehmann and Magidor in [3], although is easy to see that they are equivalent. Precisely, Lehmann and Magidor define a rational consequence relation $\mid\!\sim_*$ to be *preferable* to a rational consequence relation $\mid\!\sim'$ if **(RC1),(RC2)** hold. They then show that this notion of preference between rational consequence relations is transitive and irreflexive and go on to define a rational consequence relation $\mid\!\sim_*$ satisfying K to be the rational closure of K if it is preferable to all other rational consequence relations satisfying K.

In order to explain how it is that the rational closure might justifiably claim to be the 'simplest' or 'most reasonable' rational consequence relation satisfying K (and why the above relation between rational consequence relations might warrant the name 'preference') we need to examine the meanings of the above conditions **(RC1)** and **(RC2)**, .

First of all notice that for any rational consequence relation $\mid\!\sim$ modelled, say, by $t_1, \ldots, t_k \subseteq At$, and for any sentences $\delta, \theta \in SL$, $\delta \vee \theta \mid\!\sim \neg\theta$ means that in all the most natural worlds in which $\delta \vee \theta$ is true, θ is false. In other words, for any world in which θ is true there is a strictly more natural world in which δ is true (and θ false). Hence $\delta \vee \theta \mid\!\sim \neg\theta$ is saying that according to $\mid\!\sim$, provided $\delta \vee \theta$ holds in some world (i.e. $S_{\delta\vee\theta} \cap t_i \neq \emptyset$ for some i) δ is strictly less exceptional or more acceptable than θ. Now suppose $\mid\!\sim_*$ satisfied **(RC1)** and **(RC2)** and that $\mid\!\sim'$ was a rational consequence relation satisfying K which was distinct from $\mid\!\sim_*$. Suppose there existed sentences $\lambda, \chi \in SL$ such that $\lambda \mid\!\sim_* \chi$ and $\lambda \not\mid\!\sim' \chi$. Then an advocate for $\mid\!\sim'$ to be the most reasonable rational consequence relation satisfying K might attempt to denounce $\mid\!\sim_*$ on the grounds that $\mid\!\sim_*$ satisfied a conditional assertion which $\mid\!\sim'$ did not and hence, since $\mid\!\sim'$ satisfied K, satisfied a conditional which must be unsupported by K. However, by **(RC2)**, an advocate for $\mid\!\sim_*$ is able to defend an attack of this kind by pointing out the existence of sentences $\eta, \psi \in SL$ such that $\eta \vee \lambda \mid\!\sim' \neg\lambda$, $\eta \mid\!\sim' \psi$ and $\eta \not\mid\!\sim_* \psi$. In other words $\mid\!\sim'$ similarly satisfies a conditional assertion which is unsupported by K but also asserts that η is less exceptional than λ and so in fact draws an unsupported conclusion from an assumption η which it itself holds to be more reasonable than λ. By **(RC1)**, though, an advocate of $\mid\!\sim_*$ can also point out the existence of an unsupported conditional which $\mid\!\sim'$ satisfies and $\mid\!\sim_*$ does not and for which $\mid\!\sim'$ cannot make the defence available to $\mid\!\sim_*$ above.

Note that, for a given K, if the rational closure of K exists then it must be unique. For suppose $\mid\!\sim_*$ and $\mid\!\sim'_*$ were two distinct rational closures. Then by **(RC1)** for $\mid\!\sim_*$ there would be θ and ϕ such that $\theta \mid\!\sim'_* \phi$, $\theta \not\mid\!\sim_* \phi$ and whenever $\delta \vee \theta \mid\!\sim_* \neg\theta$ and $\delta \mid\!\sim_* \gamma$ for some $\delta, \gamma \in SL$ then $\delta \mid\!\sim'_* \gamma$. However by **(RC2)** for $\mid\!\sim'_*$ there would be $\eta, \psi \in SL$ such that $\eta \vee \lambda \mid\!\sim_* \neg\lambda$, $\eta \mid\!\sim_* \psi$ and $\eta \not\mid\!\sim'_* \psi$ — a contradiction.

In [3] Lehmann and Magidor show that for finite L the the rational closure of K always exists (their approach is rather more general than ours here because they also consider infinite languages). Indeed in this case they give two explicit constructions, one in terms of ranks of sentences (which we shall refer to again briefly later) and a second using an explicit model-theoretic construction. In the next section we will present an alternative, and rather simple, model-theoretic construction of the rational closure which will be easily generalisable to the case of negative knowledge to which we now turn.

The developments described in the preceding discussion provide, to our mind, a very elegant and satisfactory theory in the case where K consists of conditional assertions of the form

If θ *then normally* ϕ

Furthermore it is certainly true that much of our knowledge is of precisely this form. However the fact that an agent can know, or believe, that *if* θ *then normally* ϕ surely entails that he, or she, can also know, or believe, that such an assertion is *not* true, i.e. *not*$(\theta \mid\!\sim \phi)$, or as we shall write it $\theta \not\mid\!\sim \phi$, and our analysis so far has taken no account of the fact that in practice K might just as well contain negative knowledge $\psi \not\mid\!\sim \lambda$ in addition to the usual positive knowledge $\theta \mid\!\sim \phi$. Indeed, of course, such negative conditional assertions have already appeared in the RMO condition. [Notice that $\psi \not\mid\!\sim \lambda$ is certainly not the same as $\psi \mid\!\sim \neg\lambda$, indeed we can have both $\psi \mid\!\sim \lambda$ and $\psi \mid\!\sim \neg\lambda$ holding for a rational consequence relation.]

The main purpose of this paper will be to investigate the problem stated at the start of this section in the case where the knowledge base K may additionally contain negative as well as positive knowledge.

For the remainder of this paper let us suppose that K may contain negative as well as positive knowledge, say, $K = K_P \cup K_N$ (which we may sometimes alternatively write as $K_P + K_N$) where

$$K_P = \{\ \theta \mid\!\sim \phi \mid (\theta \mid\!\sim \phi) \in K\ \}$$

$$K_N = \{\ \psi \not\mid\!\sim \lambda \mid (\psi \not\mid\!\sim \lambda) \in K\ \}$$

Unlike the case earlier when $K_N = \emptyset$ and K was automatically consistent (because it was satisfied by the trivial rational consequence relation which holds between all pairs θ, ϕ) we can now no longer guarantee that K is consistent (with $\mid\!\sim$ being a rational consequence relation). For example K could be

$$\{\ \theta \mid\!\sim \psi,\ \ \theta \mid\!\not\sim \neg\phi,\ \ \theta \wedge \phi \mid\!\not\sim \psi\ \}$$

which contradicts RMO. However, we are assuming that K has been given by an agent corresponding to a rational consequence relation so unless otherwise stated we shall assume throughout that K is consistent.

Proceeding now as in our earlier discussion we are interested in answers to the following problems given that the agent has given us K:

(A) For what θ, ϕ can we be sure that the agent believes that *if* θ *then normally* ϕ ? Equivalently what is the relation $\mid\!\sim_P^K$ given by

$$\theta \mid\!\sim_P^K \phi \quad iff \quad \theta \mid\!\sim \phi \text{ holds for all rational consequence relations } \mid\!\sim \text{ satisfying } K$$

(B) For what ψ, λ can we be sure that the agent believes that *it is not the case that if* ψ *then normally* λ ? Equivalently what is the relation $\mid\!\not\sim_N^K$ given by

$$\psi \mid\!\not\sim_N^K \lambda \quad iff \quad \psi \mid\!\not\sim \lambda \text{ holds for all rational consequence relations } \mid\!\sim \text{ satisfying } K$$

(C) If the agent also asserts that K contains essentially all his/her knowledge (of this form) what should we take the agent's relation $\mid\!\sim$ to be?

The plan of the remainder of this paper is that in the next section we give an alternative model-theorectic construction of the rational closure of K in the general case that K is a mixed, positive and negative knowledge base (so giving an answer to problem C above). We next show that an algorithm presented in [3] can be adapted to yield the rational closure in this case. In the following section we show how a consideration of this algorithm enables us to provide completeness theorems for $\mid\!\sim_P^K$ and $\mid\!\not\sim_N^K$, thus giving answers to problems A and B. It is worth observing here that because of the special form of the negative conditional assertion in the premises of the RMO rule and the fact that no such negative conditional appears as a conclusion of a GM rule the GM axioms and rules alone certainly do not suffice for such completeness results. Indeed even changing the RMO rule to the more versatile

- $\dfrac{\theta \mid\!\sim \phi,\ \theta \mid\!\not\sim \psi}{\theta \wedge \neg\psi \mid\!\sim \phi}$

does not alter the position. To see this let $K = \{q \vee \neg q \mid\!\sim p, q \mid\!\not\sim p\}$. Then every rational consequence relation $\mid\!\sim$ which satisfies K also satisfies $\neg q \mid\!\sim p$ but this conclusion cannot be derived from K using the GM rules (with this new version of RMO). For suppose it could be. Then the only way $q \mid\!\not\sim p$ could be used in this derivation is via the (new) RMO rule and hence any use of this rule in the derivation could be replaced by an application of the 'rule'

- $\dfrac{q \mid\sim \theta}{q \wedge \neg p \mid\sim \theta}$.

But for the rational consequence relation $\mid\sim$ with model $t_1 = \{p \wedge q\}$, $t_2 = \{\neg p \wedge \neg q\}$ this rule (for all θ) and $q \vee \neg q \mid\sim p$ are both satisfied whilst $\neg q \mid\sim p$ fails, showing that no such derivation can exist.

At this point the reader may question why we stop at simply negations of conditional assertions, why not also consider conjunctions and disjunctions, or even more complicated combinations of conditional assertions, rather in the way Delgrande does in [13], [14] (albeit in a somewhat different system)? Our view on this is that whilst simple conjunctions of conditional assertions or their negations are obviously already handled by the present set-up, disjunctions of conditional assertions seem to us, in the context of an agent's knowledge, rather unnatural. For similar reasons we have limited our investigations in this paper to finite languages.

The Rational Closure of K

Given, as usual, K consistent let $t_{i1}, t_{i2}, \ldots, t_{ik_i} \subseteq At$ for $i = 1, 2, \ldots, r$ enumerate all normal models of rational consequence relations satisfying K. Let $k = \max\{\, k_i \mid i = 1, 2, \ldots, r\}$ (so $k \leq |At|$) and for $1 \leq j \leq k$ let $\mathcal{U}_j = \bigcup_{i=1}^{r} t_{ij}$ where we take $t_{ij} = \emptyset$ if $k_i < j \leq k$. Let $\mid\sim_*$ be the rational consequence relation modelled by $\mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_k$ or, equivalently, by $u_1, \ldots, u_k$ where we define $u_1 = \mathcal{U}_1$ and $u_j = \mathcal{U}_j - \bigcup_{i<j} \mathcal{U}_i$ for $2 \leq j \leq k$.

Theorem 2 *$\mid\sim_*$ is the rational closure of K.*

Proof. Firstly we must show that $\mid\sim_*$ satisfies K. This will follow once we have proved the following two claims for $\theta, \phi \in SL$:

(i) If $\theta \mid\sim \phi$ holds for all rational consequence relations $\mid\sim$ satisfying K then $\theta \mid\sim_* \phi$. (And hence $\mid\sim_*$ satisfies K_P.)

(ii) If $\theta \not\mid\sim_* \phi$ then there is a rational consequence relation $\mid\sim$ satisfying K such that $\theta \not\mid\sim \phi$. (And hence $\mid\sim_*$ satisfies K_N.)

To prove claim (i), suppose we have $\theta \mid\sim \phi$ for all rational consequence relations $\mid\sim$ satisfying K. If $S_\theta \cap \mathcal{U}_j = \emptyset$ for all j then $\theta \mid\sim_* \phi$ as required so let j be minimal such that $S_\theta \cap \mathcal{U}_j \neq \emptyset$. Let $\alpha \in At$ be such that $\alpha \in S_\theta \cap \mathcal{U}_j$. Then $\alpha \in S_\theta \cap t_{ij}$ for some $1 \leq i \leq r$. Hence $S_\theta \cap t_{ij} \neq \emptyset$ and furthermore j is minimal such for this i. This follows since if $S_\theta \cap t_{ij'} \neq \emptyset$ for some $j' < j$ then $S_\theta \cap \mathcal{U}_{j'} \neq \emptyset$ contradicting the minimality of j. Hence, since $t_{i1}, \ldots, t_{ik}$ models a rational consequence relation which satisfies K, we have $S_\theta \cap t_{ij} \subseteq S_\phi$ and so $\alpha \in S_\phi$. Thus $S_\theta \cap \mathcal{U}_j \subseteq S_\phi$ as required to show $\theta \mid\sim_* \phi$.

To prove claim (ii), suppose $\theta \not\mid\sim_* \phi$. First of all if $S_\theta \cap \mathcal{U}_j = \emptyset$ for all $1 \leq j \leq k$ then $S_\theta \cap t_{ij} = \emptyset$ for all $1 \leq i \leq r$ and $1 \leq j \leq k$, giving $\theta \mid\sim \phi$ for all rational consequence relations $\mid\sim$ satisfying K which suffices. So suppose $S_\theta \cap \mathcal{U}_j \neq \emptyset$ for some j and let j' be the minimal such j. As in the proof of claim (i) this gives us $S_\theta \cap t_{ij'} \neq \emptyset$ for some i and that j' is minimal such for this i. Then $S_\theta \cap t_{ij'} \subseteq S_\theta \cap \mathcal{U}_{j'} \not\subseteq S_\phi$ since $\theta \not\mid\sim_* \phi$. Hence $t_{i1}, \ldots, t_{ik}$ models a rational consequence relation which satisfies $K \cup \{\theta \not\mid\sim \phi\}$ as required.

All that remains is to prove that conditions **(RC1)** and **(RC2)** from definition 2 hold for any rational consequence relation $\mid\sim$ which satisfies K and is distinct from $\mid\sim_*$. To do this we will be making use of the model $u_1, \ldots, u_k$ for $\mid\sim_*$. Let $\mid\sim'$ be such a relation, so $\mid\sim'$ is modelled by $t_{i1}, \ldots, t_{ik}$ for some $1 \leq i \leq r$. Since $\mid\sim' \neq \mid\sim_*$, let p be minimal such that $t_{ip} \neq u_p$. We claim that $t_{ip} \subseteq u_p$ (so $t_{ip} \subset u_p$). To see this let $\alpha \in t_{ip}$. Then $\alpha \in \mathcal{U}_p$. Suppose $\alpha \in \bigcup_{j<p} \mathcal{U}_j$, then let $j < p$ be minimal such that $\alpha \in \mathcal{U}_j$. Then $\alpha \in u_j$ and so $\alpha \in t_{ij}$ since $u_j = t_{ij}$ for $j < p$. But this contradicts $t_{ip} \cap t_{ij} = \emptyset$. (Recall each $t_{i1}, \ldots, t_{ik}$ is a normal model.) Hence $\alpha \in \mathcal{U}_p - \bigcup_{j<p} \mathcal{U}_j = u_p$ as required.

Now put $\theta = \bigvee u_p$, $\phi = \bigvee t_{ip}$ (so $S_\theta = u_p$ and $S_\phi = t_{ip}$). Then, since $u_1, u_2, \ldots, u_k$ are disjoint, p is minimal such that $S_\theta \cap u_p \neq \emptyset$ and $S_\theta \cap u_p = u_p \not\subseteq t_{ip} = S_\phi$ so $\theta \not\mathrel{|\!\sim}_* \phi$. If $t_{ip} = \emptyset$ then $t_{il} = \emptyset$ for all $l \geq p$ (since we are in normal form) while for $l < p$ we have $u_p \cap t_{il} = u_p \cap u_l = \emptyset$. Hence $S_\theta \cap t_{il} = \emptyset$ for all l and so $\theta \mathrel{|\!\sim}' \phi$. If on the other hand $t_{ip} \neq \emptyset$ then p is minimal such that $S_\theta \cap t_{ip} \neq \emptyset$ and $S_\theta \cap t_{ip} = t_{ip} \subseteq t_{ip} = S_\phi$ and so again $\theta \mathrel{|\!\sim}' \phi$. Hence we have shown the first part of condition **(RC1)**.

For the second part of **(RC1)** let $\delta, \gamma \in SL$ be such that $\delta \vee \theta \mathrel{|\!\sim}_* \neg\theta$ and $\delta \mathrel{|\!\sim}_* \gamma$. We must show $\delta \mathrel{|\!\sim}' \gamma$. We have $S_{\delta\vee\theta} = S_\delta \cup S_\theta = S_\delta \cup u_p$. $\delta \vee \theta \mathrel{|\!\sim}_* \neg\theta$ implies that for l minimal such that $(S_\delta \cup u_p) \cap u_l \neq \emptyset$ we have $(S_\delta \cup u_p) \cap u_l \subseteq S_{\neg\theta} = At - u_p$. We know $l \leq p$ since $(S_\delta \cup u_p) \cap u_p = u_p \neq \emptyset$ but if $l = p$ then $u_p \subseteq At - u_p$ and so $u_p = \emptyset$ — a contradiction since $u_p \supset t_{ip}$. Hence $l < p$ and so l must also be minimal such that $S_\delta \cap u_l \neq \emptyset$. Therefore, since $t_{il} = u_l$ for $l < p$, l is also minimal such that $S_\delta \cap t_{il} \neq \emptyset$ and, since $\delta \mathrel{|\!\sim}_* \gamma$, we have $S_\delta \cap t_{il} = S_\delta \cap u_l \subseteq S_\gamma$ giving $\delta \mathrel{|\!\sim}' \gamma$. This completes the proof that **(RC1)** holds.

Now to show **(RC2)** let $\lambda, \chi \in SL$ be such that $\lambda \mathrel{|\!\sim}_* \chi$ and $\lambda \not\mathrel{|\!\sim}' \chi$. We will show that $\theta \vee \lambda \mathrel{|\!\sim}' \neg\lambda$ which, since we have already shown $\theta \mathrel{|\!\sim}' \phi$ and $\theta \not\mathrel{|\!\sim}_* \phi$ above, will suffice. Let l' be minimal such that $S_\lambda \cap t_{il'} \neq \emptyset$. We know l' exists since otherwise $\lambda \mathrel{|\!\sim}' \eta$ for all $\eta \in SL$ and hence $\lambda \mathrel{|\!\sim}' \chi$ — contradiction. We also know $l' \geq p$ since otherwise $S_\lambda \cap t_{il'} = S_\lambda \cap u_{l'} \subseteq S_\chi$ since $\lambda \mathrel{|\!\sim}_* \chi$, contradicting $\lambda \not\mathrel{|\!\sim}' \chi$. But if $l' = p$ then $S_\lambda \cap t_{ip} \subseteq S_\lambda \cap u_p \subseteq S_\chi$ since $\lambda \mathrel{|\!\sim}_* \chi$, again contradicting $\lambda \not\mathrel{|\!\sim}' \chi$. Hence $l' > p$, i.e. $S_\lambda \cap t_{il} = \emptyset$ for all $l \leq p$. Now $t_{ip} \neq \emptyset$, since otherwise we would have $t_{il'} = \emptyset$ which contradicts $S_\lambda \cap t_{il'} \neq \emptyset$. Hence p must be the minimal l_1 such that $S_{\theta\vee\lambda} \cap t_{il_1} = (S_\lambda \cup u_p) \cap t_{il_1} = (S_\lambda \cap t_{il_1}) \cup (u_p \cap t_{il_1}) \neq \emptyset$ and we have $S_\lambda \cap t_{ip} = \emptyset$, i.e. $t_{ip} \subseteq At - S_\lambda$ and $S_{\theta\vee\lambda} \cap t_{ip} = (S_\lambda \cap t_{ip}) \cup (u_p \cap t_{ip}) = \emptyset \cup t_{ip} = t_{ip}$. Hence $\theta \vee \lambda \mathrel{|\!\sim}' \neg\lambda$ as required. This completes the proof that **(RC2)** holds. $\square$

A further observation about this construction, which we shall take advantage of later is that in the original enumeration of (normal) models of rational consequence relations satisfying K we did not need to limit ourselves to normal models. Any superset would have sufficed since the effect of any non-normal model would clearly disappear in the construction of the final $\mathcal{U}_i$ because of the presence of the 'normalisation' of that model in the enumeration. Alternatively we could have worked with all infinite models $t_{\nu 1}, t_{\nu 2}, t_{\nu 3}, \ldots$, an observation which we will shortly exploit.

The form of the model of the rational closure of K constructed above adds, we believe, a further justification for the rational closure of K being the 'simplest' rational consequence relation satisfying K. For in terms of the naturalness of worlds this model assigns as many worlds as possible to the top level of naturalness, or preference, i.e. to $\mathcal{U}_1$. Given that, the model assigns as many of the remaining worlds to the next highest level of naturalness, and so on. Expressed in another way this model gives each world the highest level of naturalness that it can have (in any rational consequence relation satisfying K), it never unjustifiably condemns a world to a lower status than is a necessary.

As further evidence in favour of the rational closure's claim to be the simplest (at least in terms of information content) rational consequence relation satisfying K we now include three propositions which follow easily from the above construction. In these propositions $\mathrel{|\!\sim}_*$ stands, as usual, for the rational closure of K and $u_1, u_2, \ldots, u_k,\ t_{i1}, t_{i2}, \ldots, t_{ik_i}$ are as above. As we have already said, the intersection of all the rational consequence relations satisfying K, i.e. $\mathrel{|\!\sim}_P^K$, need not itself form a rational consequence relation. As an example of this consider the case when we have $L = \{p, q, r\}$ and $K = \{p \mathrel{|\!\sim} q\}$. Then clearly $p \mathrel{|\!\sim}_P^K q$. Also we have $p \not\mathrel{|\!\sim}_P^K \neg r$ (i.e. *not* $p \mathrel{|\!\sim}_P^K \neg r$), for consider the rational consequence relation modelled simply by t_1 where $t_1 = \{p \wedge q \wedge r\}$. If $\mathrel{|\!\sim}_P^K$ satisfied RMO then we would now conclude $p \wedge r \mid\!\sim^K_P q$. However consider the rational consequence relation $\mid\!\sim'$ modelled by u_1, u_2 where $u_1 = \{p \wedge q \wedge \neg r\}$, $u_2 = \{p \wedge \neg q \wedge r\}$. It is easy to check that we have $p \mid\!\sim' q$ but $p \wedge r \not\mid\!\sim' q$ and so $p \wedge r \not\mid\!\sim^K_P q$. Hence $\mid\!\sim^K_P$ here fails to satisfy RMO and so is not a rational consequence relation. It is possible, however, that, for certain K, $\mid\!\sim^K_P$ *might* turn out to be a rational consequence relation, in which case intuition tells us we should look no further for the simplest rational consequence relation satisfying K. Our first proposition says that the rational closure matches these intuitions.

Proposition 1 *If $\mid\!\sim^K_P$ is itself a rational consequence relation then $\mid\!\sim^K_P$ is the rational closure of K.*

Proof. Suppose $\mid\!\sim^K_P$ is a rational consequence relation with normal model $t_{i1}, t_{i2}, \ldots, t_{ik} \subseteq At$ and suppose $\mid\!\sim^K_P \neq \mid\!\sim_*$. Then let j be minimal such that $t_{ij} \neq u_j$. As in the proof of theorem 2 we then have $t_{ij} \subset u_j$ so let $\alpha \in At$ be such that $\alpha \in u_j - t_{ij}$. Then $\bigvee u_j \mid\!\sim^K_P \neg\alpha$, i.e. $\bigvee u_j \mid\!\sim \neg\alpha$ for all rational consequence relations $\mid\!\sim$ satisfying K. But $\mid\!\sim_*$ is a rational consequence relation satisfying K and $\bigvee u_j \not\mid\!\sim_* \neg\alpha$ — a contradiction. Hence $\mid\!\sim^K_P = \mid\!\sim_*$ as required. □

In what follows we shall write simply $\mid\!\sim \phi$ etc. if $\theta \mid\!\sim \phi$ for some tautology θ. Notice that by LLE whether or not this holds is independent of the particular tautology, θ, involved. The following result, with a different proof, already appeared as lemmas 5.15, 5.16 in [3] (assuming the equivalences in 1).

Proposition 2 *[Lehmann and Magidor] For $\mid\!\sim_*$ the rational closure of K:*
(i) If $\mid\!\sim_* \theta$ then $\mid\!\sim \theta$ holds for all rational consequence relations $\mid\!\sim$ satisfying K,
(ii) If $\theta \mid\!\sim_* \phi$ where ϕ is a contradiction (by RWE it does not matter which contradiction) then $\theta \mid\!\sim \phi$ holds for all rational consequence relations $\mid\!\sim$ satisfying K.

Proof. (i). We have $\mid\!\sim_* \theta$ iff whenever j is minimal such that $u_j \neq \emptyset$ then $u_j \subseteq S_\theta$. If K is not satisfied by any non-trivial rational consequence relation then the right hand side holds vacuously. If K *is* satisfied by a non-trivial rational consequence relation then $u_1 = \mathcal{U}_1 = \bigcup_{i=1}^r t_{i1} \neq \emptyset$ and

$$\begin{aligned}
\mid\!\sim_* \theta \quad &\Rightarrow \quad u_1 = \bigcup_{i=1}^{r} t_{i1} \subseteq S_\theta \\
&\Rightarrow \quad t_{i1} \subseteq S_\theta \text{ for } i = 1, \ldots, r \\
&\Rightarrow \quad \text{for } i \in \{1, \ldots, r\},\ t_{i1}, \ldots, t_{ik} \text{ models a rational consequence} \\
&\qquad \text{relation satisfying } \mid\!\sim \theta \\
&\Rightarrow \quad \mid\!\sim \theta \text{ for all rational consequence relations } \mid\!\sim \text{ satisfying } K.
\end{aligned}$$

(ii). Suppose $\theta \mid\!\sim_* \phi$ where ϕ is a contradiction. Suppose $S_\theta \cap \mathcal{U}_j \neq \emptyset$ for some $j \in \{1, \ldots, k\}$. Let j' be the least such j. We then have $S_\theta \cap \mathcal{U}_{j'} \subseteq S_\phi = \emptyset$ and hence $S_\theta \cap \mathcal{U}_{j'} = \emptyset$ — contradiction. Hence there can be no $j \in \{1, \ldots, k\}$ such that $S_\theta \cap \mathcal{U}_j \neq \emptyset$. So

$$\begin{aligned}
\theta \mid\!\sim_* \phi \quad &\Rightarrow \quad S_\theta \cap \mathcal{U}_j = \emptyset \text{ for } j = 1, \ldots, k \\
&\Rightarrow \quad S_\theta \cap t_{ij} = \emptyset \text{ for } i = 1, \ldots, r \text{ and } j = 1, \ldots, k \\
&\Rightarrow \quad \text{for } i \in \{1, \ldots, r\},\ t_{i1}, \ldots, t_{ik} \text{ models a} \\
&\qquad \text{rational consequence relation satisfying } \theta \mid\!\sim \phi \\
&\Rightarrow \quad \theta \mid\!\sim \phi \text{ for all rational consequence relations } \mid\!\sim \text{ satisfying} \\
&\qquad K.
\end{aligned}$$

$\square$

Proposition 3 *Let $\mid\!\sim_*$ be the rational closure of K and suppose $\psi \in SL$ is such that no propositional variable that appears in ψ also appears in a sentence in K or in θ. Then whenever $\theta \mid\!\sim_* \phi$ holds we also have that $\theta \wedge \psi \mid\!\sim_* \phi$ holds.*

Proof. If ψ is not satisfiable then the result is clear. Otherwise let let $\alpha \in u_i$ and let $\alpha' \in At$ be an atom which, as a conjunction of propositional variables or their negations, agrees with α on all propositional variables mentioned in sentences in K. Then adding α' to u_i would produce a model of a rational consequence relation which still satisfies K. It follows then by the construction of $u_1, u_2, \ldots, u_k$ that we must already have $\alpha' \in u_j$ for some $j \leq i$. By symmetry we must have $i = j$. Thus if i is minimal such that $S_\theta \cap u_i \neq \emptyset$ and $\alpha \in S_\theta \cap u_i$ then we can find $\alpha' \in u_i$ which agrees with α on all propositional variables not mentioned in ψ and satisfies $\alpha' \in S_\psi$. Clearly if $\alpha \in S_\phi$ then $\alpha' \in S_\phi$ and the result follows. $\square$

We now present an algorithm, based on an algorithm on p40 of [3] (see also [15] for a correction), for finding a model $\mathcal{U}'_1, \mathcal{U}'_2, \ldots$ of the rational closure of K. In the case that $K = K_P$ this algorithm reduces to the one given in [3] p40 and shows that in that case $\mathcal{U}_i = S_{\Omega_i}$ where the Ω_i are defined, simultaneously with subsets C_i of K by

$C_1 = K$

$\Omega_1 = \bigwedge \tilde{K}$, where for C a set of positive conditional assertions

$\tilde{C} = \{ \ (\theta \to \phi) \mid (\theta \mid\!\sim \phi) \in C \ \}$,

$C_{i+1} = \{ \ (\theta \mid\!\sim \phi) \in C_i \mid$ there is a proof in P of $\mid\!\sim \neg\theta$ from C_i $\}$

$= \{ \ (\theta \mid\!\sim \phi) \in C_i \mid \ \mid\!\sim \neg\theta$ holds in all rational consequence relations satisfying C_i $\}$

$\Omega_{i+1} = \bigwedge \tilde{C}_{i+1}$ for as long as $C_{i+1} \neq C_i$.

[Indeed, exactly as one would have expected, for $\theta \in SL$ the least i such that $S_\theta \cap u_i \neq \emptyset$ is precisely the *rank* of θ as defined in [3].]

The algorithm is as follows:

1. $i := 1, \ j := 1$
2. $C_j := K_P, \ D_j := K_N$
3. $\mathcal{U}^i_j := S_{\bigwedge \tilde{C}_j}$
4. $\mathcal{U}^{i+1}_j := \mathcal{U}^i_j - \bigcup\{S_\chi \mid (\chi \not\mid\!\sim \lambda) \in D_j \text{ and } S_{\chi\wedge\neg\lambda} \cap \mathcal{U}^i_j = \emptyset\}$
5. if $\mathcal{U}^{i+1}_j = \mathcal{U}^i_j$ then $\mathcal{U}'_j := \mathcal{U}^i_j$. Otherwise $i := i+1$ and go to 4.
6. $C_{j+1} := \{(\theta \mid\!\sim \phi) \in C_j \mid S_\theta \cap \mathcal{U}'_j = \emptyset\}, \ D_{j+1} := \{(\chi \not\mid\!\sim \lambda) \in D_j \mid S_{\chi\wedge\neg\lambda} \cap \mathcal{U}'_j = \emptyset\}$. If $C_{j+1} = C_j$ and $D_{j+1} = D_j$ then STOP and return $\mathcal{U}'_1, \ldots \mathcal{U}'_j$ as the answer, otherwise $j := j+1, \ i := 1$ and go to 3.

Before proving that this algorithm works we need the following lemma from [3].

Lemma 1 *Let K_P be any set of positive conditional assertions and let $\theta \in SL$. Then $\tilde{K} \models \theta$ iff $\mid\!\sim \theta$ holds for all rational consequence relations satisfying K_P.*

Theorem 3 *Given input of the form of a consistent knowledge base K, the above algorithm gives as output an (increasing) finite sequence of subsets $\mathcal{U}'_1, \ldots, \mathcal{U}'_k \subseteq At$ which form a model for the rational closure $\mid\!\sim_*$ of K.*

Proof. Let $K = K_P \cup K_N$ and let the infinite sequences $t_{\nu 1}, t_{\nu 2}, \ldots$ for $\nu \in E$ enumerate all the models of all the rational consequence relations which satisfy K. Notice that $E \neq \emptyset$ since K is consistent. We will show that in fact the algorithm returns the model $\bigcup_{\nu \in E} t_{\nu 1}, \bigcup_{\nu \in E} t_{\nu 2}, \ldots$, which, as we have already remarked after the proof of theorem 2, is a model for the rational closure of K.

We first show by induction on j that $\mathcal{U}_j^1 \supseteq t_{\nu 1}$ for all $\nu \in E$. Firstly for $j = 1$ notice that if $\alpha \in \bigcup_{\nu \in E} t_{\nu 1}$ then $\mathrel{|\!\sim} \neg\alpha$ fails for some rational consequence relation satisfying K (hence also K_P), by lemma 1, so $\tilde{K}_P \not\models \neg\alpha$, equivalently $\alpha \in S_{\bigwedge \tilde{K}_P} = \mathcal{U}_1^1$. Now suppose that $\mathcal{U}_j^1 \supseteq t_{\nu 1}$ for $\nu \in E$ and $\mathcal{U}_j^1 \neq \mathcal{U}_{j+1}^1$. If $(\chi \mathrel{|\!\not\sim} \lambda) \in D_j$ and $S_{\chi \wedge \neg\lambda} \cap \mathcal{U}_j^1 = \emptyset$ then $S_{\chi \wedge \neg\lambda} \cap t_{\nu 1} = \emptyset$ for $\nu \in E$ by inductive hypothesis so $S_\chi \cap t_{\nu 1} = \emptyset$, since $t_{\nu 1}, t_{\nu 2}, \ldots$ is a model of a rational consequence relation satisfying K_N. Hence $\mathcal{U}_j^1 - S_\chi \supseteq t_{\nu 1}$ so $\mathcal{U}_{j+1}^1 \supseteq \bigcup_{\nu \in E} t_{\nu 1}$, as required.

Now notice that if $t_1, t_2, \ldots$ is a model of a rational consequence relation satisfying $C_2 \cup D_2$ then $\mathcal{U}_1', t_2, t_3, \ldots$ is a model of a rational consequence relation satisfying $K = (C_1 \cup D_1)$. For suppose first that $(\theta \mathrel{|\!\sim} \phi) \in K_P$. If $S_\theta \cap \mathcal{U}_1' \neq \emptyset$ then $S_\theta \cup \subseteq S_\phi$ (since $\mathcal{U}_1' \subseteq \mathcal{U}_1^1 = S_{\bigwedge \tilde{K}_P} \subseteq S_{\neg\theta \vee \phi}$). On the other hand if $S_\theta \cap \mathcal{U}_1' = \emptyset$ then $(\theta \mathrel{|\!\sim} \phi) \in C_2$ so, since $t_2, t_3, \ldots$ is a model of a rational consequence relation satisfying C_2, the first term (if any) in $\mathcal{U}_1', t_2, t_3, \ldots$ not disjoint from S_θ must be some $t_i, i \geq 2$ and must satisfy $t_i \cup S_\theta \subseteq S_\phi$. Now suppose that $(\chi \mathrel{|\!\not\sim} \lambda) \in K_N$. If $S_\chi \cup \mathcal{U}_1' \neq \emptyset$ it must be the case, by the construction, that $S_{\chi \wedge \neg\lambda} \cap \mathcal{U}_1' \neq \emptyset$. Otherwise $S_\chi \cap \mathcal{U}_1' = \emptyset$ so the first term in $\mathcal{U}_1', t_2, t_3, \ldots$ with non-empty intersection with S_χ must be a t_i and must satisfy $S_{\chi \wedge \neg\lambda} \cap t_i \neq \emptyset$ since $t_2, t_3, \ldots$ is a model of a rational consequence relation satisfying $\chi \mathrel{|\!\not\sim} \lambda$. Either way then $\mathcal{U}_1', t_2, t_3, \ldots$ is a model of a rational consequence relation satisfying $\chi \mathrel{|\!\not\sim} \lambda$.

From this it follows that $\mathcal{U}_1', t_2, t_3, \ldots$ is a model of a rational consequence relation satisfying K. In particular then this model already appears amongst the $t_{\nu 1}, t_{\nu 2}, \ldots$ so $\mathcal{U}_1' = \bigcup_{\nu \in E} t_{\nu 1}$.

The above construction can, in a sense, be reversed. For suppose $\nu \in E$. Then we claim that $t_{\nu 1}, t_{\nu 2}, \ldots$ is a model of a rational consequence relation satisfying $C_2 \cup D_2$. For if $(\theta \mathrel{|\!\sim} \phi) \in C_2$ then $S_\theta \cap t_{\nu 1} = \emptyset$ (since $S_\theta \cap \mathcal{U}_1' = \emptyset$ and $t_{\nu 1} \subseteq \mathcal{U}_1'$) so the least i (if it exists) for which $S_\theta \cap t_{\nu 1} \neq \emptyset$ must be at least 2 and must satisfy $S_\theta \cap t_{\nu 1} \subseteq S_\phi$ (since $t_{\nu 1}, t_{\nu 2}, \ldots$ is a model of a rational consequence relation satisfying C_2). Similarly if $(\chi \mathrel{|\!\not\sim} \lambda) \in D_2$ then $S_\chi \cap \mathcal{U}_1' = \emptyset$ so $S_\chi \cap t_{\nu 1} = \emptyset$ and the least i for which $S_\chi \cap t_{\nu i} \neq \emptyset$ exists, and is at least 2, and satisfies $S_{\chi \wedge \neg\lambda} \cap t_{\nu i} \neq \emptyset$ (since $t_{\nu 1}, t_{\nu 2}, \ldots$ is a model of a rational consequence relation satisfying D_2.)

The importance of these observations is that we now see that the $t_{\nu 2}, t_{\nu 3}, \ldots$ for $\nu \in E$ runs through all models of rational consequence relations satisfying $C_2 \cup D_2$, so by repeating the argument that showed that $\mathcal{U}_1' = \bigcup_{\nu \in E} t_{\nu 1}$ we can show that $\mathcal{U}_2' = \bigcup_{\nu \in E} t_{\nu 2}$, $\mathcal{U}_3' = \bigcup_{\nu \in E} t_{\nu 3}, \ldots$ and that this will continue for as long as the $C_j \cup D_j$ keep decreasing. Once they stop decreasing the sets $\bigcup_{\nu \in E} t_{\nu j}$ will be constant (and clearly not influence the modeled rational consequence relation). Furthermore from the remarks following theorem 2 it follows that the $\mathcal{U}_1', \mathcal{U}_2', \ldots$ provide a model of the rational closure of K. It should also be clear (although we will not need this fact in what follows) that the $\mathcal{U}_i' = \mathcal{U}_i$ and that both sequences have the same length. $\square$

A point to notice about the above proof is that it also shows that at any intermediate stage r in the algorithm the $\mathcal{U}_1', \mathcal{U}_2', \ldots, \mathcal{U}_r'$ so far constructed model a rational consequence relation satisfying $K - (C_{r+1} \cup D_{r+1})$ and indeed this will hold even if is *not consistent*. We will use this observation in the next section.

Completeness results for the positive and negative consequences of K

In this section we demonstrate two sets of rules and axioms extending those for rational consequence which are complete for the positive, respectively negative, consequences of K. In other words these rules and axioms will have the property that $\theta \mid\!\sim \phi$ ($\theta \mid\!\not\sim \phi$) will be derivable from $K = K_P + K_N$ just if $\theta \mid\!\sim \phi$ ($\theta \mid\!\not\sim \phi$) holds of all rational consequence relations satisfying K. This then will give a characterisation of the relations $\mid\!\sim_P^K$ and $\mid\!\not\sim_N^K$, so answering our problems A and B. Obviously a set of rules and axioms which works for both the positive and negative consequences of K simultaneously can be obtained by simply combining these two sets.

Precisely our further rules of proof are:

- $\dfrac{\theta \vee \phi \mid\!\sim (\theta \wedge \eta) \vee (\neg\theta \wedge \phi \wedge \gamma),\ \phi \mid\!\not\sim \gamma}{\theta \mid\!\sim \eta}$ R1
- $\dfrac{\theta \vee \phi \mid\!\sim \neg\phi \vee \gamma,\ \phi \mid\!\not\sim \gamma}{\theta \vee \phi \mid\!\sim \neg\phi}$ R2
- $\dfrac{\phi \equiv \psi,\ \phi \mid\!\not\sim \gamma}{\psi \mid\!\not\sim \gamma}$ R3
- $\dfrac{\theta \vee \phi \mid\!\not\sim \neg\phi \vee \eta}{\phi \mid\!\not\sim \eta}$ R4
- $\dfrac{(\phi \wedge \neg\gamma) \vee \theta \mid\!\sim \neg\eta,\ \phi \mid\!\not\sim \gamma}{\theta \vee \phi \mid\!\not\sim \eta}$ R5
- $\dfrac{(\phi \wedge \neg\gamma) \vee \theta \mid\!\not\sim \eta,\ \phi \mid\!\not\sim \gamma}{\theta \vee \phi \mid\!\not\sim \eta}$ R6

These additional rules have largely been tailored to fit directly into the subsequent completeness proofs with little regard as to their elegance or transparent soundness. We address this shortcoming in the appendix where we show that, in fact, they are all direct consequences of the GM axioms and rules and their 'reversals'. [Where, for example, the rule

- $\dfrac{\theta \mid\!\not\sim \psi,\ \theta \equiv \phi}{\phi \mid\!\not\sim \psi}$

is the reversal of the left logical equivalence rule. A full list of these reversals is given in the appendix.]

Lemma 2 *The following rule is derivable from the GM rules and axioms and $\mathcal{R}1-\mathcal{R}6$:*

- $\dfrac{\theta \vee \phi \mid\!\sim \eta,\ \phi \mid\!\not\sim \gamma}{\theta \vee (\phi \wedge \neg\gamma) \mid\!\sim \eta}$ *R7*

[Notice that RMO follows from R7 in the case $\theta = (\phi \wedge \neg\gamma)$.]

Proof. By SC we have $(\phi \wedge \neg\gamma) \vee \theta \mid\!\sim \neg\neg(\theta \vee \neg\gamma)$ so with $\phi \mid\!\not\sim \gamma$ and $\mathcal{R}5$ we obtain $\theta \vee \phi \mid\!\not\sim \neg(\theta \vee \neg\gamma)$. By RMO together with the other antecedent of $\mathcal{R}7$, $\theta \vee \phi \mid\!\sim \eta$, we obtain $(\theta \vee \phi) \wedge (\theta \vee \neg\gamma) \mid\!\sim \eta$ and $\theta \vee (\phi \wedge \neg\gamma) \mid\!\sim \eta$ follows by LLE, as required. $\square$

In the proofs of the next two theorems we shall omit explicit mention of the more obvious instances of the rules RWE and LLE.

Theorem 4 *Let $K = K_P + K_N$ be consistent. Then $\chi \mathrel{|\!\sim} \lambda$ holds for all rational consequence relations satisfying K (i.e. $\chi \mathrel{|\!\sim}^K_P \lambda$) iff $\chi \mathrel{|\!\sim} \lambda$ is derivable from K using the GM rules and axioms for rational consequence augmented with $\mathcal{R}1$ and $\mathcal{R}2$.*

Proof. It is easy to check, using the usual semantics for rational consequence, that these rules are all sound so all that remains is to show that if $\chi \mathrel{|\!\sim} \lambda$ holds for all rational consequence relations satisfying K then $\chi \mathrel{|\!\sim} \lambda$ can be derived from K using the GM rules and axioms together with $\mathcal{R}1$ and $\mathcal{R}2$.

Now assuming that $\chi \mathrel{|\!\sim} \lambda$ holds in all rational consequence relations satisfying K it follows that $K + (\chi \mathrel{|\!\not\sim} \lambda)$ is inconsistent and hence that attempting to use the algorithm described in the previous section to construct 'the rational closure' of $K + (\chi \mathrel{|\!\not\sim} \lambda)$ must fail to produce a rational consequence relation satisfying $K + (\chi \mathrel{|\!\not\sim} \lambda)$. Now the only way in which this can happen is that at the point at which the $C_j + D_j$ become fixed (where $C_1 = K_P,\ D_1 = K_N + (\chi \mathrel{|\!\not\sim} \lambda)$) we still have $D_j \neq \emptyset$. For simplicity let us suppose that this occurs at the point at which

$$C_j = \{\ \theta_i \mathrel{|\!\sim} \phi_i \mid i = 1, \ldots, m\ \}$$

$$D_j = \{\ \chi_1 \mathrel{|\!\not\sim} \lambda_1,\ \chi_2 \mathrel{|\!\not\sim} \lambda_2,\ \chi_3 \mathrel{|\!\not\sim} \lambda_3\ \}$$

[It will be clear from the proof that the argument for a general D_j can be carried through.] Notice that $(\chi \mathrel{|\!\not\sim} \lambda) \in D_j$, since otherwise we would have $C_j \subseteq K_P,\ D_j \subseteq K_N$ so $C_j + D_j$ would be consistent and the algorithm would not fail.

Since the algorithm would continue to cycle at this point we may, without loss of generality, take it that the $\mathcal{U}^i_j$'s satisfy

$\mathcal{U}^1_j = S_\Omega$ where $\Omega = \bigwedge_{i=1}^m (\neg\theta_i \vee \phi_i)$

$\mathcal{U}^1_j \cap S_{\chi_1 \wedge \neg\lambda_1} = \emptyset$ so Ω and $\chi_1 \wedge \neg\lambda_1$ are mutually inconsistent, henceforth denoted by $\Omega \perp \chi_1 \wedge \neg\lambda_1$

$\mathcal{U}^2_j = \mathcal{U}^1_j - S_{\chi_1} = S_{\Omega \wedge \neg\chi_1}$

$\mathcal{U}^2_j \cap S_{\chi_2 \wedge \neg\lambda_2} = \emptyset$ so $\Omega \wedge \neg\chi_1 \perp \chi_2 \wedge \neg\lambda_2$

$\mathcal{U}^3_j = \mathcal{U}^2_j - S_{\chi_2} = S_{\Omega \wedge \neg\chi_1 \wedge \neg\chi_2}$

$\mathcal{U}^3_j \cap S_{\chi_3 \wedge \neg\lambda_3} = \emptyset$ so $\Omega \wedge \neg\chi_1 \wedge \neg\chi_2 \perp \chi_3 \wedge \neg\lambda_3$

$\mathcal{U}^4_j = \mathcal{U}^3_j - S_{\chi_3} = S_{\Omega \wedge \neg\chi_1 \wedge \neg\chi_2 \wedge \neg\chi_3}$

and finally, since none of the $(\theta_i \mathrel{|\!\sim} \phi_i)$ in C_j are dropped in forming C_{j+1},

$$\Omega \wedge \neg\chi_1 \wedge \neg\chi_2 \wedge \neg\chi_3 \perp \theta_i \qquad for\ i = 1, \ldots, m$$

Consequently we now have that

$$\chi_1 \wedge \neg\lambda_1 \models \bigvee_{i=1}^m (\theta_i \wedge \neg\phi_i) \quad (\equiv \neg\Omega) \tag{1}$$

$$\neg\chi_1 \wedge \chi_2 \wedge \neg\lambda_2 \models \bigvee_{i=1}^m (\theta_i \wedge \neg\phi_i) \tag{2}$$

$$\neg\chi_1 \wedge \neg\chi_2 \wedge \chi_3 \wedge \neg\lambda_3 \models \bigvee_{i=1}^m (\theta_i \wedge \neg\phi_i) \tag{3}$$

$$\theta_k \models \chi_1 \vee \chi_2 \vee \chi_3 \vee \bigvee_{i=1}^{m} (\theta_i \wedge \neg\phi_i) \qquad fork = 1, \ldots, m \tag{4}$$

and, of course,

$$\theta_k \mid\!\sim \phi_k \quad for\ k = 1, \ldots, m \tag{5}$$

The plan now is to use (1)-(5), D_j and the rules to derive a string of sequents, ending up with $\chi \mid\!\sim \lambda$. First notice that by semantic considerations there is a derivation, in P, of $\bigvee_{i=1}^{m} \theta_i \mid\!\sim \neg(\theta_k \wedge \neg\phi_k)$ from (5). By AND we now obtain

$$\bigvee_{i=1}^{m} \theta_i \mid\!\sim \bigwedge_{i=1}^{m} \neg(\theta_i \wedge \neg\phi_i)$$

and using (4), SC and OR, gives

$$\bigvee_{i=1}^{m} \theta_i \mid\!\sim \chi_1 \vee \chi_2 \vee \chi_3 \vee \bigvee_{i=1}^{m} (\theta_i \wedge \neg\phi_i).$$

Combining these using AND and RWE yields

$$\bigvee_{i=1}^{m} \theta_i \mid\!\sim \chi_1 \vee \chi_2 \vee \chi_3. \tag{6}$$

By a semantic argument we see that the following is a derived rule (of P)

- $\dfrac{\theta \mid\!\sim \phi, \quad \psi \mid\!\sim (\theta \wedge \neg\phi) \vee \eta}{\psi \vee \theta \mid\!\sim (\neg\psi \wedge \theta) \vee \eta}$

and applying this to (1) and (5) (with $k = 1$) and using RWE gives

$$(\chi_1 \wedge \neg\lambda_1) \vee \theta_1 \mid\!\sim ((\neg\chi_1 \vee \lambda_1) \wedge \theta_1) \vee \bigvee_{i=2}^{m} (\theta_i \wedge \neg\phi_i).$$

Applying the rule again to this conditional and (5) (with $k = 2$) gives

$$(\chi_1 \wedge \neg\lambda_1) \vee \theta_1 \vee \theta_2 \mid\!\sim ((\neg\chi_1 \vee \lambda_1) \wedge \neg\theta_1 \wedge \theta_2) \vee ((\neg\chi_1 \vee \lambda_1) \wedge \theta_1) \vee \bigvee_{i=3}^{m} (\theta_i \wedge \neg\phi_i)$$

and by RWE

$$(\chi_1 \wedge \neg\lambda_1) \vee \theta_1 \vee \theta_2 \mid\!\sim ((\neg\chi_1 \vee \lambda_1) \wedge (\theta_1 \vee \theta_2)) \vee \bigvee_{i=3}^{m} (\theta_i \wedge \neg\phi_i).$$

Continuing in this way gives

$$(\chi_1 \wedge \neg\lambda_1) \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim (\neg\chi_1 \vee \lambda_1) \wedge \bigvee_{i=1}^{m} \theta_i.$$

Hence, by RWE,

$$(\chi_1 \wedge \neg\lambda_1) \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim \neg\chi_1 \vee \lambda_1$$

and since, by SC,

$$\chi_1 \wedge \lambda_1 \mid\!\sim \neg\chi_1 \vee \lambda_1,$$

by OR and LLE,

$$\chi_1 \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim \neg\chi_1 \vee \lambda_1. \tag{7}$$

Exactly similarly using (2),(3) in place of (1) we can obtain

$$(\neg\chi_1 \wedge \chi_2) \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim \chi_1 \vee \neg\chi_2 \vee \lambda_2 \tag{8}$$

$$(\neg\chi_1 \wedge \neg\chi_2 \wedge \chi_3) \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim \chi_1 \vee \chi_2 \vee \neg\chi_3 \vee \lambda_3 \tag{9}$$

Now, by semantic considerations we see that using (6),(7),(8),(9) we can obtain

$$\chi_1 \vee \chi_2 \vee \chi_3 \mid\!\sim (\lambda_1 \wedge \chi_1) \vee (\lambda_2 \wedge \chi_2 \wedge \neg\chi_1) \vee (\lambda_3 \wedge \chi_3 \wedge \neg\chi_2 \wedge \neg\chi_1) \tag{10}$$

We now need to consider cases according to which $i \in \{1,2,3\}$ we have $(\chi_i \mid\!\not\sim \lambda_i)$ equal to $(\chi \mid\!\not\sim \lambda)$.

<u>*Case* $i = 1$</u>

In this case $(\chi_2 \mid\!\not\sim \lambda_2), (\chi_3 \mid\!\not\sim \lambda_3) \in K_N$. By (10),LLE,RWE we have

$$\theta \vee (\neg\theta \wedge \chi_3) \mid\!\sim (\theta \wedge \eta) \vee (\neg\theta \wedge \chi_3 \wedge \lambda_3),$$

where

$$\theta = \chi_1 \vee \chi_2, \quad \eta = (\lambda_1 \wedge \chi_1) \vee (\lambda_2 \wedge \chi_2 \wedge \neg\chi_1),$$

so by $\mathcal{R}1$ and $\chi_3 \mid\!\not\sim \lambda_3$ we obtain

$$\chi_1 \vee \chi_2 \mid\!\sim (\lambda_1 \wedge \chi_1) \vee (\lambda_2 \wedge \chi_2 \wedge \neg\chi_1) \tag{11}$$

and hence by a similar argument with $\chi_2 \mid\!\not\sim \lambda_2$,

$$\chi_1 \mid\!\sim \lambda_1$$

as required.

<u>*Case* $i = 2$</u>

In this case $(\chi_1 \mid\!\not\sim \lambda_1)$, $(\chi_3 \mid\!\not\sim \lambda_3) \in K_N$. By using $\mathcal{R}2$ with $(\chi_1 \mid\!\not\sim \lambda_1)$ and (7) we obtain

$$\chi_1 \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim \neg\chi_1. \tag{12}$$

Applying the derived rule (of P)

- $$\frac{\psi \vee \theta \mid\!\sim \neg\psi, \quad (\neg\psi \wedge \phi) \vee \theta \mid\!\sim \psi \vee \neg\phi \vee \eta}{\phi \vee \theta \mid\!\sim \neg\phi \vee \eta}$$

to (12),(8) and (12),(9) yields

$$\chi_2 \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim \neg\chi_2 \vee \lambda_2, \tag{13}$$

$$(\neg\chi_2 \wedge \chi_3) \vee \bigvee_{i=1}^{m} \theta_i \mid\!\sim \chi_2 \vee \neg\chi_3 \vee \lambda_3. \tag{14}$$

In addition, from the derived rule

- $\dfrac{\psi \vee \theta \mathrel{|\!\sim} \neg\psi,\ \theta \mathrel{|\!\sim} \psi \vee \phi}{\theta \mathrel{|\!\sim} \phi}$

and (12),(6) we obtain

$$\bigvee_{i=1}^{m} \theta_i \mathrel{|\!\sim} \chi_2 \vee \chi_3. \tag{15}$$

By an exactly analogous argument to that used in the case $i = 1$ to obtain (10) from (6),(7),(8),(9) we can obtain

$$\chi_2 \vee \chi_3 \mathrel{|\!\sim} (\lambda_2 \wedge \chi_2) \vee (\lambda_3 \wedge \chi_3 \wedge \neg\chi_2)$$

from (13),(14),(15) and, in turn, get $\chi_2 \mathrel{|\!\sim} \lambda_2$ from $\chi_3 \mathrel{|\!\not\sim} \lambda_3$.

<u>*Case $i = 3$*</u>

In this case we proceed just as in the case $i = 2$ to derive from $\chi_2 \mathrel{|\!\not\sim} \lambda_2$ and (13),(14)

$$\chi_3 \vee \bigvee_{i=1}^{m} \theta_i \mathrel{|\!\sim} \neg\chi_3 \vee \lambda_3, \tag{16}$$

and from $\chi_2 \mathrel{|\!\not\sim} \lambda_2$ and (13),(15)

$$\bigvee_{i=1}^{m} \theta_i \mathrel{|\!\sim} \chi_3. \tag{17}$$

By considering the semantic argument we see that there is a derivation from (16),(17) of the analogue,

$$\chi_3 \mathrel{|\!\sim} \lambda_3 \wedge \chi_3,$$

of (10) and the required $\chi_3 \mathrel{|\!\sim} \lambda_3$ now follows.

Whilst we have limited ourselves here to the case of D_j having just 3 members we hope it is clear that this method of proof provides a procedure which will work in general. □

We now turn our attention to the completeness result for the negative consequences of K.

Theorem 5 *Let $K = K_P + K_N$ be consistent. Then $\theta \mathrel{|\!\not\sim} \phi$ holds for all rational consequence relations satisfying K (i.e. $\theta \mathrel{|\!\not\sim}_N^K \phi$) iff $\theta \mathrel{|\!\not\sim} \phi$ is derivable from K using the GM rules and axioms for rational consequence together with $\mathcal{R}3$,$\mathcal{R}4$,$\mathcal{R}5$,$\mathcal{R}6$.*

Proof. The soundness of these rules for rational consequence relations is easy to check. For the converse we proceed as in the proof of the previous theorem, noting, of course, that by the lemma 2 we also have the rule R7 available. So suppose that every rational consequence relation satisfying K also satisfies $\theta \mathrel{|\!\not\sim} \phi$. Then our attempt to use the algorithm to find a rational consequence relation to satisfy $K_P + (\theta \mathrel{|\!\sim} \phi) + K_N$ must fail. Again this must occur because at some stage the C_j, D_j do not decrease and $D_j \neq \emptyset$, say for simplicity that at this stage

$$D_j = \{\ \chi_1 \mathrel{|\!\not\sim} \lambda_1,\ \chi_2 \mathrel{|\!\not\sim} \lambda_2,\ \chi_3 \mathrel{|\!\not\sim} \lambda_3\}.$$

We must also have at this stage that $(\theta \mathrel{|\!\sim} \phi) \in C_j$, otherwise $C_j + D_j \subseteq K$, so $C_j + D_j$ would be consistent and the algorithm would not fail. Let

$$C_j \cap K_P = \{\ (\theta_i \mathrel{|\!\sim} \phi_i) \mid i = 1, \ldots, m\}.$$

As in the previous theorem we may suppose that we have

$$(\neg\theta \vee \phi) \wedge \bigwedge_{i=1}^{m} (\neg\theta_i \vee \phi_i) \perp \chi_1 \wedge \neg\lambda_1 \tag{18}$$

$$(\neg\theta \vee \phi) \wedge \bigwedge_{i=1}^{m} (\neg\theta_i \vee \phi_i) \wedge \neg\chi_1 \perp \chi_2 \wedge \neg\lambda_2 \tag{19}$$

$$(\neg\theta \vee \phi) \wedge \bigwedge_{i=1}^{m} (\neg\theta_i \vee \phi_i) \wedge \neg\chi_1 \wedge \neg\chi_2 \perp \chi_3 \wedge \neg\lambda_3 \tag{20}$$

$$(\neg\theta \vee \phi) \wedge \bigwedge_{i=1}^{m} (\neg\theta_i \vee \phi_i) \wedge \neg\chi_1 \wedge \neg\chi_2 \wedge \neg\chi_3 \perp \theta_k,\ \theta \qquad for\ k = 1, \dots, m. \tag{21}$$

Proceeding as before we can now obtain from (18)-(21) and $\theta_i \mid\!\sim \phi_i$ for $i = 1, \dots, m$ that

$$\bigvee_{i=1}^{m} \theta_i \mid\!\sim \chi_1 \vee \chi_2 \vee \chi_3 \vee (\theta \wedge \neg\phi) \tag{22}$$

$$\theta \mid\!\sim \chi_1 \vee \chi_2 \vee \chi_3 \vee \bigvee_{i=1}^{m} (\theta_i \wedge \neg\phi_i) \vee \neg\phi \tag{23}$$

By a semantics argument, using (18)-(23), we can now show that there is a derivation from $\{\ (\theta_i \mid\!\sim \phi_i) \mid i = 1, \dots, m\ \}$ of

$$\begin{aligned} \theta \vee \chi_1 \vee \chi_2 \vee \chi_3 \quad \mid\!\sim \quad & [\chi_1 \wedge (\lambda_1 \vee (\theta \wedge \neg\phi))] \vee [\chi_2 \wedge \neg\chi_1 \wedge (\lambda_2 \vee (\theta \wedge \neg\phi))] \\ & \vee [\chi_3 \wedge \neg\chi_2 \wedge \neg\chi_1 \wedge (\lambda_3 \vee (\theta \wedge \neg\phi))] \\ & \vee [\neg\chi_3 \wedge \neg\chi_2 \wedge \neg\chi_1 \wedge \theta \wedge \neg\phi] \end{aligned} \tag{24}$$

[In this case the semantic argument is rather more involved than usual. A useful point to notice however is that if $t_1, t_2, \dots$ is a model of a rational consequence relation satisfying $\{\ (\theta_i \mid\!\sim \phi_i) \mid i = 1, \dots, m\ \}$ and j is minimal such that $t_j \cap S_{\theta \vee \chi_1 \vee \chi_2 \vee \chi_3} \neq \emptyset$ then for α an atom in this intersection $\alpha \notin S_{\theta_i \wedge \neg\phi_i}$, since otherwise by (22) j would also have to be minimal such that $t_j \cap S_{\theta_i} \neq \emptyset$ and $\theta_i \mid\!\sim \phi_i$ would not hold in this rational consequence relation, contradiction.]

Using LLE to substitute $\neg\chi_1 \wedge \neg\chi_2 \wedge \neg\chi_3 \wedge \theta$ for θ and then using R7 and LLE again we can now replace the left hand side of (24) by

$$\chi_1 \vee \chi_2 \vee (\chi_3 \wedge \neg\chi_2 \wedge \neg\chi_1 \wedge \neg\lambda_3) \vee (\neg\chi_1 \wedge \neg\chi_2 \wedge \neg\chi_3 \wedge \theta) \tag{25}$$

Repeating for χ_1, χ_2 and using REF,AND and RWE yields

$$(\chi_1 \wedge \neg\lambda_1) \vee (\chi_2 \wedge \neg\chi_1 \wedge \neg\lambda_2) \vee (\chi_3 \wedge \neg\chi_2 \wedge \neg\chi_1 \wedge \neg\lambda_3) \vee (\neg\chi_1 \wedge \neg\chi_2 \wedge \neg\chi_3 \wedge \theta) \mid\!\sim \theta \wedge \neg\phi \tag{26}$$

Now using $\mathcal{R}5$ we obtain from (26)

$$\chi_1 \vee (\chi_2 \wedge \neg\chi_1 \wedge \neg\lambda_2) \vee (\chi_3 \wedge \neg\chi_2 \wedge \neg\chi_1 \wedge \neg\lambda_3) \vee (\neg\chi_1 \wedge \neg\chi_2 \wedge \neg\chi_3 \wedge \theta) \not\mid\!\sim \neg\theta \vee \phi \tag{27}$$

and now using $\mathcal{R}3$ to remove the copies of $\neg\chi_1$ and applying $\mathcal{R}6$ and $\mathcal{R}3$ gives

$$\chi_1 \vee \chi_2 \vee (\chi_3 \wedge \neg\chi_2 \wedge \neg\lambda_3) \vee (\neg\chi_2 \wedge \neg\chi_3 \wedge \theta) \not\mid\!\sim \neg\theta \vee \phi. \tag{28}$$

By a similar step for χ_3 we now obtain

$$\chi_1 \vee \chi_2 \vee \chi_3 \vee \theta \not\mid\!\sim \neg\theta \vee \phi \tag{29}$$

and applying $\mathcal{R}4$ now gives $\theta \not\mid\!\sim \phi$ as required. □

Clearly we now also have a completeness result for both the positive and negative consequences simultaneously by simply combining $\mathcal{R}1$-$\mathcal{R}6$ with the GM rules and axioms. In fact we can do rather better since $\mathcal{R}1$ is actually derivable from the other rules. To see this suppose that $\theta \vee \phi \mid\!\sim (\theta \wedge \eta) \vee (\neg\theta \wedge \phi \wedge \gamma)$ and $\phi \not\mid\!\sim \gamma$. Then by R7 $\theta \vee (\phi \wedge \neg\gamma) \mid\!\sim (\theta \wedge \eta) \vee (\neg\theta \wedge \phi \wedge \gamma)$. By AND, since $\neg\theta \wedge \phi \wedge \gamma \perp \theta \vee (\phi \wedge \neg\gamma)$,

$$\theta \vee (\phi \wedge \neg\gamma) \mid\!\sim \theta \wedge \eta$$

and by the usual semantic considerations $\theta \mid\!\sim \eta$ now follows.

Conclusion

In this paper we have extended some of the results of Lehmann and Magidor [3] (for a finite language) to the case where we also have negative as well as positive conditional assertions. In addition we have demonstrated extensions of the usual GM rules and axioms which are complete for the positive and negative consequences of a mixed knowledge base. Furthermore the proofs of these completeness results are *elementary* in the sense that they are given directly within the framework of nonmonotonic logic.

Appendix

In this appendix we show how each of the rules $\mathcal{R}1 - \mathcal{R}6$ can be derived from the GM rules and axioms together with the following rules (their 'reversals'):

- $\dfrac{\theta \not\mid\!\sim \psi,\ \theta \equiv \phi}{\phi \not\mid\!\sim \psi}$ (lle)
- $\dfrac{\theta \not\mid\!\sim \psi,\ \phi \models \psi}{\theta \not\mid\!\sim \phi}$ (rwe)
- $\dfrac{\theta \vee \phi \not\mid\!\sim \psi,\ \theta \mid\!\sim \psi}{\phi \not\mid\!\sim \psi}$ (or)
- $\dfrac{\theta \not\mid\!\sim \phi \wedge \psi,\ \theta \mid\!\sim \phi}{\theta \not\mid\!\sim \psi}$ (and)
- $\dfrac{\theta \wedge \psi \not\mid\!\sim \phi,\ \theta \not\mid\!\sim \neg\psi}{\theta \not\mid\!\sim \phi}$ (rmo1)
- $\dfrac{\theta \wedge \psi \not\mid\!\sim \phi,\ \theta \mid\!\sim \phi}{\theta \mid\!\sim \neg\psi}$ (rmo2)
- $\dfrac{\theta \wedge \phi \not\mid\!\sim \psi}{\theta \not\mid\!\sim \phi \wedge \psi}$ (A)

Note how each of the above rules, apart from (A), is a reversal of one of the GM rules. The rule (A) itself is the reversal of a rule which, in the presence of AND and RWE, is clearly equivalent to CMO. The set of rules consisting of the GM rules and axioms together with the above seven rules we will henceforth denote by

$GM^{\pm}$. Before giving the main result of this section we will first present another rule which is derivable in $GM^{\pm}$ and which will be used several times in the upcoming derivations.

Lemma 3 *The following rule is derivable in $GM^{\pm}$.*

- $\dfrac{\phi \not\mid\!\sim \gamma}{\theta \vee \phi \not\mid\!\sim \neg(\theta \vee \neg\gamma)}$ *(B)*

Proof. From $\phi \not\mid\!\sim \gamma$ we get $(\theta \vee \phi) \wedge \phi \not\mid\!\sim \gamma$ by (lle). Then $\theta \vee \phi \not\mid\!\sim \phi \wedge \gamma$ by (A) and hence $\theta \vee \phi \not\mid\!\sim (\theta \vee \phi) \wedge (\neg\theta \wedge \gamma)$ by (rwe). Using (and) together with $\theta \vee \phi \mid\!\sim \theta \vee \phi$ (REF), this gives us $\theta \vee \phi \not\mid\!\sim \neg\theta \wedge \gamma$ and the desired conclusion follows by (rwe). □

Theorem 6 *The rules ($\mathcal{R}2$) – ($\mathcal{R}7$) are derivable in $GM^{\pm}$.*

Proof.

$\underline{\mathcal{R}2}$ $\dfrac{\theta \vee \phi \mid\!\sim \neg\phi \vee \gamma,\ \phi \not\mid\!\sim \gamma}{\theta \vee \phi \mid\!\sim \neg\phi}$

From $\phi \not\mid\!\sim \gamma$ using (rwe) we get $\phi \not\mid\!\sim \phi \wedge (\neg\phi \vee \gamma)$. Using (and) with this and $\phi \mid\!\sim \phi$ gives us $\phi \not\mid\!\sim \neg\phi \vee \gamma$. By (lle) we then get $(\theta \vee \phi) \wedge \phi \not\mid\!\sim \neg\phi \vee \gamma$ and using this together with $\theta \vee \phi \mid\!\sim \neg\phi \vee \gamma$ gives us the conclusion by (rmo2).

$\underline{\mathcal{R}3}$ $\dfrac{\phi \equiv \psi,\ \phi \not\mid\!\sim \gamma}{\psi \not\mid\!\sim \gamma}$

Immediate from (lle).

$\underline{\mathcal{R}4}$ $\dfrac{\theta \vee \phi \not\mid\!\sim \neg\phi \vee \eta}{\phi \not\mid\!\sim \eta}$

From $\theta \vee \phi \not\mid\!\sim \neg\phi \vee \eta$ and (lle) we get $(\theta \wedge \neg\phi) \vee \phi \not\mid\!\sim \neg\phi \vee \eta$. Using (or) together with $\theta \wedge \neg\phi \mid\!\sim \neg\phi \vee \eta$ (which itself follows from SC) this gives us $\phi \not\mid\!\sim \neg\phi \vee \eta$ and the conclusion then follows by (rwe).

$\underline{\mathcal{R}5}$ $\dfrac{(\phi \wedge \neg\gamma) \vee \theta \mid\!\sim \neg\eta,\ \phi \not\mid\!\sim \gamma}{\theta \vee \phi \not\mid\!\sim \eta}$

From $\phi \not\mid\!\sim \gamma$ we get $(\phi \wedge \gamma) \vee (\phi \wedge \neg\gamma) \not\mid\!\sim \gamma$ by (lle). Using this with (or) and $\phi \wedge \gamma \mid\!\sim \gamma$ (instance of SC) gives us $(\phi \wedge \neg\gamma) \not\mid\!\sim \gamma$. Then $(\phi \wedge \neg\gamma) \not\mid\!\sim \neg\eta \wedge \eta$ by (rwe) and so $(\theta \vee (\phi \wedge \neg\gamma)) \wedge (\phi \wedge \neg\gamma) \not\mid\!\sim \neg\eta \wedge \eta$ by (lle). From this we get $\theta \vee (\phi \wedge \neg\gamma) \not\mid\!\sim (\phi \wedge \neg\gamma) \wedge (\neg\eta \wedge \eta)$ by (A) and hence, using (rwe), $\theta \vee (\phi \wedge \neg\gamma) \not\mid\!\sim \neg\eta \wedge \eta$. Using LLE on $(\phi \wedge \neg\gamma) \vee \theta \mid\!\sim \neg\eta$ gives us $\theta \vee (\phi \wedge \neg\gamma) \mid\!\sim \neg\eta$ and hence, using (and) and $\theta \vee (\phi \wedge \neg\gamma) \not\mid\!\sim \neg\eta \wedge \eta$ we get $\theta \vee (\phi \wedge \neg\gamma) \not\mid\!\sim \eta$ and then $(\theta \vee \phi) \wedge (\theta \vee \neg\gamma) \not\mid\!\sim \eta$ by (lle). From $\phi \not\mid\!\sim \gamma$, using the derived rule of proof (B) from lemma 3, we get $\theta \vee \phi \not\mid\!\sim \neg(\theta \vee \neg\gamma)$ and using (rmo1) with $(\theta \vee \phi) \wedge (\theta \vee \neg\gamma) \not\mid\!\sim \eta$ gives us $\theta \vee \phi \not\mid\!\sim \eta$ as required.

$\underline{\mathcal{R}6}$ $\dfrac{(\phi \wedge \neg\gamma) \vee \theta \not\mid\!\sim \eta,\ \phi \not\mid\!\sim \gamma}{\theta \vee \phi \not\mid\!\sim \eta}$

From $\phi \not\mid\!\sim \gamma$ and (B) we get $\theta \vee \phi \not\mid\!\sim \neg(\theta \vee \neg\gamma)$ and from $(\phi \wedge \neg\gamma) \vee \theta \not\mid\!\sim \eta$ and (lle) we get $(\theta \vee \phi) \wedge (\theta \vee \neg\gamma) \not\mid\!\sim \eta$ and the required conclusion follows by (rmo1) and $\theta \vee \phi \not\mid\!\sim \neg(\theta \vee \neg\gamma)$.

$\underline{\mathcal{R}7}$ $\dfrac{\theta \vee \phi \mid\!\sim \eta,\ \phi \not\mid\!\sim \gamma}{\theta \vee (\phi \wedge \neg\gamma) \mid\!\sim \eta}$

From $\phi \not\mid\!\sim \gamma$ and (B) we get $\theta \vee \phi \not\mid\!\sim \neg(\theta \vee \neg\gamma)$ and this with RMO and $\theta \vee \phi \mid\!\sim \eta$ gives us $(\theta \vee \phi) \wedge (\theta \vee \neg\gamma) \mid\!\sim \eta$. The required conclusion then follows by LLE. $\Box$

As we showed in the remarks following theorem 5, the rule $\mathcal{R}1$ is derivable from $\mathcal{R}2 - \mathcal{R}7$ and so by the above theorem is also derivable in $GM^{\pm}$. We therefore have the following characterisation of the relations $\mid\!\sim^K_P$ and $\not\mid\!\sim^K_N$:

Theorem 7 *Let $K = K_P + K_N$ be consistent. Then $\theta \mid\!\sim^K_P \phi$ ($\theta \not\mid\!\sim^K_N \phi$) iff $\theta \mid\!\sim \phi$ ($\theta \not\mid\!\sim \phi$) is derivable from K in $GM^{\pm}$.*

Proof. Clearly all the rules in $GM^{\pm}$ are sound for rational consequence relations. Conversely if $\theta \mid\!\sim^K_P \phi$ then by theorem 4 $\theta \mid\!\sim \phi$ is derivable from K using the rules $\mathcal{R}1$ and $\mathcal{R}2$. By theorem 6 (and the above remarks) we can replace any such derivation by a derivation in $GM^{\pm}$. A similar argument applies to $\not\mid\!\sim^K_N$. $\Box$

References

[1] D. M. Gabbay, Theoretical foundations for non-monotonic reasoning in expert systems, in: K. R. Apt, ed, *Proceedings NATO Advanced Institute on Logics and Models of Concurrent Systems*, La Colle-sur-loup, France (Springer, Berlin, 1985) 439-457.

[2] D. Lehmann, What does a conditional knowledge base entail?, in: R. Brachman and H. J. Levesque, eds., *Proceedings First International Conference on Principles of Knowledge Representation and Reasoning,* Toronto, Ont. (1989).

[3] D. Lehmann and M. Magidor, What does a conditional knowledge base entail?, *Artificial Intelligence* **55** (1992) 1–60.

[4] Y. Shoham, A semantical approach to nonmonotonic logics, in: *Proceedings Logics in Computer Science*, Ithaca, NY (1987) 275–279.

[5] R. Reiter, A logic of default reasoning, *Artificial Intelligence* **13** (1980) 81–132.

[6] J. McCarthy, Circumscription: a form of non-monotonic reasoning, *Artificial Intelligence* **13** (1980) 27–39.

[7] D. McDermott and J. Doyle, Non-monotonic logic I, *Artificial Intelligence* **13** (1980) 41–72.

[8] R.C. Moore, Possible-world semantics for autoepistemic logic, in: *Proceedings Workshop on Non-Monotonic Reasoning,* New Paltz, NY, (1984) 369–401.

[9] S. Kraus, D. Lehmann and M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, *Artificial Intelligence* **44** (1990) 167–207.

[10] D. Makinson, General theory of cumulative inference, in: M. Reinfrank et al, eds., *Proceedings Second International Workshop on Non-Monotonic Reasoning*, Grassau, Germany (Springer lecture notes in computer science, 1988).

[11] P. Gärdenfors, *Knowledge in Flux* (M.I.T., 1988).

[12] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* (Morgan Kaufmann, Palo Alto, CA, 1988).

[13] J. P. Delgrande, A first-order logic for prototypical properties, *Artificial Intelligence* **33** (1987) 105–130.

[14] J. P. Delgrande, An approach to default reasoning based on a first-order conditional logic: Revised report, *Artificial Intelligence* **36** (1988) 63–90.

[15] D. Lehmann and M. Magidor, erratum, What does a conditional knowledge base entail?, *Artificial Intelligence*, 68 (1994), page 411.