MATH43032/63032, Nonmonotonic Logic

Jeff Paris,
Room 1.119,
jeff.paris@manchester.ac.uk

February 6, 2014
Motivating Examples

Some examples of ‘reasoning’, ‘follows’, or ‘inference’:

1. Young people are usually law abiding
   Joy riders are usually not law abiding
   Joy riders are usually young
   \[\therefore\] Young joy riders are usually not law abiding.

2. Young people are usually law abiding
   Joy riders are usually not law abiding
   Joy riders are usually young
   \[\therefore\] Young ornithologists are usually law abiding.

The purpose of this first part of the course is, in part, to study the relation of ‘follows’ as given in the above examples. Currently this is considered important for the construction of ‘intelligent’ computers capable of reasoning and reaching sensible conclusions (i.e. the same conclusions we would reach!)

We shall start with a quick revision of the Propositional or Sentential Calculus.
Chapter 1

Propositional or Sentential Calculus

For motivation consider the following example of reasoning

If someone is young then they are law abiding
If someone is a joy rider then they are not law abiding
∴ If someone is a joy rider then they are not young.

We would say that this conclusion ‘follows’ from the two given assumptions/premises – because whenever the premises are true then so is the conclusion. Indeed this ‘following’ has really nothing to do with ‘young’, ‘law abiding’ or ‘joy riding’, it is simply a result of the form of the premises and conclusion. To wit, in

If \( p \) then \( q \)
If \( r \) then not \( q \)
∴ If \( r \) then not \( p \).

this conclusion would follow from these premises no matter what declarative statements \( p, q, r \) stood for.

The study of this very basic notion of ‘follows’ is called the Propositional or Sentential Calculus. In order to do this we first isolate the important ingredients in examples like the one above i.e. the variables \( p, q, r \) standing for propositions which can be true or false, and (truth functional) connectives such as ‘if . . . then’ and ‘not’. Enough motivation, down to business.

1.1 Definitions

A language \( L \) for the sentential calculus is a non-empty set of letters, called propositional variables and usually denoted \( p, q, r \) etc. For this course we shall often assume that \( L \) is finite, for definiteness say \( L = \{p_1, p_2, \ldots, p_n\} \), unless otherwise indicated.

From \( L \) we form the set \( SL \) of sentences of \( L \) using the connectives \( \land, \lor, \rightarrow, \neg \) (where \( \land \)
CHAPTER 1. PROPOSITIONAL OR SENTENTIAL CALCULUS

stands for ‘and’, \( \lor \) stands for ‘or’, \( \rightarrow \) stands for ‘if . . . then’ (or ‘implies’) and \( \neg \) stands for ‘not’) and parentheses ‘(, )’ as follows:-

\[
SL_0 = L \\
SL_{n+1} = SL_n \cup \{(\theta \land \phi), (\theta \lor \phi), (\theta \rightarrow \phi), \neg \theta | \theta, \phi \in SL_n\} \\
SL = \bigcup_{n \in \mathbb{N}} SL_n
\]

Hence \( SL \subseteq \) set of words (i.e. finite strings of symbols) from the alphabet \( L \cup \{\land, \lor, \rightarrow, \neg, (, )\} \).

We use \( \theta, \phi, \psi \), etc. for members of \( SL \) (not necessarily distinct) and \( \Gamma, \Delta, \Omega \) etc. for subsets of \( SL \), possibly empty.

For \( \theta \in SL \) let \(|\theta|\), the length of \( \theta \), be the number of symbols in \( \theta \) viewed as a word from the alphabet \( L \cup \{\land, \lor, \rightarrow, \neg, (, )\} \). E.g. \( |\neg (p_1 \land p_2)| = 6 \).

Note. A general method for showing that every sentence has a property \( P \) is to show that the propositional variables have \( P \), i.e. \( SL_0 \subseteq P \), and if \( \theta, \phi \) have \( P \) so do \( \neg \theta, (\theta \land \phi), (\theta \lor \phi), (\theta \rightarrow \phi) \), which gives that \( SL_n \subseteq P \implies SL_{n+1} \subseteq P \). For then, by induction, \( \forall n(SL_n \subseteq P) \) so \( SL \subseteq P \). We may refer to this method as ‘proof by induction on the \( SL_n \)’. An alternative to this method is to show by induction on \(|\theta|\) that every \( \theta \in SL \) has property \( P \). More precisely one assumes that \( \phi \in P \) for every sentence \( \phi \) with \(|\phi| < |\theta|\) and then shows that \( \theta \in P \). [Note that this assumption is trivially true if \( \theta \) has the smallest possible length, 1, i.e. \( \theta \) is a propositional variable, so that gets the induction ‘started’.]

Convention. We may drop outer parentheses when considering a particular sentence in isolation. E.g. write \( \theta \rightarrow \phi \) for \( (\theta \rightarrow \phi) \).

The reason for parentheses is the following.

Theorem 1.1 (The Unique Readability Theorem). Let \( \theta \in SL \). Then exactly one of the following holds:

\[
\begin{align*}
\theta &\in L \\
\theta &= (\phi_1 \land \phi_2) \quad \text{for some} \ \phi_1, \phi_2 \in SL \\
\theta &= (\phi_1 \lor \phi_2) \quad \text{for some} \ \phi_1, \phi_2 \in SL \\
\theta &= (\phi_1 \rightarrow \phi_2) \quad \text{for some} \ \phi_1, \phi_2 \in SL \\
\theta &= \neg \phi_1 \quad \text{for some} \ \phi_1 \in SL
\end{align*}
\]

Furthermore, in each case \( \phi_1, \phi_2 \) are unique and consequently if \( \theta \in SL_{n+1} \) then \( \phi_1, \phi_2 \in SL_n \).

Proof. The following follow by induction on \( n \), where \( \theta \in SL_n \):


1.1. DEFINITIONS

(i) $\theta \in SL_0$, or $\theta$ begins with ‘(’ and ends with ‘)’, or $\theta = \neg \phi$ for some unique $\phi$ and $\phi \in SL_{n-1}$.

(ii) If $r(\theta)$ is the number of ‘(’ in $\theta$, $l(\theta)$ is the number of ‘)’ in $\theta$ then $r(\theta) = l(\theta)$.

(iii) If $\theta = \gamma \ast \tau$ where $\ast$ is one of $\land, \lor, \rightarrow$ and $\gamma$ and $\tau$ are words then $r(\gamma) < l(\gamma)$.

The result now follows. For suppose, say, $\theta = (\gamma \ast \tau) = (\alpha \cdot \beta)$ with $\gamma, \tau, \alpha, \beta \in SL$, $\ast, \cdot \in \{\land, \lor, \rightarrow\}$ and $|\gamma| < |\alpha|$, say $\alpha = \gamma \ast \nu$. Then $r(\gamma) < l(\gamma)$ by (iii) whilst $r(\gamma) = l(\gamma)$ by (ii).

Definition. A valuation (assignment) on a language $L$ is a function $V : L \rightarrow \{0, 1\}$.

Here the (truth) value 1 corresponds to ‘true’ and the value 0 to ‘false’. As given above a valuation $V$ is only defined on $L = SL_0$. We extend the domain of $V$ from $SL_n$ to $SL_{n+1}$ for $n = 0, 1, 2, \ldots$ (and hence to all of $SL$) by setting, for $\theta \in SL_{n+1} - SL_n$, $\phi, \psi \in SL_n$,

- if $\theta = \neg \phi$ then $V(\theta) = 1$ if $V(\phi) = 0$ and 0 otherwise,
- if $\theta = (\phi \land \psi)$ then $V(\theta) = 1$ if $V(\phi) = V(\psi) = 1$ and 0 otherwise,
- if $\theta = (\phi \lor \psi)$ then $V(\theta) = 0$ if $V(\phi) = V(\psi) = 0$ and 1 otherwise,
- if $\theta = (\phi \rightarrow \psi)$ then $V(\theta) = 0$ if $V(\phi) = 1$, $V(\psi) = 0$ and 1 otherwise.

Notice that by unique readability $V(\theta)$ gets exactly one value.

Write $V(\Gamma) = 1$ if $\forall \phi \in \Gamma, V(\phi) = 1$. (So $not(V(\Gamma) = 1)$ iff $\exists \phi \in \Gamma, V(\phi) = 0$.)

Definition. $\Gamma$ logically implies $\theta$ (or $\theta$ is a logical consequence of $\Gamma$), written $\Gamma \models \theta$, if for all valuations $V$ on $L$, if $V(\Gamma) = 1$ then $V(\theta) = 1$.

$\theta$ is a tautology$^1$ if $\emptyset \models \theta$, i.e. $V(\theta) = 1$ for all valuations $V$ on $L$, since, vacuously, $V(\phi) = 1$ for all $\phi \in \emptyset$. We write $\vdash \theta$ for $\emptyset \models \theta$.

Exercise 1.2. This definition appears to depend on $L$. Show that this is not the case.

Remark. $\vdash$ is a formalization of “follows”. We have derived it by giving an interpretation or meaning or semantics to $SL$ (in this case, a valuation) and saying that $\theta$ follows from $\Gamma$ if for every interpretation, if everything in $\Gamma$ is true then so is $\theta$.

Definition. $\Gamma$ is satisfiable if there is a valuation $V$ such that $V(\Gamma) = 1$. In such a case we say that the valuation $V$ satisfies $\Gamma$. (With such statements it is implicit that $\Gamma \subseteq SL$ and $V$ is a valuation on $L$.)

Lemma 1.3. (i) $\Gamma \models \theta \iff \Gamma \cup \{\neg \theta\}$ is not satisfiable.

(ii) $\Gamma, \phi \models \theta \iff \Gamma \models \phi \rightarrow \theta$.

(iii) $\Gamma, \phi, \psi \models \theta \iff \Gamma, (\phi \land \psi) \models \theta$.

(We write $\Gamma, \phi$ as a shorthand for $\Gamma \cup \{\phi\}$ etc.)

$^1$Some books use instead the confusing term theorem.
Proof.

(i) \( \Gamma \models \theta \iff \) for all valuations \( V \), if \( V(\Gamma) = 1 \) then \( V(\theta) = 1 \)
\( \iff \) there is no valuation \( V \) such that \( V(\Gamma) = 1 \) but \( V(\theta) = 0 \)
\( \iff \) there is no valuation \( V \) such that \( V(\Gamma) = 1 \) but \( V(\neg \theta) = 1 \)
\( \iff \) \( \Gamma \cup \{ \neg \theta \} \) is not satisfiable.

(ii) \( \Gamma, \phi \models \theta \iff \) for all valuations \( V \), if \( V(\Gamma) = V(\phi) = 1 \) then \( V(\theta) = 1 \)
\( \iff \) there is no valuation \( V \) such that \( V(\Gamma) = V(\phi) = 1 \)
\( \text{and } V(\theta) = 0 \)
\( \iff \) there is no valuation \( V \) such that \( V(\Gamma) = 1 \) and \( V(\phi \rightarrow \theta) = 0 \)
\( \iff \) \( \Gamma \models \phi \rightarrow \theta \).

(iii) Similar, exercise.

\[ \square \]

Remark. For \( \Gamma \) finite we may check if \( \Gamma \models \theta \) by verifying that for each of \( 2^{|L|} \) valuations \( V \) on \( L \), if \( V(\psi) = 1 \) for all \( \psi \in \Gamma \) then \( V(\theta) = 1 \), where \( L \) is some (by 1.2 it doesn’t matter which) finite language such that \( \Gamma \subseteq SL \) and \( \theta \in SL \). Various tricks may be used to shorten this computation in special cases, however there is no known algorithm for checking \( \Gamma \models \theta \) which isn’t exponential in the worst case.

Definition. We say sentences \( \theta \) and \( \phi \) are logically equivalent, written \( \theta \equiv \phi \), if for all valuations \( V \), \( V(\theta) = V(\phi) \) (i.e. \( \theta, \phi \) say the same thing), or, equivalently \( \theta \models \phi \) and \( \phi \models \theta \), or \( \models (\theta \iff \phi) \) where \( \theta \iff \phi \) is short for \( (\theta \rightarrow \phi) \land (\phi \rightarrow \theta) \), or for all valuations \( V \), \( V(\theta) = 1(0) \iff V(\phi) = 1(0) \).

Clearly this is an equivalence relation on \( SL \) (i.e. it is reflexive, symmetric and transitive). Notice also by 1.2 that \( \theta \equiv \phi \) is independent of the overlying language.

Since for a valuation \( V \), \( V(\theta \land \phi), V(\theta \lor \phi), V(\theta \rightarrow \phi), V(\neg \theta) \) are functions of \( V(\theta) \) and \( V(\phi) \), it follows that if \( \theta_1 \equiv \theta_2, \phi_1 \equiv \phi_2 \) then \( \theta_1 \land \phi_1 \equiv \theta_2 \land \phi_2, \theta_1 \lor \phi_1 \equiv \theta_2 \lor \phi_2, \theta_1 \rightarrow \phi_1 \equiv \theta_2 \rightarrow \phi_2, \neg \theta_1 \equiv \neg \theta_2 \).

As an abbreviation we write
\[ \bigwedge_{i=1}^{n} \theta_i \text{ or } \theta_1 \land \theta_2 \land \ldots \land \theta_n \]
for
\[ ((\ldots (((\theta_1 \land \theta_2) \land \theta_3) \land \ldots) \land \theta_n). \]

Similarly we write
\[ \bigvee_{i=1}^{n} \theta_i \text{ or } \theta_1 \lor \theta_2 \lor \ldots \lor \theta_n \]
for 
\[(\ldots((\theta_1 \lor \theta_2) \lor \theta_3) \lor \ldots) \lor \theta_n).\]

By induction on \(n\), for any valuation \(V\),
\[
V(\bigwedge_{i=1}^{n} \theta_i) = 1 \iff \forall 1 \leq i \leq n, V(\theta_i) = 1,
\]
\[
V(\bigvee_{i=1}^{n} \theta_i) = 0 \iff \forall 1 \leq i \leq n, V(\theta_i) = 0.
\]

It is convenient to define the empty conjunction,
\[
\bigwedge_{i=1}^{0} \theta_i,
\]
and the empty disjunction
\[
\bigvee_{i=1}^{0} \theta_i,
\]
to be \(p \lor \neg p\), \(p \land \neg p\) respectively, for some \(p \in L\). Notice that with this definition the above equivalences still hold.

It follows that, up to logical equivalence, \(\bigwedge_{i=1}^{n} \theta_i\) and \(\bigvee_{i=1}^{n} \theta_i\) are independent of the order of the \(\theta_i\) (and of the insertion of repeats). Using this, for \(R\) a finite (possibly empty) subset of \(SL\) we shall write \(\bigvee R\) for the disjunction and \(\bigwedge R\) for the conjunction of the sentences of \(R\) in some order. Since in practice we shall only be interested in such disjunctions and conjunctions up to logical equivalence, we shall leave the precise ordering unspecified.

### 1.2 Atoms

Henceforth, until otherwise indicated, suppose \(L = \{p_1, p_2, \ldots, p_n\}\). For \(p \in L\) let \(p^1 = p\) and \(p^0 = \neg p\), so for a valuation \(V\) (on \(L\)), \(\varepsilon \in \{0,1\}\),
\[
V(p^\varepsilon) = 1 \iff \begin{cases} \varepsilon = 1 \text{ and } V(p) = 1 \\ \text{or} \\ \varepsilon = 0 \text{ and } V(\neg p) = 1 \end{cases} \iff V(p) = \varepsilon.
\]

Let \(\alpha_1, \alpha_2, \ldots, \alpha_J\), where \(J = 2^n\), run through the \(2^n\) sentences of \(L\) of the form
\[
p_{1}^{\varepsilon_{1}} \land p_{2}^{\varepsilon_{2}} \land \ldots \land p_{n}^{\varepsilon_{n}}
\]
where \(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n \in \{0,1\}\). We call these \(\alpha_i\) atoms of \(L\) and denote the set of these atoms by \(\text{At}^L\). Notice that atoms only make sense for a finite language \(L\).
Then for $\alpha = p_1^{\varepsilon_1} \land p_2^{\varepsilon_2} \land \ldots \land p_n^{\varepsilon_n} \in \text{At}_L$, and a valuation $V$ on $L$,

$$V(\alpha) = 1 \iff \forall 1 \leq i \leq n, V(p_i^{\varepsilon_i}) = 1 \iff \forall 1 \leq i \leq n, V(p_i) = \varepsilon_i.$$ 

From this it follows that given this atom $\alpha$ there is a unique valuation $V$ such that $V(\alpha) = 1$, namely that valuation $V_\alpha$ such that $V_\alpha(p_i) = \varepsilon_i$ for $1 \leq i \leq n$.

Conversely for a valuation $V$ there is a unique atom $\alpha \in \text{At}_L$ such that $V(\alpha) = 1$, namely $\alpha = p_1^{V(p_1)} \land p_2^{V(p_2)} \land \ldots \land p_n^{V(p_n)}$.

So $V = V_\alpha$ for this atom $\alpha$.

For $\theta \in SL$ let

$$S_\theta = \{\alpha \in \text{At}_L \mid \alpha \models \theta\} = \{\alpha \in \text{At}_L \mid V_\alpha(\theta) = 1\},$$

since

$$\alpha \models \theta \iff V(\theta) = 1 \text{ for all valuations such that } V(\alpha) = 1 \iff V_\alpha(\theta) = 1 \text{ since } V_\alpha \text{ is the only valuation } V \text{ such that } V(\alpha) = 1.$$ 

**Theorem 1.4.** For $\theta, \phi \in SL$

(i) $\theta \equiv \bigvee S_\theta$.

(ii) $S_\theta$ is the unique subset $S$ of $\text{At}_L$ such that $\theta \equiv \bigvee S$.

(iii) For $R \subseteq \text{At}_L$, $S_{\bigvee R} = R$.

(iv) $\theta \equiv \phi \iff S_\theta = S_\phi$.

(v) $S_{\theta \land \phi} = S_\theta \cap S_\phi$.

(vi) $S_{\theta \lor \phi} = S_\theta \cup S_\phi$.

(vii) $S_{\neg \theta} = \text{At}_L - S_\theta$.

(viii) $\models \theta \iff S_\theta = \text{At}_L$.

(ix) $\models \theta \iff S_\theta \subseteq S_\phi$.

**Proof.** (i) Let $V$ be a valuation, say $V = V_\alpha$ where $\alpha \in \text{At}_L$. Then

$$V_\alpha(\theta) = 1 \iff \alpha \in S_\theta \iff V_\alpha(\beta) = 1 \text{ for some } \beta \in S_\theta \iff V_\alpha(\bigvee S_\theta) = 1 \therefore \theta \equiv \bigvee S_\theta.$$
2. ATOMS

(ii) Suppose on the contrary that $S, S' \subseteq At^L$ and $\bigvee S \equiv \theta \equiv \bigvee S'$ but $S \neq S'$; say without loss of generality $\alpha \in S - S'$. Then $V_\alpha(\bigvee S) = 1$, $V_\alpha(\bigvee S') = 0$ (since $\beta \in S' \Rightarrow \beta \neq \alpha \Rightarrow V_\alpha(\beta) = 0$) so $\bigvee S \not\equiv \bigvee S'$. Contradiction.

(iii) This follows from (ii) since $\bigvee S \bigvee R \equiv \bigvee R$.

(iv) $\theta \equiv \phi \iff \bigvee S_\theta \equiv \bigvee S_\phi$ since $\equiv$ is an equivalence relation.

(v) $\alpha \in S_{\theta \land \phi} \iff V_\alpha(\theta \land \phi) = 1$
    $\iff V_\alpha(\theta) = V_\alpha(\phi) = 1$
    $\iff \alpha \in S_\theta$ and $\alpha \in S_\phi$

\[ S_{\theta \land \phi} = S_\theta \cap S_\phi. \]

(vi) $\alpha \in S_{\theta \lor \phi} \iff V_\alpha(\theta \lor \phi) = 1$
    $\iff V_\alpha(\theta) = 1$ or $V_\alpha(\phi) = 1$
    $\iff \alpha \in S_\theta$ or $\alpha \in S_\phi$

\[ S_{\theta \lor \phi} = S_\theta \cup S_\phi. \]

(vii) $\alpha \in S_{\neg \theta} \iff V_\alpha(\neg \theta) = 1$
    $\iff V_\alpha(\theta) = 0$
    $\iff \alpha \notin S_\theta$

\[ S_{\neg \theta} = At^L - S_\theta. \]

(viii) $\vdash \theta \iff \forall V V(\theta) = 1$
    $\iff \forall \alpha \in At^L V_\alpha(\theta) = 1$ since every valuation is $V_\alpha$ for some $\alpha$
    $\iff At^L = S_\theta$.

(ix) $\theta \not\vdash \phi \iff V(\theta) = 1, V(\phi) = 0$, for some valuation $V$
    $\iff V_\alpha(\theta) = 1, V_\alpha(\phi) = 0$ for some $\alpha \in At^L$
    $\iff \alpha \in S_\theta$ and $\alpha \notin S_\phi$ for some $\alpha \in At^L$
    $\iff S_\theta \not\subseteq S_\phi$.

\[ \square \]

Remark. A corollary to this result is that there are no ‘missing’ (truth functional) connectives. That is, suppose $*$ is a new connective of $m$ arguments for which $V^*(\theta_1, \ldots, \theta_m)$ is a fixed function of $V(\theta_1), \ldots, V(\theta_m)$. Then any sentence written using $*$ is, by the above proof of (i), logically equivalent to $\bigvee S$ for some $S \subseteq At^L$. Hence adding new connectives does not increase expressive power, $\land, \lor$ and $\neg$ already provide all that is possible. (Indeed just $\land$ and $\neg$, or $\lor$ and $\neg$, actually suffice.)

Theorem 1.5 (The Disjunctive Normal Form Theorem). Every $\theta \in SL$ is logically equivalent to a sentence in disjunctive normal form (DNF) i.e. in the form

$$\bigvee_{i=1}^{m} \bigwedge_{j=1}^{k_i} \gamma_{ij}$$

where the $\gamma_{ij}$ are literals, that is propositional variables or negations of propositional variables.
Proof. Let $\theta \in SL$ and let $L' \subseteq L$ be the smallest language containing every propositional variable in $\theta$. Then, working in $SL'$, $\bigvee S_\theta$ has this DNF and $\theta \equiv \bigvee S_\theta$. By our choice of $L'$ we have also ensured that each propositional variable mentioned in this DNF actually occurs in $\theta$. \hfill \square

**Theorem 1.6** (The Conjunctive Normal Form Theorem). Every $\theta \in SL$ is logically equivalent to a sentence in conjunctive normal form i.e. in the form

$$m \bigwedge_{i=1}^{k_i} \bigvee_{j=1}^{\gamma_{ij}}$$

where the $\gamma_{ij}$ are literals.

Proof. By the DNFT for $\neg \theta$, $\neg \theta \equiv \bigvee_{i=1}^{m} \bigwedge_{j=1}^{k_i} \gamma_{ij}$ for some literals $\gamma_{ij}$. So, using Q4 of Examples 1,

$$\theta \equiv \neg \neg \theta \equiv \bigvee_{i=1}^{m} \bigwedge_{j=1}^{k_i} \neg \gamma_{ij} \equiv \bigwedge_{i=1}^{m} \bigvee_{j=1}^{k_i} \delta_{ij},$$

where

$$\delta_{ij} = \begin{cases} 
\neg p & \text{if } \gamma_{ij} = p \in L \\
 p & \text{if } \gamma_{ij} = \neg p \text{ for some } p \in L
\end{cases}$$

\hfill \square
Chapter 2

Nonmonotonic Reasoning

Much of our everyday reasoning concerns statements of the form

\[ \text{If } \theta \text{ then usually/generally/normally } \phi \]  

(2.1)

rather than the uncompromising

\[ \text{If } \theta \text{ then } \phi \]

For example in trying to decide how long it will take me to get from the ATB to Piccadilly Gardens I might think

- I usually wait less than 10 minutes for a bus at the Precinct stop
- The bus normally takes less than 15 minutes to reach Piccadilly Gardens

\[ \therefore \text{ Under normal circumstances I can get to Piccadilly Gardens within 25 minutes} \]

rather than

- I'll wait less than 10 minutes for a bus at the precinct stop
- The bus does take less than 15 minutes to reach Piccadilly Gardens

\[ \therefore \text{ I can get to Piccadilly Gardens within 25 minutes} \]

Because of the everyday importance of such reasoning, which ultimately we would like computers to replicate, many attempts were make in the 1980’s to formalize the semantics and syntax being used here, just as was done half a century earlier for the Propositional and Predicate Calculi. The attempt we shall study here is (I think) by far the most believable and is referred to as Rational Consequence.

Assertions of the form (2.1) are nowadays usually referred to as conditionals and abbreviated to \( \theta \models \phi \) (said \( \theta \) snake \( \phi \), or for this course \( \phi \) is a rational consequence of \( \theta \)).
So what we wish to do is try to formalize what it means for some \textit{conditional} \( \theta \mid \neg \phi \) to ‘follow’ from some other conditionals \( \theta_1 \mid \neg \phi_1, \ldots, \theta_n \mid \neg \phi_n \) and maybe also some negative conditionals \( \psi_1 \mid \neg \zeta_1, \ldots, \psi_m \mid \neg \zeta_m \) where \( \psi \mid \neg \zeta \) stands for \(~ It is not the case that is \).

One idea here, as with the rules of proof of the Predicate Calculus, is to first write down some simple rules, or conditions, and an axiom which we agree such such a notion of ‘follows’ should satisfy and then say that that a conditional ‘follows’ from some other set of conditions just if it \textit{can be shown on the basis of these rules}. In this case these rules of the form\(^1\)

\[
\frac{\theta_1 \mid \neg \phi_1, \ldots, \theta_n \mid \neg \phi_n, \psi_1 \mid \neg \zeta_1, \ldots, \psi_m \mid \neg \zeta_m}{\theta \mid \neg \phi}
\]

which we might think of as saying

\textit{If you think that if }\theta_i\textit{ then usually }\phi_i\textit{ for }i = 1, \ldots, n\textit{ and you think that it is not the case that if }\psi_j\textit{ then usually }\zeta_j\textit{ for }j = 1, \ldots, m\textit{ then you should think that if }\theta\textit{ then usually }\phi\textit{.}

We give these rules with a single sentence on the l.h.s. because that is the convention in the area. As will be clear, we could view \( \theta_1, \ldots, \theta_m \mid \neg \phi \) as shorthand for \( \theta_1 \land \ldots \land \theta_m \mid \neg \phi \) (so, for finite l.h.s., nothing is lost by doing this) and \( \vdash \phi \) as shorthand for \( \eta \vdash \phi \) where \( \eta \) is any tautology.

\(^1\)And sometimes some further side conditions like \( \xi \models \chi \).
2.1 The Gabbay-Makinson (GM) Conditions or Rules

left logical equivalence \( \theta \equiv \phi \quad \theta \vdash \psi \quad \phi \vdash \psi \quad \text{LLE} \)

right weakening \( \theta \vdash \phi \quad \phi \vdash \psi \quad \theta \vdash \psi \quad \text{RWE} \)

and on right \( \theta \vdash \phi \quad \theta \vdash \psi \quad \theta \vdash \phi \land \psi \quad \text{AND} \)

disjunction, or left \( \theta \vdash \psi \quad \phi \vdash \psi \quad \theta \lor \phi \vdash \psi \quad \text{DIS} \)

cautious monotonicity \( \theta \vdash \psi \quad \theta \vdash \phi \quad \theta \land \phi \vdash \psi \quad \text{CMO} \)

rational monotonicity \( \theta \vdash \psi \quad \theta \not\vdash \neg \phi \quad \theta \land \phi \vdash \psi \quad \text{RMO} \)

[Right now you might feel that there could be some more such rules that are not included here but subsequent events should convince you that is not the case.]

We also have a rule without any premises, i.e. an axiom

\[ \text{Reflexivity} \quad \theta \vdash \theta \quad \text{REF} \]

The attribute ‘nonmonotonic’ here comes from the fact that we do not have the MON rule

\[ \frac{\theta \vdash \phi}{\theta \land \psi \vdash \phi} \]

even as a derived rule. Indeed is now undesirable since for example we might well accept “If the cake contains butter then it normally tastes good” but not “If the cake contains butter and soap then it normally tastes good”.

By combining these rules we can obtain some useful derived rules:

\[ \text{Supraclassicality} \quad \theta \models \phi \quad \text{SCL} \]

\[ \frac{\theta \models \phi}{\theta \vdash \phi} \quad \frac{\theta \vdash \theta \text{ REF}}{\theta \vdash \phi} \quad \text{RWE} \]
Conditionalisation

\[
\frac{\theta \land \phi \vdash \psi}{\theta \vdash (\phi \rightarrow \psi)} \text{ CON}
\]

\[
\begin{align*}
\theta \equiv (\theta \land \phi) \lor (\theta \land \neg \phi) \\
\theta \land \phi \vdash \phi \rightarrow \psi & \quad \psi \vdash \phi \rightarrow \psi & \quad \theta \land \neg \phi \vdash \phi \rightarrow \psi & \quad \theta \land \phi \lor (\theta \land \neg \phi) \vdash \phi \rightarrow \psi \\
\theta \land \phi \vdash \phi \rightarrow \psi & \quad \theta \land \neg \phi \vdash \phi \rightarrow \psi & \quad (\theta \land \phi) \lor (\theta \land \neg \phi) \vdash \phi \rightarrow \psi & \quad \theta \lor \phi \rightarrow \psi
\end{align*}
\]

Cautious Cut

\[
\frac{\theta \vdash \phi \quad \theta \land \phi \vdash \psi}{\theta \vdash \psi} \text{ CC}
\]

\[
\begin{align*}
\theta \land \phi \vdash \psi & \quad \theta \lor \phi \rightarrow \psi & \quad CON \\
\theta \vdash \phi \rightarrow \psi & \quad \theta \vdash \phi & \quad \text{ AND} \\
\theta \vdash \phi \land (\phi \rightarrow \psi) & \quad \phi \land (\phi \rightarrow \psi) \vdash \psi & \quad \text{ RWE} \\
\theta \lor \phi \rightarrow \psi & \quad \theta \lor \phi \rightarrow \psi & \quad \theta \lor \psi
\end{align*}
\]

One might suppose at this point that we would proceed as in the Propositional and Predicate Calculi and introduce a notion of proof, subsequently proving a completeness theorem linking provability with a semantic notion of logical consequence, i.e. ‘truth preservation in all interpretations’. Unfortunately that does not work in this case, essentially because of the odd rule RMO, and instead we take a different approach.

A binary relation \( \vdash \) on \( SL \) is a rational consequence relation (rcr) if it satisfies the GM rules\(^2\) together with the axiom REF.\(^3\)

Notice that the logical consequence relation \( \models \) from the Propositional Calculus satisfies all of these and is hence a rational consequence relation (rcr). Another example of an rcr is the trivial rcr where \( \theta \vdash \phi \) holds for all \( \theta, \phi \in SL \).

We can now say that the conclusion \( \theta \vdash \phi \) follows from the premises

\[
\theta_1 \vdash \phi_1, \theta_2 \vdash \phi_2, \ldots, \theta_n \vdash \phi_n, \psi_1 \not\vdash \zeta_1, \ldots, \psi_m \not\vdash \zeta_m
\]

if every rational consequence relation which satisfies these premises must also satisfy this conclusion.\(^4\)

---

\(^2\)As usual we write \( \theta \vdash \phi \) rather than \( \langle \theta, \phi \rangle \in \vdash \).

\(^3\)Notice that \( \vdash \) is being used with two meanings, a purely syntactic one in the statement of the GM rules and to stand for a rational consequence relation, rcr.

\(^4\)More generally we could add to the premises also some logical consequence conditions such as \( \xi_1 \models \chi_1, \ldots, \xi_m \models \chi_k \).
2.2 An Example

Young people are usually law abiding \((Y \vdash L)\)
Joy riders are usually not law abiding \((J \vdash \neg L)\)
\[\therefore \text{Joy riders are usually young } (J \vdash Y)\]
\[\therefore \text{Young joy riders are usually not law abiding } (Y \land J \vdash \neg L)\].

Can we make sense of this? In what sense does the conclusion ‘follow’? An answer is: if we interpret the hypotheses as \(Y \vdash L\), \(J \vdash \neg L\), \(J \vdash Y\) where \(\vdash\) is a rational consequence relation, then \(Y \land J \vdash \neg L\) (i.e. the corresponding interpretation of the conclusion) holds. Indeed this is immediate, since \(Y \land J \vdash \neg L\) follows directly from \(J \vdash \neg L\), \(J \vdash Y\) by CMO and LLE.

Of course this example simply provides a single piece of evidence in favor of human consequence being rational. However many similar examples exist to give further support. Just as importantly, rational consequence does not (seem to) give conclusions which we feel should not follow. For example, we cannot conclude also that ‘Young joyriders are usually law abiding’, i.e. \(Y \land J \vdash L\).

But how can we show that some conclusion does not follow from certain premises? One way, following our experience in the Propositional/Predicate Calculus, would be to find a rcr which satisfied the premises but not the conclusion. For this we need to have a rich source of rcr’s to hand (right now we only have \(\models\)), a matter we now turn to.

2.3 Generating rational consequence relations

The idea behind the family of rational consequence relations that we introduce is to try to give a meaning, or semantics, to assertions such as

\[
\text{If } \theta \text{ then usually } \phi
\]

One such idea is to interpret this as meaning

‘For all the most usual (or most expected or most normal) worlds in which \(\theta\) is true, \(\phi\) is also true.’ ⋆

Since we have seen that we can think of valuations on \(L\), alternatively atoms\(^5\) of \(L\), as worlds this idea paints a picture of sets \(s_1, s_2, \ldots s_m\) of atoms with those atoms/worlds in \(s_1\) being the most usual/normal/etc., those in \(s_2\) being the next most usual, and so on down to the atoms/worlds in \(s_m\) being the least usual (but still possible). And maybe not

\(^5\text{Recall that unless otherwise stated we are working in the language } L = \{p_1, p_2, \ldots, p_n\}.$
all atoms/worlds will be included in $s_1 \cup s_2 \cup \ldots \cup s_m$, the ones missing being considered totally impossible. With this picture we would interpret $\phi$ as: Whenever $i$ is minimal such that for some $\alpha \in s_i$, $V_\alpha(\theta) = 1$ then for all $\alpha \in s_i$, if $V_\alpha(\theta) = 1$ then $V_\alpha(\phi) = 1$. Writing this relation as $\vdash_s$, where $s = s_1, s_2, \ldots, s_m$ and recalling that $V_\alpha(\theta) = 1$ is equivalent to $\alpha \in S_\theta$, this leads to the formal definition:

For $s = s_1, s_2, \ldots, s_m \subseteq At^L$,

\[ \theta \vdash_s \phi \iff \exists i \ s_i \cap S_\theta \neq \emptyset \text{ or for the least such } i, s_i \cap S_\theta \subseteq S_\phi. \]

**Example**

Let $L = \{p, q\}$, $\alpha_1 = p \land q$, $\alpha_2 = p \land \neg q$, $\alpha_3 = \neg p \land q$, $\alpha_4 = \neg p \land \neg q$. Let $s = s_1, s_2$ where $s_1 = \{\alpha_2, \alpha_4\}$, $s_2 = \{\alpha_1\}$. Then

- $S_p = \{\alpha_1, \alpha_2\}$, 1 is minimal such that $s_1 \cap S_p \neq \emptyset$ and $s_1 \cap S_p = \{\alpha_2\}$,
- $S_q = \{\alpha_1, \alpha_3\}$, 2 is minimal such that $s_2 \cap S_p \neq \emptyset$ and $s_2 \cap S_q = \{\alpha_1\}$,
- $S_{\neg q} = \{\alpha_2, \alpha_4\}$, 1 is minimal such that $s_1 \cap S_p \neq \emptyset$ and $s_1 \cap S_{\neg q} = \{\alpha_2, \alpha_4\}$.

Hence:

- $p \vdash_s \neg q$ since $s_1 \cap S_p = \{\alpha_2\} \subseteq S_{\neg q}$,
- $q \vdash_s p$ since $s_2 \cap S_q = \{\alpha_1\} \subseteq S_p$,
- $\neg p \land q \vdash_s p$ since $S_{\neg p \land q} = \{\alpha_3\}$ and $s_1 \cap \{\alpha_3\} = s_2 \cap \{\alpha_3\} = \emptyset$.

**Theorem 2.1.** For $s = s_1, s_2, \ldots, s_m \subseteq At^L$, $\vdash_s$ is a rational consequence relation.

**Proof.** (We write $\vdash$ for $\vdash_s$). We need to check that REF + LLE–RMO hold. We check each case separately:

**REF** Clearly either $\forall i \ s_i \cap S_\theta = \emptyset$ or for the least $i$ such that $s_i \cap S_\theta \neq \emptyset$, $s_i \cap S_\theta \subseteq S_\theta$. So $\theta \vdash \theta$ either way.

**LLE** Suppose $\theta \equiv \phi$, $\theta \vdash \psi$. Note that the definition of $\theta \vdash \psi$ only depends on $S_\theta$ and $S_\psi$. Since $S_\theta = S_\phi$ (as $\theta \equiv \phi$), we get $\phi \vdash \psi$.

**RWE** Suppose $\theta \vdash \phi$, $\phi \vdash \psi$. If $\forall i \ s_i \cap S_\theta = \emptyset$ then $\theta \vdash \psi$. Otherwise, let $i$ be minimal such that $s_i \cap S_\theta \neq \emptyset$. Then $s_i \cap S_\theta \subseteq S_\phi$ (since $\theta \vdash \phi$) $\subseteq S_\psi$ (since $\phi \vdash \psi \implies S_\phi \subseteq S_\psi$). So $\theta \vdash \psi$ as required.

**AND** Suppose $\theta \vdash \phi$ and $\theta \vdash \psi$. If $\forall i \ s_i \cap S_\theta = \emptyset$ then $\theta \vdash \phi \land \psi$. Otherwise let $i$ be minimal such that $s_i \cap S_\theta \neq \emptyset$. Therefore $s_i \cap S_\theta \subseteq S_\phi, S_\psi$ (by $\theta \vdash \phi$, $\theta \vdash \psi$) so $s_i \cap S_\theta \subseteq S_\phi \cap S_\psi = S_{\phi \land \psi}$. Hence $\theta \vdash \phi \land \psi$ as required.
2.3. GENERATING RATIONAL CONSEQUENCE RELATIONS

DIS Suppose \( \theta \vdash \psi \) and \( \phi \vdash \psi \). We have to show \( \theta \lor \phi \vdash \psi \) and we may assume that there is a smallest \( i \) with \( s_i \cap S_{\theta \lor \phi} \neq \emptyset \). Recall that \( S_{\theta \lor \phi} = S_\theta \cup S_\phi \), hence it is enough to show that \( s_i \cap S_\theta \subseteq S_\psi \) (by symmetry, this also shows \( s_i \cap S_\phi \subseteq S_\psi \)).

If \( s_i \cap S_\theta = \emptyset \) there is nothing to do. If \( s_i \cap S_\theta \neq \emptyset \), then the index \( i \) is smallest with \( s_i \cap S_\theta \neq \emptyset \) and from \( \theta \vdash \psi \) we get \( s_i \cap S_\theta \subseteq S_\psi \).

CMO Suppose \( \theta \vdash \phi \) and \( \theta \vdash \psi \). We have to show \( \theta \land \phi \vdash \psi \) and we may assume that there is a smallest \( i \) with \( s_i \cap S_{\theta \land \phi} \neq \emptyset \). Recall that \( S_{\theta \land \phi} = S_\theta \cap S_\phi \). Since \( \theta \vdash \phi \), the index \( i \) is also smallest with \( s_i \cap S_\theta \neq \emptyset \). Hence from \( \theta \vdash \psi \) we get \( s_i \cap S_\theta \subseteq S_\psi \), which implies \( s_i \cap S_{\theta \land \phi} \subseteq S_\psi \).

RMO Assume \( \theta \vdash \psi \) and \( \theta \not\vdash \phi \). We have to show \( \theta \land \neg \phi \vdash \psi \) and we may assume that there is a smallest \( i \) with \( s_i \cap S_{\theta \land \neg \phi} \neq \emptyset \). Since \( \theta \vdash \psi \) it is enough to show that the index \( i \) is smallest with \( s_i \cap S_\theta \neq \emptyset \). Otherwise, pick \( j < i \) smallest such that \( s_j \cap S_\theta \neq \emptyset \). Since \( \theta \not\vdash \neg \phi \) we have \( s_j \cap S_\theta \nsubseteq S_{\neg \phi} \). Recall that \( S_{\neg \phi} = AtL - S_\phi \). Hence \( s_j \cap S_\theta \nsubseteq S_{\neg \phi} \) says \( s_j \cap S_\theta \cap S_\phi = \emptyset \) in contradiction to \( j < i \) and the choice of \( i \).

This result is extremely useful because it allows us to construct rational consequence relations to show that certain conclusions do not follow by the GM rules from certain premises (i.e. do not hold for all rcr’s satisfying these premises).

Convention. For \( \vdash \) a rational consequence relation, we write \( \vdash \theta \) to mean that \( \eta \vdash \theta \) for some/any tautology \( \eta \). Notice that by LLE the particular choice of tautology is irrelevant, since \( \eta_1 \equiv \eta_2 \) for any tautologies \( \eta_1, \eta_2 \).

Remark. Notice that if \( \vec{s} = \emptyset, \emptyset, \ldots, \emptyset \), or \( \vec{s} \) is the empty sequence then \( \vdash \vec{s} \) is the trivial rcr such that for all \( \theta, \phi, \theta \vdash \phi \).

Exercise: What is \( \vec{s} \) when \( \vdash \vec{s} = \vec{?} \)? If \( \vec{s} = s_1 \) then what is \( \vdash \vec{s} \)?

2.3.1 Joyriders

Recall the ‘Joyriders’ example (Example 2.2 on page 15). We have \( Y \vdash L, J \vdash \neg L \) and \( J \vdash Y \). To show that \( Y \land J \vdash L \) does not follow from these and the GM rules, it is enough to find \( \vec{s} \) such that

\[
Y \vdash_{\vec{s}} L, \quad J \vdash_{\vec{s}} \neg L, \quad J \vdash_{\vec{s}} Y, \quad Y \land J \not\vdash_{\vec{s}} L
\]

The idea here is that amongst young people, joyriders are exceptional.

\[
s_1 = \{Y \land L \land \neg J\} \quad s_2 = \{Y \land \neg L \land J\}
\]
\[ S_Y = \{Y \land L \land J, Y \land L \land \neg J, Y \land \neg L \land J, Y \land \neg L \land \neg J\} \]
\[ S_J = \{Y \land L \land J, Y \land \neg L \land J, \neg Y \land L \land J, \neg Y \land \neg L \land J\} \]
\[ S_{Y \land J} = \{Y \land L \land J, Y \land \neg L \land J\} = S_Y \cap S_J \]

So the least \( i \) such that \( s_i \cap S_Y \neq \emptyset \) is 1
the least \( i \) such that \( s_i \cap S_J \neq \emptyset \) is 2
the least \( i \) such that \( s_i \cap S_{Y \land J} \neq \emptyset \) is 2

\[ s_1 \cap S_Y = \{Y \land L \land \neg J\} \subseteq S_L \quad \text{so} \quad Y \vdash_{\exists} L \]
\[ s_2 \cap S_J = \{Y \land \neg L \land J\} \subseteq S_Y, S_{\neg L} \quad \text{so} \quad J \vdash_{\exists} \neg L, J \vdash_{\exists} Y \]
but \( s_2 \cap S_{Y \land J} = \{Y \land \neg L \land J\} \nsubseteq S_L \quad \text{so} \quad Y \land J \not\vdash_{\exists} L \]

So MON fails.

### 2.3.2 Penguins

Birds normally fly
Penguins normally don’t fly
Penguins are birds
∴ Birds are normally not penguins.

\[ b \vdash f \]
\[ b \vdash \neg f \]
\[ p \vdash \neg \eta \quad (\eta \text{ a tautology}) \]
\[ p \land \neg b \vdash \neg \eta \]
\[ b \vdash \neg p \]

Here we have interpreted ‘Penguins are birds’ by ‘In all the most usual worlds in which penguins are not birds, a contradiction holds’. Notice this is rather stronger than \( p \vdash b \) ‘Penguins are usually birds’.

To see that this conclusion follows for \( \vdash \) an rcr notice that

\[ p \vdash \neg b \vdash \neg \eta \quad \text{from} \quad p \land \neg b \vdash \neg \eta \text{ by CON} \]
\[ \therefore p \vdash b \quad \text{since} \quad \neg b \rightarrow \neg \eta \vdash b \text{ and RWE} \]
\[ \therefore p \land b \vdash \neg f \quad \text{by CMO with} \quad p \vdash \neg f \]
\[ \therefore b \vdash p \rightarrow \neg f \quad \text{by CON} \]
\[ \therefore b \vdash f \land (p \rightarrow \neg f) \quad \text{by AND with} \quad b \vdash f \]
\[ \therefore b \vdash \neg p \quad \text{by RWE since} \quad f \land (p \rightarrow \neg f) \vdash \neg p \]

On the other hand we cannot conclude from these hypotheses and the assumption that \( \vdash \) is an rcr, that \( b \land \neg f \vdash p \) (i.e. non-flying birds are usually penguins). To see this, notice that for

\[ s_1 = \{b \land \neg p \land \neg f\}, \quad s_2 = \{b \land \neg p \land \neg f\}, \quad s_3 = \{b \land p \land \neg f\} \]
\[ b \vdash_{\exists} f, \quad p \vdash_{\exists} \neg f, \quad p \land \neg b \vdash_{\exists} \neg \eta \quad (\text{since} \quad s_i \cap S_{p \land \neg b} = \emptyset \text{ for all} \ i) \]

but \( b \land \neg f \not\vdash_{\exists} p \) since 2 is the least \( i \) such that \( s_i \cap S_{b \land \neg f} \neq \emptyset \) and \( s_2 \cap S_{b \land \neg f} \nsubseteq S_p \).
2.4 The Representation Theorem for rcr's

**Theorem 2.2.** Every rcr on SL, ‹, is of the form ‹s for some \( s = s_1, s_2, \ldots, s_m \subseteq At^L \), and conversely (by Theorem 2.1).

**Remark.** It is not the case that given an rcr ‹ there is a unique \( s \) such that ‹ = ‹s. To see this, notice that if we remove \( s_i \) from (or add \( s_i \) to) \( s \) such that \( \bigcup_{j<i} s_j \supseteq s_i \) then ‹s is unchanged since \( i \) can never be minimal (nor unique) s.t. \( s_i \cap S_\theta \neq \emptyset \).

Similarly if we replace \( s_k \) in ‹s by \( s_k - \bigcup_{j<k} s_j \) then ‹s will not change, since \( s_k \cap S_\theta \neq \emptyset \) and \( s_j \cap S_\theta = \emptyset \) for all \( j < k \) and for such \( k, \theta \), \( s_k \cap S_\theta \subseteq S_\varphi \) if and only if \( (s_k - \bigcup_{j<k} s_j) \cap S_\theta \subseteq S_\varphi \).

From these observations we see that we can take the \( s_i \) to be disjoint and non-empty (then we do have uniqueness).

**Proof of Theorem 2.2**

**Idea.** If we already knew that ‹ = ‹s (with the \( s_i \) disjoint and non-empty for all \( i \in \{1, ..., m\} \)) we could recover \( s_1 \) by noting that for \( R \subseteq At^L \):

\[
\begin{align*}
s_1 \subseteq R & \iff s_1 \subseteq S_V R \\
& \iff s_1 \cap At^L \subseteq S_V R \\
& \iff s_1 \cap S_\eta \subseteq S_V R \text{ for } \eta \text{ a tautology} \\
& \iff \vdash \bigvee R.
\end{align*}
\]

So

\[
s_1 = \bigcap \{R \subseteq At^L \mid \vdash \bigvee R \}.
\]

Similarly, given that we have recovered \( s_1, \ldots, s_{k-1} \) \((k \leq m)\) we can recover \( s_k \) by noticing that for \( R \subseteq At^L \),

\[
\begin{align*}
s_k \subseteq R & \iff s_k \cap (At^L - \bigcup_{i=1}^{k-1} s_i) \subseteq R \text{ since } s_k \cap s_i = \emptyset \text{ for } i < k \\
& \iff s_k \cap S_{-\bigvee_{i=1}^{k-1} s_i} \subseteq R \text{ since } S_{-\bigvee_{i=1}^{k-1} s_i} = At^L - \bigcup_{i=1}^{k-1} s_i \\
& \iff -\bigvee_{i=1}^{k-1} s_i \vdash \bigvee R \text{ since } R = S_V R \text{ and } k \text{ is clearly minimal such that } s_k \cap S_{-\bigvee_{i=1}^{k-1} s_i} \neq \emptyset \text{ (since } k \neq 0 \text{ since } k \leq m) \\
& \therefore s_k = \text{smallest set } R \subseteq At^L \text{ s.t. } -\bigvee_{i=1}^{k-1} s_i \vdash \bigvee R \\
& = \bigcap \{R \subseteq At^L \mid -\bigvee_{i=1}^{k-1} s_i \vdash \bigvee R \}.
\end{align*}
\]
Returning to the proof proper, of course we don’t know yet that $\vdash = \vdash_{\xi}$. However we do know what $s_1, s_2, \ldots, s_m$ have to be if they exist, so the idea is to try the $s_i$ defined by this recovery process in any case. That is, for $k = 1, 2, \ldots$.

$$
\begin{align*}
s_1 &= \bigcap \{ R \subseteq At^L \mid \vdash \bigvee R \} \\
    &= \bigcap \{ R \subseteq At^L \mid \neg \bigvee \bigcup_{i \leq 1} s_i \vdash \bigvee R \}
\end{align*}
$$

since $\neg \bigvee \bigcup_{i \leq 1} s_i = \neg \bigvee \emptyset$ is a tautology

$$
\begin{align*}
s_k &= \bigcap \{ R \subseteq At^L \mid \neg \bigvee \bigcup_{i \leq k} s_i \vdash \bigvee R \}
\end{align*}
$$

Notice that by AND, $\{ R \subseteq At^L \mid \neg \bigvee \bigcup_{i \leq k} s_i \vdash \bigvee R \}$ is closed under intersections, so since this set is finite it contains the intersection of all members of the set, i.e. $s_k$. Thus again,

$$
s_k = \text{smallest } R \subseteq At^L \text{ s.t. } \neg \bigvee \bigcup_{i \leq k} s_i \vdash \bigvee R
$$

As things stand we have infinitely many $s_k$. In order to cut down to a finite number we first show:

**Claim 1.** the $s_k$ are disjoint, i.e. $s_k \cap s_j = \emptyset$ for $j \neq k$.

Suppose $j < k$. First notice that

$$
\neg \bigvee_{i \leq k} s_i \equiv \bigvee \left( At^L - \bigcup_{i \leq k} s_i \right), \text{ equivalently } S_{\neg \bigvee_{i \leq k} s_i} = At^L - \bigcup_{i \leq k} s_i
$$

(as follows from $S_{\neg \bigvee R} = At^L - S_{\bigvee R} = At^L - R$ for all sets $R \subseteq At^L$; this will be used frequently in what follows).

Therefore by SCL, $\neg \bigvee_{i \leq k} s_i \vdash \bigvee (At^L - \bigcup_{i \leq k} s_i)$ so by the definition of $s_k$, $s_k \subseteq At^L - \bigcup_{i \leq k} s_i$ and $s_k \cap s_j \subseteq (At^L - \bigcup_{i \leq k} s_i) \cap s_j \subseteq (At^L - s_j) \cap s_j = \emptyset$, as required.

Since there are only $2^n$ atoms, claim 1 implies that at least one of $s_1, s_2, \ldots, s_{2^n+1}$ must be empty. Let $m$ be minimal such that $s_{m+1} = \emptyset$.

**Claim 2.** $\vdash_{\xi} = \vdash$, where $\tilde{s} = s_1, \ldots, s_m$

**Proof.** First suppose that $\theta \vdash \phi$. We must show that $\theta \vdash_{\xi} \phi$. If $s_i \cap S_\theta = \emptyset$ for $i = 1, \ldots, m$
we are done, so suppose instead that \( k \) is minimal such that \( s_k \cap S_\theta \neq \emptyset \).

\[
\therefore \quad S_\theta \cap \bigcup_{i<k} s_i = \emptyset
\]

\[
\implies \quad S_\theta \subseteq At^L - \bigcup_{i<k} s_i = S_\neg \bigcup_{i<k} s_i
\]

\[
\therefore \quad \theta \models \neg \bigcup_{i<k} s_i
\]

so \( \theta \vdash \neg \bigcup_{i<k} s_i \) by SCL

\[
\therefore \quad \theta \land \neg \bigcup_{i<k} s_i \models \phi
\]

\[
\implies \quad \neg \bigcup_{i<k} s_i \models \theta \rightarrow \phi
\]

so \( \neg \bigcup_{i<k} s_i \models \theta \rightarrow \phi \) by RWE

\[
\implies \quad \bigcup_{i<k} s_i \subseteq S_{\neg \theta \rightarrow \phi}
\]

\[
\implies \quad s_k \cap S_\theta \subseteq S_{\neg \theta \rightarrow \phi}
\]

\[
= S_{\theta \land (\theta \rightarrow \phi)}
\]

\[
\subseteq S_\phi
\]

as required, since \( \theta \land (\theta \rightarrow \phi) \models \phi \)

To prove the converse suppose that \( \theta \models \neg \phi \).

**Case 1.** \( S_\theta \cap s_i = \emptyset \) for \( i = 1, \ldots, m \).

Since \( s_{m+1} = \emptyset \),

\[
\neg \bigcup_{i \leq m} s_i \models \bigvee \emptyset
\]

so \( \neg \bigcup_{i \leq m} s_i \models \theta \land \neg \bigcup_{i \leq m} s_i \models \phi \) by RWE (since \( \bigvee \emptyset \) is a contradiction so \( \bigvee \emptyset \models \theta \), \( \bigvee \emptyset \models \phi \)).

Hence by CMO

\[
\theta \land \neg \bigcup_{i \leq m} s_i \models \phi \tag{\dagger}
\]

The assumption in case 1 says \( S_\theta \subseteq At^L - \bigcup_{i \leq m} s_i \). Hence \( S_{\theta \land \neg \bigcup_{i \leq m} s_i} = S_\theta \cap \neg \bigcup_{i \leq m} s_i = S_\theta \cap (At^L - \bigcup_{i \leq m} s_i) \) (by \( \dagger \)) = \( S_\theta \), so \( \theta \equiv \theta \land \neg \bigcup_{i \leq m} s_i \) and \( \theta \models \phi \) follows from (\dagger) using LL.

**Case 2.** There is a minimal \( k \leq m \) such that \( S_\theta \cap s_k \neq \emptyset \).

Hence \( s_k \not\subseteq S_{\neg \theta} \) so \( \neg \bigcup_{i<k} s_i \not\models \bigvee S_{\neg \theta} \). Since since \( \neg \theta \equiv \bigvee S_{\neg \theta} \), RWE gives

\[
\neg \bigcup_{i<k} s_i \not\models \neg \theta.
\] (2.2)

By minimality of \( k \), we have

\[
S_\theta \subseteq At^L - \bigcup_{i<k} s_i,
\]

so with (\( \dagger \)) we get \( \theta \models \neg \bigcup_{i<k} s_i \) and therefore

\[
\theta \equiv \theta \land \neg \bigcup_{i<k} s_i.
\] (2.3)
Since \( \theta \vdash_{\vec{s}} \phi \) and \( k \) is minimal such that \( s_k \cap S_{\theta} \neq \emptyset \), we have \( s_k \cap S_{\theta} \subseteq S_{\theta} \), hence \( S_{\theta} \cap S_{\theta} = s_k \cap S_{\theta} \subseteq S_{\theta} \). It follows \( \bigvee s_k \wedge \theta \models \phi \), hence \( \bigvee s_k \models \theta \rightarrow \phi \). Using \( \neg \bigvee \bigcup_{i<k} s_i \vdash \bigvee s_k \) (by definition of \( s_k \)) and RWE, we obtain 
\[
\neg \bigvee_{i<k} s_i \vdash \theta \rightarrow \phi.
\]
This with (2.2) and RMO gives \( \theta \wedge \neg \bigvee_{i<k} s_i \vdash \theta \rightarrow \phi \). With (2.3) and LLE we get
\[
\theta \models \theta \rightarrow \phi.
\]
By AND with \( \theta \models \theta \) (REF), \( \theta \models \theta \wedge (\theta \rightarrow \phi) \), and hence \( \theta \models \phi \) as required, by RWE using \( \theta \wedge (\theta \rightarrow \phi) \models \phi \). 

**Applying the Representation Theorem.**

We can use Theorem 2.2 to show both that some rule does not hold for all rcr’s (by finding \( \vec{s} \) such that the rule fails for \( \models_{\vec{s}} \)) and to show that a rule does hold for all rcr’s – by showing that it holds for all \( \models_{\vec{s}} \) and then appealing to the Representation Theorem. We give an example of this latter by showing (again) that
\[
\begin{align*}
 b & \models f, \ p \models \neg f, \ p \wedge \neg b \models \bot \\
 b & \models \neg p 
\end{align*}
\]
holds for all rcr’s.

Let \( \models \) be a rcr, so by Theorem 2.2, \( \models = \models_{\vec{s}} \) for some \( \vec{s} \). Suppose that \( b \models_{\vec{s}} f, \ p \models_{\vec{s}} \neg f, \ p \wedge \neg b \models_{\vec{s}} \bot \). If \( S_b \cap s_i = \emptyset \) for all \( i \) then certainly \( b \models_{\vec{s}} \neg p \) as required. Otherwise let \( k \) be minimal such that \( s_k \cap S_b \neq \emptyset \). Then since \( b \models_{\vec{s}} f \),
\[
 s_k \cap S_b \subseteq S_f. \tag{2.4}
\]

Now suppose on the contrary that \( b \not\models_{\vec{s}} \neg p \). Then \( s_k \cap S_b \not\subseteq S_{\neg p} \) so there must be some atom \( \alpha \) such that \( \alpha \in s_k \cap S_b \) and \( \alpha \not\in S_{\neg p} \), in other words \( \alpha \in S_p \). Hence
\[
 s_k \cap S_b \cap S_p \neq \emptyset. \tag{2.5}
\]

If \( k \) was minimal such that \( s_k \cap S_p \neq \emptyset \) we’d have
\[
 s_k \cap S_p \subseteq S_{\neg f} \tag{2.6}
\]
from the given \( p \models_{\vec{s}} \neg f \) so from (2.4), (2.6),
\[
s_k \cap S_p \cap S_b \subseteq S_f \cap S_{\neg f} = S_{f \wedge \neg f} = \emptyset
\]
contradicting (2.5). Hence the least $j$ such that $s_j \cap S_p \neq \emptyset$ must be less than $k$. But in that case $s_j \subseteq (A^{L} - S_b) = S_{\neg b}$ by minimality of $k$. Hence $j$ will also be minimal such that

$$s_j \cap S_{p \land \neg b} = s_j \cap S_p \cap S_{\neg b} \neq \emptyset$$

and so from the given $p \land \neg b \not\models \bot$,

$$\emptyset \neq s_j \cap S_{p \land \neg b} \subseteq S_{\bot} = \emptyset,$$

contradiction!

We conclude that the assumption $b \not\models \neg p$ must fail and hence that $b \not\models \neg p$, as required.
Chapter 3

Rational Closure

3.1 Statement of the problem

By any of the two methods given so far (by direct derivation from the GM rules or by using the Representation Theorem) we can conclude that the ‘rule’

\[
\begin{array}{c}
\text{K}_0 \\
Y \vdash L & J \vdash \neg L & J \vdash Y \\
\end{array}
\]

\[
Y \land J \vdash \neg L
\]

holds for all rcr’s — equivalently if we assume (in the context of such a knowledge base \(K_0^1\)) that an agent’s notion of consequence is a rational consequence relation then knowing that an agent believes \(K_0\) we can conclude that the agent believes \(Y \land J \vdash \neg L\).

But what about \(Y \land O \vdash L\)? \((O = \text{ornithologist})\). Can we also conclude that such an agent believes \(Y \land O \vdash L\) knowing that s/he believes \(K_0\)? The answer to this is no — and indeed nor should we be able to, after all this particular agent might be identifying ‘ornithologist’ with ‘oologist’! (Formally we can show that \(Y \land O \not\vdash L\) doesn’t follow from \(K_0\) by noticing that if \(s_1 = \{Y \land L \land \neg J \land \neg O\}, s_2 = \{Y \land \neg L \land J \land O\}\) then \(\vdash_{s_1} s_2\) satisfies \(K_0\) but not \(Y \land O \vdash L\).

So why did the conclusion \(Y \land O \vdash L\) look reasonable? Possibly because we tacitly assumed that \(K_0\) sums up all the agent’s (relevant) knowledge (so \(O, \neg O\) are independent of \(Y, J, L\) — otherwise any connection would have been explicitly included in \(K_0\)). In other words we were assuming that the agent had no (relevant) beliefs or knowledge not already included in \(K_0\), equivalently that the agent’s notion of consequence is the ‘simplest rcr’ satisfying \(K_0\).

So, can we formalise what we mean by the simplest rcr satisfying \(K_0\)?

\(^1\)The assertions \(\theta \vdash \phi\) in such a knowledge base are sometimes referred to as defeasible assertions or alternatively conditionals. In this course I shall try to stick with this latter.
### 3.2 Definition of the rational closure

#### An attempt for the definition of the rational closure

Our first stab at producing the rational closure $\vdash_{\text{min}}$ of a finite conditional knowledge base $K = \{ \theta_i \vdash \phi_i \mid i = 1, \ldots, m \}$ might be to say that if any $\vdash$ which satisfied $K$ failed to satisfy $\theta \vdash \phi$ then $\vdash_{\text{min}}$ should also fail to satisfy this conditional. After all if $\vdash_{\text{min}}$ is supposed to be the simplest, or least informative, rcr satisfying $K$ then it should not assert that $\theta \vdash_{\text{min}} \phi$ when this was not forced for all rcr’s satisfying $K$.

This idea formalizes as:

$$\theta \vdash_{\text{min}} \phi \iff \theta \vdash \phi \text{ for all rcr’s } \vdash \text{ satisfying } K,$$

equivalently $\vdash_{\text{min}} = \bigcap \{ \vdash \mid \vdash \text{ is an rcr satisfying } K \}.$$

Unfortunately for rcr’s $\vdash_{\text{min}}$ so defined is not in general a rational consequence relation (— so this attempt has failed). To see that $\vdash_{\text{min}}$ is not in general an rcr consider again $K_0 = \{ Y \vdash L, J \vdash Y, J \vdash \neg L \}$. We have already seen an example $\vdash_{\vec{s}}$ of an rcr satisfying $K_0$ such that $Y \land O \not\vdash_{\vec{s}} L$ ($s_1 = \{ Y \land L \land \neg J \land \neg O \}, s_2 = \{ Y \land \neg L \land J \land O \}$). Also if $t_1 = \{ Y \land L \land \neg J \land O \}, t_2 = \{ Y \land \neg L \land J \land O \}$ then $\vdash_{\vec{t}}$ satisfies $K_0$ and $Y \not\vdash_{\vec{t}} \neg O$. So for $\vdash_{\text{min}}$ as defined above for $K = K_0$ we see that

$$Y \vdash_{\text{min}} L \quad \text{since } \vdash_{\text{min}} \text{ satisfies } K_0$$
$$Y \not\vdash_{\text{min}} \neg O \quad \text{since } Y \not\vdash_{\vec{t}} \neg O$$

but

$$Y \land O \not\vdash_{\text{min}} L \quad \text{since } Y \land O \not\vdash_{\vec{s}} L$$

so $\vdash_{\text{min}}$ fails to verify RMO and is not an rcr.

[The key reason for the failure of this method to work is because some of the defining rules (namely RMO) for rcrs involve negations of conditionals, the assertion that certain conditionals don’t hold. If we dropped the offending rule RMO then this device of taking the intersection does work. Consequence relations satisfying the rules for rational consequence less RMO are called Preferential Consequence Relations. So given a ‘positive’ $K = \{ \theta_i \vdash \phi_i \mid i = 1, \ldots, m \}$ there is a set-theoretically smallest preferential consequence relation satisfying $K$, namely the intersection of all preferential consequence relations satisfying $K$.]

#### 3.2 Definition of the rational closure

The idea here is to construct a simplest rcr satisfying $K$ as $\vdash_{\vec{u}}$ in such a way that $\vec{u}$ introduces as few ‘inessential preferences’ as possible. In other words, $u_1$ should be as large as possible (subject to the requirement that $\vdash_{\vec{u}}$ satisfies $K$). Then $u_2$ should be as large as possible given that it is disjoint from $u_1$ (and $\vdash_{\vec{u}}$ satisfies $K$) and so on. (At this point it is not even clear that such unique largest $u_1, u_2, \ldots$ exist.)
In fact we can construct such a $\bar{u}$ rather more simply as shown below.

<table>
<thead>
<tr>
<th>$s_1^i$</th>
<th>$s_2^i$</th>
<th>$\ldots$</th>
<th>$s_n^i$</th>
<th>( \cup_i s_i^i = t_1 )</th>
<th>$u_1 = t_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_1^j$</td>
<td>$s_2^j$</td>
<td>$\ldots$</td>
<td>$s_n^j$</td>
<td>( \cup_i s_i^j = t_2 )</td>
<td>$u_2 = t_2 - t_1$</td>
</tr>
<tr>
<td>$s_1^k$</td>
<td>$s_2^k$</td>
<td>$\ldots$</td>
<td>$s_n^k$</td>
<td>( \cup_i s_i^k = t_3 )</td>
<td>$u_3 = t_3 - (t_1 \cup t_2)$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ldots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$s_1^{m'}$</td>
<td>$s_2^{m'}$</td>
<td>$\ldots$</td>
<td>$s_n^{m'}$</td>
<td>( \cup_i s_i^{m'} = t_n )</td>
<td>( u_n = t_n - \bigcup_{j&lt;n} t_j )</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\ldots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
</tbody>
</table>

Notice that $u_1 \cup u_2 \cup \ldots \cup u_n = t_1 \cup t_2 \cup \ldots \cup t_n$.

Clearly since the $u_n$ are disjoint subsets of $At^L$ at most $2^n$ of them can be non-empty. Let $m$ be minimal such that $u_{m+r} = \emptyset$ for all $r > 0$ and set $\bar{u} = u_1, u_2, \ldots, u_m$ and $\bar{t} = t_1, t_2, \ldots, t_m$. From the remark after 2.2 we know that $\models_{\bar{u}} = \models_{\bar{t}}$.

**Claim.** $\models_{\bar{u}}$ satisfies $K$.

**Proof.** Let $\theta \models \phi$ be in $K$. If $S_\theta \cap t_i = \emptyset$ for all $i = 1, \ldots, m$ then $\theta \models_{\bar{t}} \phi$ as required. Otherwise let $j$ be minimal such that $S_\theta \cap t_j \neq \emptyset$. Let $\alpha \in S_\theta \cap t_j$. Then $\alpha \in s_j^n$ for some $n$, since $t_j = \bigcup_n s_j^n$. Moreover $s_j^n \cap S_\theta = \emptyset$ for $r < j$, since $s_j^n \subseteq t_r$ and $S_\theta \cap t_r = \emptyset$ for $r < j$. So since $\alpha \in S_\theta \cap s_j^n$, $j$ is also minimal such that $S_\theta \cap s_j^n \neq \emptyset$. Then since $\models_{\bar{u}}$ satisfies $K$ (so $\theta \models_{\bar{u}} \phi$), $\alpha \in S_\theta \cap s_j^n \subseteq S_\phi$. Therefore since $\alpha$ was an arbitrary element of $S_\theta \cap t_j$, $S_\theta \cap t_j \subseteq S_\phi$ and hence $\theta \models_{\bar{t}} \phi$ as required. \[\square\]

Now quite clearly $u_1, u_2, \ldots$ satisfy our initial requirements; $u_1$ is the largest $s_1^n$ of any $s^n$ such that $\models_{\bar{u}}$ satisfies $K$. (Notice that since $\models_{\bar{u}}$ satisfies $K$, $\bar{u}$ itself is in the list of $s^n$.) Again if $s_1^n = u_1$, $s_2^n$ was disjoint from $s_1^n$ then $u_2 = t_2 - t_1 \supseteq s_2^n - u_1 = s_1^n = s_2^n$, and so on.

The rcr $\models_{\bar{u}}$ is called the **rational closure of $K$** and is widely viewed as the top candidate for the title of ‘simplest rcr satisfying $K$.’ In terms of our earlier discussion then, if we believe that $K$ is all of an agent’s knowledge then we are arguing that the agent’s rcr should be the rational closure of $K$.

In the case of $K_0 = \{Y \models L, J \models Y, J \models \neg L\}$ (in the language $\{Y, L, J, O\}$) the $u_i$ come
3.2. DEFINITION OF THE RATIONAL CLOSURE

out to be:

\[
\begin{align*}
u_1 &= \left\{ \begin{array}{l}
Y \land L \land \neg J \land O \\
Y \land L \land \neg J \land \neg O \\
\neg Y \land L \land \neg J \land O \\
\neg Y \land L \land \neg J \land \neg O \\
\neg Y \land \neg L \land \neg J \land O \\
\neg Y \land \neg L \land \neg J \land \neg O
\end{array} \right\} \\
u_2 &= \left\{ \begin{array}{l}
Y \land \neg L \land J \land O \\
Y \land \neg L \land J \land \neg O \\
Y \land \neg L \land \neg J \land O \\
Y \land \neg L \land \neg J \land \neg O
\end{array} \right\} \\
u_3 &= \{ \text{all remaining atoms} \}
\end{align*}
\]

And so \( Y \land O \vdash_{\mathcal{R}} L \) (as well as \( Y \land \neg O \vdash_{\mathcal{R}} L \))

An Alternative Characterisation of Rational Closure

The above characterisation of the rational closure of \( K \) is in terms of ‘how it is constructed.’ We could instead have characterized the rational closure in terms of some properties it possesses (although in that case we would need to show that an rcr exists to satisfy those properties).

One such characterization was given by Lehmann and Magidor.

**Theorem 3.1.** The rational closure of a (finite) knowledge base \( K \) is that rcr \( \vdash_0 \) satisfying \( K \) such that for any other rcr \( \vdash \) satisfying \( K \) different from \( \vdash_0 \):

\((R1)\) If for some \( \theta \) and \( \phi \) it is true that \( \theta \vdash_0 \phi \) and \( \theta \not\models \phi \) then there are \( \theta' \), \( \phi' \) such that \( \theta' \lor \theta \vdash \neg \theta \), \( \theta' \vdash \phi' \) and \( \theta' \not\models \phi' \).

\((R2)\) There are \( \theta \) and \( \phi \) such that \( \theta \vdash \phi \), \( \theta \not\models \phi \) and for all \( \theta' \) and \( \phi' \) if \( \theta' \lor \theta \vdash \neg \theta \) and \( \theta' \vdash_0 \phi' \) then \( \theta' \vdash \phi' \).

Roughly (R1) is saying that if \( \vdash_0 \) thinks \( \phi \) follows from \( \theta \) but \( \vdash \) asserts that this is an extravagant conclusion because \( \vdash \) does not think that \( \phi \) follows from \( \theta \) then \( \vdash_0 \) can point out that there are \( \theta' \), \( \phi' \) such that \( \vdash \) thinks \( \theta' \) is more reasonable than \( \theta \) (i.e. holds in worlds which are more normal than any in which \( \theta \) holds — this is summed up by \( \theta' \lor \theta \vdash \neg \theta \) but \( \vdash \) is being, by his/her own standards, more extravagant in holding that \( \phi' \) follows from \( \theta' \) (unnecessarily \( \vdash_0 \) thinks since \( \theta' \not\models \phi' \)) than \( \vdash_0 \) was in holding that \( \phi \) follows from \( \theta \).

(R2) is saying that \( \vdash \) cannot use this same argument in reverse against \( \vdash_0 \).

From this it follows that it is not possible for there to be two rcr’s \( \vdash_0 \) satisfying (R1) and (R2), so for the purposes of the proof it is enough to show that the rational closure \( \vdash_0 \) as defined above satisfies (R1) and (R2).

**Proof.** First notice that if \( u_i = \emptyset \) with \( u_{i+1} \neq \emptyset \) for \( i < m \) then \( u_1, u_2, \ldots, u_{i-1}, u_{i+1}, \ldots, u_m \) would have been in the above list so \( \emptyset \neq u_{i+1} \subseteq t_i \subseteq \{ \alpha \mid \alpha \in t_j \text{ for some } j \leq i \} = u_1 \cup u_2 \cup \cdots \cup u_i \), contradicting the disjointness of the \( u_i \). Hence \( u_i \neq \emptyset \) for \( i \leq m \).
Now suppose that \( \vdash_s \neq \vdash_a \) also satisfies \( K \). We may assume that \( s = s_1, s_2, \ldots, s_v \) are disjoint and non-empty. To check (R1) suppose that \( \theta \vdash_a \phi \), \( \theta \not\vdash_s \phi \). Let \( j \) be minimal such that \( S_\theta \cap s_j \neq \emptyset \) (such a \( j \) must exist, otherwise \( \theta \vdash_s \phi \)).

Then, by \( \theta \not\vdash_s \phi \), \( S_\theta \cap s_j \not\subseteq S_\phi \), say \( \alpha \in S_\theta \cap s_j \cap S_\neg \phi \). Then \( \alpha \in t_j \) so \( \alpha \in u_r \) for some \( r \leq j \). Hence \( S_\theta \cap u_r \neq \emptyset \), so since \( \theta \vdash_a \phi \) if \( q \) is minimal such that \( S_\theta \cap u_q \neq \emptyset \) then \( S_\theta \cap u_q \subseteq S_\phi \) and we must have \( q < r \). Thus \( u_q \neq s_q \) since \( \emptyset = s_q \cap S_\theta \). Let \( h \) be minimal such that \( u_h \neq s_h \) so that \( h \leq q < r \). If there were some \( \beta \in s_h \) such that \( \beta \not\in u_h \) then \( \beta \in t_h \), so \( \beta \in u_p \) for some \( p < h \) (as \( \beta \in u_1 \cup u_2 \cup \cdots \cup u_h \) and \( \beta \not\in u_h \)). Hence \( u_p \neq s_p \) since \( \beta \in u_p \) and \( \beta \not\in s_p \) because \( \beta \in s_h \) and the \( s_i \) are disjoint, a contradiction. So \( u_h \supset s_h \).

Noticing that \( s_h \neq \emptyset \) we now see that \( \theta \cup \theta_h \vdash_a \neg \theta \), because \( h < r \), \( \neg \theta \vdash_a \theta_h \vdash_s \theta \), because \( s_h \cap S_\theta = s_h \) and \( s_i \cap S_\theta = \emptyset \) for \( i < h \), and \( \theta_h \vdash_a \theta \vdash \theta_h \), because \( u_h \cap S_\theta = u_h \not\subseteq s_h \), we now see that (R1) holds with \( \theta' = \theta_h \), \( \phi' = \theta \).

To check (R2) let \( h \) be minimal such that \( u_h \neq s_h \) (or \( h = v + 1 \) and \( s_h \) is not even defined). Notice that it is not possible that \( h = m + 1 \) because in that case \( s_{m+1} \) would have contained an element not in \( s_1 \cup \cdots \cup s_m = u_1 \cup \cdots \cup u_m \) so \( u_{m+1} \neq \emptyset \). As above it must be that \( u_h \supset s_h \) (where \( s_h = \emptyset \) for \( h = v + 1 \)) so \( \theta \cup u_h \vdash_s \theta \vdash_s \theta \cup u_h \), \( \theta \vdash_s \theta \vdash_s \theta \vdash_s \theta \), \( \theta \vdash_s \theta \vdash_s \theta \vdash_s \theta \) (either way). Now suppose \( \theta' \vdash_s \theta \vdash_s \theta \vdash_s \theta \) but \( \theta \not\vdash_s \theta' \). Let \( q \) be minimal such that \( S_{\theta'} \cap u_q \neq \emptyset \). Clearly \( q \leq h \) and since

\[
\emptyset \neq u_h \cap S_{\theta'} \cap u_h = u_h \not\subseteq A_{\theta'} - u_h = S_{\neg \theta'} ,
\]

it cannot be that \( q = h \), so \( q < h \). But in that case \( s_i = u_i \) for \( i \leq q \) so \( q \) is also minimal such that \( s_q \cap S_{\theta'} = \emptyset \). But then \( u_q \cap S_{\theta'} = s_q \cap S_{\theta'} \) is both a subset of \( S_{\theta'} \) from \( \theta \vdash_a \theta' \) and not a subset of \( S_{\theta'} \) from \( \theta' \vdash_s \theta' \), contradiction.

\[\square\]

### 3.3 An Algorithm for Computing the Rational Closure

In this section we shall look at an algorithm for constructing the rational closure of a conditional knowledge base \( K = \{ \theta_i \vdash \phi_i \mid i = 1, 2, \ldots, r \} \) (as \( \vdash_s \)), where the \( \theta_i, \phi_i \in SL \). Of course in a sense the simple ‘union’ construction we gave earlier for forming the rational closure \( \vdash_s \) of \( K \) can be turned into such an algorithm though it would be horribly impractical. A superior algorithm has been given by Lehmann-Magidor and, independently, Pearl, from whom the common name ‘Z-algorithm’ derives.

Hardly surprisingly the idea behind this algorithm is the same as that behind the ‘union’ construction. Namely put as many atoms as possible into \( s_1 \), then put as many of the remaining atoms as possible into \( s_2 \), and so on until any further \( s_i \) are empty.

Before giving the algorithm it is useful to introduce a little notation: For \( \alpha \in A_L \) we shall say that
3.3. AN ALGORITHM FOR COMPUTING THE RATIONAL CLOSURE

\[ \alpha \text{ confirms } \theta \vdash \phi \text{ if } \alpha \in S_{\theta \land \phi} \]
and
\[ \alpha \text{ refutes } \theta \vdash \phi \text{ if } \alpha \in S_{\theta \land \neg \phi}. \]

The Z-algorithm for constructing the rational closure of a conditional knowledge base \( K \) computes sequences
\[ s_1, s_2, \ldots \subseteq At^L \]
\[ K = K_0 \supseteq K_1 \supseteq K_2 \supseteq \cdots, \]
\[ At^L = A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots \]
as follows:

- Initially set \( K_0 = K \), \( A_0 = At^L \).
- Given \( K_i \subseteq K \) and \( A_i \subseteq At^L \) set
  \[ s_{i+1} = \{ \alpha \in A_i \mid \alpha \text{ does not refute any } \theta \vdash \phi \text{ in } K_i \} \]
    \[ = \{ \alpha \in A_i \mid \alpha \not\in S_{\theta \land \neg \phi} \text{ for any } \theta \vdash \phi \text{ in } K_i \} \]
    \[ = \{ \alpha \in A_i \mid \alpha \in S_{\neg \theta \lor \phi} \text{ for all } \theta \vdash \phi \in K_i \} \]
    \[ = A_i \cap \bigcap_{\theta \vdash \phi \in K_i} S_{\neg \theta \lor \phi}. \]
- If \( s_{i+1} = \emptyset \), then stop. Otherwise set
  \[ A_{i+1} = A_i - s_{i+1} \]
  \[ K_{i+1} = \{ \theta \vdash \phi \in K_i \mid \text{ no } \alpha \in s_{i+1} \text{ confirms } \theta \vdash \phi \} \]
  \[ = \{ \theta \vdash \phi \in K_i \mid s_{i+1} \cap S_{\theta \land \phi} = \emptyset \}. \]

Notice that \( K = K_0 \supseteq K_1 \supseteq K_2 \supseteq \cdots, At^L = A_0 \supseteq A_1 \supseteq A_2 \supseteq \cdots \) and the \( s_i \) are disjoint subsets of the finite set of atoms, so necessarily the algorithm terminates.

To appreciate the effect of this algorithm notice that if \( \alpha \not\in s_1 \) then there must be some \( \theta \vdash \phi \) in \( K_0 (= K) \) which \( \alpha \) refutes. Hence if \( \vdash_\ell \) satisfies \( K \) then \( \alpha \not\in t_1 \), otherwise, by the above remarks there would have to be a \( j < 1 \) and \( \alpha' \in t_j \) confirming \( \theta \vdash \phi \), which is impossible since there are no \( t_j \) before \( t_1 \). Hence \( s_1 \) is as large as it can be without immediately negating the possibility that there are some further \( s_2, s_3, \ldots \) for which \( \vdash_\ell \) satisfies \( K \).

Similarly if \( \vdash_\ell \) satisfies \( K \), \( s_1 = t_1, \alpha \in A_1 \) (i.e. \( \alpha \not\in s_1 \)) then \( \alpha \not\in s_2 \) means that \( \alpha \) must refute some \( \theta \vdash \phi \) in \( K_1 \) (i.e. not already confirmed by some \( \alpha' \in s_1 = t_1 \)). Again then by the above remarks \( \alpha \not\in s_2 \) since otherwise \( \theta \not\vdash_\ell \phi \). Hence yet again \( s_2 \) is as large as it
can be without immediately negating the possibility that there some further $s_3, s_4, \ldots$ for which $\vdash_{g}$ satisfies $K$, and so on.

Essentially then this algorithm aims to make the $s_i$ as large as possible without at any stage (obviously) destroying the hope that we may be able to extend the sequence of $s_i$ to an $\bar{s}$ such that $\vdash_{g}$ satisfies $K$. In view of the ‘union’ construction of the rational closure of $K$ we might hope that this would succeed. After all, the resulting $\bar{s}$ couldn’t be any bigger so if this process of never ‘immediately going wrong’ eventually meant it ‘went right’ and gave $\bar{s}$ such that $\vdash_{g}$ satisfied $K$ then ‘surely’ $\vdash_{g}$ must be the rational closure of $K$.

We now turn this informal hand waving into a proper proof that the process works, more precisely that if $\bar{u}$ is $u_1, u_2, \ldots, u_m$ with the $u_j$ as in the union construction disjoint and non-empty, then $u_j = s_j$ for $j = 1, 2, \ldots, m$ and $s_{m+1} = \emptyset$ the point at which the algorithm terminates. (Of course we can lop this empty $s_{m+1}$ off the end of $\bar{s}$ without changing $\vdash_{g}$.)

Let $\bar{s}$ be $s_1, s_2, \ldots, s_r+1$ with $s_{r+1} = \emptyset$ the point at which the algorithm terminates. We first show that $\vdash_{g}$ satisfies $K$. Let $\theta \vdash \phi$ be in $K$. If $s_i \cap S_{\theta} = \emptyset$ for all $i$ then $\theta \vdash_{g} \phi$. Otherwise let $i$ be minimal such that $s_i \cap S_{\theta} \neq \emptyset$ and let $\alpha \in s_i \cap S_{\theta}$. Suppose $\alpha \notin S_{\phi}$, i.e. $\alpha \in S_{\neg \phi}$. Then $\alpha$ refutes $\theta \vdash \phi$. By definition of $s_i$ then, $\theta \vdash \phi$ is not in $K_{i-1}$. However $\theta \vdash \phi$ is in $K = K_0$ so let $j$ be maximal such that $\theta \vdash \phi$ is in $K_j$. Notice $j < i - 1$ since the $K_m$ are decreasing. Then $\theta \vdash \phi$ is not in $K_{j+1}$ so by definition there must be some $\alpha' \in s_{j+1}$ which confirms $\theta \vdash \phi$, i.e. $\alpha' \in S_{\theta}$ and $\alpha' \in S_{\phi}$. But $j + 1 < i$, contradicting the minimality of $i$. This shows that $\vdash_{g}$ satisfies $K$.

We are now set up to show that $u_j = s_j$ for $j = 1, 2, \ldots, m$ and $s_{m+1} = \emptyset$ (so in consequence $r = m$). Suppose $j \leq m + 1$ and $s_i = u_i$ for $i < j$. By construction of $u_j$ (before chopping off the terminal empty sets) and the fact that the $s_i$ are disjoint and $\vdash_{g}$ satisfies $K$, $s_j$ being non-empty and containing an atom not already in $u_j$ is impossible. Hence either $j = m + 1$ and $s_{m+1} = \emptyset$ (in which case there’s nothing more to prove) or $j \leq m$ and $s_j \subseteq u_j$.

Suppose $s_j \subseteq u_j$, say $\alpha \in u_j - s_j$. Notice that

$$\alpha \notin \bigcup_{i < j} u_i = \bigcup_{i < j} s_i = \bigcup_{i < j} (A_{i-1} - A_i) = A_0 - A_{j-1} = A^L - A_{j-1}$$

so $\alpha \notin A_{j-1}$. Since $\alpha \notin s_j$ there must be some $\theta \vdash \phi$ in $K_{j-1}$ which $\alpha$ refutes. Now since $\alpha \in u_j$ and $\vdash_{g}$ satisfies $K$, by the earlier discussion there must be some $k < j$ and $\alpha' \in u_k$ which confirms $\theta \vdash \phi$. But then, since $s_i = u_i$ for $i < j$, $\alpha' \in s_k$ so even if $\theta \vdash \phi$ were in $K_{k-1}$ it would not be in $K_k$ (since $\alpha' \in s_k$ confirms it) and hence (since $k \leq j - 1$ and $K_k \supseteq K_{j-1}$) cannot be in $K_{j-1}$, contradiction!

This shows that $u_j = s_j$ (and incidentally that the algorithm doesn’t stop too early) and completes the proof that $\vdash_{g}$, with or without the terminal empty $s_{q+1}$, is the rational closure of $K$. 


3.3.1 Example

Let $K = \{ p \vdash q, q \vdash r, p \land q \vdash \neg r \}$. We use the Z-algorithm to find the rational closure of $K$.

To simplify things let $\alpha_1, \alpha_2, \alpha_3, \alpha_4, \ldots, \alpha_8$ list the atoms

\[
\begin{align*}
\alpha_1 &= p \land q \land r \\
\alpha_2 &= p \land q \land \neg r \\
\alpha_3 &= p \land \neg q \land r \\
\alpha_4 &= p \land \neg q \land \neg r \\
\vdots &= \vdots \\
\alpha_8 &= \neg p \land \neg q \land \neg r
\end{align*}
\]

in that (obvious) order.

Thus,

\[
\begin{align*}
S_{p \lor q} &= \{\alpha_1, \alpha_2, \alpha_5, \alpha_6, \alpha_7, \alpha_8\}, \\
S_{q \lor r} &= \{\alpha_1, \alpha_3, \alpha_4, \alpha_5, \alpha_7, \alpha_8\}, \\
S_{(p \lor q) \land \neg r} &= S_{p \lor q} \land S_{q \lor r} = \{\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\}, \\
S_{(p \land q) \land \neg r} &= \{\alpha_2\}.
\end{align*}
\]

Following the steps of the algorithm we obtain

\[
\begin{align*}
K_0 &= \{ p \vdash q, q \vdash r, p \land q \vdash \neg r \}, \\
A_0 &= \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\}, \\
s_1 &= A_0 \cap S_{p \lor q} \cap S_{q \lor r} \cap S_{(p \lor q) \land \neg r} \\
&= \{\alpha_1, \alpha_2, \alpha_5, \alpha_6, \alpha_7, \alpha_8\} \cap \{\alpha_1, \alpha_3, \alpha_4, \alpha_5, \alpha_7, \alpha_8\} \cap \{\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\} \\
&= \{\alpha_5, \alpha_7, \alpha_8\}
\end{align*}
\]

$s_1 \cap S_{p \land q} = s_1 \cap S_{(p \land q) \land \neg r} = \emptyset$, $s_1 \cap S_{q \land r} = \{\alpha_5\}$, so only confirmation is $q \vdash r$ by $\alpha_5$.

\[
\begin{align*}
K_1 &= \{ p \vdash q, p \land q \vdash \neg r \}, \\
A_1 &= A_0 - s_1 \\
&= \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_6\}, \\
s_2 &= A_1 \cap S_{p \lor q} \cap S_{(p \lor q) \land \neg r} \\
&= \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_6\} \cap \{\alpha_1, \alpha_2, \alpha_5, \alpha_6, \alpha_7, \alpha_8\} \cap \{\alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7, \alpha_8\} \\
&= \{\alpha_2, \alpha_6\}
\end{align*}
\]

$s_2 \cap S_{p \land q} = s_2 \cap S_{(p \land q) \land \neg r} = \{\alpha_2\}$, so both $p \vdash q$ and $(p \land q) \vdash \neg r$ are confirmed by $\alpha_2$.

---

2Unless otherwise stated assume in examples such as this that $L$ consists just of the propositional variables explicitly mentioned, in this case $p, q, r$. 
\[ K_2 = \emptyset, \]
\[ A_2 = A_1 - s_2 \]
\[ = \{\alpha_1, \alpha_3, \alpha_4\}, \]
\[ s_3 = A_2 \]
\[ = \{\alpha_1, \alpha_3, \alpha_4\}, \]

and there is nothing left to confirm at this stage.

Clearly since the \( s_i \) are disjoint and all the atoms have now been used up the next step will give \( s_4 = \emptyset \) and the algorithm will terminate. Formally we obtain

\[ K_3 = K_2 = \emptyset, \quad A_3 = A_2 - s_3 = \emptyset, \quad s_4 = A_3 = \emptyset. \]

(It is not always the case that the algorithm terminates because all the atoms have been used up — clearly, since we can have rational closures \( \vdash u \) where \( \bigcup_i u_i \subset \text{At}^L \).)

To sum up then, the rational closure of \( K \) is given by \( \vdash \bar{s} \) where \( \bar{s} \) is

\[ \{\alpha_5, \alpha_7, \alpha_8\}, \{\alpha_2, \alpha_6\}, \{\alpha_1, \alpha_3, \alpha_4\}. \]

### 3.4 Criticisms of Rational Closure

(i) Consider the knowledge base

\[ K = \{b \vdash f, p \vdash \neg f, p \land \neg b \vdash \bot, b \vdash w\} \]

where \( \bot \) is any contradiction, and \( b, p, f, w \) stand for ‘bird’, ‘penguin’, ‘flies’, and ‘has wings’.

Then the rational closure, \( \vdash_0 \), of \( K \) gives \( p \K_0 w \), i.e. does not satisfy \( p \vdash_0 w \) despite the fact that one might feel that the general property, ‘wingedness’ of birds should be inherited even by the exceptional class, penguins, of birds.

A similar example where an exceptional subclass fails, according to the rational closure, to inherit a general property of the class is given by the following knowledge base

\[ K = \{s \vdash t, s \vdash f\}, \]

where \( s, t, f \) stand for ‘Swedish’, ‘tall’, ‘fair’ respectively. In this case the rational closure, \( \vdash_0 \), of \( K \) fails to satisfy that ‘short Swedes are usually fair’, i.e. \( \neg t \land s \vdash f \), despite the fact that one might have felt the general property, fairness, of Swedes should have been inherited by the exceptional subclass of short Swedes.

Whilst both these examples (more generally the failure of inheritance to exceptional subclasses) have been cited as criticisms of the rational closure it would seem that
the question of the desirability, if any, of such inheritance cannot be argued for by examples like these where there is in addition also a great deal of unstated knowledge. (Like we already know penguins have wings.) Indeed the dangers of giving any weight to such examples can be clearly seen if we replace ‘has wings’ by ‘builds a nest’. Would we want to conclude (erroneously) that ‘penguins usually build nests’? [Similarly if instead we replaced ‘tall’ by ‘not born in Ethiopia’ in the Swedes example.]

(ii) Our second example of a criticism of the rational closure is rather more intriguing since it draws a conclusion we might feel is unjustified (rather than as above failing to draw a conclusion which we might feel is justified). Let

$$K = \{ f \land m \models w, \ l \models \neg w \}$$

where $f, m, w, l$ stand for ‘factory worker’, ‘manager’, ‘well-off’, and ‘drives a Lada’. Then for the rational closure, $\models_0$ of $K$, we do not have $f \land m \land l \models_0 w$. In other words the rational closure does not conclude that managers who work in factories and drive Ladas are well-off, despite the fact that without the additional information about the Lada this was in the knowledge base $K$. However suppose we now extend $K$ to $K \cup \{ f \models \neg w \}$. That is, we add to $K$ the knowledge that factory workers are usually not well-off.

Surprisingly if we now take the rational closure $\models_1$ of this new knowledge base we find that

$$f \land m \land l \models_1 w,$$

despite the fact that the new evidence would seem to further support the conclusion that managers who work in factories and drive Ladas are normally not well-off!
1. By using the Representation Theorem for Rational Consequence Relations show that the rule

\[
\frac{\psi \not\vdash \varphi, \; \psi \land \neg \theta \not\vdash \varphi}{\psi \land \theta \not\vdash \varphi}
\]

holds for all rational consequences relations but that the rule

\[
\frac{\psi \not\vdash \varphi}{\psi \land \theta \not\vdash \varphi}
\]

fails for some rational consequence relation (and choice of $\psi, \theta, \varphi$).

[10 marks]
Throughout we work with a finite language $L$.

1. (p13) Show that the following rules are derivable from the GM rules:

   \[
   \begin{array}{c}
   \theta \vdash \phi \\
   \phi \vdash \theta \\
   \theta \vdash \psi
   \end{array}
   \]

   \[
   \begin{array}{c}
   \phi \vdash \psi
   \end{array}
   \] (Equivalence)

   \[
   \begin{array}{c}
   \theta \vdash (\phi \rightarrow \psi) \\
   \theta \vdash \phi
   \end{array}
   \]

   \[
   \begin{array}{c}
   \theta \vdash \psi
   \end{array}
   \] (MP, Modus Ponens)

2. (p15) For $\vec{s} = s_1, s_2, \ldots, s_m \subseteq At^L$, show that $\models_{\vec{s}}$ satisfies the ‘Consistency Preservation Rule’:

   \[
   \begin{array}{c}
   \models \eta \\
   \theta \not\models \neg \eta
   \end{array}
   \]

   \[
   \begin{array}{c}
   \theta \not\models \neg \eta
   \end{array}
   \]

   if and only if $\bigcup_{i=1}^{m} s_i = At^L$.

3. (p17) Show that the rule

   \[
   \begin{array}{c}
   \theta \vdash \phi
   \end{array}
   \]

   \[
   \begin{array}{c}
   \neg \phi \vdash \neg \theta
   \end{array}
   \]

   cannot be derived from the GM rules.
Solutions for Example Sheet 1

Example 1  Show that the following rules are derivable from the GM rules:

| $\theta \vdash \phi$  | $\phi \vdash \theta$  | $\theta \vdash \psi$  | (Equivalence) |
| $\theta \vdash (\phi \rightarrow \psi)$  | $\theta \vdash \phi$  | $\theta \vdash \psi$  | (MP, Modus Ponens) |

Solution

| $\theta \vdash \phi$  | $\phi \vdash \psi$  | CMO |
| $\theta \land \phi \vdash \psi$  | $\phi \vdash \theta$  | CC |

| $\theta \vdash \phi \rightarrow \psi$  | $\theta \vdash \phi$  | $\theta \vdash \psi$  | AND |
| $\theta \vdash \phi \land (\phi \rightarrow \psi)$  | $\phi \land (\phi \rightarrow \psi) \models \psi$  | RWE |

Example 2 For $\vec{s} = s_1, s_2, \ldots, s_m \subseteq At^L$, show that $\vdash_{\vec{s}}$ satisfies the Consistency Preservation Rule:

$\models \eta$  $\theta \not\models -\eta$

if and only if $\bigcup_{i=1}^{m} s_i = At^L$.

Solution First suppose $\bigcup_{i=1}^{m} s_i \neq At^L$. Assume $\models \eta$ and pick $\alpha \in At^L \setminus \bigcup_{i=1}^{m} s_i$. Then $\alpha \not\models -\eta$, since $S_\alpha = \{\alpha\} \not\subseteq \emptyset = S_{-\eta}$. On the other hand $\alpha \models \neg \eta$, since $s_i \cap S_\alpha = \emptyset$ for all $i$. Hence $\alpha$ violates the Consistency Preservation Rule.

Conversely suppose $\bigcup_{i=1}^{m} s_i = At^L$ and let $\theta \in SL$ with $\theta \not\models \eta$, $\eta$ a tautology. This simply means $S_\eta \neq \emptyset$. Since $\bigcup_{i=1}^{m} s_i = At^L$, there is a smallest $i \in \{1, \ldots, m\}$ with $S_\theta \cap s_i \neq \emptyset$. As $S_{-\eta} = \emptyset$ we obtain $\theta \not\models_{\vec{s}} \neg \eta$. Hence $\vdash_{\vec{s}}$ satisfies the Consistency Preservation Rule.

Example 3 Show that the rule $\frac{\theta \vdash \phi}{\neg\phi \vdash \neg\theta}$ cannot be derived from the GM rules.

Solution Take atoms $\alpha \neq \beta$. Let $s_1 = \{\alpha\}$, $s_2 = \{\beta\}$ and $\vec{s} = (s_1, s_2)$. By definition of $\vdash_{\vec{s}}$ we get $\alpha \vee \beta \models_{\vec{s}} \alpha$ but $\neg \alpha \not\models_{\vec{s}} (\alpha \vee \beta)$ (note that $S_{-(\alpha \vee \beta)} = At^L \setminus \{\alpha, \beta\}$) Hence the rule can not be deduced from the GM-rules.
1. (p18) Is it true that if \( \vdash_1, \vdash_2 \) are rcr’s such that

\[ \vdash_1 \theta \iff \vdash_2 \theta \text{ for all } \theta \in SL \]

then \( \vdash_1 = \vdash_2 \)? (Compare this with mcr’s.)

2. (p21) Suppose that your knowledge base consists of:

- The poor normally work \((p \vdash w)\)
- Poor students normally don’t work \((p \land s \vdash \neg w)\)
- Teenagers are normally students \((t \vdash s)\)
- Teenagers are not normally rich \((t \not\vdash \neg p)\)

Assuming (as usual) that \( \vdash \) is a rational consequence relation, show that the poor are normally not teenagers \((p \not\vdash t)\) and the poor are normally not students \((p \not\vdash s)\).

Show that you cannot conclude from this knowledge base that students are normally not poor \((s \not\vdash \neg p)\), nor that students are normally poor \((s \not\vdash p)\).

3. (p21) By using the Representation Theorem for rcr’s, show that the following rules hold for all rcr’s:

\[
\begin{align*}
\text{(a)} & \quad \phi \vdash \eta \\
\text{(b)} & \quad \theta \lor \phi \not\vdash -\phi \lor \eta \\
\text{(c)} & \quad \phi \lor \theta \not\vdash -\lambda \lor \phi \not\vdash \gamma
\end{align*}
\]

4. (p21) Suppose that \( \vdash_1, \vdash_2 \) are non-trivial rcr’s on \( SL_1, SL_2 \) respectively, where \( L_1, L_2 \) are disjoint languages. Show that there is a rational consequence relation \( \vdash \) on \( SL \), where \( L = L_1 \cup L_2 \), such that for \( i = 1, 2 \),

\[ \theta \vdash \phi \iff \theta \vdash_i \phi \quad \text{for all } \theta, \phi \in SL_i \]
Solutions for Example Sheet 2

Example 1
Is it true that if $\mathcal{R}_1$, $\mathcal{R}_2$ are rcr's such that

$$\mathcal{R}_1 \theta \iff \mathcal{R}_2 \theta \text{ for all } \theta \in SL$$

then $\mathcal{R}_1 = \mathcal{R}_2$? (Compare this with mcr's.)

**Solution** The answer is no. Notice that if $\vec{s} = s_1, \ldots, s_m$ ($m \geq 1, s_1 \neq \emptyset$) then for $\theta \in SL$ we have $\mathcal{R}_\vec{s} \theta \iff s_1 \subseteq S_\theta$ (in more detail,

$$\mathcal{R}_\eta \theta \iff \eta \vdash \theta \text{ for some/any tautology } \eta \iff S_\eta \cap s_1 \subseteq S_\theta \iff s_1 \subseteq S_\theta$$

since for a tautology $\eta$, $S_\eta = \text{At}_L$). Hence if $\vec{t} = t_1, \ldots, t_q$ ($q \geq 1, t_1 \neq \emptyset$) is such that

$$\mathcal{R}_\vec{t} \theta \iff \mathcal{R}_\vec{\eta} \theta$$

But merely $s_1 = t_1$ does not force $\mathcal{R}_\vec{t} \eta$; take e.g. $L = \{p, q\},$ $s_1 = t_1 = \{p \land q\}$, $s_2 = \{\neg p \land q\},$ $t_2 = \{\neg p \land \neg q\};$ then

$$\neg p \vdash \neg q \text{ and } \neg p \not\vdash \neg q.$$ 

(For mcr's the answer is yes since for a mcr $\vdash$ we have $\Gamma \vdash \phi \iff \wedge \Gamma \vdash \phi \iff \vdash \wedge \Gamma \rightarrow \phi$.)

Example 2
Suppose that your knowledge base consists of:

- The poor normally work ($p \vdash w$)
- Poor students normally don’t work ($p \land s \vdash \neg w$)
- Teenagers are normally students ($t \vdash s$)
- Teenagers are not normally rich ($t \not\vdash \neg p$)

Assuming (as usual) that $\vdash$ is a rational consequence relation, show that the poor are normally not teenagers ($p \vdash \neg t$) and the poor are normally not students ($p \vdash \neg s$). Show that you cannot conclude from this knowledge base that students are normally not poor ($s \vdash \neg p$), nor that students are normally poor ($s \vdash p$).

**Solution** We are given

1. ($p \vdash w$)
2. ($p \land s \vdash \neg w$)
3. ($t \vdash s$)
4. ($t \not\vdash \neg p$)

From (2) by CON, $p \vdash s \rightarrow \neg w$ and with (1) and AND, $p \vdash w \land (s \rightarrow \neg w)$. Since $w \land (s \rightarrow \neg w) \models \neg s$, by RWE $p \vdash \neg s$. From (3) and (4) by RMO $p \land t \vdash s$ so by CON...
$p \models t \to s$ and with $p \models \neg s$ and AND, $p \models \neg s \land (t \to s)$. Since $\neg s \land (t \to s) \models \neg t$, by RWE $p \models \neg t$.

For the last part, it suffices (by Theorem 2.1) to find $\bar{s} = s_1, \ldots, s_m \subseteq At^L$, $\bar{i} = t_1, \ldots, t_q \subseteq At^L$ (where $L = \{p, w, s, t\}$) such that (1)-(4) hold for $\models^{\bar{s}}$, $\models^{\bar{i}}$ and $s \not\models^{\bar{s}} p$, $s \not\models^{\bar{i}} p$. In fact this can be achieved if we take $\bar{s} = i$, and

$$s_1 = \{p \land w \land \neg s \land \neg t\}, \quad s_2 = \{p \land \neg w \land s \land \neg t, \quad \neg p \land \neg w \land s \land \neg t\}.$$ 

**Example 3** By using the Representation Theorem for rcr's, show that the following rules hold for all rcr's:

1. $\phi \models \eta$ 
   $\theta \lor \phi \models \neg \phi \lor \eta$

2. $\theta \not\models \psi$ 
   $\phi \not\models \psi$

3. $\phi \lor \theta \models \neg \lambda$ 
   $\phi \not\models \lambda$

**Solution** By the representation theorem 2.2 every rcr is of the form $\models^{\bar{s}}$ for some $\bar{s} = s_1, \ldots, s_m \subseteq At^L$ so to show that (i) - (iii) hold for every rcr it suffices to show in each case that if the hypotheses of the rule hold for $\models^{\bar{s}}$ then so does the conclusion.

(i) Assume $\phi \models^{\bar{s}} \eta$. If for all $i$, $s_i \cap S_{\theta \lor \phi} = \emptyset$ then $\theta \lor \phi \models^{\bar{s}} \neg \phi \lor \eta$, so assume that there is some $i$ such that $s_i \cap S_{\theta \lor \phi} \neq \emptyset$ and let $i$ be the least such. Let

$$\alpha \in s_i \cap S_{\theta \lor \phi} = s_i \cap (S_{\theta} \cup S_{\phi}) = (s_i \cap S_{\theta}) \cup (s_i \cap S_{\phi}).$$

If $\alpha \in s_i \cap S_{\theta}$ then $s_i \cap S_{\phi} \neq \emptyset$ and $i$ must be minimal such that this holds (since for $j < i$, $s_j \cap S_{\phi} \subseteq s_j \cap (S_{\theta} \cup S_{\phi}) = s_j \cap S_{\theta \lor \phi} = \emptyset$ by the choice of $i$). Consequently, since $\phi \models^{\bar{s}} \eta$, $s_i \cap S_{\phi} \subseteq S_{\eta}$ so $\alpha \in S_{\eta} \subseteq S_{\neg \phi} \cup S_{\eta} = S_{\neg \phi \lor \eta}$. Otherwise, if $\alpha \not\in s_i \cap S_{\theta}$ then $\alpha \in s_i \cap S_{\theta}$ (since $\alpha \in (s_i \cap S_{\theta}) \cup (s_i \cap S_{\phi})$), so, trivially, also $\alpha \in s_i$ and $\alpha \not\in S_{\phi}$. Hence $\alpha \in S_{\neg \phi} \subseteq S_{\neg \phi \lor \eta}$. Either way $\alpha \in S_{\neg \phi \lor \eta}$ so $s_i \cap S_{\theta \lor \phi} \subseteq S_{\neg \phi \lor \eta}$ and $\theta \lor \phi \models^{\bar{s}} \neg \phi \lor \eta$ follows.

(ii) Assume $\theta \not\models^{\bar{s}} \psi$, $\phi \not\models^{\bar{s}} \psi$. Clearly there must be some $i$ such that $s_i \cap S_{\theta} \neq \emptyset$ (otherwise $s_i \cap S_{\theta} = \emptyset$ for all $i$ so $\theta \models^{\bar{s}} \psi$). Let $i$ be the least such and similarly let $j$ be minimal such that $s_j \cap S_{\phi} \neq \emptyset$. Then since $s_k \cap S_{\theta \lor \phi} = (s_k \cap S_{\theta}) \cup (s_k \cap S_{\phi})$, $s_k \cap S_{\theta \lor \phi}$ will be empty for $k < \min\{i, j\}$. Without loss of generality, assume that $i = \min\{i, j\}$ so $i$ is minimal such that $s_i \cap S_{\theta \lor \phi} \neq \emptyset$. Then since $\theta \not\models^{\bar{s}} \psi$, there is $\alpha \in s_i \cap S_{\theta}$ such that $\alpha \not\in S_{\psi}$. But then $\alpha \in s_i \cap S_{\theta \lor \phi} \subseteq s_i \cap S_{\theta}$ so $\theta \lor \phi \not\models^{\bar{s}} \psi$ as required.
(iii) Assume $\phi \lor \theta \vdash \neg \lambda, \phi \nvdash \gamma$. As in (ii) there must be $i$ such that $s_i \cap S_{\phi} \neq \emptyset$ so since $S_{\phi} \subseteq S_{\phi \lor \theta}$, there is $i$ such that $s_i \cap S_{\phi \lor \theta} \neq \emptyset$. Let $i$ be the least such and let $\alpha \in s_i \cap S_{\phi \lor \theta}$. Then, since $\phi \lor \theta \vdash \neg \lambda, \alpha \in S_{\neg \lambda}$ so $\alpha \notin S_{\lambda}$ and $\phi \lor \theta \nvdash \lambda$ as required.

**Example 4** Suppose that $\vdash_1, \vdash_2$ are non-trivial rcr’s on $SL_1, SL_2$ respectively, where $L_1, L_2$ are disjoint languages. Show that there is a rational consequence relation $\vdash$ on $SL$, where $L = L_1 \cup L_2$, such that for $j = 1, 2$,

$$\theta \vdash \phi \iff \theta \vdash_j \phi \quad \text{for all } \theta, \phi \in SL_i$$

**Solution** Without loss of generality, to simplify the notation, we may assume that

$$At^L = \{\alpha \land \beta \mid \alpha \in At^{L_1} \land \beta \in At^{L_2}\}.$$

Let $\vdash_j = \vdash_{s_j}$ where $s_j = s_{j1}, s_{j2}, \ldots, s_{jn_j} \subseteq At^{L_j}$ ($j = 1, 2$). Since $\vdash_j$ is non-trivial, not all the $s_{ji}$ are empty. So removing if necessary the empty ones we may assume all the $s_{ji} \neq \emptyset$. We may also assume $n_1 = n_2$ (say) since if for example $n_1 < n_2$ then we can add $n_2 - n_1$ copies of $s_{1_{n_1}}$ to the end of $s_{11}, s_{21}, \ldots, s_{n_1}$ without changing the corresponding rcr.

Now for $i = 1, 2, \ldots, n$ define $s_i = \{\alpha \land \beta \mid \alpha \in s_{i1} \land \beta \in s_{i2}\}$. Then for $\theta, \phi \in SL_1 (\subseteq SL)$ and $\alpha \in s_{i1}, \beta \in s_{i2}$,

$$\alpha \land \beta \in s_i \cap S_{\phi}^L \iff \alpha \in s_{i1} \land \beta \in s_{i2} \land V_{\alpha \land \beta}^L(\theta) = 1$$

$$\iff \alpha \in s_{i1} \land \beta \in s_{i2} \land V_{\alpha}^L(\theta) = 1$$

($Exercise, \ by \ induction \ on \ k \ for \ \theta \in (SL_1)_k$)

$$\iff \alpha \in s_{i1} \cap S_{\theta}^{L_1} \land \beta \in s_{i2}^L.$$

Hence

$$s_i \cap S_{\phi}^L = \{\alpha \land \beta \mid \alpha \in s_{i1} \cap S_{\theta}^{L_1} \land \beta \in s_{i2}\}$$

and since $s_{i2} \neq \emptyset$, $\exists i(s_i \cap S_{\phi}^L \neq \emptyset) \iff \exists i(s_{i1} \cap S_{\theta}^{L_1} \neq \emptyset)$. Furthermore, if such an $i$ exists then the minimal $i$ such that $s_i \cap S_{\phi}^L \neq \emptyset$ equals the minimal $i$ such that $s_{i1} \cap S_{\theta}^{L_1} \neq \emptyset$ and for this least $i$

$$s_i \cap S_{\phi}^L \subseteq S_{\phi}^L \iff V_{\alpha \land \beta}^L(\theta) = 1 \text{ for all } \alpha \land \beta \in s_i \cap S_{\phi}^L$$

$$\iff V_{\alpha}^L(\theta) = 1 \text{ for all } \alpha \in s_{i1} \cap S_{\theta}^{L_1}, \beta \in s_{i2}^L$$
\[\Leftrightarrow \alpha \in S_{\phi}^{L_1} \text{ for all } \alpha \in s_{i}^{1} \cap S_{\theta}^{L_1} \text{ (since } s_{i}^2 \neq \emptyset)\]

\[\Leftrightarrow s_{i}^{1} \cap S_{\theta}^{L_1} \subseteq S_{\phi}^{L_1}.\]

By substituting these equivalents into the definition of \(\theta \vdash \phi\) we now obtain that for \(\theta, \phi \in SL_1\),

\[\theta \vdash_{s} \phi \Leftrightarrow \theta \vdash_{s^1} \phi \Leftrightarrow \theta \vdash_{1} \phi.\]

By symmetry the result also holds for \(\vdash_{2}\) so \(\vdash_{s}\) is the required rcr on SL.
1. (p21) By using the Representation Theorem (Theorem 2.2) for rcr’s, show that the following rules hold for all rcr’s.

   (a) \( \theta \lor \phi \vdash \bot \) where \( \bot \) is any contradiction, i.e. \( S_\bot = \emptyset \).

   (b) \( \theta \lor \phi \vdash \psi \) \( \phi \not\vdash \gamma \)

   \( \theta \lor (\phi \land \neg \gamma) \vdash \psi \)

2. (p21) Show that for any rcr \( \vdash \) (on \( SL \)) there exists a unique \( \vec{s} = s_1, s_2, \ldots, s_m \subseteq At^L \) such that

   (a) \( \vdash_{\vec{s}} = \vdash \)

   (b) the \( s_i \) are disjoint and non-empty. (Harder.)

3. (p21) Let \( L, L' \) be languages such that \( L \subseteq L' \) and let \( \vdash' \) be an rcr on \( SL' \). Show that the relation \( \vdash \) defined on \( SL \) by \( \theta \vdash \phi \iff \theta \vdash' \phi \) for \( \theta, \phi \in SL \) (\( \subseteq SL' \)) is a rational consequence relation. Given that \( \vdash' = \vdash_{\vec{s}} \) where \( \vdash_{\vec{s}} = s_1, s_2, \ldots, s_m \subseteq At^L' \), find \( \vec{t} = t_1, t_2, \ldots, t_m \subseteq At^L \) (directly from \( \vec{s} \)) such that \( \vdash = \vdash_{\vec{t}} \).

4. (p30) Use the Z-algorithm to find the rational closure of \( K = \{ p \vdash \neg q, \vdash p \rightarrow q \} \).

   \[ \text{[Take } L = \{p, q\}. \]
Solutions for Example Sheet 3

Example 1

By using the Representation Theorem 2.2 for rational consequence relations, show that the following rules hold for all rational consequence relations.

\[
\begin{align*}
(i) \quad & \theta \lor \phi \vdash \bot \\
& \theta \vdash \bot \\
(ii) \quad & \theta \lor \phi \vdash \psi \quad \phi \not\vdash \gamma \\
& \theta \lor (\phi \land \neg \gamma) \vdash \psi
\end{align*}
\]

**Solution** By the representation theorem 2.2 every rational consequence relation is of the form \( \vdash_s \) for some \( s = s_1, s_2, \ldots, s_m \subseteq At^L \), so it suffices to check that these rules hold for such an \( \vdash_s \).

(i) Suppose \( \theta \lor \phi \vdash_s \bot \). If \( s_i \cap S_{\theta \lor \phi} \neq \emptyset \) for some \( i \) then, by \( \theta \lor \phi \vdash_s \bot \) for the least such \( i \), \( s_i \cap S_{\theta \lor \phi} \subseteq S_\bot = \emptyset \), contradiction. Consequently, \( s_i \cap S_{\theta \lor \phi} = \emptyset \) for all \( i \) so \( s_i \cap S_\theta = \emptyset \) for all \( i \) since \( S_\theta \subseteq S_\theta \lor S_\phi = S_{\theta \lor \phi} \). It follows that \( \theta \vdash_s \bot \) as required.

(ii) Suppose \( \theta \lor \phi \vdash_s \psi \) and \( \phi \not\vdash_s \gamma \). If \( s_i \cap S_\phi = \emptyset \) for all \( i \) then \( \phi \not\vdash_s \gamma \), contradiction. So let \( i_0 \) be minimal such that \( s_{i_0} \cap S_\phi \neq \emptyset \). If \( s_{i_0} \cap S_\phi \subseteq S_\gamma \) then \( \phi \vdash_s \gamma \), contradiction, so there must be some \( \alpha \in s_{i_0} \cap S_\phi \), \( \alpha \notin S_\gamma \), equivalently, \( \alpha \in S_{\neg \gamma} \). It follows that \( \emptyset \neq s_{i_0} \cap S_\phi \cap S_{\neg \gamma} = s_{i_0} \cap S_{\phi \land \neg \gamma} \) and clearly \( i_0 \) is the least such since for \( k < i_0 \) we have \( s_k \cap S_{\phi \land \neg \gamma} \subseteq s_k \cap S_\phi = \emptyset \) (by the choice of \( i_0 \)). Also,

\[
s_{i_0} \cap S_{\theta \lor (\phi \land \neg \gamma)} = s_{i_0} \cap (S_\theta \lor S_{\phi \land \neg \gamma}) \supseteq s_{i_0} \cap S_{\phi \land \neg \gamma} \neq \emptyset
\]

so the minimal \( j_0 \) such that \( s_{j_0} \cap S_{\theta \lor (\phi \land \neg \gamma)} \neq \emptyset \) exists and \( j_0 \leq i_0 \). Clearly \( s_{j_0} \cap S_{\theta \lor \phi} \neq \emptyset \) (since \( \theta \lor (\phi \land \neg \gamma) \vdash \theta \lor \phi \), \( S_{\theta \lor \phi} \supseteq S_{\theta \lor (\phi \land \neg \gamma)} \)) and \( j_0 \) is the least such since if \( k < j_0 \) then

\[
s_k \cap S_\theta \subseteq s_k \cap S_{\theta \lor (\phi \land \neg \gamma)} = \emptyset \quad (\text{since } k < j_0),
\]

\[
s_k \cap S_\phi = \emptyset \quad (\text{since } k < j_0 \leq i_0),
\]

\[
s_k \cap S_{\theta \lor \phi} = s_k \cap (S_\theta \lor S_\phi) = (s_k \cap S_\theta) \cup (s_k \cap S_\phi) = \emptyset.
\]

Consequently, since \( \theta \lor \phi \vdash_s \psi \),

\[
S_\psi \supseteq s_{j_0} \cap S_{\theta \lor \phi} \supseteq s_{j_0} \cap S_{\theta \lor (\phi \land \neg \gamma)}
\]

so \( \theta \lor (\phi \land \neg \gamma) \vdash_s \psi \) as required.

Example 2 Show that for any rational consequence relation \( \vdash \) (on \( SL \)) there exists a unique \( \vec{s} = s_1, s_2, \ldots, s_m \subseteq At^L \) such that \( \vdash = \vdash_\vec{s} \) and the \( s_i \) are disjoint and non-empty.
Solution Let $\vdash$ be a rational consequence relation. By the Representation Theorem 2.2 and comments following it, let $\vec{s} = s_1, s_2, \ldots, s_m \subseteq A^L$ be such that $\vdash = \vdash_\vec{s}$ and the $s_i$ are disjoint and non-empty.

(Details: By the theorem 2.2 some $\vec{s} = s_1, s_2, \ldots, s_m \subseteq A^L$ exists such that $\vdash = \vdash_\vec{s}$ holds. Replacing each $s_k$ in $\vec{s}$ by $s_k - \bigcup_{j<k} s_j$ does not change $\vdash_\vec{s}$, since

$$s_k \cap S_\emptyset \neq \emptyset \text{ and } s_j \cap S_\emptyset = \emptyset \text{ for all } j < k$$

$$\iff (s_k - \bigcup_{j<k} s_j) \cap S_\emptyset \neq \emptyset \text{ and } (s_j - \bigcup_{i<j} s_i) \cap S_\emptyset = \emptyset \text{ for all } j < k$$

and for such $k, \emptyset$ we have $s_k \cap S_\emptyset = (s_k - \bigcup_{j<k} s_j) \cap S_\emptyset$ so

$$s_k \cap S_\emptyset \subseteq S_\emptyset \iff (s_k - \bigcup_{j<k} s_j) \cap S_\emptyset \subseteq S_\emptyset.$$ 

In the resulting sequence the sets of atoms are disjoint since if $i < k$ then

$$\left(s_i - \bigcup_{j<i} s_j\right) \cap \left(s_k - \bigcup_{j<k} s_j\right) \subseteq s_i \cap (s_k - s_i) = \emptyset.$$

Also, removing empty sets from the resulting sequence does not change $\vdash_\vec{s}$ and hence we may assume the $s_i$ disjoint and non-empty.)

Suppose that $\vec{t} = t_1, t_2, \ldots, t_q, \vec{t} \neq \vec{s}$ is also such that $\vdash_\vec{t} = \vdash$ and the $t_i$ are disjoint and non-empty. Without loss of generality, assume $m \leq q$. Let $\vec{s}^+$ be the result of adding $s_i = \emptyset$ for $i = m + 1, \ldots, q$ to $\vec{s}$. Then $\vdash_\vec{s}^+ = \vdash_\vec{t}$ and there is some $i \leq q$ such that $t_i \neq s_i$ (since either $m = q$ and $t_1, \ldots, t_m \neq s_1, \ldots, s_m$ or $m < q$ in which case $t_{m+1} \neq \emptyset$, $s_{m+1} = \emptyset$). Let $i$ be the least such, and let $\alpha \in t_i$, $\alpha \not\in s_i$ (the case of $\alpha \in s_i$, $\alpha \not\in t_i$ is similar, easier). We have $t_i \cap S_{\bigvee s_i \vee t_i} = t_i \cap (s_i \cup t_i) = t_i \neq \emptyset$ and for $j < i$

$$t_j \cap S_{\bigvee s_i \vee t_i} = t_j \cap (s_i \cup t_i) = (t_j \cap s_i) \cup (t_j \cap t_i) = \emptyset$$

since $t_j = s_j$ (as $j < i$) and the $s_k$ are disjoint and the $t_k$ are disjoint. Hence $\bigvee s_i \vee t_i \not\vdash \lnot \alpha$ since $\alpha \not\in S_{-\alpha}$ but $\alpha \in t_i = t_i \cap S_{\bigvee s_i \vee t_i}$. If $s_i \neq \emptyset$ (so $i \leq m$) then as above

$$s_i \cap S_{\bigvee s_i \vee t_i} = s_i \neq \emptyset, \quad s_j \cap S_{\bigvee s_i \vee t_i} = \emptyset \quad \text{for } j < i$$

from which we now obtain $\bigvee s_i \vee t_i \not\vdash_\vec{s} \lnot \alpha$ since in this case $\alpha \not\in s_i \cap S_{\bigvee s_i \vee t_i}$, contradicting $\vdash_\vec{t} = \vdash_\vec{s}$.

Example 3 Let $L, L'$ be languages such that $L \subseteq L'$ and let $\vdash'$ be a rational consequence relation on $SL'$. Show that the relation $\vdash$ defined on $SL$ by $\emptyset \vdash \phi \iff \emptyset \vdash' \phi$ for $\emptyset, \phi \in SL \subseteq SL'$ is a rational consequence relation. Given that $\vdash' = \vdash_\vec{s}$ where
\( \vec{s} = s_1, s_2, \ldots, s_m \subseteq At^{L'} \), find \( \vec{t} = t_1, t_2, \ldots, t_m \subseteq At^L \) (directly from \( \vec{s} \)) such that \( \vdash = \vdash_{\vec{t}} \).

**Solution** Without loss of generality assume \( L = \{ p_1, \ldots, p_m \}, L' = \{ p_1, \ldots, p_k \} \) \( (k \geq m) \). \( \vdash \) is a rational consequence relation since the GM rules and the REF axiom hold for \( \vdash' \) when the sentences in the rules and REF are restricted to being in SL \( (\subseteq SL') \) trivially, and hence hold for \( \vdash \). Now suppose \( \vdash' = \vdash_{\vec{s}} \). For \( \beta = \bigwedge_{i=1}^{k} p_i^{\vec{s}} \in At^{L'} \) let \( \tilde{\beta} = \bigwedge_{i=1}^{m} p_i^{\vec{t}} \in At^L \). Notice that for \( \theta, \phi \in SL \)

\[
\beta \in S_{\phi}^{L'} \\
\downarrow \\
V_{\beta}(\theta) = 1 \text{ where } V_{\beta} \text{ is the valuation on } L' \text{ given by } V_{\beta}(p_i) = \varepsilon_i, \ i = 1, \ldots, k \\
\downarrow \text{(Exercise)}
\]

\[
\tilde{\beta} \in S_{\phi}^{L}
\]

Let \( t_i = \{ \tilde{\beta} | \beta \in s_i \subseteq At^L \} \) for \( i = 1, \ldots, m \) and let \( \theta, \phi \in SL \). Then

\[
S_{\phi}^{L'} \cap s_i = \{ \beta \in At^{L'} | \beta \in s_i \& \beta \in S_{\phi}^{L'} \} \\
= \{ \beta \in At^{L'} | \beta \in s_i \& \tilde{\beta} \in S_{\phi}^{L} \}
\]

\[
S_{\phi}^{L} \cap t_i = \{ \alpha \in At^L | \alpha \in t_i \& \alpha \in S_{\phi}^{L} \} \\
= \{ \alpha \in At^L | \alpha = \tilde{\beta} \text{ for some } \beta \in s_i \& \alpha \in S_{\phi}^{L} \} \\
= \{ \tilde{\beta} \in At^L | \beta \in s_i \& \tilde{\beta} \in S_{\phi}^{L} \}
\]

It follows that

\[
S_{\phi}^{L'} \cap s_i \neq \emptyset \iff S_{\phi}^{L} \cap t_i \neq \emptyset \tag{E3.1}
\]

and

\[
S_{\phi}^{L'} \cap s_i \subseteq S_{\phi}^{L} \iff S_{\phi}^{L'} \cap s_i \cap (At^{L'} - S_{\phi}^{L'}) = \emptyset \\
\iff S_{\theta \land \neg \phi} \cap s_i = \emptyset \\
\iff S_{\theta \land \neg \phi} \cap t_i = \emptyset \quad \text{(by(E3.1) for } \theta \land \neg \phi \in SL) \\
\iff S_{\phi}^{L} \cap t_i \subseteq S_{\phi}^{L}.
\]

Hence, directly from the definition of \( \vdash_{\vec{s}}, \vdash_{\vec{t}} \) for \( \theta, \phi \in SL \),
\[ \theta \models \phi \iff \theta \models_{\mathfrak{r}} \phi. \]

**Example 4** Use the Z-algorithm to find the rational closure of \( \{ p \models \neg q, \models p \rightarrow q \} \).

[Take \( L = \{p, q\} \).]

**Solution**

\[ K_0 = K = \{ p \models \neg q, \models p \rightarrow q \}, \quad A_0 = \{ p \land q, p \land \neg q, \neg p \land q, \neg p \land \neg q \}. \]

So

\[
\begin{align*}
u_1 & = \{ \alpha \in A_0 \mid \alpha \notin S_{p \land \neg q}, \alpha \notin S_{\neg (p \rightarrow q)} \} \\
& = A_L \cap S_{\neg (p \land \neg q)} \cap S_{\neg (p \rightarrow q)} \\
& = S_{\neg p \lor q} \cap S_{\neg p \lor q} \\
& = \{ \neg p \land q, \neg p \land \neg q \}.
\end{align*}
\]

Both \( \neg p \land q \) and \( \neg p \land \neg q \in u_1 \) confirm \( \models p \rightarrow q \in K_0 \) but nothing in \( u_1 \) confirms \( p \models \neg q \in K_0 \). Hence

\[ K_1 = \{ p \models \neg q \}, \quad A_1 = A_0 - u_1 = \{ p \land q, p \land \neg q \}, \]

and

\[
\begin{align*}
u_2 & = \{ \alpha \in A_1 \mid \alpha \notin S_{p \land \neg q} \} \\
& = \{ \alpha \in \{ p \land q, p \land \neg q \} \mid \alpha \notin S_{p \land q} \} \\
& = \{ p \land \neg q \}.
\end{align*}
\]

\( p \land \neg q \in u_2 \) confirms \( p \models \neg q \in K_1 \) so since this is the only element of \( K_1, K_2 = \emptyset \). \( A_2 = A_1 - u_2 = \{ p \land q \} \). Since \( K_2 = \emptyset \) the single atom \( p \land q \in A_2 \) is not in any \( S_{\theta \land \neg \phi} \) for \( \theta \models \phi \in K_2 \) (because there aren’t any \( \theta \models \phi \in K_2 \)!) so \( u_3 = \{ p \land q \} \).

Now \( A_3 = A_2 - u_3 = \emptyset \) so \( u_4 = \emptyset \) since \( u_4 \subseteq A_3 \). Hence the required rational closure of \( K \) is \( \models_{\mathfrak{u}} \) where

\[ \mathfrak{u} = \{ \neg p \land q, \neg p \land \neg q \}, \quad \{ p \land \neg q \}, \quad \{ p \land q \}. \]