Defaults and Revision in Structured Theories*

Mark Ryan†

Abstract

Starting from a logic which specifies how to make deductions from a set of sentences (a 'flat theory'), a way to generalise this to a partially ordered bag of sentences (a 'structured theory') is given. The partial order is used to resolve conflicts. If ϕ occurs below ψ then ψ is accepted only insofar as it does not conflict with ϕ.

We start with a language L, a set of interpretations $\mathcal{M}$ and a satisfaction relation $\Vdash \subseteq \mathcal{M} \times L$. The key idea is to define, for each structured theory, a pre-order on interpretations. Models of the structured theory are defined to be maximal interpretations in the ordering. They are shown to exist if the logic $\langle L, \mathcal{M}, \Vdash \rangle$ is compact.

A revision operator is defined, which takes a structured theory and a sentence and returns a structured theory. The consequence relation has the properties of weak monotonicity, weak cut and weak reflexivity with respect to this operator, but fails their strong counterparts.

1 Introduction

Ordering sentences in a theory presentation may be used to specify how conflicts between sentences are resolved. This idea has applications in artificial intelligence (default reasoning) as well as in software specification. We show how to find consequences of such *structured* theory presentations and how to revise them with new and potentially conflicting information while retaining consistency. Most of the paper is devoted to the question of how to reason with structured theory presentations.

This paper is entirely about finite sets of sentences, possibly with structure. Perhaps it would be more correct to call them "theory presentations". Structured theories should also then be called "structured theory presentations", but this is too long-winded. Therefore, 'theory' and 'structured theory' are used to abbreviate 'theory presentation' and 'structured theory presentation'.

The paper is arranged as follows. First we give examples of the intended behaviour of structured theories (section 2). Then, in section 3, the logical setting is introduced. Section 4 is the main section, and is about the definitions required to produce the desired behaviour, and their properties. Section 5 shows how to revise structured theories with new and possibly conflicting information, and discusses properties of the revision operator. Next, in section 6, we make comparisons with other work on defaults and theory revision. Finally, some applications are outlined in the last section.

2 Motivating examples

Intuitively, a structured theory is a finite set of sentences equipped with a partial order (but because the same sentence can occur several times, each in a different place of the ordering, the formal definition is more complicated (see section 4)). If the sentences are mutually consistent then it is safe to ignore the partial order. The models of such a structured theory are just the models of the set of sentences. But if the sentences conflict, sentences lower in the ordering are to be treated as having greater weight or priority. This does not mean that a sentence high in the ordering can be ignored, even if it conflicts with sentences below it; some 'components' of it may still be needed in determining the models of the theory. One of the principle aims of this paper is to formalise this notion of component. The following examples illustrate this discussion, showing the intended behaviour of structured theories. The reader can check them against his or her intuitions. All of them work out successfully in the theory described in this paper. For each example, first we give the structured theory; then the flat theory to which it is equivalent.

*A shorter version of this paper appeared in the proceedings of the IEEE conference on **Logic in Computer Science**, Amsterdam, July 1991. This revision is dated 14 September 1991.

†Department of Computing, Imperial College, London SW7 2BZ. E-mail: mdr@doc.ic.ac.uk. Phone: +44 71 589 5111 ext. 5074.

Example 2.1

$$\begin{array}{c} p \\ \uparrow \\ \neg p \end{array} \quad \equiv \quad \neg p$$

This theory says: we want $\neg p$ (remember, the bottom sentences are the most important), and, subject to that, we want as much of the import of p as possible. Since p is atomic, we can't extract anything of it which does not conflict with $\neg p$, so all we can deduce is $\neg p$. But note, this analysis is not valid if p is replaced by an arbitrary ϕ, as example 2.2 shows.

Of course, the partial order is important here. If it was the other way around, the structured theory would be equivalent to p; if the two sentences were incomparable in the ordering, nothing interesting could be deduced:

$$\begin{array}{c} \neg p \\ \uparrow \\ p \end{array} \quad \equiv \quad p; \qquad (p \quad \neg p) \quad \equiv \quad \top$$

The left-hand equation shows that p dominating $\neg p$ is equivalent to p. On the right, we see that if p and $\neg p$ are incomparable in the order then we must remain agnostic about p. The idea is to extract what we can from a structured theory without allowing contradictions. Notice that this means that a structured theory in which all the sentences are incomparable is not the same as the flat theory formed from the same sentences.

Example 2.2

$$\begin{array}{c} p\wedge q \\ \uparrow \\ \neg p\vee\neg q \end{array} \quad \equiv \quad p\leftrightarrow\neg q$$

This is similar to the first example, since $\neg p\vee\neg q$ is identical to $\neg(p\wedge q)$ in the underlying logic (classical propositional logic in this case). We want $\neg(p\wedge q)$, and subject to that we want as much of $p\wedge q$. What we can have is either p or q but not both.

Example 2.3

$$\begin{array}{c} p\wedge q \\ \uparrow \\ \neg p \end{array} \quad \equiv \quad \neg p\wedge q$$

We want $\neg p$, and subject to that, as much of $p\wedge q$ as possible. $p\wedge q$ does conflict with $\neg p$, so we can't have it all. But we can have the q component. Of course the ordering is significant:

$$\begin{array}{c} \neg p \\ \uparrow \\ p\wedge q \end{array} \quad \equiv \quad p\wedge q; \qquad (\neg p \quad p\wedge q) \quad \equiv \quad q$$

Example 2.4

$$\begin{array}{c} p\vee q \\ \uparrow \\ \neg q \end{array} \quad \equiv \quad p\wedge\neg q$$

Here, since $p\vee q$ and $\neg q$ are consistent with each other, we can simply have them both and it doesn't matter how they are ordered:

$$\begin{array}{c} \neg q \\ \uparrow \\ p\vee q \end{array} \quad \equiv \quad p\wedge\neg q; \qquad (\neg q \quad p\vee q) \quad \equiv \quad p\wedge\neg q$$

Example 2.5

$$\begin{array}{c} \forall x.\, p(x) \\ \uparrow \\ \exists x.\, \neg p(x) \end{array} \quad \equiv \quad \begin{array}{l} \exists x.\, (\neg p(x)\wedge \\ \quad \forall y.\, (x\neq y\rightarrow p(y))) \end{array}$$

Of course, the ordering matters. If the two sentences $\forall x.\, p(x)$ and $\exists x.\, \neg p(x)$ are incomparable in the ordering, then one can conclude that there is one element whose claim to the property p is disputed, but that all other elements have the property p.

$$\begin{array}{l} (\forall x.\, p(x) \quad \exists x.\, \neg p(x)) \\ \qquad \equiv \quad \exists x.\, \forall y.\, (x\neq y\rightarrow p(y)) \end{array}$$

Example 2.6

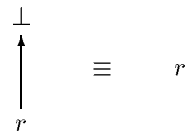
$$\begin{array}{ccc} p & & q \\ & \nwarrow \quad \nearrow & \\ & \neg(p\wedge q) & \end{array} \quad \equiv \quad p\leftrightarrow\neg q$$

There seems no reason to treat this differently from example 2.2. In general, is it possible to squash trees into linear orders in this way? The following example answers this question negatively.

Example 2.7

$$\begin{array}{ccc} p & & \neg p\wedge q \\ & \nwarrow \quad \nearrow & \\ & r & \end{array} \quad \equiv \quad q\wedge r$$

It is not possible to reduce non-linear partial orders to linear ones by zipping them up with $\wedge$s, since



The intuitions for non-linear partial orders seem to depend on whether the branches share non-logical language or not. This is important in specification theory applications (section 7).

These examples serve as a benchmark for the development of the system for dealing with structured theories given in this paper. The other important criterion to apply to a system of structured theories is that of *independence of the underlying logic*. This means that the system should not change the meanings of the connectives or introduce hacks which interfere with the mechanism of the underlying logic. Rather, it should be defined 'on top' of it. For example, substitution of logical equivalents at any point of a structured theory should not change its meaning, as mentioned in the discussion of example 2.2.

3 Prerequisites

The definitions we give in section 4 apply to any logic which is given in terms of language interpretations and a satisfaction relation, subject to being able to define the standard notion of positive and negative occurrences of non-logical symbols. Such logics include classical, intuitionistic and modal logics, in their propositional and predicate forms; Horn clause logic; equational logic, action logic and a host of others. We keep to this level of generality (that of concrete *institutions*—see [3]) for most of this paper as far as the definitions and results are concerned. The examples are mostly from classical propositional and classical predicate logic.

In this section, some standard definitions are given to which it will be useful to refer later.

Definition 3.1 A *language* L is (i) a finite set of logical connectives, (ii) a (possibly sorted) collection of non-logical symbols and (iii) a set of rules for forming L-sentences. L considered as a set is the set of L-sentences.

Definition 3.2 A *interpretation system* $\langle \mathcal{M}, \Vdash \rangle$ for a language L is a set $\mathcal{M}$ of *interpretations* and a relation (called *satisfaction*) $\Vdash \subseteq \mathcal{M} \times L$.

Example 3.3 (Classical propositional logic) L has (i) the connectives $\{\wedge, \vee, \rightarrow, \leftrightarrow, \neg, \bot, \top\}$; a set atoms($L$) of propositional atoms; and (iii) the following rules for sentence formation: if $p \in$ atoms(L) and ϕ and ψ are sentences then $\top$, $\bot$, p, $\neg\phi$, $\phi\wedge\psi$, $\phi\vee\psi$, $\phi\rightarrow\psi$ and $\phi\leftrightarrow\psi$ are all sentences. $\mathcal{M}$ consists of assignments of truth values to propositional atoms; if $M \in \mathcal{M}$ then $M : \text{atoms}(L) \rightarrow \{\mathbf{t}, \mathbf{f}\}$. The satisfaction relation is defined in the standard way:

$$\begin{array}{l}
M \Vdash \top \\
M \nVdash \bot \\
M \Vdash p \text{ iff } M(p) = \mathbf{t} \text{ and } p \in \text{atoms}(L) \\
M \Vdash \neg\phi \text{ iff } M \nVdash \phi \\
M \Vdash \phi\wedge\psi \text{ iff } M \Vdash \phi \text{ and } M \Vdash \psi \\
M \Vdash \phi\vee\psi \text{ iff } M \Vdash \phi \text{ or } M \Vdash \psi \\
M \Vdash \phi\rightarrow\psi \text{ iff } M \Vdash \phi \text{ implies } M \Vdash \psi \\
M \Vdash \phi\leftrightarrow\psi \text{ iff } (M \Vdash \phi \text{ iff } M \Vdash \psi)
\end{array}$$

Example 3.4 (Classical predicate logic) L has (i) each of the connectives of example 3.3 plus $\{\forall, \exists\}$; (ii) a set of predicate symbols, each with an arity $n \geqslant 0$, a set of function symbols, also each with an arity $n \geqslant 0$, and a set of variables; and (iii) the following rules for term formation, formula formation, and sentence formation:

- if x is a variable, f a function symbol with arity n and $t_1, \ldots, t_n$ are terms then x and $f(t_1, \ldots, t_n)$ are terms.
- if $t_1, \ldots, t_n$ are terms, p a predicate symbol with arity n, and ϕ and ψ are formulas and x is a variable then $p(t_1, \ldots, t_n)$, $\top$, $\bot$, $\neg\phi$, $\phi\wedge\psi$, $\phi\vee\psi$, $\phi\rightarrow\psi$, $\phi\leftrightarrow\psi$, $\exists x.\,\phi$ and $\forall x.\,\phi$ are formulas.
- if ϕ is a formula with no free variables (standard definition) then ϕ is a sentence.

Each $M \in \mathcal{M}$ has (a) a domain of individuals D_M; (b) for each predicate symbol p with arity n, a subset $M[\![p]\!]$ of D_M^n (which means $D_M \times \ldots \times D_M$, n times); (c) for each function symbol f with arity n a function $M[\![f]\!]$ from D_M^n to D_M; and (d) for each variable x an element $M[\![x]\!]$ of D_M.

$M[\![\cdot]\!]$ is extended to terms by

$$M[\![f(t_1, \ldots, t_n)]\!] = M[\![f]\!](M[\![t_1]\!], \ldots, M[\![t_n]\!])$$

for each function symbol f with arity n.

For each variable x of L, an equivalence relation $\sim_x \subseteq \mathcal{M} \times \mathcal{M}$ is defined as follows: $M \sim_x N$ if $D_M = D_N$ and for each predicate symbol p and function symbol f, $M[\![p]\!] = N[\![p]\!]$ and $M[\![f]\!] = N[\![f]\!]$ and for each variable y with the possible exception of x, $M[\![y]\!] = N[\![y]\!]$. That is to say, M and N are alike in every way except possibly in how they assign the variable x.

The satisfaction relation is defined as follows: if χ is of the form $\top$, $\bot$, $\neg\phi$, $\phi\wedge\psi$, $\phi\vee\psi$, $\phi\rightarrow\psi$, or $\phi\leftrightarrow\psi$, then $M \Vdash \chi$ according to example 3.3. Otherwise,

$$\begin{aligned}
&M \Vdash p(t_1,\ldots,t_n) \text{ if} \\
&\quad \langle M[\![t_1]\!],\ldots,M[\![t_n]\!]\rangle \in M[\![p]\!] \\
&M \Vdash \forall x.\,\phi \text{ if} \\
&\quad N \Vdash \phi \text{ for each } N \text{ s.t. } M \sim_x N \\
&M \Vdash \exists x.\,\phi \text{ if} \\
&\quad N \Vdash \phi \text{ for some } N \text{ s.t. } M \sim_x N
\end{aligned}$$

We now return to standard definitions and a result:

Definition 3.5 A (flat) theory over a language L, or an L-theory, is a set of L-sentences.

Notice that, as already mentioned, we do not require theories to be consequence-closed sets of sentences.

Definition 3.6 Let Φ be a theory. Then $M \Vdash \Phi$ if $M \Vdash \phi$ for each $\phi \in \Phi$.

Definition 3.7 ϕ is a *consequence* of Φ, or Φ *entails* ϕ, written $\Phi \vDash \phi$, if for each $M \in \mathcal{M}$, $M \Vdash \Phi$ implies $M \Vdash \phi$.

Simple though these definitions are, there are some well known consequences.

Proposition 3.8 Let L be a language and $\vDash$ the consequence relation defined from an interpretation system $\langle \mathcal{M}, \Vdash\rangle$. The following properties hold of $\vDash$:

1. Inclusion: $\Phi, \phi \vDash \phi$
2. Monotonicity: $\dfrac{\Phi \vDash \psi}{\Phi, \phi \vDash \psi}$
3. Cut: $\dfrac{\Phi, \phi \vDash \psi \quad \Psi \vDash \phi}{\Phi, \Psi \vDash \psi}$

As usual, Φ, Ψ and Φ, ϕ abbreviate $\Phi \cup \Psi$ and $\Phi \cup \{\phi\}$ respectively. The horizontal rule means: if the top sequent holds then so does the bottom one.

The last standard definition to consider is that of positive and negative occurrences of non-logical symbols in formulas. The exact definition depends on the connectives and their interpretations.

Example 3.9 Let L, $\mathcal{M}$ and $\Vdash$ be classical propositional logic (example 3.3) with $p \in \text{atoms}(L)$.

- p occurs positively in p.
- If p occurs positively (negatively) in ϕ then it occurs negatively (positively) in $\neg\phi$.
- If p occurs positively (negatively) in ϕ or in ψ then it occurs positively (negatively) in $\phi\wedge\psi$ and $\phi\vee\psi$.
- If p occurs negatively (positively) in ϕ or positively (negatively) in ψ then it occurs positively (negatively) in $\phi\rightarrow\psi$.
- If p occurs at all in ϕ or ψ then it occurs both positively and negatively in $\phi\leftrightarrow\psi$.
- p does not occur in either $\top$ or $\bot$.

Example 3.10 In the case of predicate logic, if p is a predicate symbol and $t_1,\ldots,t_n$ are terms then p occurs positively in $p(t_1,\ldots,t_n)$. Each of the clauses for the propositional connectives above applies. Moreover, if p occurs positively (negatively) in ϕ then it occurs positively (negatively) in $\forall x.\,\phi$ and $\exists x.\,\phi$.

4 Consequence for Structured theories

The purpose of this section is to define satisfaction for structured theories, so that consequence for structured theories can be defined by definition 3.7. As before we assume we are working with a fixed language L and interpretation system $\langle \mathcal{M}, \Vdash\rangle$.

Intuitively, a structured theory is a finite collection of sentences equipped with a partial order. But to cover the case that the same sentence occurs several times in different places in the theory, it is necessary to posit a 'carrier set' on which the order is defined and whose points are labelled by sentences.

Definition 4.1 A structured theory ? over a language L is a tuple $\langle X, \leqslant, F\rangle$ where

1. X is a finite set (the carrier set).
2. $\leqslant$ is a partial order on X.
3. F is a function mapping X to L-sentences.

The letters Φ and Ψ were used for "flat" theories (definition 3.5); we shall use ? and Δ for structured theories.

The intuitive meaning of the ordering is: if $x \leqslant y$ then the sentence $F(x)$ has greater priority (or more influence) than $F(y)$. This information is used when $F(x)$ and $F(y)$ conflict.

We want to define the models of a structured theory, that is, to extend the satisfaction relation to structured theories analogously to its extension to flat theories in definition 3.6. Let ? $= \langle X, \leqslant, F\rangle$ be a structured theory over $\langle L, \mathcal{M}, \Vdash\rangle$. If all the sentences of ? are mutually consistent, then the models of ? are just the models of that set of sentences. The interesting case is when sentences in ? are inconsistent with each other and we have to use the ordering to resolve the conflict. In this case we cannot hope to satisfy all the sentences but models of ? should satisfy as many of them as possible, taking account of their ordering.

The technique to be adopted is to order interpretations of L according to ?, so that those higher up the ordering are better at satisfying ?. This ordering is written $\sqsubseteq^{\Gamma}$. $M \sqsubseteq^{\Gamma} N$ means N is at least as good as M at satisfying ?. Models of ? are then taken to be the interpretations which are maximal according to $\sqsubseteq^{\Gamma}$.

The remainder of this section is structured as follows. In subsection 4.1, we establish the need for another family of orders on $\mathcal{M}$, one for each sentence ϕ. The ordering corresponding to ϕ is written $\sqsubseteq_{\phi}$. Subsection 4.2 deals with examples and the definition of this ordering. In subsection 4.3, $\sqsubseteq^{\Gamma}$ is defined in terms of $\sqsubseteq_{\phi}$. The existence of models of structured theories is shown in subsection 4.4. Finally, subsection 4.5 summarises the definitions.

4.1 First ideas

As we have said, the task is to define an ordering $\sqsubseteq^{\Gamma}$ in terms of a structured theory ?. Models of ? are then defined to be the maximal elements of this ordering. The main question addressed in this paper is how $\sqsubseteq^{\Gamma}$ is defined. If ? were not itself ordered, this task would be easier. For example, one might say $M \sqsubseteq^{\Gamma} N$ if N satisfies all the sentences of ? that M does. But ? is ordered, and our definition must take account of that. Consider again the interpretations M and N. If $M \sqsubseteq^{\Gamma} N$, but there is a sentence ϕ in ? such that M satisfies ϕ and N does not, then there must be a more important sentence ψ which is satisfied by N but not by M. Thus we might be tempted to define $\sqsubseteq^{\Gamma}$ as follows:

Proposal 4.2 $M \sqsubseteq^{\Gamma} N$ if $\forall x \in X. M \Vdash F(x)$ and $N \nVdash F(x)$ implies $\exists y \leqslant x. M \nVdash F(y)$ and $N \Vdash F(y)$.

To see that this is wrong, consider the structured theory given in example 2.3. In terms of definition 4.1, this is the structured theory $\langle X, \leqslant, F\rangle$ in classical propositional logic with the propositional atoms $\{p, q\}$, given by:

- $X = \{1, 2\}$ with $1 \leqslant 1$, $1 \leqslant 2$ and $2 \leqslant 2$
- $F(1) = \neg p$ and $F(2) = p \wedge q$

Graphically it is represented in figure 1(i). In such "theory" diagrams, the arrows mean $\leqslant$; an arrow from $F(x)$ to $F(y)$ means $x \leqslant y$. A model of this theory is an interpretation which satisfies $\neg p$ and as much of $p \wedge q$ as it can. Let $\langle \mathcal{M}, \Vdash\rangle$ be the usual interpretation system for this logic (see example 3.3). An interpretation M of $\mathcal{M}$ is specified by whether it satisfies the atoms p and q. Let us write 10 for the interpretation which satisfies p but not q; 11, 01 and 00 are defined analogously.

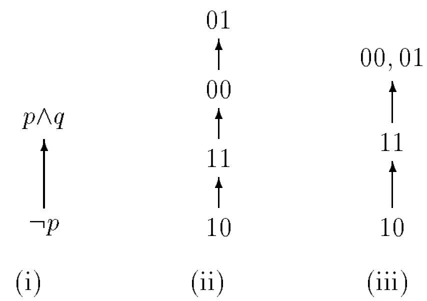


Figure 1: A structured theory and candidate interpretation orderings

Intuitively we expect the interpretation 01 to be the only model of ?. To see this, notice that it must be either 00 or 01 since $\neg p$ is the most important sentence of ?. Of these two 01 is better at satisfying ? overall because, while neither of them satisfy $p \wedge q$, it at least satisfies half of $p \wedge q$. Further reasoning along these lines results in the conclusion that figure 1(ii) is the correct interpretation ordering for the theory in question. There, the arrows mean $\sqsubseteq^{\Gamma}$.

But since neither of the interpretations 01 and 00 fully satisfy $p \wedge q$, and proposal 4.2 just looks at what sentences are satisfied by the various interpretations, the proposal cannot distinguish between 01 and 00. In fact, according to the proposal $\sqsubseteq^{\Gamma}$ is the order given in figure 1(iii). 01 and 00 are both maximal in this ordering, so both would be models of ? according to the proposal.

The problem is that we were not able to take account of the fact that, while neither 01 nor 00 satisfy $p \wedge q$, 01 is actually better at it than 00; at least it satisfies q, which is a consequence of $p \wedge q$. This thought leads us to the idea that, given a sentence and an interpretation, there is more we can say than whether the interpretation satisfies the sentence or not. We can compare two interpretations as to the degree to which they satisfy the sentence.

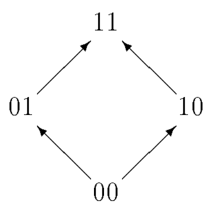
This intuition, about degrees of satisfaction, is formalised in the following subsection. The idea is to define an ordering $\sqsubseteq_{\phi}$ on interpretations (for each sentence ϕ) and use that to define $\sqsubseteq^{\Gamma}$. In subsection 4.3 we return to the question of ordering interpretations to give models of structured theories.

4.2 Ordering interpretations by a single sentence

Given a sentence ϕ and an interpretation M, we are interested in how well M satisfies ϕ. If $M \Vdash \phi$, then this is the best one could hope for; M satisfies ϕ to the fullest possible extent. But if $M \not\Vdash \phi$, all is not lost. It may still satisfy some of the consequences of ϕ.

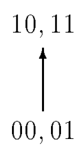
Before considering possible definitions for $\sqsubseteq_\phi$, it is worth looking at some examples to see how it should behave. The reader can check that the maximal interpretations for each sentence are precisely the models of the sentence according to the underlying logic. The aim of $\sqsubseteq_\phi$ is to order the interpretations which do not satisfy ϕ according to how *nearly* they do.

Example 4.3 Consider again classical propositional logic with the atoms $\{p, q\}$. The interpretations are $\{00, 01, 10, 11\}$ as before. If ϕ is $p \wedge q$ then $\sqsubseteq_\phi$ is as follows:



The point is that even if an interpretation doesn't satisfy $p \wedge q$, it can do better than $\neg p \wedge \neg q$.

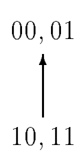
Example 4.4 If ϕ is just p, then $\sqsubseteq_\phi$ is as follows:



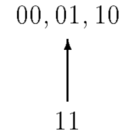
Either a model satisfies p or it doesn't. Notice that in general $\sqsubseteq_\phi$ is a pre-order, i.e. reflexive and transitive, but not necessarily antisymmetric. For here, 10 and 11 are equivalent as far as satisfying p is concerned, but they are not equal.

Example 4.5 If ϕ is $\top$ or $\bot$, then the ordering is just the one in which everything is equivalent, for no model is any better at satisfying $\top$ (or $\bot$) than any other. In the case of $\top$ it is because they all satisfy it. In the case of $\bot$ it is because, while none satisfy it, neither is any model any better than any other.

Example 4.6 The ordering according to $\neg p$ is simply that of p turned upside down:

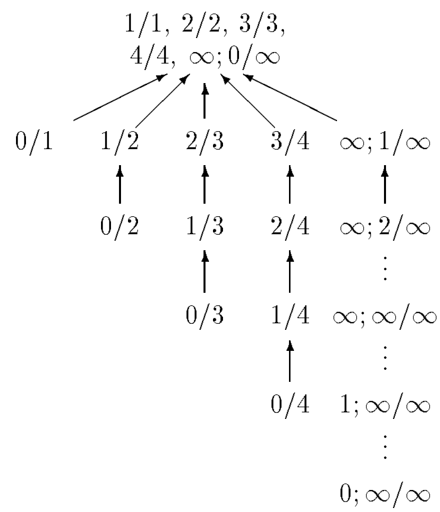


But the ordering for $\neg(p \wedge q)$ (or, equivalently, $\neg p \vee \neg q$) bears little resemblance to that for $p \wedge q$:



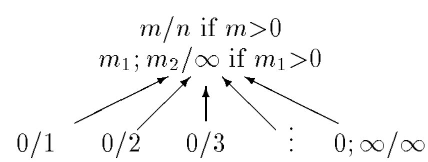
It should be clear that the ordering is only concerned with the interpretations which fail to satisfy the sentence in question.

Example 4.7 This example shows the behaviour in classical predicate logic. Suppose the language contains the single unary predicate p, and ϕ is $\forall x.\, p(x)$. Then the ordering $\sqsubseteq_\phi$ looks roughly like this:



'Roughly', because there are many bits missing. Notation: m/n denotes the class of interpretations with n elements of which m satisfy p. $m_1; m_2/\infty$ denotes that class with infinitely many elements of which there are m_1 satisfying p and m_2 not.

Example 4.8 If ϕ is $\exists x.\, p(x)$:



The intuition to be gained from these examples is the following. The degree to which an interpretation satisfies a sentence has something to do with the *consequences* of the sentence which it satisfies. For example, while 10 does not satisfy $p \wedge q$ it at least satisfies p, which is a consequence. In example 4.7 the same analysis applies. 2/3 more nearly satisfies $\forall x.\, p(x)$ than

1/3, for it satisfies a consequence of $\forall x.\, p(x)$ which the latter fails, namely:

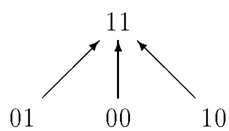
$$\exists x_1 x_2 x_3.\,(x_1{\neq}x_2{\wedge}x_2{\neq}x_3{\wedge}x_3{\neq}x_1) \\ \to\exists y_1 y_2.(y_1{\neq}y_2{\wedge}p(y_1){\wedge}p(y_2))$$

which says 'if there are three elements in the domain then there are two elements in the domain which satisfy p': Thus one might consider the following definition for $\sqsubseteq_\phi$:

Proposal 4.9 $M \sqsubseteq_\phi N$, if for each ψ,

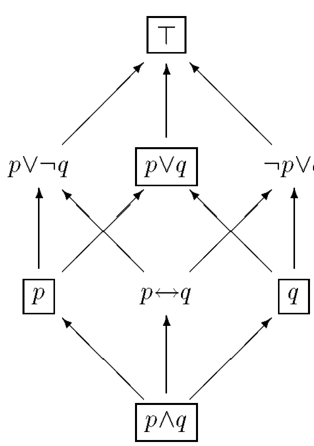
$$\phi \vDash \psi \Rightarrow (M \Vdash \psi \Rightarrow N \Vdash \psi)$$

However, one can immediately see that not all the consequences of ϕ are appropriate to take into account in the definition of $\sqsubseteq_\phi$. Consider again example 4.3. p, $p{\leftrightarrow}q$ and q are all consequences of $p{\wedge}q$, but none of each other. Therefore proposal 4.9 gives the following for $\sqsubseteq_{p\wedge q}$:



This is wrong according to example 4.3. Indeed, it turns out that under this definition $\sqsubseteq_\phi$ always has a height of just 2. To be precise, if $\sqsubseteq_\phi$ is defined in this way and the underlying logic has the property that for each interpretation there is a sentence which picks it out uniquely up to isomorphism (classical propositional logic over a finite language has this property), then $M \sqsubseteq_\phi N$ implies $N \Vdash \phi$ or $M = N$. To see this, suppose $M \sqsubseteq_\phi N$ and let χ be the sentence which characterises M. Since $\phi \vDash \phi{\vee}\chi$ and $M \Vdash \phi{\vee}\chi$, it must be that $N \Vdash \phi{\vee}\chi$, i.e. $N \Vdash \phi$ or $N = M$.

The problem is that not all the consequences of ϕ should be taken into consideration in deciding whether $M \sqsubseteq_{p\wedge q} N$. In the case of $p{\wedge}q$, only the consequences in boxes are appropriate.



What distinguishes these consequences of $p{\wedge}q$ is that they are *monotonic* in p and q. That is to say, if a model M satisfies such a consequence ψ, then so does the model N obtained from M by increasing the 'extension' of p or of q. To define this we need to define positive and negative occurrences. As stated previously, we assume that these are given by the underlying logic (examples 3.9 and 3.10).

Definition 4.10 If ϕ is an L-sentence and p a non-logical symbol in L,

1. ϕ is *monotonic in* p if it is equivalent to a sentence in which p does not occur negatively.
2. ϕ is *anti-monotonic in* p if it is equivalent to a sentence in which p does not occur positively.
3. ϕ^+ and ϕ^- are the sets of symbols in which ϕ is monotonic and anti-monotonic respectively.

The justification for this terminology is as follows. One may define the *extension* of a non-logical symbol p in a model to be the set of tuples or worlds which satisfy p in the model. (In the propositional case, if p is true in a model then its extension is defined to be the singleton $\{*\}$, otherwise it is $\varnothing$.) Extensions are naturally ordered by inclusion. Let us write $M \leqslant^p N$ if M and N are exactly alike except that N has possibly a greater p-extension than M. It follows that ϕ is monotonic in p iff $(M \leqslant^p N \Rightarrow (M \Vdash \phi \Rightarrow N \Vdash \phi))$, *i.e.* increasing p-extension in a model preserves ϕ-satisfaction. Similarly, ϕ is anti-monotonic in p iff $(N \leqslant^p M \Rightarrow (M \Vdash \phi \Rightarrow N \Vdash \phi))$.

Thus, the monotonicities of ϕ is a pair $\langle\phi^+, \phi^-\rangle$ of sets of non-logical symbols such that, if in any model of ϕ the extension of any symbol of the first set is increased, or the extension of any in the second set is decreased, the resulting interpretation is still a model of ϕ.

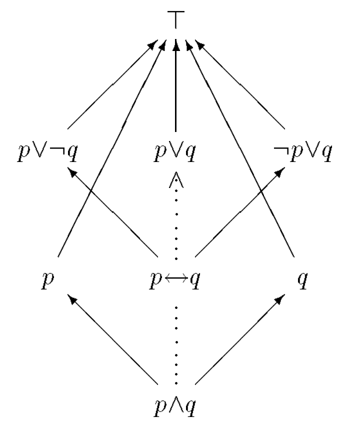
Example 4.11 Let $(L, \mathcal{M})$ be classical propositional logic over $\{p, q\}$.

ϕ	ϕ^+	ϕ^-
$\top, \bot$	$\{p,q\}$	$\{p,q\}$
p	$\{p,q\}$	$\{q\}$
q	$\{p,q\}$	$\{p\}$
$p{\wedge}q,\ p{\vee}q$	$\{p,q\}$	$\varnothing$
$p{\to}q$	$\{q\}$	$\{p\}$
$p{\leftrightarrow}q$	$\varnothing$	$\varnothing$

Definition 4.12 A consequence ψ of ϕ is a *natural consequence* (written $\phi \vDash^\natural \psi$) if it preserves the monotonicities of ϕ:

$$\phi \vDash^\natural \psi \quad \text{if} \quad \phi \vDash \psi\,,\ \ \phi^+ \subseteq \psi^+ \ \text{and}\ \phi^- \subseteq \psi^-$$

Example 4.13 The relation $\models^\natural$ among the ordinary consequences of $p \wedge q$ is shown below.



Thus: $p\wedge q \models^\natural p$ and $p \wedge q \models^\natural p\wedge q$, but $p\wedge q \not\models^\natural p \leftrightarrow q$ and $p \not\models^\natural p \vee q$. Moreover, $\bot \models^\natural \phi$ iff $\phi = \bot$ or $\phi = \top$.

Natural consequence is something like relevant consequence; it stops us adding irrelevant disjuncts in our conclusions. In passing, note the following properties of $\models^\natural$:

- Exchange; Contraction; Reflexivity.
- Monotonicity: if $\mathrm{lang}(\Phi) \cap \mathrm{lang}(\psi) = \varnothing$ and $\Phi \models^\natural \psi$ then $\Phi, \phi \models^\natural \psi$
- Cut: if $\mathrm{lang}(\Phi) \cap \mathrm{lang}(\Psi) = \varnothing$ and $\Phi, \phi \models^\natural \psi$ and $\Psi \models^\natural \phi$ then $\Phi, \Psi \models^\natural \psi$.

Finally we can define $\sqsubseteq_\phi$. The definition is just like proposal 4.9, but with $\models^\natural$ instead of $\models$.

Definition 4.14 $M \sqsubseteq_\phi N$, if for each ψ,

$$\phi \models^\natural \psi \Rightarrow (M \Vdash \psi \Rightarrow N \Vdash \psi)$$

Proposition 4.15 For each L-sentence ϕ, $\sqsubseteq_\phi$ is a pre-order.

Proof Reflexivity is obvious. For transitivity, suppose $L \sqsubseteq_\phi M \sqsubseteq_\phi N$, and let ψ be such that $\phi \models^\natural \psi$ and $L \Vdash \psi$. Then, since $L \sqsubseteq_\phi M$, $M \Vdash \psi$. And since $M \sqsubseteq_\phi N$, $N \Vdash \psi$. ♡

We finish this subsection with a few definitions and results to reassure us that everything is according to plan:

Definition 4.16

1. $M \sqsubset_\phi M'$ if $M \sqsubseteq_\phi M'$ and $M' \not\sqsubseteq_\phi M$.
2. $M \equiv_\phi M'$ if $M \sqsubseteq_\phi M'$ and $M' \sqsubseteq_\phi M$.
3. M is $\sqsubseteq_\phi$-maximal if for every $N \in \mathcal{M}$, $M \not\sqsubset_\phi N$.
4. M is $\sqsubseteq_\phi$-maximum if for every $N \in \mathcal{M}$, $N \sqsubseteq_\phi M$.

Lemma 4.17 M is $\sqsubseteq_\phi$-maximum iff ($M \Vdash \phi$ or $\phi = \bot$).

Proof (If) If $\phi = \bot$ then every interpretation is $\sqsubseteq_\phi$-maximum. If $M \Vdash \phi$ then $M \Vdash \psi$ whenever $\phi \models^\natural \psi$. Therefore, $N \sqsubseteq_\phi M$ for any N.

(Only if) Suppose $\phi \neq \bot$ and $M \not\Vdash \phi$. We show that M is not $\sqsubseteq_\phi$-maximum. Let $N \Vdash \phi$. We show that $M \sqsubset_\phi N$. (i) $M \sqsubseteq_\phi N$, since by first part N is $\sqsubseteq_\phi$-maximum. (ii) $N \not\sqsubseteq_\phi M$, since $\phi \models^\natural \phi$, $N \Vdash \phi$ and $M \not\Vdash \phi$. ♡

Lemma 4.18 If $M \not\Vdash \phi$ and $N \Vdash \phi$ then $M \sqsubset_\phi N$.

Proof (i) $M \sqsubseteq_\phi N$ since N is $\sqsubseteq_\phi$-maximum by lemma 4.17. (ii) $N \not\sqsubseteq_\phi M$, for $\phi \models^\natural \phi$, $N \Vdash \phi$ and $M \not\Vdash \phi$. ♡

4.3 Ordering interpretations by a structured theory

Now finally we can define the interpretation ordering induced by ?. The definition captures the flavour of proposal 4.2, which is that if a sentence in ? makes the 'wrong' choice of two interpretations then there is a sentence with greater priority which makes the 'right' choice. But now, the choice that the sentence ϕ makes is determined by $\sqsubseteq_\phi$. First some notation: $\sqsubseteq_x$ shall abbreviate $\sqsubseteq_{F(x)}$; similarly for $\equiv_x$ and $\sqsubset_x$.

Definition 4.19 $M \sqsubseteq^\Gamma N$ if for each $x \in X$, $M \not\sqsubseteq_x N$ implies there exists $y \leqslant x$ such that $M \sqsubset_y N$.

More notation: $M \sqsubset^\Gamma N$ iff $M \sqsubseteq^\Gamma N$ and $N \not\sqsubseteq^\Gamma M$; $M \sqsupseteq^\Gamma N$ iff $N \sqsubseteq^\Gamma M$; $M \sqsupset^\Gamma N$ if $N \sqsubset^\Gamma M$.

Lemma 4.20 $M \sqsubseteq^\Gamma N$ iff $\forall x \in X. (M \not\sqsubseteq_x N$ implies $\exists y \leqslant x. M \sqsubset_y N$ and $\forall z < y. M \equiv_z N)$.

Proof (If) Immediate. (Only if) Suppose $M \sqsubseteq^\Gamma N$ and $M \not\sqsubseteq_x N$ for some x. Let $X' = \{y \in X \mid M \sqsubset_y N$ and $y \leqslant x\}$. $X' \neq \varnothing$ since $M \sqsubseteq^\Gamma N$, and X' is finite since X is finite. Let y be a minimal point in X'. Then $M \sqsubset_y N$, and if $z < y$ then $z \notin X'$, so $M \not\sqsubset_z N$. Either $M \not\sqsubseteq_z N$ or $M \equiv_z N$. If $M \not\sqsubseteq_z N$ then $\exists z' \leqslant z. z' \in X'$, a contradiction since then $z' < y$. Therefore, $M \equiv_z N$. ♡

Proposition 4.21 $\sqsubseteq^\Gamma$ is a pre-order.

Proof Reflexivity is obvious. For transitivity, suppose $L \sqsubseteq^\Gamma M \sqsubseteq^\Gamma N$, and let $L \not\sqsubseteq_x N$. We shall show $L \sqsubset_y N$ for some $y \leqslant x$.

Suppose $L \sqsubseteq_x M$. Either $M \sqsubseteq_x N$ or $M \not\sqsubseteq_x N$. If $M \sqsubseteq_x N$ then $L \sqsubseteq_x N$, a contradiction. If $M \not\sqsubseteq_x N$, let $y_2 \leqslant x$ be such that $M \sqsubset_{y_2} N$ and $M \sqsubseteq_z N$ for

$z \leqslant y_2$ (lemma 4.20). If $L \not\sqsubseteq_{y_2} M$, then let $y \leqslant y_2$ be such that $L \sqsubset_y M$. Then $y \leqslant x$ and $L \sqsubset_y N$ follows from $L \sqsubset_y M$ and $M \sqsubseteq_y N$. If $L \sqsubseteq_{y_2} M$, set $y = y_2$. Then $y \leqslant x$, and $L \sqsubset_y N$ follows from $L \sqsubseteq_y M$ and $M \sqsubset_y N$.

On the other hand, suppose $L \not\sqsubseteq_x M$ and let $y_1 \leqslant x$ be such that $L \sqsubset_{y_1} M$ and $L \sqsubseteq_z M$ for all $z \leqslant y_1$ (lemma 4.20). Again, consider separately the two cases $M \sqsubseteq_{y_1} N$ and $M \not\sqsubseteq_{y_1} N$. If $M \sqsubseteq_{y_1} N$, set $y = y_1$. Then $y \leqslant x$, and $L \sqsubset_y N$ follows from $L \sqsubset_y M$ and $M \sqsubseteq_y N$. If $M \not\sqsubseteq_{y_1} N$ then let $y \leqslant y_1$ be such that $M \sqsubset_y N$. Then $y \leqslant x$, and $L \sqsubset_y N$ follows from $L \sqsubseteq_y M$ and $M \sqsubset_y N$. ♡

The definition of $\Vdash$ can now be extended to structured theories analogously to definition 3.6 in the expected way, as in proposal 4.2.

Definition 4.22 Let ? be a structured theory over L, and M an element of $\mathcal{M}$. Then $M \Vdash$? if M is $\sqsubseteq^\Gamma$-maximal.

This gives the interpretation ordering of figure 1(iii) for the theory of figure 1(i). Moreover, for each example in examples 2.1 to 2.7, one can show that the models of the structured theory according to definition 4.22 are exactly those of the flat theory according to definition 3.6.

Definition 4.22 further overloads $\Vdash$. (To determine whether $M \Vdash A$, we have to check whether A is a sentence, a flat theory or a structured theory and use definitions 3.2, 3.6 or 4.22 accordingly.) Finally, consequence is defined in the standard way:

Definition 4.23 Let ? be a structured L-theory and ϕ an L-sentence. ? $\vDash \phi$ if for each $M \in \mathcal{M}$, $M \Vdash$? implies $M \Vdash \phi$.

Now we give some results to continue to get the feel for the behaviour of structured theories. Naturally we expect that the minimum sentence (if there is one) is satisfied by models of the theory:

Definition 4.24 ϕ is *minimum* in ? $= \langle X, \leqslant, F\rangle$ if $\langle X, \leqslant\rangle$ has a minimum point 0 and $F(0) = \phi$.

Proposition 4.25 Let ? $= \langle X, \leqslant, F\rangle$ be a structured theory and M an element of $\mathcal{M}$ such that $M \Vdash$?. If ϕ is minimum in ? and $\phi \neq \bot$ then $M \Vdash \phi$.

Proof Let 0 be the minimum point in X. $F(0) = \phi$. Suppose for a contradiction that $M \nVdash \phi$. Since $\phi \neq \bot$, let $N \Vdash \phi$. By lemma 4.18, $M \sqsubset_0 N$; in particular, $N \not\sqsubseteq_0 M$. We show $M \nVdash$? by showing $M \sqsubset^\Gamma N$. To show $M \sqsubseteq^\Gamma N$, suppose x is such that $M \not\sqsubseteq_x N$. Let $y = 0$. Then $y \leqslant x$ and $N \not\sqsubseteq_y M$. To show $N \not\sqsubseteq^\Gamma M$, let $x = 0$. $N \not\sqsubseteq_x M$. If $y \leqslant x$, then $y = 0$ since 0 is minimum. $M \sqsubseteq_y N$ since $M \sqsubset_0 N$. ♡

4.4 Existence of models of structured theories

As stated, models of a structured theory ? are $\sqsubseteq^\Gamma$-maximal interpretations of the language of ?. When is it possible to find such maximal interpretations? In this section we show that, if the underlying logic is compact, *every* structured theory has a model.

First, it is worth noting that there are simple cases of structured theories with no models when compactness fails.

Example 4.26 Let ? be the theory

$$\forall x.\, p(x)$$

$$\uparrow$$

domain is infinite $\wedge$
$[\![p]\!]$ is finite

The bottom sentence is satisfied by an interpretation with an infinite domain of individuals of which only finitely many satisfy the predicate p. But the top sentence says that *all* the individuals must satisfy p. This is a theory in second order predicate logic; it is not possible to express finiteness of the interpretation of a predicate or infiniteness of the domain in first order logic.

There are no models of this theory, because every candidate model M can be improved to obtain an interpretation which is closer to being a model, *ad infinitum*. That is to say, for all $M \in \mathcal{M}$ there is an $N \in \mathcal{M}$ such that $M \sqsubset^\Gamma N$. To see this, suppose M pretends to be a model of ?.

- If the domain of individuals of M is finite, then construct N by adding infinitely many new individuals which do not satisfy p.
- If $M[\![p]\!]$ is infinite, then construct N from M by using the same domain but removing all but finitely many elements from $[\![p]\!]$.
- If $M[\![p]\!]$ is finite but the domain is infinite, then N is obtained by adding one more element to $[\![p]\!]$.

In each of these cases, $M \sqsubset^\Gamma N$.

Now we turn to the proof that if the underlying logic is compact (which second-order logic is not), then every structured theory has a model. The proof strategy is to use Zorn's lemma to find $\sqsubseteq^\Gamma$-maximal interpretations.

Let L be a language and $\langle \mathcal{M}, \Vdash \rangle$ its interpretation system, and let $? = \langle X, \leqslant, F \rangle$ be a structured theory over L.

Definition 4.27 The logic $\langle L, \mathcal{M}, \Vdash \rangle$ is *compact* if for all sets of sentences $\Phi \subseteq L$, Φ has a model if each of its finite subsets has a model.

Definition 4.28 For each M, N in $\mathcal{M}$, the (M,N)-frontier, written $\mathrm{fr}(M,N)$, is the set of minimal elements of the set $\{x \in X \mid M \not\equiv_x N\}$.

Lemma 4.29 For all $M, N \in \mathcal{M}$ and $x \in X$, either $M \equiv_x N$ or $\exists y \leqslant x.\, y \in \mathrm{fr}(M,N)$.

Proof $\{x \in X \mid M \not\equiv_x N\}$ is finite since X is, so it has minimal elements. ♡

Lemma 4.30 $M \sqsubset^\Gamma N$ iff $\mathrm{fr}(M,N) \neq \varnothing$ and $\forall x \in \mathrm{fr}(M,N).\, M \sqsubset_x N$.

Proof (If) First we show $M \sqsubseteq^\Gamma N$. Suppose $x \in X$ with $M \not\sqsubseteq_x N$. By lemma 4.29, $\exists y \in \mathrm{fr}(M,N)$ with $y \leqslant x$. By hypothesis, $M \sqsubset_y N$. Next, we show $N \not\sqsubseteq^\Gamma M$. Let $x \in \mathrm{fr}(M,N)$. Then $N \not\sqsubseteq_x M$, but for each $y < x$, $M \equiv_y N$.

(Only if) If $\mathrm{fr}(M,N) = \varnothing$ then $M \equiv^\Gamma N$, a contradiction. Let $x \in \mathrm{fr}(M,N)$. Either $M \not\sqsubseteq_x N$ or $N \not\sqsubseteq_x M$. In the former case, $\exists y \leqslant x$ with $M \sqsubset_y N$; since $x \in \mathrm{fr}(M,N)$, y must equal x. In the latter case, $N \not\sqsubseteq_x M$ and if $M \sqsubseteq_x N$ then $M \sqsubset_x N$. Therefore, in both cases $M \sqsubset_x N$ as required. ♡

Lemma 4.31 Let $\mathcal{N}$ be a non-empty chain in $\mathcal{M}$ with no maximal element (i.e. for every $M, N \in \mathcal{N}$, if $M \neq N$ then $M \sqsubset^\Gamma N$ or $N \sqsubset^\Gamma M$; and for each $M \in \mathcal{N}$ there is an $N \in \mathcal{N}$ such that $M \sqsubset^\Gamma N$). There is a non-empty set $Y \subseteq X$ and a non-empty chain $\mathcal{L} \subseteq \mathcal{N}$ such that

1. For each $a \in Y$ and $M, N \in \mathcal{L}$, if $M \sqsubset^\Gamma N$ then $M \sqsubseteq_a N$; and
2. For each $a \in Y$ and $M \in \mathcal{L}$ there exists $P \in \mathcal{L}$ such that $M \sqsubset^\Gamma P$ and $M \sqsubset_a P$.

Proof Let $X' = \{x \in X \mid \forall M \in \mathcal{N}\ \exists M_1, M_2 \in \mathcal{N}\ (M \sqsubset^\Gamma M_1 \sqsubset^\Gamma M_2 \text{ and } x \in \mathrm{fr}(M_1, M_2))\}$.

If $X = X'$, let $\mathcal{L} = \mathcal{N}$. Otherwise, for each $x \in X \perp X'$ let M_x be such that, for all $M_1, M_2 \in \mathcal{N}$, if $M_x \sqsubset^\Gamma M_1 \sqsubset^\Gamma M_2$ then $x \notin \mathrm{fr}(M_1, M_2)$. That such an M_x can be found follows immediately from the definition of X'. Let $M_X = \max(\{M_x \mid x \in X \perp X'\})$; and let $\mathcal{L} = \{M \in \mathcal{N} \mid M_X \sqsubseteq^\Gamma M\}$. $\mathcal{L} \neq \varnothing$ since $M_X \in \mathcal{L}$.

Thus, whether $X = X'$ or not, we have that $\mathcal{L} \neq \varnothing$. Also, $\mathcal{L}$ is upwards closed (i.e. for all $M, N \in \mathcal{N}$, $M \in \mathcal{L}$ and $M \sqsubseteq^\Gamma N$ imply $N \in \mathcal{L}$). Let $M_1, M_2 \in \mathcal{L}$ with $M_1 \neq M_2$. Then either $M_1 \sqsubseteq^\Gamma M_2$ or $M_2 \sqsubseteq^\Gamma M_1$. In either case, $\mathrm{fr}(M_1, M_2) \neq \varnothing$. But, $\mathrm{fr}(M_1, M_2) \subseteq X'$, so $X' \neq \varnothing$. Let Y be the minimal points of X'.

1. Suppose $a \in Y$, $M, N \in \mathcal{L}$, and $M \sqsubset^\Gamma N$. If $a \in \mathrm{fr}(M,N)$ then $M \sqsubset_a N$. If $a \notin \mathrm{fr}(M,N)$ and $M \not\sqsubseteq_a N$ then $\exists y \in \mathrm{fr}(M,N).\, y \leqslant a$ by lemma 4.29, so $a \notin Y$, a contradiction.
2. Suppose $a \in Y$ and $M \in \mathcal{L}$. Since $a \in X'$, $\exists M_1, M_2.\, M \sqsubseteq^\Gamma M_1 \sqsubseteq^\Gamma M_2$ and $a \in \mathrm{fr}(M_1, M_2)$. Since $M \sqsubseteq^\Gamma M_1 \sqsubseteq^\Gamma M_2$, $M \sqsubseteq_a M_1 \sqsubseteq_a M_2$; and since $a \in \mathrm{fr}(M_1, M_2)$, we have $M \sqsubseteq_a M_1 \sqsubset_a M_2$. Let $P = M_2$. ♡

Lemma 4.32 If $\langle L, \mathcal{M}, \Vdash \rangle$ is compact then for each $M \in \mathcal{M}$, there exists $N \in \mathcal{M}$ such that $M \sqsubseteq^\Gamma N$ and N is $\sqsubseteq^\Gamma$-maximal.

Proof Let $M \in \mathcal{M}$. We show that $\{N \mid M \sqsubseteq^\Gamma N\}$ has maximal elements. Let $\mathcal{N}$ be a non-empty chain in that set. By **Zorn's lemma** it suffices to show that every such chain has an upper bound. If $\mathcal{N}$ has a maximal element, that element is also an upper bound. Suppose, then, that $\mathcal{N}$ does not have a maximal element. Let Y and $\mathcal{L}$ be as given by lemma 4.31. Let $Z = Y \cup \{x \in X \mid \forall y \in Y.\, y \not\leqslant x\}$. We now show that for each $x \in Z$ and $M, N \in \mathcal{L}$, $M \sqsubseteq^\Gamma N$ implies $M \sqsubseteq_x N$. If $x \in Y$, this follows from lemma 4.31 part 1. If $x \in Z \perp Y$, then $\forall y \in Y.\, y \not\leqslant x$ by definition of Z. By lemma 4.29, $\exists y' \leqslant x.\, y' \in \mathrm{fr}(M,N) \subseteq X'$, so $\exists y \in Y.\, y \leqslant y'$, a contradiction.

For each $M \in \mathcal{L}$ let M^* be $\{\psi \mid M \Vdash \psi \text{ and } \exists x \in Z.\, F(x) \vDash^\natural \psi\}$. M^* has a model, since it has M as a model. Also, $M \sqsubset^\Gamma N$ implies $M^* \subseteq N^*$. For suppose $\psi \in M^*$. Then $M \Vdash \psi$, and there is an $x \in Z$ s.t. $F(x) \vDash^\natural \psi$. Since $M \sqsubseteq_x N$, we have $N \Vdash \psi$. Therefore, $\psi \in N^*$.

Let $\Phi = \bigcup_{M \in \mathcal{L}} M^*$. Φ has a model, since every M^* and therefore every finite subset of Φ has a model, and the underlying logic is compact. Let $K \Vdash \Phi$. It remains to show that $\forall M \in \mathcal{L}.\, M \sqsubseteq^\Gamma K$, *i.e.* that K is an upper bound. Since $\mathcal{L}$ is a non-empty upwards-closed subchain of $\mathcal{N}$, it is sufficient to consider the case $M \in \mathcal{L}$. Let $M \in \mathcal{L}$. The fact that $M^* \subseteq \Phi$ implies that for each $x \in Z$, $M \sqsubseteq_x K$. Suppose $M \not\sqsubseteq_x K$. Then $x \notin Z$. We require that $M \sqsubset_y K$ for some $y \leqslant x$. Since $x \notin Z$, $\exists y \in Y.\, y \leqslant x$. We now show that $M \sqsubset_y K$ for every $y \in Y$, completing the proof. By lemma 4.31, pick P such that $M \sqsubset^\Gamma P$ and $M \sqsubset_y P$. It suffices to show that $P \sqsubseteq_y K$. Suppose $F(y) \vDash^\natural \psi$ and $P \Vdash \psi$. Then $\psi \in P^*$, so $\psi \in \Phi$, so $K \Vdash \psi$. ♡

As an immediate corollary, we get:

Proposition 4.33 Every structured theory ? over a compact logic has a model.

Proof By lemma 4.32, $\sqsubseteq^{\Gamma}$ has maximal elements. ♡

A consequence of this result is that contradictions can never be derived from a structured theory, not even the contradictory one! Indeed, *nothing* can be derived from the theory with one sentence which is $\bot$. That is because every interpretation is a model of that theory. This may come as a surprise, but really it is quite rational.

Proposition 4.34 If $? \vDash \phi$ then $\phi \neq \bot$.

Proof Let $M \Vdash ?$. Since $M \Vdash \phi$, $\phi \neq \bot$. ♡

4.5 Summary of definitions

In this subsection we summarise the position so far. We started with a logic given in terms of a language and a set of interpretations in the standard way. Structured theories consist of a poset of points, each one labelled by a sentence in the language (definition 4.1). To define the models of structured theories, we first define, for each sentence ϕ in the language, an ordering on the interpretations written $\sqsubseteq_{\phi}$ (definition 4.14). $M \sqsubseteq_{\phi} N$ intuitively means that N satisfies ϕ at least as well as M. To define $\sqsubseteq_{\phi}$, we need the notion of natural consequence (definition 4.12). Then we define the ordering $\sqsubseteq^{\Gamma}$ (definition 4.19). $M \sqsubseteq^{\Gamma} N$ intuitively means that N is as good as M at satisfying ?, taking account of ?'s own ordering. Finally, models of ? are the $\sqsubseteq^{\Gamma}$-maximal elements, whose existence is guaranteed by lemma 4.32, and consequence is defined in the standard way (definition 4.23).

5 Revision of Structured theories

Unlike the case for *flat* theories (see [2] and section 6 of this paper), revising *structured* theories with new and potentially conflicting information is easy. If ϕ is to be incorporated into ?, the resulting theory is ? with a new bottom element labelled by ϕ. Proposition 4.25 guarantees that the revision is successful (*i.e.* that $? * \phi \vDash \phi$ unless $\phi = \bot$).

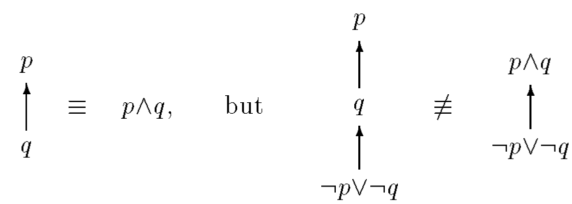
Definition 5.1 Let $? = \langle X, \leqslant, F\rangle$ be a structured theory over L and ϕ an L-sentence. Pick any name a not in X. The theory $? * \phi$ is $\langle X', \leqslant', F'\rangle$ where

1. $X' = X \cup \{a\}$,
2. $\leqslant' = \leqslant \cup \{(a,x) \mid x \in X'\}$, and
3. $F'(x) = \begin{cases} \phi & \text{if } x = a \\ F(x) & \text{otherwise} \end{cases}$

Definition 5.2 Structured theories ? and Δ over L are *extensionally equivalent*, written $? \equiv \Delta$, if for each $M \in \mathcal{M}$,

$$M \Vdash ? \text{ iff } M \Vdash \Delta$$

Example 5.3



This is quite a coarse relation which does not capture the "intensions" of ? and Δ. $? \equiv \Delta$ does not imply $? * \phi \equiv \Delta * \phi$, as the example shows. However, if $? \vDash \phi$ we would not expect that revising ? by ϕ should change the set of models:

Proposition 5.4 If $? \vDash \phi$ then $? \equiv ? * \phi$.

Proof Let $? = \langle X, \leqslant, F\rangle$ and $X' = X \cup \{a\}$. Suppose $M \Vdash ?$ and $M \nVdash ? * \phi$, *i.e.* $M \sqsubset^{\Gamma * \phi} N$ for some N. We will show $M \sqsubset^{\Gamma} N$, contradicting $M \Vdash ?$. First, notice that $M \Vdash \phi$ follows from $? \vDash \phi$ and $M \Vdash ?$.

1. $M \sqsubseteq^{\Gamma} N$. Suppose $M \not\sqsubseteq_x N$ for some $x \in X$. Since $M \sqsubseteq^{\Gamma * \phi} N$ and $x \in X'$, $\exists y \in X'. M \sqsubset_y N$. Moreover, $y \neq a$ because $M \sqsubset_{\phi} N$ implies $M \nVdash \phi$ (lemma 4.17), a contradiction. Therefore, $y \in X$.
2. $N \not\sqsubseteq^{\Gamma} M$. Since $N \not\sqsubseteq^{\Gamma * \phi} M$, there is $x \in X'$ such that $N \not\sqsubseteq_x M$ and for all $y \in X'$ with $y \leqslant x$, $N \not\sqsubset_y M$. Moreover, $x \neq a$ since $N \sqsubseteq_{\phi} M$ (which follows from $M \Vdash \phi$ and lemma 4.17). Therefore, $x \in X$. Since $X \subseteq X'$, it follows that $\forall y \in X. y \leqslant x$ implies $N \not\sqsubset_y M$.

Conversely, suppose $M \Vdash ? * \phi$ and $M \nVdash ?$, *i.e.* $M \sqsubset^{\Gamma} N$, some N. Suppose (lemma 4.32) that N is maximal in the set $\{N \mid M \sqsubset^{\Gamma} N\}$, *i.e.* $N \Vdash ?$. By proposition 4.25 we have that $M \Vdash \phi$. Therefore $M \equiv_{\phi} N$. We now show $M \sqsubset^{\Gamma * \phi} N$, thus proving $M \nVdash ? * \phi$, a contradiction.

1. $M \sqsubseteq^{\Gamma * \phi} N$. Suppose $x \in X'$ with $M \not\sqsubseteq_x N$. Since $M \sqsubseteq_{\phi} N$, $x \neq a$, i.e., $x \in X$. Pick $y \leqslant x$ with $M \sqsubset_y N$. Then $y \in X'$.
2. $N \not\sqsubseteq^{\Gamma * \phi} M$. Since $N \not\sqsubseteq^{\Gamma} M$, $\exists x \in X. N \not\sqsubseteq^{\Gamma} M$ and $\forall y \in X$ with $y \leqslant x$, $N \not\sqsubset_y M$. Moreover, $x \in X'$, so it suffices to show that $\forall y \in X'$, $y \leqslant x$ implies $N \not\sqsubset_y M$, i.e. that $N \not\sqsubset M$, which follows from $M \equiv_{\phi} N$. ♡

We also obtain weak analogues of proposition 3.8:

Proposition 5.5

1. Weak inclusion: if $\phi \neq \bot$ then $? * \phi \models \phi$
2. Weak monotonicity: $\dfrac{? \models \phi \quad ? \models \psi}{? * \phi \models \psi}$
3. Weak cut: $\dfrac{? * \phi \models \psi \quad ? \models \phi}{? \models \psi}$

These principles are accepted as being requirements which a default system should have (see for example [6]).

6 Comparison with other work

6.1 Default logic

The idea of ordering interpretations and considering maximal elements (or minimal elements, depending on how the ordering is oriented) is the basic idea in *circumscription* [8]. There, one reasons with a theory in predicate logic together with a set of predicates to be minimised, and considers only the models of the theory which have minimal extensions of the specified predicates. Y. Shoham [11] first ordered interpretations according to more general criteria, and this idea is now widespread in the literature [6, 7, 12].

Our contribution is to extend the idea to *structured theories* over an arbitrary logic. Partially ordered theories have been studied before; for example, in [13], a computational approach is taken for a Prolog-like language. This paper provides a more general setting.

6.2 Theory revision

In [2], Gärdenfors lists eight postulates which a revision operator $*$ should have. In that work the revision operator takes an ordinary (flat) theory closed under consequence and a sentence. and returns an ordinary closed theory. The revision operator of definition 5.1 takes a structured theory and a sentence, returning a structured theory. Therefore direct comparison is not possible. Instead, we rewrite Gärdenfors' axioms systematically into a form against which we can check our approach. The result is that our definitions satisfy all the axioms except in their treatment of contradictions. The full results with proofs are described elsewhere [9]; here is a summary.

As well as the $*$ revision operator, in Gärdenfors' work there is a $+$ operator which simply adds the sentence and closes under consequence. $K + \phi$ is $\{\psi \mid K \cup \{\phi\} \models \psi\}$. K is a consequence-closed theory. The eight postulates are:

K*1	$K * \phi$ is a consequence-closed theory
K*2	$\phi \in K * \phi$
K*3	$K * \phi \subseteq K + \phi$
K*4	$K + \phi \subseteq K * \phi$ if $\neg\phi \notin K$
K*5	$K * \phi = L$ iff $\phi = \bot$
K*6	$K * \phi = K * \psi$ whenever $\models \phi \leftrightarrow \psi$
K*7	$K * (\phi \wedge \psi) \subseteq (K * \phi) + \psi$
K*8	$(K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$ iff $\neg\psi \notin K * \phi$

Notation: Given a structured theory $?$, let $?^f$ be a flat theory such that $?^f \equiv ?$. Thus, f is an operator which non-deterministically flattens a theory, keeping exactly the same models. (The question of whether there always is an equivalent flat theory is still under investigation. Proposition 4.33 is a necessary condition, and examples 2.1 to 2.7 show there usually is.) Given two sets of sentences Φ_1 and Φ_2 over a logic $\models$, let $\Phi_1 \leqslant \Phi_2$ if $\{\psi \mid \Phi_1 \models \psi\} \subseteq \{\psi \mid \Phi_2 \models \psi\}$.

Gärdenfors' axioms are re-written in the following way:

- $K + \phi$ is rewritten to $?^f \cup \phi$.
- $K * \phi$ is rewritten to $(? * \phi)^f$.
- other occurrences of K are rewritten to $?^f$.

Under this procedure the axioms become:

K*1	$? * \phi$ is a structured theory
K*2	$? * \phi \models \phi$
K*3	$(? * \phi)^f \leqslant ?^f \cup \{\phi\}$
K*4	$?^f \cup \{\phi\} \leqslant (? * \phi)^f$ if $? \not\models \neg\phi$
K*5	$(? * \phi)^f = L$ iff $\phi = \bot$
K*6	$? * \phi \equiv ? * \psi$ if $\models \phi \leftrightarrow \psi$
K*7	$(? * (\phi \wedge \psi))^f \leqslant (? * \phi)^f \cup \{\psi\}$
K*8	$(? * \phi)^f \cup \{\psi\} \leqslant (? * (\phi \wedge \psi))^f$ if $? * \phi \not\models \neg\psi$

In the context of structured theories, axioms K*1, K*2, K*3, K*6 and K*7 all hold true. K*4 is true only if $\phi \neq \bot$ (for recall that $? * \bot \equiv ?$). Similarly, K*8 is true only if $\phi \wedge \psi \neq \bot$. Half of K*5 is true only because $(? * \phi)^f$ is never L.

7 Applications

In the introduction two application areas were mentioned, specification theory and AI. Also, two types of activity have been considered, reasoning with defaults and theory revision. There is space here only for a very brief look at how reasoning with defaults can be applied in the areas mentioned, and an even briefer look at specification revision at the end.

7.1 Defaults

In AI, logics which handle defaults are used to allow consequences to be drawn which are not strictly warranted by the facts at hand. One can consider this to mean that the sentences expressing the defaults have a weaker 'strength' or 'priority' than the known facts. The well-known example about birds and penguins is a situation in which several defaults are available with a known heuristic for priorising them. Therefore it can be seen as a theory in which sentences have different strengths.

The main topic in default reasoning is: how should *conflicting* defaults be handled? In most frameworks for default reasoning there is a distinction between 'factual' information and 'default' information, and little doubt about what to do when a default conflicts with a fact. The difficulty arises when a default conflicts with another default. The example about birds and penguins may be presented as follows:

Facts: $\forall x(p(x) \rightarrow b(x))$.

Defaults: $\forall x(p(x) \rightarrow \neg f(x))$ and $\forall x(b(x) \rightarrow f(x))$.

Let Φ be this theory. The task is to show:

$$\Phi, b(t) \models f(t) \quad \text{and} \quad \Phi, p(t') \models \neg f(t')$$

One can distinguish two approaches in the literature for solving this problem. The first says: there is no solution to the problem as it stands; the second default must be rewritten to something like

$$\forall x(b(x) \rightarrow f(x) \text{ unless } p(x))$$

The 'unless' connective is to resolve the conflict. The following formalisms are examples of this approach: circumscription; Reiter's default logic; Poole's default logic. There are many others. The equivalents of the 'unless' connective are, respectively: abnormality predicates; the consistency check in default rules; constraints.

The second approach says: this problem *has* got a solution in the way it is presented. One must invoke the *specificity heuristic* for determining that the first of the two defaults must take precedence when both are applicable. Examples of this approach include: Veltman's update semantics [12]; inheritance networks [5].

Both of these approaches can be taken in structured theories.

7.1.1 'Unless'-like connectives

The claim in this section is: *the various 'unless'-like connectives in the literature are there in order to priorise defaults*, and hence to resolve conflicts between them. Structured theories are simply a cleaner way of doing this.

$$\begin{array}{c} b(x) \rightarrow f(x) \\ \uparrow \\ p(x) \rightarrow \neg f(x) \\ \uparrow \\ p(x) \rightarrow b(x) \\ b(t) \end{array} \models f(t); \qquad \begin{array}{c} b(x) \rightarrow f(x) \\ \uparrow \\ p(x) \rightarrow \neg f(x) \\ \uparrow \\ p(x) \rightarrow b(x) \\ p(t') \end{array} \models \neg f(t')$$

Elsewhere [9] we show that there are rules for translating default theories in all the usual formalisms with 'unless'-like connectives into structured theories.

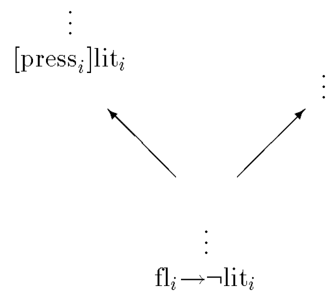
7.1.2 Inheritance

On the above analysis, the way of representing the fact that all penguins are (by definition) birds was $\forall x(p(x) \rightarrow b(x))$. The key idea for *this* section is: consider the *theory* of birds and the *theory* of penguins. They have the (propositional) axioms f and $\neg f$ respectively. The theory of penguins *inherits* the theory of birds, as penguins are a special case of birds. In terms of structured theories, this means that the theory of penguins is obtained by *revising* the theory of birds with the new axioms which apply to penguins—in this case, $\neg f$. It therefore looks like

$$\begin{array}{c} f \\ \uparrow \\ \neg f \end{array}$$

from which, of course, we deduce $\neg f$ (example 2.1).

This idea has great applicability in specification theory. Consider for example a lift (elevator) system, whose components consist of doors, buttons, indicator lights and so on. Here we will focus on just one tiny aspect of its operation—the events which illuminate and extinguish its indicator lights. (See [10] for more details.) The lift is made of several buttons, as well as many other components. Crudely speaking, all lifts are buttons (with loads of extra stuff). This structuring is used to order the axioms of the lift. Lights are illuminated by pressing the buttons; therefore one might expect the axiom $[\text{press}_i]\text{lit}_i$, which says that after the ith button is pressed the ith light comes on. (This is an axiom in modal-action logic; see *e.g.* [4, 1, 10] for further details.) It is also true about lifts that the lights are off whenever the lift is at the relevant floor, even if the button for that floor has just been pressed. Thus: $\text{fl}_i \rightarrow \neg\text{lit}_i$. These axioms conflict, so which is right? The answer of course is that both are right; $[\text{press}_i]\text{lit}_i$ is a true default for light/button combinations (they certainly do this in isolation), but it can be overridden by axioms with greater strength in the specification. The structured theory looks like this:



This shows that [press$_i$]lit$_i$ gets overridden by fl$_i$→¬lit$_i$ if there is a conflict. The full treatment of this example is lengthy and will be given elsewhere. The crucial point to note here is that the ordering of sentences in the structured theory comes from the inheritance hierarchy. The other point to note is that inheritance hierarchies are generally non-linear and the question of whether the branches share non-logical language or not is determined by the mode of interaction of the components. Hence the remark about shared languages in example 2.7.

7.2 Specification revision

Revising old specifications is a typical way of making new ones, in software engineering and elsewhere. For example, a Metro car was conceived as a Mini but with a bigger engine and some other changes; the important point is that most, but not all, of the features of the Metro are just those of the Mini. Of course the new features introduced will in general conflict with what was there before; if they don't, then the 'revision' is merely a matter of enrichment. Typically one does not know what features of the old specification have to be abandoned to ensure the consistency of the new one. Revision in terms of structured theories means placing the new sentences in the most important position, thereby weakening the status of previous sentences there. The formal definition was given in section 5.

Acknowledgements

Many thanks to my supervisor, Steve Vickers. The following people also provided invaluable help: Abbas Edalat, Ian Hodkinson, Lex Holt, and Martin Sadler. Paul Taylor's TEX diagram macros were used.

References

[1] J. Fiadeiro, C. Sernadas, T. Maibaum, and G. Saake. Proof-theoretic semantics of object-oriented specification constructs. In *Object-oriented Databases: Analysis, Design and Construction*. North-Holland, 1990.

[2] P. Gärdenfors. *Knowledge in Flux*. MIT Press, 1988.

[3] J. A. Goguen and R. M. Burstall. A study in the foundations of programming methodology: Specifications, institutions, charters and parchments. In *Category Theory and Computer Programming*. Springer-Verlag, 1986.

[4] Robert Goldblatt. *Axiomatising the Logic of Computer Programming*, volume 130 of *Lecture Notes in Computer Science*. Springer Verlag, 1982.

[5] John F. Horty, Richmond H. Thomason, and David S. Touretzky. A skeptical theory of inheritance in nonmonotonic semantic networks. In *AAAI-87*, 1987.

[6] D. Lehmann. What does a conditional knowledge base entail? In *Knowledge Representation*. Morgan Kaufmann, 1989. Toronto.

[7] D. Makinson. General patterns in non-monotonic reasoning. In C. Hogger D. Gabbay and J. Robinson, editors, *Handbook of Logic in Artificial Intelligence*. Oxford University Press, 1991. forthcoming.

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

[9] Mark Ryan. *Specification Revision and Defaults*. PhD thesis, Imperial College, 1991? forthcoming.

[10] Mark Ryan, José Fiadeiro, and Tom Maibaum. Sharing actions and attributes in modal action logic. In *Theoretical Aspects of Computer Software*, 1991.

[11] Y. Shoham. *Reasoning about Change*. MIT Press, 1986.

[12] Frank Veltman. Defaults in update semantics I. In Hans Kamp, editor, *Conditionals, Defaults and Belief Revision*, pages 28–64. Center for Cognitive Science, Edinburgh, 1990. DYANA deliverable R2.5A.

[13] D. Vermier, P. Geerts, and D. Nute. A logic for defeasible perspectives. Technical Report 89-14, Universiteit Antwerpen, 1989.