PRE-ENTS, ENTS AND GENERALISED RATIONAL CONSEQUENCE

A thesis submitted to the University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Science

November 2002

By
Richard Booth
Department of Mathematics

Contents

Abstract

This thesis is concerned mainly with results involving the pre-ent and ent models of belief formation. These models, pre-ents being the more general of the two, were put forward by Paris and Vencovská as a possible explanation of how an intelligent agent *could* conceivably be acting in forming numerical beliefs in various propositions. We prove a result which establishes, in succinct terms, the essential difference between the two classes. This result may be interpreted as saying that, starting from the class of pre-ents, if we restrict attention to that subclass of pre-ents containing those pre-ents which satisfy a certain, natural, property, then we are led automatically to the class of ents. We then move on to trying to find instances of consequence relations which can arise from the pre-ent model, and use this model to characterise a class of relations which we call fully transitive natural consequence relations. This class contains as a subclass the class of rational consequence relations defined by Kraus, Lehmann and Magidor.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

Copyright

Acknowledgements

I would like to thank Professor Jeff Paris, for being an excellent supervisor from whom I have learned much, and the E.P.S.R.C. for their financial support. Lastly I acknowledge the enormous contribution of my parents in helping me get this far. This thesis is dedicated to them.

Chapter 1

Introduction

1.1 Motivation for Pre-Ents and Ents

Knowledge engineers (K.E.'s), faced with the task of building an expert system, must begin with a knowledge collection stage which usually involves soliciting, from a suitable human expert, i.e., one who shares the same domain of knowledge as the required expert system, a set of facts or rules which the human expert is supposed to use when forming judgements and beliefs. This set, or knowledge base, has traditionally been assumed to take the form of a finite collection of statements. For example, if the K.E. seeks to design a system to provide diagnoses for patients visiting a health clinic, then s/he may ask a real human doctor to supply a list of statements like the following:

Disease A is very uncommon

Symptom B is an indicator of disease C, though not a strong one

etc.

Once a knowledge base like this has been obtained, the usual practice is then, in consultation with the human expert, to translate phrases such as "very uncommon" in the above into numbers which reflect the numerical degrees of belief of the expert. These numbers will be real numbers between 0 and 1, with 1 corresponding to certainty, 0 corresponding to certainty in the negation and $\frac{1}{2}$ to indifference. For example the above two statements could be translated as

$$Bel(\text{patient has disease } A) = 0.01$$
$$Bel(\text{patient has disease } C \mid \text{patient displays symptom } B) = 0.6$$

or, possibly,

$$\begin{aligned} Bel(\text{patient displays symptom } B \textit{ and } \text{has disease } C) &= \\ &= 0.6 \times Bel(\text{patient displays symptom } B). \end{aligned}$$

Here the function Bel is taken to be the real doctor's personal belief function (or conditional belief function), and the patient referred to is an entirely random, as yet unseen, visitor to the clinic. In this way the builder of the required expert system obtains a finite set S of (customarily linear) equations, or constraints, over the rationals involving the beliefs of various propositions, thus completing the knowledge collection stage of the system building process. What remains for the K.E. to do to complete the construction of the expert system is then to employ some "inference engine" which can use the knowledge contained in S to generate new conclusions. For example, in the medical setting described above, the system may be presented with a patient exhibiting a certain combination of symptoms and will then, on the basis both of these symptoms and the given S, form numerical beliefs regarding the possible different diagnoses. These beliefs should, preferably, be close to the beliefs that the original human expert would give in the same situation. The choice of which inference engine to employ will, to an extent, be governed by which additional assumptions the K.E. would like to put on the function Bel. (Note that it is extremely unlikely that S alone will determine uniquely what function Bel should be.) For example Bel may be assumed to satisfy the axioms of probability (see, for example, [3], [13]), or may be taken to be a Dempster-Shafer belief function [17], or to be a valuation in fuzzy logic (see, for example, [5]). The choice of restrictions on Bel are normally justified by considerations of how a rational, intelligent agent *ideally* acts in forming beliefs. (See [10] for a critique of each of the three approaches named above, and also some examples of possible choices for inference engine, also [18] and [9].)

The preceding paragraph provides a summary of the task of the K.E. We now turn to some of the assumptions which underlie the above programme. Firstly, although it is rather infrequently made explicit, it is often assumed that the knowledge base S which is solicited from the expert does not just *represent* the expert's knowledge but essentially *is* the expert's knowledge. Furthermore it is also implicitly assumed, with the rationality-motivated conditions on Bel in place, that the K.E.'s choice of inference engine essentially corresponds to the actual inference process that the expert him/herself uses to draw further conclusions from his/her knowledge base S. Unfortunately there are a number of criticisms which may be aimed at these assumptions. The first one is that the statements of belief supplied by the expert, i.e., the set of equations in S, usually turn out to be seriously inconsistent with whatever the rationality-motivated conditions on Bel are taken to be. For example, if Bel is assumed to be a probability function, then it will often be the case that there is no probability function which satisfies S, and similarly if Bel is taken to be a Dempster-Shafer function or a fuzzy logic valuation. (This may indicate that the notions of rationality captured by these extra conditions represent *ideals* which are rarely, if ever, achieved in reality.) Secondly, the chosen inference engine will usually require calculations and inferences which we ourselves, as intelligent human agents, are generally rather poor at. What is more the methods involved in the engine will, assuming the widely accepted hypothesis from computational complexity theory that RP $\neq$ NP, usually be computationally infeasible (see [10], [8]) so that, even if one were to forgive the expert system using extra-human processes so long as it still, somehow, reached reliable conclusions, many of the inference engines described in the literature would still leave something to be desired.

The *ent* and *pre-ent* models of belief were conceived by Paris and Vencovská in [12] with the intention of providing an alternative model of an agent's knowledge base, quite different to the usual model S described above, and of providing also an interpretation for the agent's belief function Bel. This interpretation is based on modelling the expert, with the result that Bel becomes a *derived* function, devoid of any a priori restrictions. Although these models are borne out of considerations of how the agent might actually be acting in forming beliefs, Paris and Vencovská are quick to point out ([12] p199) that

> "...we are not claiming this is the way human beings actually act, only that it is a way some entities (who might perhaps consider themselves intelligent) might conceivably be acting."

Roughly, the idea behind ents (i.e., ent models – henceforth we will normally drop the word "model") is that, when asked about his belief in a proposition, the ent's answer will be provided by the extent to which his knowledge of the world *supports* a state of affairs in which that proposition, as opposed to its negation, is true. More precisely the ent constructs imaginary, partial worlds in which the sentence is decided (positively or negatively) by combining fragments of past cases or, as we shall call them, scenarios. The belief given to the sentence is then identified with the proportional weight of these partial worlds in which it is decided positively.

In [12], the authors define the class of ents via a wider class of belief-forming "entities" called *pre-ents*. Although, as we shall describe, ents are really a refinement of pre-ents, many of our results will be directed specifically towards this more general class. We leave the formal definitions of these models until Chapter

2. We shall now give the overall plan of our material.

1.2 Plan of This Thesis

In Chapter 2 we formally define pre-ents and ents and give a summary of the properties of their resultant belief functions Bel. We shall also see how these models go some way to avoiding the criticisms of such as probability functions given in Section 1.1 before ending the chapter with a short review of probability functions. Chapter 3 is devoted to the study of some important binary relations which arise from the pre-ent and ent models. Reproduced from [12] we have the syntactic characterisation of the binary relation given by $Bel(\theta) = Bel(\phi)$ for all pre-ents. In other words we answer the question of which pairs of sentences are always treated as equivalent by any pre-ent. From this result we extract a similar characterisation of the weaker relation given by $Bel(\theta) \leq Bel(\phi)$ for all pre-ents, i.e., we answer the question of which pairs of sentences, in the world of pre-ents, should be considered as logical consequences of each other. Then, in Chapter 4, we give a detailed proof of a theorem which was first stated (without proof) in [11] which shows the essential difference, as far as their resultant belief functions are concerned, between the pre-ent and ent models. This result can be read as showing how the ent model is more general than it might first appear, in that it says that if we require of pre-ents a certain desirable property then we are lead automatically to the class of ents. It is in this chapter that we first widen our framework to include non-standard real numbers, and we again make use of this setting in the final two chapters of this thesis. The work contained in these chapters is motivated by a desire to find instances of non-monotonic consequence relations arising from pre-ents and ents. In Chapter 5 we try and fail to find instances of the *rational* consequence relations of [16] and [7]. Instead we define a new class of consequence relation, more general than rational consequence, which we call *natural* consequence relations and show how instances of *this* type of relation do occur naturally in pre-ents. We also provide another example of a family of natural consequence relation based on the framework of *permatoms* and show how the class of rational consequence relations can be characterised inside this framework. This last result is essentially the same as the one given in [7], though there are differences in the method of proof. In Chapter 6 we attempt to find an analogous characterisation for the natural case. We almost succeed, since to obtain our result we need to augment the set of rules with which we define natural consequence with a further rule. Any natural consequence relation which satisfies this extra rule we call a *fully transitive* natural consequence relation. Thus what we produce in Chapter 6 is not a characterisation of natural consequence but a characterisation of fully transitive natural consequence. We end Chapter 6 by showing how, by adding a further rule to the rules for fully transitive natural consequence, we obtain a class of relations which may be characterised in terms of *ents*.

1.3 The Propositional Setting

In this section we provide some basic background notation that will be used throughout this thesis. Our setting will be the propositional calculus. We shall always assume that L, L', etc. denote finite propositional languages, i.e., finite sets of propositional variables. Furthermore, unless explicitly stated otherwise, we take $L = \{p_1, \ldots, p_n\}$, though we will often also use p, q, r, etc. to denote propositional variables. Given a language L, we let SL denote the set of sentences built up from the variables in L in the usual way using the connectives $\neg, \wedge$ and $\vee$ (we assume $\rightarrow$ and $\leftrightarrow$ are defined via these connectives in the standard way). We use θ, ϕ, ψ, etc. to denote sentences. We use the symbol $\top$ to denote the sentence $p' \vee \neg p'$ and $\bot$ to denote $p' \wedge \neg p'$ where p' is some fixed arbitrary propositional variable in L. Given $p \in L$ we define $p^1 = p$ and $p^0 = \neg p$ and we call any sentence of the form p^ϵ for some $\epsilon \in \{0, 1\}$ a *literal* over L. We define the set of *atoms* over L, At^L, by

$$At^L = \{p_1^{\epsilon_1} \wedge p_2^{\epsilon_2} \wedge \ldots \wedge p_n^{\epsilon_n} \mid \epsilon_i \in \{0, 1\} \text{ for } i = 1, 2, \ldots, n\}.$$

We use lower-range Greek letters α, β, etc. to denote atoms. So the set of atoms consists of all conjunctions of literals α over L in which each variable in L appears precisely once and always in the same position relative to the other variables appearing in α. We let $\vdash \subseteq \mathcal{P}SL \times SL$ denote the binary relation of classical logical consequence (with the usual abuses of notation such as writing $\theta \vdash \phi$ for $\{\theta\} \vdash \phi$) and we let $\equiv \subseteq SL \times SL$ stand for the binary relation of classical logical equivalence on SL (so, for all $\theta, \phi \in SL$, $\theta \equiv \phi$ iff $\vdash \theta \leftrightarrow \phi$ iff both $\theta \vdash \phi$ and $\phi \vdash \theta$). We call θ a *tautology* if $\vdash \theta$ and a *contradiction* if $\vdash \neg\theta$. For each $\theta \in SL$ we define a set $S_\theta \subseteq At^L$ by

$$S_\theta = \{\alpha \in At^L \mid \alpha \vdash \theta\}.$$

Then, by the disjunctive normal form theorem, $\theta \equiv \bigvee S_\theta$ (irrespective of the order we take the atoms in S_θ to be in). We also have $\theta \vdash \phi$ iff $S_\theta \subseteq S_\phi$, $\vdash \theta$ iff $S_\theta = At^L$, $S_{\neg\theta} = At^L - S_\theta$, $S_{\theta \wedge \phi} = S_\theta \cap S_\phi$ and $S_{\theta \vee \phi} = S_\theta \cup S_\phi$.

With our background notation in place, we now move on to formally defining the pre-ent and ent models of belief.

Chapter 2

Pre-Ents and Ents

2.1 Introduction

In this chapter we set up the definitions of pre-ents and ents. We start with the more general class of pre-ents in Section 2.2, where we give an example of a pre-ent together with a summary of the properties of their resultant belief functions. We also indicate their connection with probability functions and sketch some of the advantages which they hold over probability functions, Dempster-Shafer functions, etc., in connection with the discussion in Section 1.1. We end that section by pointing out some of the failings of pre-ents before showing, in Section 2.3, how we can rectify some of these failings by restricting our attention to ents. We end this chapter in Section 2.4 with a brief description of the main properties of probability functions.

2.2 Pre-Ents

Before we define pre-ents we need to give the formal definition of scenario.

Definition 2.1 *We define a* scenario *(over L) to be a consistent subset of literals over L. The set of all scenarios (over L) will be denoted by WL.*

We shall use s, t, u, etc. to denote scenarios. Let us now give the definition, as given in [12], of a pre-ent over L.

Definition 2.2 *A* pre-ent over L *is a function* $G : L \times WL \times WL \rightarrow [-1, 1]$ *such that, for each* $p \in L,\ s, t \in WL$*,*

- *(i)* $G_p(s, t) > 0$ *implies* $p \in t \supseteq s$*,* $G_p(s, t) < 0$ *implies* $\neg p \in t \supseteq s$*.*
- *(ii)* $\sum_t |G_p(s, t)| = 1$*.*
- *(iii)* $p \in s$ *implies* $G_p(s, s) = 1$*,* $\neg p \in s$ *implies* $G_p(s, s) = -1$*.*

The idea behind this definition is that, for each $p \in L$ and each $s, t \in WL$, if s represents the current knowledge that the pre-ent G has about the world then the number $|G_p(s, t)|$ represents the likelihood that G, when called upon to imagine a scenario in which p is decided one way or the other, will imagine t. By condition (ii) in the above definition these likelihoods are in fact probabilities. Condition (i) says that any scenario which either does not decide p or does not contain the currently held scenario s has no chance of being imagined while condition (iii) says that if p is already decided at s (which will sometimes be written $\pm p \in s$) then no act of imagination is necessary to decide p. The use of negative values is just a useful way of indicating both the probability and the way in which p is decided.

A pre-ent G is extended to a map from $SL \times WL \times WL$ into $[-1, 1]$ inductively as follows:

$$G_{\neg\theta}(s, t) = -G_\theta(s, t)$$

$$G_{(\theta\wedge\phi)}(s, t) = \begin{cases} \sum_r G_\theta(s, r) \cdot G_\phi(r, t) & \text{if } t \vdash \theta \\ G_\theta(s, t) & \text{if } t \vdash \neg\theta \\ 0 & \text{otherwise} \end{cases}$$

$$G_{(\theta\vee\phi)}(s,t) = \begin{cases} -\sum_r G_\theta(s,r)\cdot G_\phi(r,t) & \text{if } t\vdash\neg\theta \\ G_\theta(s,t) & \text{if } t\vdash\theta \\ 0 & \text{otherwise} \end{cases}$$

The motivation behind these definitions is that, for any $\theta \in SL$, $|G_\theta(s,t)|$ should be the probability that the pre-ent G, given that s represents the current knowledge that G has about the world, will imagine the scenario t when called upon to imagine a scenario which decides θ one way or the other, with $G_\theta(s,t)$ being negative just if t decides θ negatively, i.e., $t \vdash \neg\theta$. In the case of $(\theta \wedge \phi)$, G will first imagine a scenario r which decides θ. If this scenario decides θ negatively then it must decide $(\theta \wedge \phi)$ negatively and so G stops here (this corresponds to the middle case in the above definition of $G_{(\theta\wedge\phi)}$). If r decides θ positively then G goes on to imagine a further scenario which extends r and decides the second conjunct ϕ. A similar process underlies the definition applying to the disjunction $(\theta\vee\phi)$. It may be checked, via an inductive (on θ) proof (see [12] Theorem 2.1(b)) that, under these definitions, for any $\theta \in SL$ and $s,t \in WL$,

$$G_\theta(s,t) > 0 \text{ implies } s \subseteq t \vdash \theta \text{ and } G_\theta(s,t) < 0 \text{ implies } s \subseteq t \vdash \neg\theta. \tag{2.1}$$

Definition 2.3 *Given a pre-ent G over L as described above, we define the function $Bel_s^G : SL \to [0,1]$, relative to the scenario $s \in WL$, by setting, for each $\theta \in SL$*

$$Bel_s^G(\theta) = \sum_{t\vdash\theta} G_\theta(s,t).$$

So, following the above discussion, $Bel_s^G(\theta)$ is the probability that G, given that s represents the facts about the world that G knows to be true, will, when called upon to imagine a scenario which decides θ one way or the other, imagine a scenario in which θ is true. When no confusion will arise we will omit the superscript G. In addition we will write Bel for $Bel_\emptyset$ and identify this as the pre-ent's "belief function".

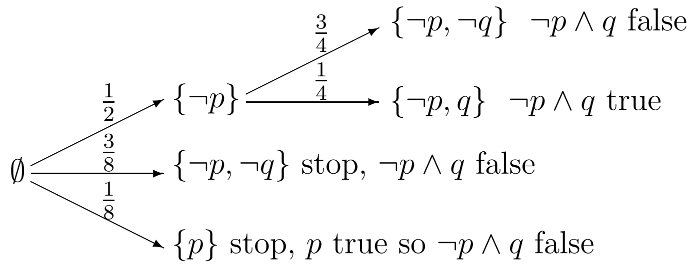
Let us straight away consider an example of the preceding definitions "in action".

Example 2.4 For this example we assume that $L = \{p, q\}$. Let G be the pre-ent over L completely specified (in that all its other values are determined by the definition of pre-ent) by the following values:

s	$\emptyset$	$\emptyset$	$\emptyset$	q	q	$\neg q$
t	p	$\neg p$	$\neg p, \neg q$	p, q	$\neg p, q$	$p, \neg q$
$G_p(s,t)$	$\frac{1}{8}$	$-\frac{1}{2}$	$-\frac{3}{8}$	$\frac{1}{5}$	$-\frac{4}{5}$	1

s	$\emptyset$	$\emptyset$	p	$\neg p$	$\neg p$
t	q	$\neg p, \neg q$	p, q	$\neg p, q$	$\neg p, \neg q$
$G_q(s,t)$	$\frac{1}{4}$	$-\frac{3}{4}$	1	$\frac{1}{4}$	$-\frac{3}{4}$

Then the following graph illustrates how G acts in computing a value for $Bel(\neg p \wedge q)$.



G begins in the situation represented by $\emptyset$ at the root of the above tree diagram, i.e., G knows nothing at all about the world. To determine his belief in $\neg p \wedge q$, G firstly constructs a scenario in which the first conjunct $\neg p$ is decided. This is what happens at the first branching in the diagram. There are three possibilities here for G. The first possibility is that he imagines $\{p\}$ with probability $\frac{1}{8}$ (corresponding to the bottom branch). In this case, since $\neg p$ is false, he has decided

$\neg p \wedge q$ false and so stops there. The second possibility is that he imagines $\{\neg p, \neg q\}$ with probability $\frac{3}{8}$ (corresponding to the middle branch). In this case G decides the first conjunct positively and so stops to decide the second conjunct q, but he realises that he is already now in a scenario in which q is decided negatively and so stops there with the whole sentence $\neg p \wedge q$ decided negatively. Finally G may imagine $\{\neg p\}$ with probability $\frac{1}{2}$ (the top branch). In this case, as in the second one just described, the first conjunct is decided positively and so G then turns his attention to the second conjunct. This time G must construct a further scenario to decide whether q is true. The two possibilities for this scenario are $\{\neg p, \neg q\}$, which has probability $\frac{3}{4}$ of being imagined from $\{\neg p\}$, and $\{\neg p, q\}$, which has probability $\frac{1}{4}$ of being imagined. In the former scenario q, and therefore the sentence $\neg p \wedge q$, is decided false while in the latter scenario the sentence is decided true. Summing the probabilities of reaching a tip in the graph at which $\neg p \wedge q$ is decided positively gives

$$Bel(\neg p \wedge q) = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8}.$$

A similar diagram can be drawn to show that, for the G of Example 2.4, we have

$$Bel(q \wedge \neg p) = \frac{1}{5}$$

and so, unlike, for example, probability functions, pre-ents do not generally treat $\wedge$ (or, for that matter, $\vee$) commutatively. The following theorem, which appeared (minus (e)) as Theorem 2.4 in [12], tells us some basic properties which Bel_s for arbitrary $s \in WL$ *does* satisfy.

Theorem 2.5 *For a pre-ent G over L, $s \in WL$ and $\theta, \phi \in SL$,*

(a) $Bel_s(\theta \vee \phi) = Bel_s(\neg(\neg\theta \wedge \neg\phi))$,

(b) $Bel_s(\theta \wedge \phi) = Bel_s(\neg(\neg\theta \vee \neg\phi))$,

(c) $Bel_s(\theta) + Bel_s(\neg\theta) = 1$,

(d) $Bel_s(\theta \vee \phi) = Bel_s(\theta) + Bel_s(\neg\theta \wedge \phi)$,

(e) $Bel_s(\theta) = Bel_s(\theta \wedge \phi) + Bel_s(\theta \wedge \neg\phi)$.

Proof. The reader is referred to [12] for the proofs of (a)–(d). Part (e) can also be proved using results in [12] but, for completeness, we shall give here a direct proof. We have

$$\begin{aligned} Bel_s(\theta \wedge \phi) &= \sum_{t \vdash \theta \wedge \phi} G_{\theta \wedge \phi}(s,t) \\ &= \sum_{t \vdash \theta \wedge \phi} \sum_{r} G_\theta(s,r) \cdot G_\phi(r,t) \\ &\quad \text{by the definition of } G_{\theta \wedge \phi}(s,t) \text{, since } t \vdash \theta. \end{aligned}$$

Now, for any $r, t \in WL$ such that $t \vdash \theta \wedge \phi$, we have, by (2.1) above, that $G_\theta(s,r) \cdot G_\phi(r,t) \neq 0$ implies $s \subseteq r \subseteq t$ and either $r \vdash \theta$ or $r \vdash \neg\theta$. But, in this case, if $r \vdash \neg\theta$ then, since $r \subseteq t$, we would have $t \vdash \neg\theta$ contradicting the consistency (by definition of scenario) of t. Hence we must have $r \vdash \theta$ and we may write

$$\begin{aligned} Bel_s(\theta \wedge \phi) &= \sum_{t \vdash \theta \wedge \phi} \sum_{r \vdash \theta} G_\theta(s,r) \cdot G_\phi(r,t) \\ &= \sum_{r \vdash \theta} G_\theta(s,r) \sum_{t \vdash \theta \wedge \phi} G_\phi(r,t). \end{aligned}$$

Now, given $r \in WL$ such that $r \vdash \theta$ and $t \in WL$, we have, again using (2.1), that $G_\phi(r,t) \neq 0$ implies $r \subseteq t$ and hence that also $t \vdash \theta$. Hence, in the above summation, we can drop the condition that $t \vdash \theta$ since it holds anyway for all t which make a non-zero contribution to the sum. Hence

$$Bel_s(\theta \wedge \phi) = \sum_{r \vdash \theta} G_\theta(s,r) \sum_{t \vdash \phi} G_\phi(r,t) = \sum_{r \vdash \theta} G_\theta(s,r) \cdot Bel_r(\phi).$$

By similar reasoning we can show

$$Bel_s(\theta \wedge \neg\phi) = \sum_{r \vdash \theta} G_\theta(s,r) \cdot Bel_r(\neg\phi).$$

Hence

$$\begin{aligned} Bel_s(\theta \wedge \phi) + Bel_s(\theta \wedge \neg\phi) &= \sum_{r \vdash \theta} G_\theta(s,r) \cdot Bel_r(\phi) + \sum_{r \vdash \theta} G_\theta(s,r) \cdot Bel_r(\neg\phi) \\ &= \sum_{r \vdash \theta} G_\theta(s,r) \cdot \{Bel_r(\phi) + Bel_r(\neg\phi)\} \\ &= \sum_{r \vdash \theta} G_\theta(s,r) \qquad \text{by (c) of this theorem} \\ &= Bel_s(\theta) \end{aligned}$$

as required. □

Note that an immediate corollary of part (e) of the above Theorem is that, for any pre-ent G and $s \in WL$, for all $\theta, \phi \in SL$ we have

$$Bel_s(\theta \wedge \phi) \leq Bel_s(\theta).$$

The properties listed in the above theorem 2.5 are all also satisfied by any probability function. The following theorem (Theorem 2.5 of [12]) reveals the conditions under which Bel_s may be identified with such a function.

Theorem 2.6 *For any pre-ent G over L and $s \in WL$, Bel_s is a probability function on L iff one of the following two equivalent conditions hold:*
(1) For all $\theta, \phi \in SL$, $Bel_s(\theta \wedge \phi) = Bel_s(\phi \wedge \theta)$,
(2) For all $\theta, \phi \in SL$, $Bel_s(\theta \vee \phi) = Bel_s(\phi \vee \theta)$. □

The following two theorems (2.6 and 2.9 in [12]) complete the basic properties of Bel_s for a pre-ent.

Theorem 2.7 *Let $\theta \in SL$ and $s \in WL$. Then $s \vdash \theta$ iff $Bel_s(\theta) = 1$ for all pre-ents over L. In particular $\vdash \theta$ iff $Bel(\theta) = 1$ for all pre-ents over L.* □

Theorem 2.8 *Let $\theta \in SL$. Then θ is satisfiable iff there exists a pre-ent G over L such that $Bel(\theta) = 1$.* □

As they stand, pre-ents have a number of attractive features. One of these relates to testing the consistency of a set of constraints S such as described in Section 1.1. In general it is computationally infeasible to test whether a given set of constraints S is consistent with Bel being a probability function, Dempster-Shafer function, etc. By Theorem 2.10 of [12], provided the propositions θ_i appearing in S are of short bounded length, it *is* feasible to test whether there is a pre-ent satisfying S and, if so, to describe such a pre-ent. Another advantage relates to one of the arguments given against such as probability functions in Section 1.1. As indicated there, there is a problem concerning the infeasibility of the methods proposed to generate new beliefs from S. In particular it is infeasible, given a consistent S and $\theta \in SL$, to compute an approximation to a value for $Bel(\theta)$ which is consistent with S. However this is not the case for a pre-ent, in fact it is feasible to find, at least with high probability of success, a good approximation to $Bel(\theta)$ (or, more generally, $Bel_s(\theta)$) in time linear in the length of θ. (For the precise details see [12].)

The belief functions yielded by pre-ents may be thought of as having much in common with probability functions, with the big difference being that they do not necessarily treat $\wedge$ (or $\vee$) commutatively – they are generally sensitive to the order in which they receive, or review, pieces of information. Indeed this sensitivity can, in some cases, lead to a pre-ent giving $\theta \wedge \phi$ belief value 0 while giving $\phi \wedge \theta$ belief value > 0, or even value 1, for some $\theta, \phi \in SL$. For consider the pre-ent G over a language L such that $L \supseteq \{p, q\}$ for which we have $G_p(\emptyset, \{\neg p, q\}) = -1$ and $G_q(\emptyset, \{p, q\}) = 1$. For this G we have $Bel^G(p \wedge q) = 0$ (indeed $Bel^G(p) = 0$) while $Bel^G(q \wedge p) = 1$. According to the following proposition, this means that the set of sentences believed with value 1 by a given pre-ent need not be closed under logical consequences.

Proposition 2.9 *Let* $Bel : SL \to [0, 1]$ *be a function yielded by some pre-ent*

(i.e., $Bel = Bel^G_\emptyset$ for some pre-ent G over L). Then the following are equivalent:
(i). For all $\theta, \phi \in SL$, if $Bel(\theta \wedge \phi) = 0$ then $Bel(\phi \wedge \theta) = 0$.
(ii). For all $k \geq 0$ and for all $\theta_1, \ldots, \theta_k, \psi \in SL$, if $\theta_1, \ldots, \theta_k \vdash \psi$ and $Bel(\theta_i) = 1$ for $i = 1, \ldots, k$, then $Bel(\psi) = 1$.

Proof. To show that (i) implies (ii), suppose the function Bel is given by some pre-ent G. Let $\theta_1, \ldots, \theta_k, \psi \in SL$ be such that $\theta_1, \ldots, \theta_k \vdash \psi$ and suppose $Bel(\theta_i) = 1$ for $i = 1, \ldots, k$. We must show $Bel(\psi) = 1$. We first notice that

$$\theta_1, \ldots, \theta_k \vdash \psi \text{ iff } \vdash \neg\theta_1 \vee \neg\theta_2 \vee \ldots \vee \neg\theta_k \vee \psi.$$

Hence, from Theorem 2.7, we have

$$Bel(\neg\theta_1 \vee \neg\theta_2 \vee \ldots \vee \neg\theta_k \vee \psi) = 1$$

and so it suffices to show that, for any $\theta, \phi \in SL$, if $Bel(\theta) = 1 = Bel(\neg\theta \vee \phi)$ then $Bel(\phi) = 1$. But if $Bel(\theta) = 1$ then $Bel(\neg\theta) = 0$ by Theorem 2.5(c) which, in turn, gives $Bel(\neg\theta \wedge \neg\phi) = 0$ using Theorem 2.5(e). Applying the condition (i) then gives us $Bel(\neg\phi \wedge \neg\theta) = 0$. Meanwhile, by Theorem 2.5(a), $Bel(\neg\theta \vee \phi) = 1$ implies $Bel(\neg(\neg\neg\theta \wedge \neg\phi)) = 1$, which implies $Bel(\neg\neg\theta \wedge \neg\phi) = 0$ by Theorem 2.5(c). Applying condition (i) here gives us $Bel(\neg\phi \wedge \neg\neg\theta) = 0$. So we have

$$\begin{aligned} Bel(\neg\phi) &= Bel(\neg\phi \wedge \neg\theta) + Bel(\neg\phi \wedge \neg\neg\theta) \qquad \text{from Theorem 2.5(e)} \\ &= 0 \end{aligned}$$

which gives $Bel(\phi) = 1$ by Theorem 2.5(c) as required.
To show that (ii) implies (i) suppose $Bel(\theta \wedge \phi) = 0$. Then, using Theorem 2.5(c), we have $Bel(\neg(\theta \wedge \phi)) = 1$. Now, since $\neg(\theta \wedge \phi) \vdash \neg(\phi \wedge \theta)$, we may apply condition (ii) to obtain $Bel(\neg(\phi \wedge \theta)) = 1$ which in turn gives $Bel(\phi \wedge \theta) = 0$ as required. □

It is easy to see that the above proposition remains true if we replace "$Bel = Bel^G_\emptyset$ for some pre-ent G over L" by "$Bel = Bel^G_s$ for some pre-ent G over L and some arbitrary $s \in WL$".

The fact that a pre-ent's belief function can fail to validate either one of the conditions of Proposition 2.9 can be seen as a drawback of pre-ents, since both are certainly desirable properties to have. Luckily there is a way in which we can force these properties to hold while keeping the attractive features of pre-ents – we simply restrict attention to the subclass of pre-ents called the ents. As we shall see in the next section, ents also enjoy some other advantages over pre-ents.

2.3 Ents

The idea behind ents is similar to that behind pre-ents. An ent consists of a store of scenarios with each scenario being assigned a *potential* which represents the ease with which it springs to mind. Suppose that $s \in WL$ represents all that the ent z knows about the world. When called upon to imagine a scenario in which the variable p is decided one way or the other the ent will pull from his store of scenarios a scenario t which is *consistent* with s (rather than just extends s as with pre-ents) and which satisfies $\pm p \in t$ (we assume $\pm p \notin s$ – of course z need do nothing at this stage if it is asked about p and is already aware of the truth or falsity of p). The ent then enlarges his currently known facts from s to $s \cup t$. The main departure from pre-ents is now that (given t consistent with s, $\pm p \in t$) the likelihood of the scenario t being chosen at this instance does not depend on the currently held scenario s or the particular propositional variable p being decided. In order for this to yield a pre-ent we require that such a t always exists with $z_t > 0$. The precise definition is as follows:

Definition 2.10 *An* ent *over L is a map $z : WL \to [0, \infty)$ such that, for all*

$s \in WL$ and $p \in L$, if p is not decided by s then there is some $t \in WL$ consistent with s, deciding p, and such that $z_t > 0$.

An ent z then yields a pre-ent G^z, which in turn yields the belief function Bel^z, simply by setting, for $\pm p \notin s$ and $s \subseteq t \in WL$,

$$G_p^z(s,t) = \begin{cases} \dfrac{\sum\{z_r \mid s \cup r = t\}}{\sum\{z_r \mid s \cup r \text{ consistent and } \pm p \in r\}}, & \text{if } p \in t, \\ \dfrac{-\sum\{z_r \mid s \cup r = t\}}{\sum\{z_r \mid s \cup r \text{ consistent and } \pm p \in r\}}, & \text{if } \neg p \in t, \\ 0, & \text{if } \pm p \notin t. \end{cases}$$

We of course set $G_p^z(s,t) = 0$ if $s \not\subseteq t$ and $G_p^z(s,s) = 1$ ($G_p^z(s,s) = -1$) if $p \in s$ ($\neg p \in s$). Note that the condition expressed in Definition 2.10 ensures that the above denominators are never equal to zero and so we can be sure that the function Bel^z is well-defined. In fact, this condition may be said to be a little too strong for this purpose, and we will have cause to relax it later in this thesis (see Chapter 4 Section 4.3). As an example of an ent, consider that ent z given by the following tableau.

s	$\{p\}$	$\{\neg p, \neg q\}$	$\{\neg p\}$	$\{q\}$
z_s	1	3	4	1

Then it is straightforward to check that z yields the pre-ent G of Example 2.4 in the preceding section. It follows that restricting attention to ents does nothing to alleviate the problem of the non-commutativity of $\wedge$ and $\vee$. However whether or not this is actually a "problem" is debatable. After all, it is true that human beings, as one example of intelligent agents, do not always treat sentences of the form $\theta \wedge \phi$ the same as $\phi \wedge \theta$ (such as when $\wedge$ is given a reading which includes causal or temporal aspects – see [12] for an example). In addition, if we were to insist on commutativity then, by Theorem 2.6, this would force Bel to be a probability function and thus leave Bel open to the criticisms of Section 1.1.

Despite this possible fallibility in ents, they do enjoy some important advantages over pre-ents. First of all specifying an ent only requires polynomial storage space. This contrasts with the situation regarding pre-ents, which require space exponential in the size of the language to specify them. Secondly (by Theorem 2.16 of [12]) we now have closure under logical consequence of sentences believed with certainty:

Theorem 2.11 *Let z be an ent over L and let $\theta_1, \ldots, \theta_k, \psi \in SL$. If $Bel^z(\theta_i) = 1$ for $i = 1, \ldots, k$ and $\theta_1, \ldots, \theta_k \vdash \psi$ then $Bel^z(\psi) = 1$.* □

(In fact, as we shall see in Chapter 4, as far as their resultant belief functions are concerned, this is essentially the *only* difference between ents and pre-ents.) Finally the way ents are represented – as a collection of scenarios together with their associated potentials – would seem to lend itself naturally to a process in which the ent learns about the external world by simply absorbing its experiences as it goes along. For example if the ent "witnesses" a situation in which p and q are true, then it could increase the potential it gives to the scenario $\{p, q\}$ by some fixed amount. (For a concrete example of a possible learning strategy for ents, see [4].) Thus, if we were to make the assumption that the provider of the constraint set S from Section 1.1 was an ent, then these scenarios, with their associated potentials, might be thought of as the building blocks of the expert's statements of knowledge S, and so it is they, and not S, which perhaps more truly represent the "knowledge base" of the expert. We close this review of pre-ents and ents by pointing out that there are some clear similarities between ents and case-based reasoning, as described for example in [6].

2.4 Probability Functions

At numerous points in this work (indeed several times already in the preceding sections) we shall compare the belief functions of pre-ents and ents with probability functions. In this short section we give the formal definition of probability functions, define the notion of conditional probability and state a simple representation result for probability functions. The definition of a probability function, given relative to a given propositional language L, is as follows:

Definition 2.12 *A* probability function over L *is a function* $F : SL \to [0, 1]$ *which satisfies the following axioms, for all* $\theta, \phi \in SL$*:*
(P1) If $\vdash \theta$ *then* $F(\theta) = 1$.
(P2) If $\vdash \neg(\theta \wedge \phi)$ *then* $F(\theta \vee \phi) = F(\theta) + F(\phi)$.

An important consequence of these axioms (see [10] for a proof) is that any probability function F over L satisfies, for all $\theta, \phi \in SL$,

$$\theta \vdash \phi \text{ implies } F(\theta) \leq F(\phi),$$

(and so $\theta \equiv \phi$ implies $F(\theta) = F(\phi)$).

Definition 2.13 *Given a probability function* $F : SL \to [0, 1]$ *and* $\theta, \phi \in SL$ *such that* $F(\theta) > 0$*, we define the* conditional probability *(relative to* F*) of* ϕ given θ *by,*

$$F(\phi \mid \theta) = \frac{F(\theta \wedge \phi)}{F(\theta)}.$$

It is easy to see that, for a fixed $\theta \in SL$ such that $F(\theta) > 0$, the function $F(\cdot \mid \theta)$ is also a probability function.

If $F : SL \to [0, 1]$ is a probability function over L then, for all $\theta \in SL$, since $\theta \equiv \bigvee S_\theta$ and $\vdash \neg(\alpha \wedge \bigvee S)$ for any $S \subseteq At^L$ such that $\alpha \notin S$, we may repeatedly apply axiom P2 to get

$$F(\theta) = F(\bigvee S_\theta) = \sum_{\alpha \in S_\theta} F(\alpha).$$

While we also have

$$\sum_{\alpha \in At^L} F(\alpha) = F(\bigvee At^L) = 1$$

by P1, since $\vdash \bigvee At^L$. Hence any probability function over L is completely determined by its values on the atoms of L. Conversely, if $F : SL \to [0,1]$ is any arbitrary function on SL which satisfies, for all $\theta \in SL$,

$$F(\theta) = \sum_{\alpha \in S_\theta} F(\alpha) \text{ and } \sum_{\alpha \in At^L} F(\alpha) = 1$$

then it is straightforward to show that F satisfies axioms P1-2 and so is a probability function over L. Hence we have a simple representation result for probability functions over a language L.

Chapter 3

The Logic of Pre-Ents and Ents

3.1 Introduction

This chapter is concerned mainly with certain special binary relations on SL which arise from pre-ents. These relations will be fundamental in Chapters 5 and 6. In the first section we begin by looking at the relation $\overset{\cdot}{\sim}$ given by $G_\theta = G_\phi$ for all pre-ents. The syntactic characterisation of this relation given by Paris and Vencovská in [12] helped them to give a similar characterisation for the weaker relation $\overset{\cdot\cdot}{\sim}$ given by $Bel(\theta) = Bel(\phi)$ for all pre-ents (which, as it happens, is equivalent to saying $Bel(\theta) = Bel(\phi)$ for all *ents*). Both these characterisations will be given in the next section, along with some simple examples of the type of syntactic manipulation of sentences which we perform under them. We shall also prove a couple of closure conditions which $\overset{\cdot\cdot}{\sim}$ satisfies and give a result from [4] which shows how we can express the relation of classical logical consequence $\vdash$ in terms of $\overset{\cdot\cdot}{\sim}$. In Section 3.3 we show how each sentence can be reduced to a type of "normal form" to which it is equivalent for pre-ents (and ents) and show that the belief functions yielded by pre-ents are completely determined by the values they give to all conjunctions of literals from distinct propositional variables. Finally, in Section 3.4, we build on the results described in Section 3.2 by giving a syntactic characterisation of the relation $\mid\!\ddot{\sim}$ given by $Bel(\theta) \leq Bel(\phi)$ for all pre-ents.

3.2 Logical Equivalence for Pre-Ents and Ents

In this section we examine the question of which pairs of sentences are equivalent for pre-ents (and ents), i.e., given, $\theta, \phi \in SL$, when is it the case that $Bel(\theta) = Bel(\phi)$ for all pre-ents (and ents). In [12] Paris and Vencovská give a syntactic characterisation of the set of pairs of sentences for which this holds. They achieve this via the following syntactic characterisation (Theorem 2.7 of [12]) of the stronger relation $G_\theta = G_\phi$ for all pre-ents over L.

Theorem 3.1 *Let the relation* $\dot{\sim} \subseteq SL \times SL$ *be defined by, for* $\theta, \phi \in SL$, $\theta \dot{\sim} \phi$ *iff* $G_\theta = G_\phi$ *for all pre-ents* G *over* L. *Then* $\dot{\sim}$ *is the (unique) smallest relation* $\sim$ *on* SL *which satisfies:*
(i). If $\theta_1 \sim \phi_1$ *and* $\theta_2 \sim \phi_2$ *then* $(\theta_1 \wedge \theta_2) \sim (\phi_1 \wedge \phi_2)$, $(\theta_1 \vee \theta_2) \sim (\phi_1 \vee \phi_2)$ *and* $\neg\theta_1 \sim \neg\phi_1$.
(ii). $\sim$ *is an equivalence relation on* SL.

(iii). $\theta \sim \neg\neg\theta$, $\quad \theta \wedge (\phi \wedge \psi) \sim (\theta \wedge \phi) \wedge \psi$,
$\neg(\theta \wedge \phi) \sim \neg\theta \vee \neg\phi$, $\quad \theta \vee \phi \sim \theta \vee (\neg\theta \wedge \phi)$,
$\theta \wedge \neg\theta \sim \neg\theta \wedge \theta$, $\quad \theta \wedge \theta \sim \theta$,
$\theta \wedge (\phi \wedge \theta) \sim \theta \wedge \phi$, $\quad (\phi \vee \neg\phi) \vee \theta \sim \phi \vee \neg\phi$,
$\theta \wedge (\phi \vee \psi) \sim (\theta \wedge \phi) \vee (\theta \wedge \psi)$,
$(\theta \vee \phi) \wedge \psi \sim (\theta \wedge \psi) \vee (\neg\theta \wedge (\phi \wedge \psi))$. □

Note that the relation $\dot{\sim}$ defined in the above theorem was given relative to an underlying language L. However, given the above result, it should be clear that whether or not $\theta \dot{\sim} \phi$ holds is actually independent of what we take this underlying language to be. In other words, if $\theta, \phi \in SL_1 \cap SL_2$ for different languages L_1 and L_2, then $G_\theta = G_\phi$ for all pre-ents over L_1 iff $G_\theta = G_\phi$ for all pre-ents over L_2. Also, as is shown in [12], Theorem 3.1 remains true when we replace "for all pre-ents over L" by "for all *ents* over L". From the base set of axioms and inference rules given in Theorem 3.1 we may derive the following (see [12] Lemma A.7 for proof):

Proposition 3.2 *Let $\sim$ be any binary relation on SL satisfying (i) – (iii) from Theorem 3.1. Then $\sim$ satisfies the following:*

(a) $\theta \sim \neg\neg\theta$,

(b) $\neg(\theta \wedge \phi) \sim \neg\theta \vee \neg\phi$,

(c) $\theta \wedge \neg\theta \sim \neg\theta \wedge \theta$,

(d) $\theta \wedge (\phi \wedge \theta) \sim \theta \wedge \phi$,

(e) $\theta \wedge (\phi \vee \psi) \sim (\theta \wedge \phi) \vee (\theta \wedge \psi)$,

(f) $(\theta \vee \phi) \wedge \psi \sim (\theta \wedge \psi) \vee (\neg\theta \wedge (\phi \wedge \psi))$,

(g) $\theta \wedge (\phi \wedge \psi) \sim (\theta \wedge \phi) \wedge \psi$,

(h) $\theta \vee \phi \sim \theta \vee (\neg\theta \wedge \phi)$,

(j) $\theta \wedge \theta \sim \theta$,

(k) $(\phi \vee \neg\phi) \vee \theta \sim \phi \vee \neg\phi$,

(l) $\neg(\theta \vee \phi) \sim \neg\theta \wedge \neg\phi$,

(m) $\theta \vee \neg\theta \sim \neg\theta \vee \theta$,

(n) $\theta \vee \theta \sim \theta$,

(o) $(\phi \wedge \neg\phi) \wedge \theta \sim \phi \wedge \neg\phi$,

(p) $\theta \vee (\phi \wedge \psi) \sim (\theta \vee \phi) \wedge (\theta \vee \psi)$,

(q) $(\phi \wedge \psi) \vee \theta \sim (\phi \vee \theta) \wedge (\neg\phi \vee (\psi \vee \theta))$,

(r) $\theta \vee (\phi \vee \psi) \sim (\theta \vee \phi) \vee \psi$,

(s) $\theta \wedge \phi \sim \theta \wedge (\neg\theta \vee \phi)$,

(t) $(\theta \wedge \phi) \vee (\theta \wedge \neg\theta) \sim \theta \wedge \phi$,

(u) $(\theta \wedge \phi) \vee (\neg\theta \wedge \psi) \sim (\neg\theta \wedge \psi) \vee (\theta \wedge \phi)$. □

Out of the above list, rules (a)–(k) are just the rules listed under (iii) in Theorem 3.1 repeated for convenience while (l)–(s) are just dual versions of some of those rules. One thing to note about the above rules is the non-appearance of $\theta \wedge \phi \sim \phi \wedge \theta$ or $\theta \vee \phi \sim \phi \vee \theta$. However we do have a restricted form of commutativity via the rule (u). In future proofs in this thesis concerning the relation $\dot\sim$ we shall drop explicit mention of the uses of (i) and (ii) from Theorem 3.1 in proofs, while, in view of (g) and (r) above, we shall often omit multiple parentheses from conjunctions and disjunctions of more than two sentences. Also, in view of (a) from the above proposition, we shall treat θ and $\neg\neg\theta$ interchangeably from now on. As an example of a proof involving the properties described above we now show the following derived rule which will turn out to provide a useful short-cut in proving one or two of our subsequent results.

Proposition 3.3 *Let $\sim$ be any binary relation on SL satisfying (i) – (iii) from Theorem 3.1. Then $\sim$ satisfies:*

$$\theta \wedge (\phi \wedge \psi) \sim \neg(\theta \wedge \neg\phi) \wedge (\theta \wedge \psi).$$

Proof. In the proof below the letters on the right-hand side correspond to the relevant properties of $\sim$ from Proposition 3.2 which we are using at each step.

$$\begin{array}{rcll}
\theta \wedge (\phi \wedge \psi) & \sim & (\theta \wedge (\phi \wedge \psi)) \vee (\neg\theta \wedge \theta) & \text{(t),(c)} \\
& \sim & (\theta \wedge (\phi \wedge \psi)) \vee (\neg\theta \wedge \theta \wedge \psi) & \text{(o),(c)} \\
& \sim & (\neg\theta \wedge \theta \wedge \psi) \vee (\theta \wedge (\phi \wedge \psi)) & \text{(u)} \\
& \sim & (\neg\theta \wedge \theta \wedge \psi) \vee (\theta \wedge \phi \wedge \theta \wedge \psi) & \text{(d)} \\
& \sim & (\neg\theta \vee \phi) \wedge (\theta \wedge \psi) & \text{(f)} \\
& \sim & \neg(\theta \wedge \neg\phi) \wedge (\theta \wedge \psi) & \text{(b)}
\end{array}$$

□

The characterisation of the relation $\dot\sim$ given in Theorem 3.1 paves the way for the following characterisation of the more general relation of $Bel(\theta) = Bel(\phi)$

for all pre-ents. It should be noted that the following theorem appeared in [12] (as Theorem 2.8), although the last condition under part (v) (i.e. $\theta \sim \theta \vee \neg\phi_1$ for ϕ_1 any classical tautology) was omitted from there in error.

Theorem 3.4 *Let the relation* $\ddot{\sim} \subseteq SL \times SL$ *be defined by, for* $\theta, \phi \in SL$, $\theta \,\ddot{\sim}\, \phi$ *iff* $Bel(\theta) = Bel(\phi)$ *for all pre-ents over* L. *Then* $\ddot{\sim}$ *is the (unique) smallest relation* $\sim$ *on* SL *which extends* $\dot{\sim}$ *and satisfies:*
(iv). $\sim$ *is an equivalence relation.*
(v). For $\theta \in SL$ *and* $\phi_1, \phi_2 \in SL$ *such that* $\vdash \phi_i$ *for* $i = 1, 2$ *we have* $\phi_1 \sim \phi_2$, $\neg\phi_1 \sim \neg\phi_2$, $\theta \sim \theta \wedge \phi_1$ *and* $\theta \sim \theta \vee \neg\phi_1$. □

Once again it should be clear that the definition of $\ddot{\sim}$ is independent of the underlying language and again, as is shown in [12], Theorem 3.4 remains true when we replace "for all pre-ents over L" by "for all ents over L". Note that, for any relation $\sim$ which extends $\dot{\sim}$ and satisfies (iv) and (v) above, we have $\vdash \neg\phi$ implies $\theta \sim \theta \vee \phi$. This follows since $\vdash \neg\phi$ implies $\theta \sim \theta \vee \neg\neg\phi$ by (v), while $\theta \vee \neg\neg\phi \,\dot{\sim}\, \theta \vee \phi$.

For a simple example of a proof involving $\ddot{\sim}$ we give the following result, which will be used freely in some of our later proofs. Again we omit explicit mention whenever we use the fact that $\ddot{\sim}$ is an equivalence relation.

Proposition 3.5 *Let* $\sim$ *be any binary relation on* SL *extending* $\dot{\sim}$ *and satisfying (iv) - (v) from Theorem 3.4. Then* $\sim$ *satisfies*

$$(\theta \vee \phi) \wedge \theta \sim \theta.$$

Proof. We have

$$\begin{aligned} (\theta \vee \phi) \wedge \theta \ &\dot{\sim}\ (\theta \wedge \theta) \vee (\neg\theta \wedge \phi \wedge \theta) && \text{from Proposition 3.2(f)} \\ &\sim\ \theta \wedge \theta && \text{since } \vdash \neg(\neg\theta \wedge \phi \wedge \theta) \\ &\dot{\sim}\ \theta && \text{from Proposition 3.2(j).} \end{aligned}$$

Hence, remembering that $\sim$ extends $\dot{\sim}$, we have $(\theta \vee \phi) \wedge \theta \sim \theta$ as required. $\square$

Note that, since the relation $\equiv$ of logical equivalence satisfies the conditions of Theorem 3.4, we have that, for all $\theta, \phi \in SL$, $\theta \ddot{\sim} \phi$ implies $\theta \equiv \phi$. However the converse is false since, as we have already seen, it is not necessarily the case that $\theta \wedge \phi \ddot{\sim} \phi \wedge \theta$. Another point to be wary about is that it is *not* necessarily the case (unlike for the relation $\dot{\sim}$) that $\theta_1 \ddot{\sim} \theta_2$ implies $\theta_1 \wedge \phi \ddot{\sim} \theta_2 \wedge \phi$. We do, however, have $\theta \ddot{\sim} \phi$ implies $\neg\theta \ddot{\sim} \neg\phi$ (as is easily seen from the fact that $Bel(\neg\theta) = 1 - Bel(\theta)$ for all pre-ents) and also the following closure conditions on $\ddot{\sim}$ (I am grateful to J. B. Paris for providing the idea behind the proof of this result):

Proposition 3.6 *Let $\theta, \phi_1, \phi_2 \in SL$. Then $\phi_1 \ddot{\sim} \phi_2$ implies $\theta \wedge \phi_1 \ddot{\sim} \theta \wedge \phi_2$ and $\theta \vee \phi_1 \ddot{\sim} \theta \vee \phi_2$.*

Proof. Fix $\theta \in SL$ and define a relation $\sim_\theta \subseteq SL \times SL$ by setting, for all $\phi_1, \phi_2 \in SL$, $\phi_1 \sim_\theta \phi_2$ iff $\theta \wedge \phi_1 \ddot{\sim} \theta \wedge \phi_2$. We will show that the relation $\sim_\theta$ satisfies all the conditions of Theorem 3.4 which will mean, since that theorem tells us that $\ddot{\sim}$ is the *smallest* binary relation on SL which satisfies those conditions, that $\phi_1 \ddot{\sim} \phi_2$ implies $\phi_1 \sim_\theta \phi_2$ for all $\phi_1, \phi_2 \in SL$ as required. Firstly we have that $\sim_\theta$ extends $\dot{\sim}$, since $\phi_1 \dot{\sim} \phi_2$ implies $\theta \wedge \phi_1 \dot{\sim} \theta \wedge \phi_2$ implies (since $\ddot{\sim}$ extends $\dot{\sim}$) $\theta \wedge \phi \ddot{\sim} \theta \wedge \phi_2$. Secondly it is easy to check that $\sim_\theta$ is an equivalence relation on SL (since $\ddot{\sim}$ is). Finally we must show that $\sim_\theta$ satisfies condition (v) from Theorem 3.4. So let $\lambda, \chi_1, \chi_2 \in SL$ be such that $\vdash \chi_i$ for $i = 1, 2$. Then $\chi_1 \sim_\theta \chi_2$ follows since we already have $\theta \ddot{\sim} \theta \wedge \chi_i$ for $i = 1, 2$, while $\neg\chi_1 \sim_\theta \neg\chi_2$ follows since $\vdash \chi_i$ implies $\vdash \neg(\theta \wedge \neg\chi_i)$ for $i = 1, 2$ and so $\theta \wedge \neg\chi_1 \ddot{\sim} \theta \wedge \neg\chi_2$. To show $\lambda \sim_\theta \lambda \wedge \chi_1$ we have $\theta \wedge \lambda \ddot{\sim} (\theta \wedge \lambda) \wedge \chi_1 \dot{\sim} \theta \wedge (\lambda \wedge \chi_1)$. Lastly to show $\lambda \sim_\theta \lambda \vee \neg\chi_1$ we have $\theta \wedge \lambda \ddot{\sim} (\theta \wedge \lambda) \vee \neg\neg(\theta \wedge \neg\chi_1) \dot{\sim} (\theta \wedge \lambda) \vee (\theta \wedge \neg\chi_1) \dot{\sim} \theta \wedge (\lambda \vee \neg\chi_1)$ as required.

Since $\theta \vee \phi_1 \dot{\sim} \neg(\neg\theta \wedge \neg\phi_1)$, the second part of the proposition now follows easily from the first part and the fact that, for all ψ_1, ψ_2, $\psi_1 \ddot{\sim} \psi_2$ implies $\neg\psi_1 \ddot{\sim} \neg\psi_2$. □

We end the present section with a result given by Gladstone in [4] (Lemma 3) which provides a link between $\ddot{\sim}$ and the classical logical consequence relation $\vdash$.

Theorem 3.7 *Let $\theta, \phi \in SL$. Then $\theta \vdash \phi$ iff $\theta \ddot{\sim} \theta \wedge \phi$.* □

3.3 Normal Forms and Trees

Theorems 3.1 and 3.4 were actually proved in [12] by utilising two types of "normal form" that exist for each sentence θ. These forms, especially the second one corresponding to the relation $\ddot{\sim}$, will often be very useful in what follows. Indeed we shall use them to help us characterise the relation $\mid\!\ddot{\sim}$ in Section 3.4. First of all we shall give the normal form which corresponds to $\dot{\sim}$. The following lemma appeared as Lemma A.8 in [12].

Lemma 3.8 *For each $\theta \in SL$ there exists literals $p_{i,j}^{\epsilon_{i,j}}$ for $i = 1, \ldots, m$, $j = 1, \ldots, e(i)$, with $e(i) \geq 2$, such that for any relation $\sim \subseteq SL \times SL$ which satisfies (i), (ii) and (iii) from Theorem 3.1,*

(1) $\theta \sim \bigvee_{i \leq m} \bigwedge_{j \leq e(i)} p_{i,j}^{\epsilon_{i,j}}$.

(2) For $i = 1, \ldots, m$, $p_{i,e(i)} = p_{i,e(i)-1}$. If $\epsilon_{i,e(i)} = \epsilon_{i,e(i)-1}$ then we call $\bigwedge_{j \leq e(i)} p_{i,j}^{\epsilon_{i,j}}$ a positive *clause, otherwise a* negative *clause.*

(3) For $i = 1, \ldots, m$ and $k < j < e(i)$, $p_{i,j} \neq p_{i,k}$.

(4) For $1 \leq i < k \leq m$ there is $j < e(i), e(k)$ such that $p_{i,j}^{\epsilon_{i,j}} = p_{i,j}$, $p_{k,j}^{\epsilon_{k,j}} = \neg p_{i,j}$ and $p_{i,r}^{\epsilon_{i,r}} = p_{k,r}^{\epsilon_{k,r}}$ for $1 \leq r < j$.

(5) For $1 \leq i \leq m$ and $j < e(i)$ there is $1 \leq k \leq m$ such that $p_{i,j} = p_{k,j}$, $\epsilon_{i,j} \neq \epsilon_{k,j}$ and $p_{i,r}^{\epsilon_{i,r}} = p_{k,r}^{\epsilon_{k,r}}$ for $1 \leq r < j$. □

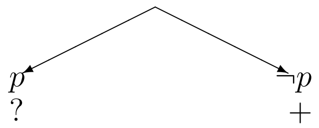
The sequences $p_{i,1}^{\epsilon_{i,1}}, \ldots, p_{i,e(i)-1}^{\epsilon_{i,e(i)-1}}$ above may be thought of as the paths through a binary tree such that no path contains a repeated propositional variable and each node has just two edges out of it labelled p and $\neg p$ for some $p \in L$. The $p_{i,e(i)}^{\epsilon_{i,e(i)}}$ are joined to the end of these paths in order to label them positive or negative. This idea should become clearer when we take in an example below.

As is shown in [12], for each θ, the sentence $\bigvee_{i \leq m} \bigwedge_{j \leq e(i)} p_{i,j}^{\epsilon_{i,j}}$ in the above lemma is unique and is given the following name:

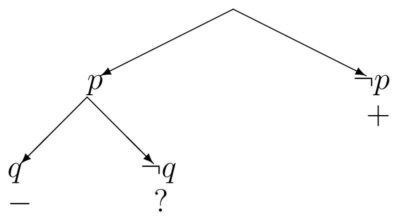
Definition 3.9 *Given $\theta \in SL$, the unique sentence $\bigvee_{i \leq m} \bigwedge_{j \leq e(i)} p_{i,j}^{\epsilon_{i,j}}$ from Lemma 3.8 is denoted by $cT(\theta)$. We denote the set of positive clauses (without their last repeated literals) of $cT(\theta)$ by $cT(\theta)^+$.*

Thus we have, for all $\theta, \phi \in SL$, $\theta \dot{\sim} cT(\theta)$ and (from [12]) $\theta \dot{\sim} \phi$ iff $cT(\theta) = cT(\phi)$. The full inductive process for finding $cT(\theta)$ for any given $\theta \in SL$ may be found in [12]. We content ourselves here with providing an example which will hopefully illustrate how this can be done.

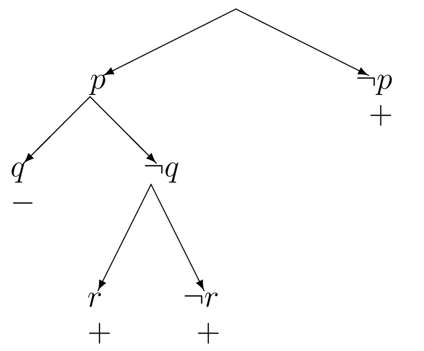
Example 3.10 Let $L = \{p, q, r\}$ and let $\theta = \neg p \vee (\neg q \wedge (r \vee \neg r))$. We will construct the binary tree which corresponds to $cT(\theta)$. We begin reading θ from left to right until we come to a propositional variable. The first variable we find is p, so we begin at the root of our tree with two branches, the left one leading to a node labelled p and the right one leading to a node labelled $\neg p$. Taking the right branch first, we see that if $\neg p$ is true then, since the main connective of θ is $\vee$ and since we now know the first disjunct is true, we know the whole sentence θ is true and so we may stop here and label this path as being positive. If we take the left branch and suppose p is true, then the first disjunct of θ is false and so we must continue to the second disjunct $\neg q \wedge (r \vee \neg r)$ of θ to decide whether θ is true. Thus our current position may be represented by the following tree diagram:



So, given we are at the node labelled with p, we must now read $\neg q \wedge (r \vee \neg r)$ from left to right until we find a propositional variable. The first variable we come across is q, so we draw two edges out of p, the left one leading to a node labelled q and the right one leading to a node labelled $\neg q$. On the left branching we see that $\neg q$ is decided negatively and so, since the principal connective of the second disjunct of θ is a $\wedge$, that second disjunct is also decided negatively and so the whole sentence θ is decided negatively at this point leading us to stop here and label this path negative. On the right branching $\neg q$ is decided positively and so, since θ is still undecided, we must carry on to the second conjunct $r \vee \neg r$. Now our current position is as follows:



Now, given that we have reached the node labelled $\neg q$ via the node labelled p we carry on moving left to right through $r \vee \neg r$ looking for the next propositional variable. This variable is r and so we have two edge leading out from $\neg q$, the left one leading to a node labelled r and the right one leading to a node labelled $\neg r$. If we take the left branching to r then we see that the first disjunct in $r \vee \neg r$ is decided positively and so $r \vee \neg r$ is decided positively. Hence the whole sentence θ is decided positively and so this path gets labelled positive. If we take the right branching to $\neg r$ then the first disjunct in $r \vee \neg r$ is decided negatively and so we must move across to the second disjunct to see which way that is decided. In this case we see without any further deliberation that it is decided positively and so this path also gets labelled positive. Thus all our paths now have a label and we arrive at the following tree representation of $cT(\theta)$:



Thus $cT(\theta) = (p \wedge q \wedge \neg q) \vee (p \wedge \neg q \wedge r \wedge r) \vee (p \wedge \neg q \wedge \neg r \wedge \neg r) \vee (\neg p \wedge \neg p)$ and $cT(\theta)^{+} = \{p \wedge \neg q \wedge r, p \wedge \neg q \wedge \neg r, \neg p\}$.

One thing to note about cT is that, since the relation $\equiv$ extends $\ddot{\sim}$ and so also $\dot{\sim}$, and since $\theta \,\dot{\sim}\, cT(\theta)$ for all $\theta \in SL$, if θ is a tautology then all the clauses of $cT(\theta)$ will be positive, while if θ is a contradiction then all the clauses will be negative.

We shall utilise the cT-tree representation in Chapter 6. However for the most part we shall deal in this thesis with the second representation for θ which was also given in [12]. Before we get to it we give another definition.

Definition 3.11 *Given $\theta \in SL$, we shall say that θ is* contingent *iff both $\not\models \theta$ and $\not\models \neg\theta$.*

So θ is contingent iff it is neither a tautology or a contradiction.

In [12] (Lemma A.12) it is shown how each contingent $\theta \in SL$ can be reduced to a kind of normal form to which it is logically equivalent for pre-ents.

Lemma 3.12 *For each contingent $\theta \in SL$ there exists literals $p_{i,j}^{\epsilon_{i,j}}$ for $i = 1, \ldots, m$, $j = 1, \ldots, e(i)$, with $e(i) \geq 2$, such that for any relation $\sim \subseteq SL \times SL$ which extends $\dot{\sim}$ and satisfies (iv) and (v) of Theorem 3.4, properties (1)-(5) from Lemma 3.8 are satisfied together with the additional property:*

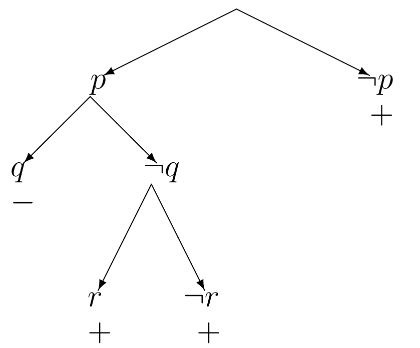
(6) If i, k are such that $1 \leq i < k \leq m$, $e(i) = e(k)$, $p_{i,j}^{\epsilon_{i,j}} = p_{k,j}^{\epsilon_{k,j}}$ for $j < e(i) - 1$ and $p_{i,e(i)-1}^{\epsilon_{i,e(i)-1}} = p_{k,e(k)-1}^{1-\epsilon_{k,e(k)-1}}$, then $p_{i,e(i)}^{\epsilon_{i,e(i)}} = p_{k,e(k)}^{\epsilon_{k,e(k)}}$ (so just one of these clauses is positive). □

As is the case in Lemma 3.8 the sentence $\bigvee_{i\leq m}\bigwedge_{j\leq e(i)} p_{i,j}^{\epsilon_{i,j}}$ in the above lemma is unique and is given a special name:

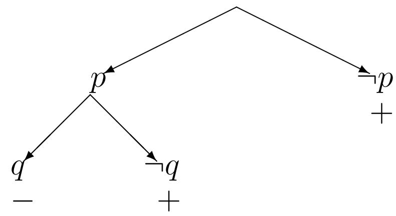
Definition 3.13 *Given $\theta \in SL$ such that θ is contingent, the unique sentence $\bigvee_{i\leq m}\bigwedge_{j\leq e(i)} p_{i,j}^{\epsilon_{i,j}}$ from Lemma 3.12 is denoted by $rT(\theta)$. We denote the set of positive clauses (without their last repeated literals) of $rT(\theta)$ by $rT(\theta)^+$.*

(Note that $rT(\theta)$ is undefined if θ is non-contingent.) Thus we have, for all contingent $\theta, \phi \in SL$, $\theta \ddot{\sim} rT(\theta)$ and (again from [12]) $\theta \ddot{\sim} \phi$ iff $rT(\theta) = rT(\phi)$. Also, since, as we have already pointed out, the relation $\ddot{\sim}$ is independent of the underlying language, so too is $rT(\theta)$. Constructing $rT(\theta)$ for θ a contingent sentence amounts to, firstly, constructing $cT(\theta)$ and then repeatedly "pruning" until no two paths of the same length which are the same until their last literal have the same label. Let us try and make this clear by extending our earlier example.

Example 3.14 Let L and θ be as in Example 3.10. Then we already have a tree representation of $cT(\theta)$:



Now from this we can see that the two paths $p \wedge \neg q \wedge r$ and $p \wedge \neg q \wedge \neg r$ have the same length and agree everywhere but their last literal and that both have the same label (i.e., positive). Hence we prune the cT-tree by replacing these two paths by the single path $p \wedge \neg q$ which, since the two paths were labelled positive, is labelled positive. This leaves us with



Since, in this new tree, no two paths of the same length which are the same until their last literal have the same label, we may stop here and so we have found the tree corresponding to $rT(\theta)$. Hence $rT(\theta) = (p \wedge q \wedge \neg q) \vee (p \wedge \neg q \wedge \neg q) \vee (\neg p \wedge \neg p)$ and $rT(\theta)^+ = \{p \wedge \neg q, \neg p\}$.

As a special case of the above definitions, it should be clear following the above example that, for τ a conjunction of literals from distinct propositional variables in L, $rT(\tau)^+ = \{\tau\}$.

We now intend to show how $Bel(\theta)$ can be expressed as a sum of the $Bel(\tau)$ for $\tau \in rT(\theta)^+$. The following result is needed.

Proposition 3.15 *Let $m \geq 1$ and, for each $i = 1, \ldots, m$, let θ_i be a conjunction of literals from L:*

$$\theta_i = \bigwedge_{j=1}^{e(i)} p_{i,j}^{\epsilon_{i,j}} \qquad e(i) \geq 1,\ i = 1, \ldots, m.$$

Suppose the θ_i satisfy the following property:

For any $1 \leq i < k \leq m$ there exists $j \leq e(i), e(k)$ such that $p_{i,j} = p_{k,j}$, $\epsilon_{i,j} = 1 - \epsilon_{k,j}$ and $p_{i,r}^{\epsilon_{i,r}} = p_{k,r}^{\epsilon_{k,r}}$ for all $r < j$.

Then, for arbitrary $\phi_1, \ldots, \phi_m \in SL$, we have the following:

(i). $\bigvee_{i=1}^m (\theta_i \wedge \phi_i) \dot{\sim} \bigvee_{i=1}^m (\theta_{\sigma(i)} \wedge \phi_{\sigma(i)})$ for any permutation σ on $\{1, \ldots, m\}$.

(ii). $Bel(\bigvee_{i=1}^m (\theta_i \wedge \phi_i)) = \sum_{i=1}^m Bel(\theta_i \wedge \phi_i)$ for any pre-ent over L.

Proof. (i). Since any permutation is a composition of transpositions it suffices to show that, for any $k \in \{1, \ldots, m-1\}$,

$$\left(\bigvee_{i=1}^{k-1} (\theta_i \wedge \phi_i)\right) \vee (\theta_k \wedge \phi_k) \vee (\theta_{k+1} \wedge \phi_{k+1}) \vee \left(\bigvee_{i=k+2}^{m} (\theta_i \wedge \phi_i)\right) \ \dot{\sim}$$

$$\dot{\sim} \quad \left(\bigvee_{i=1}^{k-1}(\theta_i \wedge \phi_i)\right) \vee (\theta_{k+1} \wedge \phi_{k+1}) \vee (\theta_k \wedge \phi_k) \vee \left(\bigvee_{i=k+2}^{m}(\theta_i \wedge \phi_i)\right).$$

This, in turn, will be proved if we can show

$$(\theta_k \wedge \phi_k) \vee (\theta_{k+1} \wedge \phi_{k+1}) \mathrel{\dot{\sim}} (\theta_{k+1} \wedge \phi_{k+1}) \vee (\theta_k \wedge \phi_k).$$

By assumption, we have that there exists $j \leq e(k), e(k+1)$ such that $p_{k,j} = p_{k+1,j}$, $\epsilon_{k,j} = 1 - \epsilon_{k+1,j}$ and $p_{k,r}^{\epsilon_{k,r}} = p_{k+1,r}^{\epsilon_{k+1,r}}$ for all $r < j$. Let $\tau = \bigwedge_{r=1}^{j-1} p_{k,r}^{\epsilon_{k,r}}$ and $p^{\epsilon} = p_{k,j}^{\epsilon_{k,j}}$. Then

$$\theta_k = \tau \wedge p^{\epsilon} \wedge \delta_k \text{ and } \theta_{k+1} = \tau \wedge p^{1-\epsilon} \wedge \delta_{k+1}$$

where δ_k and δ_{k+1} are the (possibly empty) conjunctions of the remaining literals in θ_k and θ_{k+1} respectively. Hence

$$\begin{aligned}
(\theta_k \wedge \phi_k) \vee (\theta_{k+1} \wedge \phi_{k+1}) &= (\tau \wedge p^{\epsilon} \wedge \delta_k \wedge \phi_k) \vee (\tau \wedge p^{1-\epsilon} \wedge \delta_{k+1} \wedge \phi_{k+1}) \\
&\dot{\sim} \tau \wedge ((p^{\epsilon} \wedge \delta_k \wedge \phi_k) \vee (p^{1-\epsilon} \wedge \delta_{k+1} \wedge \phi_{k+1})) \\
&\dot{\sim} \tau \wedge ((p^{1-\epsilon} \wedge \delta_{k+1} \wedge \phi_{k+1}) \vee (p^{\epsilon} \wedge \delta_k \wedge \phi_k)) \\
&\qquad \text{(using (u) from Proposition 3.2)} \\
&\dot{\sim} (\tau \wedge p^{1-\epsilon} \wedge \delta_{k+1} \wedge \phi_{k+1}) \vee (\tau \wedge p^{\epsilon} \wedge \delta_k \wedge \phi_k) \\
&= (\theta_{k+1} \wedge \phi_{k+1}) \vee (\theta_k \wedge \phi_k)
\end{aligned}$$

as required.

(ii). We prove this part by induction on m. Trivially the result holds for $m = 1$ so let us assume $m > 1$ and that the result is true for all $k < m$. Let $j \geq 1$ be minimal such that, for some $1 < r \leq m$, we have $p_{r,j}^{\epsilon_{r,j}} \neq p_{1,j}^{\epsilon_{1,j}}$. Let $\tau = \bigwedge_{i=1}^{j-1} p_{1,i}^{\epsilon_{1,i}}$ and let $p^{\epsilon} = p_{1,j}^{\epsilon_{1,j}}$. Then our assumption about the θ_i implies that, for each $i = 1, \ldots, m$ each θ_i is of the form

$$\theta_i = \tau \wedge p^{\upsilon_i} \wedge \delta_i$$

for some $v_i \in \{0,1\}$ and some (possibly empty) conjunction of literals δ_i. Let $I = \{i \in \{1,\ldots,m\} \mid v_i = \epsilon\}$ and let $I^C = \{1,\ldots,m\} - I$. Then we have

$$\begin{aligned}
\bigvee_{i=1}^{m}(\theta_i \wedge \phi_i) &\dot\sim \bigvee_{i\in I}(\theta_i \wedge \phi_i) \vee \bigvee_{i\in I^C}(\theta_i \wedge \phi_i) \qquad \text{by (i) proved above} \\
&= \bigvee_{i\in I}(\tau \wedge p^{\epsilon} \wedge \delta_i \wedge \phi_i) \vee \bigvee_{i\in I^C}(\tau \wedge p^{1-\epsilon} \wedge \delta_i \wedge \phi_i) \\
&\dot\sim (\tau \wedge p^{\epsilon} \wedge \bigvee_{i\in I}(\delta_i \wedge \phi_i)) \vee (\tau \wedge p^{1-\epsilon} \wedge \bigvee_{i\in I^C}(\delta_i \wedge \phi_i)) \\
&= (\tau \wedge p^{\epsilon} \wedge \psi_1) \vee (\tau \wedge p^{1-\epsilon} \wedge \psi_2)
\end{aligned}$$

where we define

$$\psi_1 = \bigvee_{i\in I}(\delta_i \wedge \phi_i) \text{ and } \psi_2 = \bigvee_{i\in I^C}(\delta_i \wedge \phi_i).$$

Hence, for any pre-ent over L,

$$\begin{aligned}
Bel(\bigvee_{i=1}^{m}(\theta_i \wedge \phi_i)) &= Bel((\tau \wedge p^{\epsilon} \wedge \psi_1) \vee (\tau \wedge p^{1-\epsilon} \wedge \psi_2)) \\
&= Bel(\tau \wedge p^{\epsilon} \wedge \psi_1) + Bel(\neg(\tau \wedge p^{\epsilon} \wedge \psi_1) \wedge (\tau \wedge p^{1-\epsilon} \wedge \psi_2))
\end{aligned} \tag{3.1}$$

from Theorem 2.5(d). We will now show that

$$\neg(\tau \wedge p^{\epsilon} \wedge \psi_1) \wedge (\tau \wedge p^{1-\epsilon} \wedge \psi_2) \ddot\sim \tau \wedge p^{1-\epsilon} \wedge \psi_2.$$

To see this we have, from Proposition 3.3

$$\neg(\tau \wedge p^{\epsilon} \wedge \psi_1) \wedge (\tau \wedge p^{1-\epsilon} \wedge \psi_2) \dot\sim \tau \wedge \neg(p^{\epsilon} \wedge \psi_1) \wedge (p^{1-\epsilon} \wedge \psi_2) \tag{3.2}$$

while

$$\begin{aligned}
\neg(p^{\epsilon} \wedge \psi_1) \wedge (p^{1-\epsilon} \wedge \psi_2) &\dot\sim (p^{1-\epsilon} \vee \neg\psi_1) \wedge (p^{1-\epsilon} \wedge \psi_2) \\
&\dot\sim (p^{1-\epsilon} \wedge p^{1-\epsilon} \wedge \psi_2) \vee (p^{\epsilon} \wedge \neg\psi_1 \wedge p^{1-\epsilon} \wedge \psi_2) \\
&\qquad \text{(from Proposition 3.2(f))} \\
&\dot\sim (p^{1-\epsilon} \wedge \psi_2) \vee (p^{\epsilon} \wedge \neg\psi_1 \wedge p^{1-\epsilon} \wedge \psi_2) \\
&\qquad \text{(from Proposition 3.2(j))} \\
&\ddot\sim p^{1-\epsilon} \wedge \psi_2 \qquad (\text{since } \vdash \neg(p^{\epsilon} \wedge \neg\psi_1 \wedge p^{1-\epsilon} \wedge \psi_2))
\end{aligned}$$

which gives, using Proposition 3.6,

$$\tau \wedge \neg(p^{\epsilon} \wedge \psi_1) \wedge (p^{1-\epsilon} \wedge \psi_2) \,\ddot{\sim}\, \tau \wedge p^{1-\epsilon} \wedge \psi_2.$$

Combining this with (3.2) gives the required equivalence. Hence, going back to (3.1) we may now write

$$\begin{aligned} Bel(\bigvee_{i=1}^{m}(\theta_i \wedge \phi_i)) &= Bel(\tau \wedge p^{\epsilon} \wedge \psi_1) + Bel(\tau \wedge p^{1-\epsilon} \wedge \psi_2) \\ &= Bel(\bigvee_{i \in I}(\tau \wedge p^{\epsilon} \wedge \delta_i \wedge \phi_i)) + Bel(\bigvee_{i \in I^C}(\tau \wedge p^{1-\epsilon} \wedge \delta_i \wedge \phi_i)) \\ &= Bel(\bigvee_{i \in I}(\theta_i \wedge \phi_i)) + Bel(\bigvee_{i \in I^C}(\theta_i \wedge \phi_i)). \end{aligned}$$

Since I and I^C are both strict subsets of $\{1, \ldots, m\}$ we may now apply our inductive hypothesis and write

$$Bel(\bigvee_{i \in I}(\theta_i \wedge \phi_i)) = \sum_{i \in I} Bel(\theta_i \wedge \phi_i) \text{ and } Bel(\bigvee_{i \in I^C}(\theta_i \wedge \phi_i)) = \sum_{i \in I^C} Bel(\theta_i \wedge \phi_i)$$

from which the result now follows. □

Corollary 3.16 *Let $\theta \in SL$ be contingent. Then, for any pre-ent over any language containing L,*

$$Bel(\theta) = \sum_{\tau \in rT(\theta)^+} Bel(\tau).$$

Proof. We have $\theta \,\ddot{\sim}\, rT(\theta)$ (independently of the underlying language). Let $\tau_1, \ldots, \tau_k$ be the positive clauses in $rT(\theta)$ and let $\gamma_1, \ldots, \gamma_l$ be the negative clauses. Then, applying Proposition 3.15(i) we may write

$$\theta \,\ddot{\sim} \bigvee_{i=1}^{k} \tau_i \vee \bigvee_{i=1}^{l} \gamma_i.$$

Each γ_i, since it is a negative clause, is of the form $\delta_i \wedge p_i^{\epsilon_i} \wedge p_i^{1-\epsilon_i}$ for some (possibly empty) conjunction of literals δ_i, $p_i \in L$ and $\epsilon_i \in \{0, 1\}$. Hence we have $\vdash \neg \bigvee_{i=1}^{l} \gamma_i$ and so

$$\theta \,\ddot{\sim} \bigvee_{i=1}^{k} \tau_i.$$

Also, since $q \wedge q \dot{\sim} q$ for any literal q, this equivalence will remain true if we delete the last repeated literal from each τ_i. Thus

$$\theta \ddot{\sim} \bigvee_{\tau \in rT(\theta)^+} \tau.$$

Now, by property (4) of Lemma 3.12 we know that the $\tau \in rT(\theta)^+$ satisfy the condition of Proposition 3.15 and so we have, for any pre-ent over L,

$$Bel(\theta) = Bel(\bigvee_{\tau \in rT(\theta)^+} \tau) = \sum_{\tau \in rT(\theta)^+} Bel(\tau).$$

as required. □

As a result of Corollary 3.16 we may now see that, for any pre-ent G over L, the function Bel^G is specified completely once its values on all conjunctions of literals from distinct propositional variables in L are known.

Before moving on to the next section we give another property of $rT(\theta)$ (which is not shared by $cT(\theta)$). Note that given two (possibly empty) conjunctions of literals over L $\sigma = \bigwedge_{j \leq r} p_j^{\epsilon_j}$ and $\rho = \bigwedge_{j \leq d} q_j^{\delta_j}$ we shall say that σ is an *initial segment* of ρ if $p_j^{\epsilon_j} = q_j^{\delta_j}$ for $j = 1, \ldots, r$.

Proposition 3.17 *Let $\theta \in SL$ be contingent. Let $\tau = \bigwedge_{j \leq r} q_j^{\delta_j}$ $(r \geq 0)$ be an initial segment of a clause $\bigwedge_{j \leq d} q_j^{\delta_j}$ of $rT(\theta)$ such that $r < d - 1$. Then τ is an initial segment of both a positive clause of $rT(\theta)$ and a negative clause of $rT(\theta)$.*

Proof. Let us assume that d is maximal such that τ is an initial segment of $\bigwedge_{j \leq d} q_j^{\delta_j}$ for some clause $\rho = \bigwedge_{j \leq d} q_j^{\delta_j}$ of $rT(\theta)$. Let us then assume that ρ is a *positive* clause of $rT(\theta)$ (i.e. that $q_{d-1}^{\delta_{d-1}} = q_d^{\delta_d}$). Then we need to find a *negative* clause which also has τ as an initial segment. By property (5) of Lemma 3.12 there exists a clause χ (of length $\geq d$) of $rT(\theta)$ which contains, as an initial segment, the sentence

$$\bigwedge_{j \leq d-2} q_j^{\delta_j} \wedge q_{d-1}^{1-\delta_{d-1}}.$$

Now, since $r < d-1$, the clause χ also has τ as an initial segment and so, by the maximality of d, χ must be of length d and so be of the form

$$\bigwedge_{j \leq d-2} q_j^{\delta_j} \wedge q_{d-1}^{1-\delta_{d-1}} \wedge q_{d-1}^{\upsilon}$$

for some $\upsilon \in \{0,1\}$. By property (6) of Lemma 3.12 we must have $\upsilon = \delta_d = \delta_{d-1}$ and hence χ must be a negative clause of $rT(\theta)$ as required. If we assume that ρ is a *negative* clause (i.e. that $q_{d-1} = q_d$ but $\delta_{d-1} \neq \delta_d$) then we can use exactly the same reasoning as the above to find a *positive* clause of $rT(\theta)$ which has τ as an initial segment. $\square$

Note that, putting τ to be the empty conjunction of literals in the above lemma gives us that, for any contingent θ, $rT(\theta)$ has both at least one positive clause and at least one negative clause.

3.4 Logical Consequence for Pre-Ents

In Section 3.2 we gave Paris and Vencovská's syntactic characterisation of the relation $\ddot{\sim}$. Our aim in this section is to use this characterisation as a springboard to find a similar representation for the binary relation which we define below.

Definition 3.18 *Let L be a language. We define the binary relation $\mid\!\ddot{\sim} \subseteq SL \times SL$ by, for all $\theta, \phi \in SL$, $\theta \mid\!\ddot{\sim} \phi$ iff $Bel(\theta) \leq Bel(\phi)$ for all pre-ents over L.*

Obviously we have $\theta \ddot{\sim} \phi$ iff both $\theta \mid\!\ddot{\sim} \phi$ and $\phi \mid\!\ddot{\sim} \theta$. Thus the relation $\mid\!\ddot{\sim}$ may be thought of as being a "half" of the relation $\ddot{\sim}$. We shall need the help of the following proposition which provides an expression of $\mid\!\ddot{\sim}$ in terms of rT-trees.

Proposition 3.19 *Let $\theta, \phi \in SL$ be contingent sentences. Then $\theta \mid\!\ddot{\sim} \phi$ implies that every $\tau \in rT(\theta)^+$ has an initial segment which is an element of $rT(\phi)^+$.*

Proof. Let θ, ϕ be two contingent sentences and let us suppose there existed an element $\tau \in rT(\theta)^+$ such that no initial segment of τ was an element of $rT(\phi)^+$. Say $\tau = \bigwedge_{j \le e} p_j^{\epsilon_j}$. Our result will be proved if we can produce a pre-ent G for which $Bel^G(\theta) > Bel^G(\phi)$. In order to do this let $a \le e$ be maximal such that, for some clause (positive or negative, and including its last repeated propositional variable) $\rho = \bigwedge_{j \le d} q_j^{\delta_j}$ of $rT(\phi)$ we have $p_j^{\epsilon_j} = q_j^{\delta_j}$ for $j \le a$. In other words, let $a \le e$ be maximal such that some clause ρ of $rT(\phi)$ contains $\bigwedge_{j \le a} p_j^{\epsilon_j}$ as an initial segment. We need to examine two different cases.

Case (i): $d = a + 1$

(This case can clearly only happen if $a \ge 1$.) In this case ρ must have the form

$$\rho = \bigwedge_{j \le a} p_j^{\epsilon_j} \wedge q_{a+1}^{\delta_{a+1}}$$

where $q_{a+1} = p_a$. Now if also $\delta_{a+1} = \epsilon_a$ then ρ would be a positive clause of $rT(\phi)$ and so $\bigwedge_{j \le a} p_j^{\epsilon_j} \in rT(\phi)^+$. But this would imply that τ had an initial segment (namely $\bigwedge_{j \le a} p_j^{\epsilon_j}$) which was an element of $rT(\phi)^+$ – contradiction. Hence we must have $\delta_{a+1} = 1 - \epsilon_a$, i.e., ρ is a negative clause of $rT(\phi)$. Let G be any pre-ent for which

$$G_{p_1}(\emptyset, t) = (-1)^{1-\epsilon_1}, \text{ where } t = \{p_1^{\epsilon_1}, \ldots, p_e^{\epsilon_e}\}.$$

Then, for any such G, we have $Bel^G(\theta) = 1$ and $Bel^G(\phi) = 0$. To see this, note that

$$\begin{aligned} Bel^G(\tau) &= Bel^G(\bigwedge_{j \le e} p_j^{\epsilon_j}) \\ &= \sum_{s_1 \subseteq s_2 \subseteq \cdots \subseteq s_e} G_{p_1^{\epsilon_1}}(\emptyset, s_1) G_{p_2^{\epsilon_2}}(s_1, s_2) \cdots G_{p_e^{\epsilon_e}}(s_{e-1}, s_e) \\ &= G_{p_1^{\epsilon_1}}(\emptyset, s) G_{p_2^{\epsilon_2}}(s, s) \cdots G_{p_e^{\epsilon_e}}(s, s) \\ &= 1. \end{aligned}$$

Hence, using Corollary 3.16,

$$Bel^G(\theta) = \sum_{\tau' \in rT(\theta)^+} Bel^G(\tau') \geq Bel^G(\tau) = 1$$

and so $Bel^G(\theta) = 1$. To show $Bel^G(\phi) = 0$ it is enough, by Corollary 3.16, to show that $Bel^G(\gamma) = 0$ for all $\gamma \in rT(\phi)^+$. So let $\gamma \in rT(\phi)^+$. Remembering that $\rho = \bigwedge_{j \leq a} p_j^{\epsilon_j} \wedge p_a^{1-\epsilon_a}$ is a negative clause of $rT(\phi)$ we know, by property (4) of Lemma 3.4, that γ must be of the form $\bigwedge_{j \leq t-1} p_j^{\epsilon_j} \wedge p_t^{1-\epsilon_t} \wedge \nu$ for some $1 \leq t \leq a$ and some (possibly empty) conjunction of literals ν. Since, as we have already shown, $Bel^G(\bigwedge_{j \leq e} p_j^{\epsilon_j}) = 1$, and since $Bel(\lambda \wedge \chi) \leq Bel(\lambda)$ for all pre-ents over L and all $\lambda, \chi \in SL$, we have

$$Bel^G(\bigwedge_{j \leq t-1} p_j^{\epsilon_j} \wedge p_t^{\epsilon_t}) = 1 = Bel^G(\bigwedge_{j \leq t-1} p_j^{\epsilon_j}).$$

Hence, using Theorem 2.5(e),

$$\begin{aligned} Bel^G(\bigwedge_{j \leq t-1} p_j^{\epsilon_j} \wedge p_t^{1-\epsilon_t}) &= Bel^G(\bigwedge_{j \leq t-1} p_j^{\epsilon_j}) - Bel^G(\bigwedge_{j \leq t-1} p_j^{\epsilon_j} \wedge p_t^{\epsilon_t}) \\ &= 1 - 1 = 0. \end{aligned}$$

This gives us

$$Bel^G(\gamma) = Bel^G(\bigwedge_{j \leq t-1} p_j^{\epsilon_j} \wedge p_t^{1-\epsilon_t} \wedge \nu) \leq Bel^G(\bigwedge_{j \leq t-1} p_j^{\epsilon_j} \wedge p_t^{1-\epsilon_t}) = 0,$$

i.e, $Bel^G(\gamma) = 0$ as required.

Case (ii): $d > a + 1$

So we have

$$\rho = \bigwedge_{j \leq a} p_j^{\epsilon_j} \wedge \bigwedge_{a+1 \leq j \leq d} q_j^{\delta_j}.$$

Since $d > a + 1$ we may apply Proposition 3.17 to establish that $\bigwedge_{j \leq a} p_j^{\epsilon_j}$ is an initial segment of some negative clause $\rho' = \bigwedge_{j \leq a} p_j^{\epsilon_j} \wedge \bigwedge_{a+1 \leq j \leq l} r_j^{\upsilon_j}$ of $rT(\phi)$. We claim now that $r_{a+1} \neq p_{a+1}$. To see this suppose otherwise. Then $\rho' =$

$\bigwedge_{j \leq a} p_j^{\epsilon_j} \wedge p_{a+1}^{\upsilon_{a+1}} \wedge \bigwedge_{a+2 \leq j \leq l} r_j^{\upsilon_j}$ and so, by Lemma 3.12(5), there must be a clause of $rT(\phi)$ which contains $\bigwedge_{j \leq a} p_j^{\epsilon_j} \wedge p_{a+1}^{\epsilon_{a+1}}$ as an initial segment. This contradicts the maximality of a and so we must have $r_{a+1} \neq p_{a+1}$ as required. Now let $t_1 = \{p_1^{\epsilon_1}, \ldots, p_a^{\epsilon_a}\}$, $t_2 = \{p_{a+1}^{\epsilon_{a+1}}, \ldots, p_e^{\epsilon_e}\}$ and $t_3 = \{r_{a+1}^{\upsilon_{a+1}}, \ldots, r_{l-1}^{\upsilon_{l-1}}\}$ and define G to be any pre-ent which satisfies:

$$\begin{aligned} & G_{p_1}(\emptyset, t_1) = (-1)^{1-\epsilon_1} \text{ if } a > 0, \\ & G_{p_{a+1}}(t_1, t_1 \cup t_2) = (-1)^{1-\epsilon_{a+1}} \\ \text{and} \quad & G_{r_{a+1}}(t_1, t_1 \cup t_3) = (-1)^{1-\upsilon_{a+1}}. \end{aligned}$$

Then, in a similar way to that in case (i) above, we can show that for any such G we have $Bel^G(\theta) = 1$ and $Bel^G(\phi) = 0$ as required. □

We shall see later that, in fact, the converse of the above Proposition 3.19 also holds, thus providing an alternative characterisation of $\theta \mid\!\ddot{\sim} \phi$ (at least in the case where θ, ϕ are contingent), this time in terms of rT-trees.

Our approach to finding an axiomatic characterisation for $\mid\!\ddot{\sim}$ is based on extending the list of axioms given via Theorems 3.1 and 3.4 for $\ddot{\sim}$. It turns out that we need to add only one axiom and then close under transitivity. In the next two lemmas we give some further rules which follow from this extension.

Lemma 3.20 *Let $\lhd \subseteq SL \times SL$ be any binary relation on SL which extends $\ddot{\sim}$ (and hence also $\dot{\sim}$), is transitive, and satisfies, for all $\theta, \phi \in SL$, $\theta \wedge \phi \lhd \theta$. Then, for all $\theta, \phi \in SL$ and $\lambda \in SL$ such that $\vdash \lambda$, $\lhd$ satisfies the following*

1. $\theta \lhd \theta \vee \phi$
2. $\neg\lambda \lhd \theta$
3. $\theta \lhd \lambda$.

Proof. (1). Since $\vdash \neg(\phi \wedge \neg\phi)$ we have

$$\theta \quad \ddot{\sim} \quad \theta \vee (\phi \wedge \neg\phi)$$

$$\begin{aligned} &\dot{\sim} \quad (\theta \vee \phi) \wedge (\theta \vee \neg\phi) \\ &\lhd \quad \theta \vee \phi \end{aligned}$$

Hence, since $\lhd$ extends $\ddot{\sim}$ (and $\dot{\sim}$) and is transitive, we have $\theta \lhd \theta \vee \phi$ as required.

(2). Since $\vdash \neg(\theta \wedge \neg\theta)$ we have

$$\neg\lambda \mathrel{\ddot{\sim}} \neg\neg(\theta \wedge \neg\theta) \mathrel{\dot{\sim}} \theta \wedge \neg\theta \lhd \theta$$

as required.

(3). Using (1) proved above we have $\theta \lhd \theta \vee \neg\theta$. Then, since $\vdash \theta \vee \neg\theta$ we have $\theta \vee \neg\theta \mathrel{\ddot{\sim}} \lambda$ and the result follows. □

Lemma 3.21 *Let $\lhd \subseteq SL \times SL$ be any binary relation on SL which extends $\ddot{\sim}$, is transitive, and satisfies, for all $\theta, \phi \in SL$, $\theta \wedge \phi \lhd \theta$. Let $\tau_1, \ldots, \tau_m$ $(m \geq 1)$ be conjunctions of literals over L which satisfy the condition of Proposition 3.15 and let $\phi_1, \ldots, \phi_m \in SL$ be arbitrary sentences. Then*

$$\left(\bigvee_{i=1}^{k-1} (\tau_i \wedge \phi_i)\right) \vee (\tau_k \wedge \phi_k) \vee \left(\bigvee_{i=k+1}^{m} (\tau_i \wedge \phi_i)\right) \lhd$$
$$\lhd \quad \left(\bigvee_{i=1}^{k-1} (\tau_i \wedge \phi_i)\right) \vee \tau_k \vee \left(\bigvee_{i=k+1}^{m} (\tau_i \wedge \phi_i)\right)$$

Proof. For ease of exposition let us set $\theta_1 = \bigvee_{i=1}^{k-1}(\tau_i \wedge \phi_i)$ and $\theta_2 = \bigvee_{i=k+1}^{m}(\tau_i \wedge \phi_i)$. So we must show

$$\theta_1 \vee (\tau_k \wedge \phi_k) \vee \theta_2 \;\lhd\; \theta_1 \vee \tau_k \vee \theta_2$$

We have

$$\begin{aligned} \theta_1 \vee (\tau_k \wedge \phi_k) \vee \theta_2 \quad &\dot{\sim} \quad (\tau_k \wedge \phi_k) \vee \theta_1 \vee \theta_2 \qquad \text{by Proposition 3.15(i)} \\ &\dot{\sim} \quad (\tau_k \vee \theta_1 \vee \theta_2) \wedge (\neg\tau_k \vee \phi_k \vee \theta_1 \vee \theta_2) \\ &\qquad\qquad \text{by Proposition 3.2(q)} \\ &\lhd \quad \tau_k \vee \theta_1 \vee \theta_2 \qquad \text{by our assumption about } \lhd \\ &\dot{\sim} \quad \theta_1 \vee \tau_k \vee \theta_2 \qquad \text{by Proposition 3.15(i).} \end{aligned}$$

Hence, since $\lhd$ extends $\dot{\sim}$ and is transitive, we have our required conclusion. $\Box$

We remark that the proof of the above Lemma 3.21 only requires $\lhd$ to extend $\dot{\sim}$, and not necessarily $\ddot{\sim}$.
We now give our syntactic representation of the relation $\mathrel{|}\!\ddot{\sim}$.

Theorem 3.22 *The relation $\mathrel{|}\!\ddot{\sim}$ is the (unique) smallest relation $\lhd$ on SL which extends $\ddot{\sim}$, is transitive and which satisfies, for all $\theta, \phi \in SL$, $\theta \wedge \phi \lhd \theta$.*

Proof. First of all it is clear that $\mathrel{|}\!\ddot{\sim}$ satisfies the conditions of the theorem since it is easy to see that $\mathrel{|}\!\ddot{\sim}$ extends $\ddot{\sim}$ and is transitive, while $\theta \wedge \phi \mathrel{|}\!\ddot{\sim} \theta$ follows directly from Theorem 2.5(e). Hence the main work to be done in the proof lies in showing that, for any relation $\lhd \subseteq SL \times SL$ which also satisfies all those conditions, we have, for all $\theta, \phi \in SL$, $\theta \mathrel{|}\!\ddot{\sim} \phi$ implies $\theta \lhd \phi$. So let $\lhd$ be such a relation and suppose $\theta \mathrel{|}\!\ddot{\sim} \phi$. We examine several separate cases.

Case (i): $\vdash \neg\theta$

In this case we have $\neg\neg\theta \lhd \phi$ by property 2 of Lemma 3.20 and $\theta \,\dot{\sim}\, \neg\neg\theta$. Hence, since $\lhd$ extends $\ddot{\sim}$ (and hence also $\dot{\sim}$) and is transitive, we conclude $\theta \lhd \phi$ as required.

Case (ii): $\vdash \theta$

In this case we have $Bel(\theta) = 1$ for all pre-ents and so $\theta \mathrel{|}\!\ddot{\sim} \phi$ implies also $Bel(\phi) = 1$ for all pre-ents and so, by Theorem 2.7, we must also have $\vdash \phi$. Hence $\theta \lhd \phi$ by property 3 of Lemma 3.20.

Cases (i) and (ii) jointly take care of the situation where θ is non-contingent. The remaining cases look at the situation where θ is contingent. Note that if θ is contingent then it cannot be the case that $\vdash \neg\phi$ since this, together with $\theta \mathrel{|}\!\ddot{\sim} \phi$, would imply $\vdash \neg\theta$.

Case (iii)(a): θ is contingent and $\vdash \phi$

In this case we have $\theta \lhd \phi$ from property 3 of Lemma 3.20.

<u>*Case (iii)(b)*: both θ and ϕ are contingent</u>

In this case $rT(\theta)$ and $rT(\phi)$ are both well-defined and, by Proposition 3.19, $\theta \mid\!\ddot{\sim} \phi$ implies that every $\tau \in rT(\theta)^+$ has an initial segment which is an element of $rT(\phi)^+$. We have (from the proof of Corollary 3.16)

$$\theta \ddot{\sim} \bigvee_{\tau \in rT(\theta)^+} \tau \tag{3.3}$$

Suppose $rT(\phi)^+ = \{\gamma_1, \ldots, \gamma_m\}$, where $\gamma_1, \ldots, \gamma_l$ $(1 \leq l \leq m)$ are those elements in $rT(\phi)^+$ which appear as an initial segment of at least one $\tau \in rT(\theta)^+$. For each $i = 1, \ldots, l$, denote by $rT(\theta)^+/i$ the set of $\tau \in rT(\theta)^+$ which contain γ_i as an initial segment. Then

$$\bigvee_{\tau \in rT(\theta)^+} \tau \dot{\sim} \bigvee_{i=1}^{l} \bigvee_{\tau \in rT(\theta)^+/i} \tau. \tag{3.4}$$

For each i and each $\tau \in rT(\theta)/i$ there exists a (possibly empty) conjunction of literals ρ_τ such that $\tau = \gamma_i \wedge \rho_\tau$. Thus we may write

$$\bigvee_{i=1}^{l} \bigvee_{\tau \in rT(\theta)^+/i} \tau = \bigvee_{i=1}^{l} \bigvee_{\tau \in rT(\theta)^+/i} \gamma_i \wedge \rho_\tau \dot{\sim} \bigvee_{i=1}^{l} (\gamma_i \wedge \bigvee_{\tau \in rT(\theta)^+/i} \rho_\tau). \tag{3.5}$$

Hence, given that $\lhd$ extends $\ddot{\sim}$ (and hence also $\dot{\sim}$) and is transitive, what we have shown thus far, combining equations (3.3), (3.4) and (3.5) above, is that

$$\theta \lhd \bigvee_{i=1}^{l} (\gamma_i \wedge \bigvee_{\tau \in rT(\theta)^+/i} \rho_\tau).$$

For clarity let us set, for each $i = 1, \ldots, l$,

$$\delta_i = \bigvee_{\tau \in rT(\theta)^+/i} \rho_\tau.$$

Note that δ_i may, possibly, be null. Then

$$\theta \quad \lhd \quad \bigvee_{i=1}^{l} (\gamma_i \wedge \delta_i)$$

$$\begin{aligned}
&\lhd \quad \bigvee_{i=1}^{l} \gamma_i \qquad \text{by repeated application of Lemma 3.21} \\
&\lhd \quad (\bigvee_{i=1}^{l} \gamma_i) \vee (\bigvee_{i=l+1}^{m} \gamma_i) \qquad \text{by property 1 of Lemma 3.20} \\
&\dot{\sim} \quad \bigvee_{\tau \in rT(\phi)^+} \tau \\
&\ddot{\sim} \quad \phi.
\end{aligned}$$

Hence, again since $\lhd$ is transitive and extends $\ddot{\sim}$, we have $\theta \lhd \phi$ as required. □

As is the case for $\dot{\sim}$ and $\ddot{\sim}$, we may now see that the definition of $|\ddot{\sim}$ is independent of the underlying language. We do not show in this thesis whether $\theta \mid\ddot{\sim} \phi$ iff $Bel(\theta) \leq Bel(\phi)$ for all *ents* (note the "only if" direction is trivial) although we do believe that by modifying the proof of Proposition 3.19 to showing the existence of an *ent* such that $Bel^z(\theta) > Bel^z(\phi)$ this result can be shown to be true. Note that what we have shown in case (iii)(b) of the above proof is that, for contingent $\theta, \phi \in SL$, if every $\tau \in rT(\theta)^+$ has an initial segment which is an element of $rT(\phi)^+$ then $\theta \lhd \phi$ for any binary relation $\lhd$ on SL which satisfies the conditions of Theorem 3.22. In particular, then, $\theta \mid\ddot{\sim} \phi$. Hence we have proved the following improvement on Proposition 3.19:

Proposition 3.23 *Let $\theta, \phi \in SL$ be contingent sentences. Then $\theta \mid\ddot{\sim} \phi$* **iff** *every $\tau \in rT(\theta)^+$ has an initial segment which is an element of $rT(\phi)^+$.* □

We shall meet yet another characterisation of $|\ddot{\sim}$ in Chapter 5. Before that we shall now show what is the difference, as regards properties of their respective belief functions, between pre-ents and ents.

Chapter 4

From Pre-Ents to Ents

4.1 Introduction

We saw in Section 2.3 (Theorem 2.11) that for any ent z over a language L the set of $\theta \in SL$ such that $Bel^z(\theta) = 1$ is closed under logical consequences, equivalently (by Proposition 2.9), for any sentences $\theta, \phi \in SL$, if $Bel^z(\theta \wedge \phi) = 0$ then $Bel^z(\phi \wedge \theta) = 0$. We also saw that this rather desirable result does not hold, in general, for pre-ents. The aim of this chapter is to show that, essentially, and as far as their resultant belief functions (by which, recall, we mean their associated function $Bel_\emptyset$) are concerned, the validation of this property is the *only* place where ents and pre-ents differ. More precisely this chapter is devoted to proving the following (which was first stated, but not proved, in [11]):

Theorem 4.1 *Given a language $L = \{p_1, \ldots, p_n\}$, if the function $Bel : SL \to [0,1]$ is given by a pre-ent over L and if, for all $\theta, \phi \in SL$, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$, then there exists an ent z (over a larger language than L) such that, for all $\theta \in SL$, $Bel^z(\theta) = Bel(\theta)$.*

The reader will notice that, in this theorem, we allow our required ent to be over a larger language than L, i.e., a language which contains L as a subset. In fact this is typical of many of Paris and Vencovská's results on ents. It seems that, in the world of ents and pre-ents, the underlying language is often taken to be open-ended. It is believed, though not proved, that the Theorem 4.1 fails to hold if we require z to be defined over the same language L. It is also unknown whether or not we may replace Bel everywhere in the theorem by Bel_s for some arbitrary scenario s. Note that to show this Theorem 4.1, since, as we have seen in Chapter 2, any function on SL given by a pre-ent over L is determined by the values it gives to all conjunctions of literals from distinct propositional variables in L, it will be enough to prove the existence of an ent z such that, for all such conjunctions $q_1 \wedge \ldots \wedge q_j$, we have $Bel^z(q_1 \wedge \ldots \wedge q_j) = Bel(q_1 \wedge \ldots \wedge q_j)$. In particular, if we take $n = 1$ in Theorem 4.1, i.e., if we suppose our language L consists of just a single propositional variable p, then we are required to find an ent z such that

$$Bel^z(p) = Bel(p) \text{ and } Bel^z(\neg p) = Bel(\neg p).$$

In fact we may do this straight away (even without the extra freedom of defining z over a language larger than L) by defining z via the following tableau:

s	$\{p\}$	$\{\neg p\}$
z_s	$Bel(p)$	$Bel(\neg p)$

It is easy to check that z defined above does indeed give the correct values to $Bel^z(p)$ and $Bel^z(\neg p)$. Thus we can see already that Theorem 4.1 is true in the case when $n = 1$.

We shall prove the general case of Theorem 4.1 in three stages. A complication which will arise is that the function which we produce (z_∞ in the upcoming proof), although it will bear a very close resemblance to an ent and will compute beliefs just like one, will not actually be an ent! The difference being that there will, in fact, be scenarios s for which there exists a propositional variable p such that

$\pm p \notin s$ and there is no scenario consistent with s that both decides p *and* has non-zero potential. However, it will be true that, in computing beliefs (starting from $\emptyset$), going from scenario to scenario, the function z_∞ will never be led to any scenario which suffers from this problem and so it will still yield a well-defined belief function Bel^{z_∞}. This leads us, in Section 4.3 to consider a wider class of function (to which z_∞ belongs) than the class of ents — the class of *almost-ents*. The first two stages of the proof are devoted to showing that Theorem 4.1 is true when "ent" is replaced in the statement of the theorem by "almost-ent". After introducing some key notation in Section 4.4 we describe the first stage in Section 4.5, where we show that the theorem is true in the even more general setting in which we allow our required almost-ent to have potentials which are non-standard real numbers. These potentials will, in fact, be formal power series in an indeterminate λ (which may alternatively be thought of as a positive infinitesimal – Section 4.2 will be devoted to the introduction of these concepts as they are to be applied in this thesis.) Then, in the second stage to be described in Section 4.6, we show that this particular extra freedom can be dispensed with by showing how these non-standard potentials may be replaced with equivalent standard ones. This turns out to be a question of showing how we may take our indeterminate λ to be a real number, provided it is sufficiently small, and all our power series will converge. Finally in Section 4.7 we show how our constructed almost-ent z_∞ may be converted into an equivalent ent. This ent will still be defined over a language larger than L, though this language will be different from the one over which our constructed almost-ent is defined. One important point to notice about the proof of Theorem 4.1 is that, although the function Bel is given by a pre-ent over L, absolutely no mention is made in the proof of this pre-ent. The proof is "at the level" of Bel in that only properties of Bel are used. One consequence of this is that we could if we wanted take the pre-ent itself to be defined over a larger language than L, although of course the ent we produce would still only agree with Bel on SL. We shall exploit this point in Theorem 6.23 at the end of Chapter 6. We end the present section by giving an equivalent form to one of the hypotheses of Theorem 4.1 which will prove to be more useful in the upcoming proof.

Lemma 4.2 *Let L be a language and let the function $Bel : SL \to [0,1]$ be given by a pre-ent over L. Then the following are equivalent:*
(i). For all $\theta, \phi \in SL$, if $Bel(\theta \wedge \phi) = 0$ then $Bel(\phi \wedge \theta) = 0$.
(ii). For all $j \geq 1$ and all $\theta_1, \ldots, \theta_j \in SL$, if $Bel(\theta_1 \wedge \ldots \wedge \theta_j) = 0$ then $Bel(\bigwedge S) = 0$ whenever $S \subseteq SL$ is such that $\{\theta_i \mid i = 1, \ldots, j\} \subseteq S$.

Proof. That (ii) implies (i) is clear. For the converse direction suppose (i) holds. Let $\theta_1, \ldots, \theta_j \in SL$ for some $j \geq 1$ be such that $Bel(\theta_1 \wedge \ldots \wedge \theta_j) = 0$ and let $S \subseteq SL$ be such that $\{\theta_i \mid i = 1, \ldots, j\} \subseteq S$. We must show that $Bel(\bigwedge S) = 0$. We have $Bel(\theta_1 \wedge \ldots \wedge \theta_j) = 0$ implies $Bel(\neg(\theta_1 \wedge \ldots \wedge \theta_j)) = 1$. Hence, by Proposition 2.9, since $\neg(\theta_1 \wedge \ldots \wedge \theta_j) \vdash \neg \bigwedge S$, we have also $Bel(\neg \bigwedge S) = 1$ and so $Bel(\bigwedge S) = 0$ as required. □

4.2 Introducing Non-Standard Potentials

In this section we generalise our setting to include the possibility that our pre-ents and ents (and the soon-to-be-defined almost-ents) will have potentials, and possibly even compute beliefs, which are non-standard real numbers. (For a full treatment of the subject of non-standard analysis the interested reader is referred to [15].) We shall exploit the framework developed here also in later chapters. To achieve this generalisation we will introduce a new symbol λ to the real numbers and extend the ordered field $\mathbb{R}$ to the ordered field $\mathbb{R}((\lambda))$ consisting of all fractions of power series over $\mathbb{R}$ in λ. By the end of this section we will be in a position to modify our definition of pre-ent by, instead of interpreting them as functions from $L \times SL \times SL$ into $[-1, 1]$, interpreting them as functions into $[-1, 1]^{(\lambda)}$ which denotes the set of values in $\mathbb{R}((\lambda))$ which lie between -1 and 1 according to the ordering in $\mathbb{R}((\lambda))$. Likewise we can modify the definition of ent so that the potentials of the ent are now values in $[0, \infty)^{(\lambda)}$ – the set of values of $\mathbb{R}((\lambda))$ which are greater than or equal to zero according to the ordering in $\mathbb{R}((\lambda))$. In order to define the field $\mathbb{R}((\lambda))$ we start from the set of *formal power series over* $\mathbb{R}$ *in the indeterminate* λ, which we denote by $\mathbb{R}[[\lambda]]$:

$$\begin{aligned} \mathbb{R}[[\lambda]] &= \{a_0 + a_1\lambda + a_2\lambda^2 + \ldots \mid a_i \in \mathbb{R} \text{ for } i = 0, 1, \ldots\} \\ &= \{\sum_{i=0}^{\infty} a_i\lambda^i \mid a_i \in \mathbb{R} \text{ for } i = 0, 1, \ldots\}. \end{aligned}$$

Note that, at the moment, λ is just a symbol. We do not (yet) want it to take any actual real-numbered value whatsoever. Given $a, b \in \mathbb{R}[[\lambda]]$ such that $a = \sum_{i=0}^{\infty} a_i\lambda^i$ and $b = \sum_{i=0}^{\infty} b_i\lambda^i$ we identify the two elements as being equal as follows:

$$a = b \text{ iff } a_i = b_i \text{ for all } i = 0, 1, 2, \ldots$$

so, essentially, the elements of $\mathbb{R}[[\lambda]]$ are nothing more than infinite sequences of real numbers. Given $a = \sum_{i=0}^{\infty} a_i\lambda^i \in \mathbb{R}[[\lambda]]$ if there exists m such that $a_i = 0$ for all $i > m$ then we call a a *polynomial* in λ and write $a = a_0 + a_1\lambda + \cdots + a_m\lambda^m = \sum_{i=0}^{m} a_i\lambda^i$. In addition if $a_m \neq 0$ then we will say that m is the *degree* of a. We will often use $P(\lambda), Q(\lambda)$ etc. to denote polynomials in λ. Furthermore, given a polynomial in λ, we usually drop mention of any a_i which are zero. For example we will write $1 + 3\lambda^2$ instead of $1 + 0\lambda + 3\lambda^2$. Given $b \in \mathbb{R}$ we identify b with (the polynomial) $b \in \mathbb{R}[[\lambda]]$, so essentially we have $\mathbb{R} \subseteq \mathbb{R}[[\lambda]]$. We define binary operations of addition, $+_{[\lambda]}$, and multiplication, $\cdot_{[\lambda]}$, on $\mathbb{R}[[\lambda]]$ as expected:

$$\sum_{i=0}^{\infty} a_i\lambda^i +_{[\lambda]} \sum_{i=0}^{\infty} b_i\lambda^i = \sum_{i=0}^{\infty} (a_i + b_i)\lambda^i$$

where $+$, of course, is the usual addition on $\mathbb{R}$, and

$$(\sum_{i=0}^{\infty} a_i\lambda^i) \cdot_{[\lambda]} (\sum_{i=0}^{\infty} b_i\lambda^i) = \sum_{i=0}^{\infty} c_i\lambda^i$$

where, for each $i = 0, 1, \ldots,$

$$c_i = \sum_{j=0}^{i} a_j b_{i-j}$$

where $a_j b_{i-j}$ denotes the product in $\mathbb{R}$ of a_j and b_{i-j}. Alternatively we may write the i^{th} coefficient of this product as

$$c_i = \sum_{j+k=i} a_j b_k$$

where the sum is understood to be over all pairs $\langle j, k\rangle$ of natural numbers such that $j + k = i$. By an easy induction we may generalise this and say that the i^{th} coefficient of a product of k (≥ 2) terms

$$(\sum_{i=0}^{\infty} a_i^{(1)}\lambda^i) \cdot_{[\lambda]} (\sum_{i=0}^{\infty} a_i^{(2)}\lambda^i) \cdot_{[\lambda]} \cdots \cdot_{[\lambda]} (\sum_{i=0}^{\infty} a_i^{(k)}\lambda^i)$$

is given by

$$\sum_{j_1+j_2+\cdots+j_k=i} a_{j_1}^{(1)} a_{j_2}^{(2)} \cdots a_{j_k}^{(k)}$$

where the sum is understood to be over all k-tuples of natural numbers $\langle j_1, j_2, \ldots, j_k\rangle$ such that $j_1 + j_2 + \cdots + j_k = i$. We define an ordering $<_{[\lambda]}$ on $\mathbb{R}[[\lambda]]$ by setting

$$\sum_{i=0}^{\infty} a_i\lambda^i <_{[\lambda]} \sum_{i=0}^{\infty} b_i\lambda^i \quad \text{iff} \quad \begin{array}{l}\text{there exists } k \text{ such that } a_k \neq b_k \text{ and} \\ \text{for the least such } k,\ a_k < b_k\end{array}$$

where $<$ denotes the usual ordering on $\mathbb{R}$. Hence we have $0 <_{[\lambda]} \lambda$ (i.e., $0+0\lambda <_{[\lambda]} 0 + 1\lambda$) and $\lambda <_{[\lambda]} b$ for any $b \in \mathbb{R}$ such that $b > 0$. Thus λ looks like a positive infinitesimal. Under the above definitions it easy to check that $\mathbb{R}[[\lambda]]$ forms an ordered ring (with, naturally, $0 \in \mathbb{R}$ and $1 \in \mathbb{R}$ acting as additive and multiplicative identity respectively and $-\sum_{i=0}^{\infty} a_i\lambda^i = \sum_{i=0}^{\infty}(-a_i)\lambda^i$). It is also easy to see that the operations $+_{[\lambda]}$ and $\cdot_{[\lambda]}$ and the relation $<_{[\lambda]}$ extend the corresponding operations and relation in $\mathbb{R}$. In view of this we will now drop the subscript $[\lambda]$ and, for example, use $+$ for both the usual addition on $\mathbb{R}$ and the addition on $\mathbb{R}[[\lambda]]$.

From $\mathbb{R}[[\lambda]]$ we define $\mathbb{R}((\lambda))$ to be the *field of fractions* of $\mathbb{R}[[\lambda]]$:

$$\mathbb{R}((\lambda)) = \{\frac{a}{b} \mid a, b \in \mathbb{R}[[\lambda]], b > 0\}.$$

We define equality in this set by

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} \text{ iff } a_1 \cdot b_2 = a_2 \cdot b_1,$$

and we identify $a \in \mathbb{R}[[\lambda]]$ with $\frac{a}{1} \in \mathbb{R}((\lambda))$ so we have $\mathbb{R} \subseteq \mathbb{R}[[\lambda]] \subseteq \mathbb{R}((\lambda))$. To give an example of an equality in $\mathbb{R}((\lambda))$ we have

$$\frac{1}{1+\lambda} = \sum_{i=0}^{\infty}(-1)^i\lambda^i, \text{ i.e., } \frac{1}{1+\lambda} = \frac{\sum_{i=0}^{\infty}(-1)^i\lambda^i}{1}$$

since

$$\begin{aligned}(1+\lambda)\cdot\sum_{i=0}^{\infty}(-1)^i\lambda^i &= 1\cdot\sum_{i=0}^{\infty}(-1)^i\lambda^i + \lambda\cdot\sum_{i=0}^{\infty}(-1)^i\lambda^i \\ &= \sum_{i=0}^{\infty}(-1)^i\lambda^i + \sum_{i=1}^{\infty}(-1)^{i-1}\lambda^i \\ &= 1.\end{aligned}$$

We further extend the operations $+$ and $\cdot$ to $\mathbb{R}((\lambda))$ by setting

$$\frac{a_1}{b_1} + \frac{a_2}{b_2} = \frac{a_1 \cdot b_2 + a_2 \cdot b_1}{b_1 \cdot b_2},$$

and

$$\frac{a_1}{b_1} \cdot \frac{a_2}{b_2} = \frac{a_1 \cdot a_2}{b_1 \cdot b_2}.$$

(Note that, for $b_1, b_2 \in \mathbb{R}[[\lambda]]$, $b_i > 0$ for $i = 1, 2$ implies $b_1 \cdot b_2 > 0$ so both the right-hand sides above are certainly elements of $\mathbb{R}((\lambda))$.) We extend the ordering $<$ to $\mathbb{R}((\lambda))$ by

$$\frac{a_1}{b_1} < \frac{a_2}{b_2} \text{ iff } a_1 \cdot b_2 < a_2 \cdot b_1$$

and use $a \leq b$ to mean that either $a < b$ or $a = b$. Under the above definitions $\mathbb{R}((\lambda))$ becomes an ordered field.

Given the construction of $\mathbb{R}((\lambda))$, we now make the following important definition.

Definition 4.3 *Let $a \in \mathbb{R}((\lambda))$ and let $m \in \mathbb{N}$. Then we say that a is* of the order $O(\lambda^m)$, *written $a = O(\lambda^m)$, if $a = \sum_{i=0}^{\infty} a_i \lambda^i$ and $a_i = 0$ for $i < m$.*

Intuitively $a = O(\lambda)$ means that a is infinitesimally small, $a = O(\lambda^2)$ means that a is infinitesimally small even in comparison with the $O(\lambda)$-elements of $\mathbb{R}((\lambda))$, $a = O(\lambda^3)$ means that a is infinitesimally small even in comparison with the $O(\lambda^2)$-elements and so on. Note that, under this definition, we have $a = O(\lambda^m)$ implies $a = O(\lambda^{m-1})$. So, following this definition, our earlier example shows us

$$\frac{1}{1+\lambda} = O(\lambda^0) = O(1) \text{ and } \frac{1}{1+\lambda} \neq O(\lambda^m) \text{ for all } m > 1.$$

It should be noted that, given $a \in \mathbb{R}((\lambda))$, it is possible that $a \neq O(\lambda^m)$ for all $m \in \mathbb{N}$ since it might be the case that a cannot be written in the form of a power series (for example if $a = \frac{1}{\lambda}$). On the other hand we have $0 = O(\lambda^m)$ for *all* $m \in \mathbb{N}$ although, as the following proposition makes clear, 0 is the *only* element in $\mathbb{R}((\lambda))$ for which this holds.

Proposition 4.4 *Let $a \in \mathbb{R}((\lambda))$. If, for all $m = 0, 1, 2, \ldots$ we have $a = O(\lambda^m)$ then $a = 0$.*

Proof. Suppose $a = \sum_{i=0}^{\infty} a_i \lambda_i$. If $a \neq 0$ then we must have $a_k \neq 0$ for some $k = 0, 1, \ldots$. Assume k is minimal such that this occurs, so then $a = O(\lambda^k)$, but $a \neq O(\lambda^{k+1})$. □

Corollary 4.5 *Let $a, b \in \mathbb{R}((\lambda))$. If, for all $m = 0, 1, 2, \ldots$ we have $a - b = O(\lambda^m)$ then $a = b$.* □

The next two propositions are concerned with the arithmetic of the field $\mathbb{R}((\lambda))$. They will be used repeatedly (often without explicit mention) throughout the proof of Theorem 4.1.

Proposition 4.6 *Let $a \in \mathbb{R}((\lambda))$ be such that $a = O(\lambda^k)$ for some $k > 0$. Then*

$$\frac{1}{1+a} = 1 + b$$

for some $b \in \mathbb{R}((\lambda))$ such that $b = O(\lambda^k)$. We can express this result in an abbreviated form by

$$\frac{1}{1+O(\lambda^k)} = 1 + O(\lambda^k).$$

Proof. Suppose $a = \sum_{i=0}^{\infty} a_i \lambda^i$ where $a_i = 0$ for all $i < k$. We will show that, in fact,

$$\frac{1}{1+a} = c = \sum_{i=0}^{\infty} c_i \lambda^i$$

where $c_0 = 1$, $c_i = 0$ for $0 < i < k$ and, for $i \geq k$,

$$c_i = -\sum_{j=k}^{i} a_j c_{k-j}.$$

This will clearly suffice.

We have

$$c = \frac{1}{1+a} \text{ iff } c(1+a) = 1 \text{ iff } c + ac - 1 = 0$$

and so to check the validity of our claim we must show that the i^{th} coefficient of $c + ac - 1$, which we here denote by $(c + ac - 1)_i$, is zero for all i. For $i = 0$ we have

$$(c + ac - 1)_0 = c_0 + a_0 c_0 - 1 = 0$$

as required, since $a_0 = 0$ and $c_0 = 1$. For $0 < i < k$ we have

$$(c + ac - 1)_i = c_i + \sum_{j=0}^{i} a_j c_{i-j} = 0$$

as required, since $a_j = 0$ for all $j < k$ and $c_i = 0$ for $0 < i < k$. Finally for $i \geq k$ we have

$$\begin{aligned}(c + ac - 1)_i &= c_i + \sum_{j=0}^{i} a_j c_{i-j} \\ &= c_i + \sum_{j=k}^{i} a_j c_{i-j} \qquad \text{since } a_j = 0 \text{ for all } j < k \\ &= -\sum_{j=k}^{i} a_j c_{k-j} + \sum_{j=k}^{i} a_j c_{k-j} \\ &= 0\end{aligned}$$

as required. □

Proposition 4.7 *Let $a, b \in \mathbb{R}((\lambda))$ be such that $a = O(\lambda^k)$ and $b = O(\lambda^j)$ for some $k, j \geq 0$. Then*
(i). $a + b = O(\lambda^y)$ where $y = \min\{k, j\}$, that is, $O(\lambda^k) + O(\lambda^j) = O(\lambda^y)$,
(ii). $a \cdot b = O(\lambda^{k+j})$, that is, $O(\lambda^k) \times O(\lambda^j) = O(\lambda^{k+j})$,
(iii). $\frac{a}{b} = O(\lambda^{k-j})$, that is, $\frac{O(\lambda^k)}{O(\lambda^j)} = O(\lambda^{k-j})$, so long as $k \geq j$ and $b \neq O(\lambda^{j+1})$.

Proof. Let us suppose that

$$a = \sum_{i=0}^{\infty} a_i \lambda^i \text{ and } b = \sum_{i=0}^{\infty} b_i \lambda^i$$

where we know $a_i = 0$ for all $i < k$ and $b_i = 0$ for all $i < j$.
(i). This is clear since $i < y = \min\{k, j\}$ implies $i < k$ and $i < j$ which implies $a_i = b_i = 0$ and so the i^{th} coefficient $(a + b)_i$ of $a + b$ is equal to $a_i + b_i$ is equal to zero.
(ii). To show $a \cdot b = O(\lambda^{k+j})$ we must show that the i^{th} coefficient $(a \cdot b)_i$ of $a \cdot b$ is zero for all $i < k + j$. We know

$$(a \cdot b)_i = \sum_{s+t=i} a_s b_t.$$

Then $i < k + j$ and $s + t = i$ implies either $s < k$ or $t < j$. Either way we must have $a_s b_t = 0$ and so $(a \cdot b)_i = 0$ as required.

(iii). Let us assume for this part that $k \geq j$ and $b \neq O(\lambda^{j+1})$, i.e., $b_j \neq 0$. Then we may write

$$\begin{aligned}
\frac{a}{b} &= \frac{a_k\lambda^k + a_{k+1}\lambda^{k+1} + \dots}{b_j\lambda^j + b_{j+1}\lambda^{j+1} + \dots} \\
&= \frac{\frac{a_k}{b_j}\lambda^{k-j} + \frac{a_{k+1}}{b_j}\lambda^{k-j+1} + \dots}{1 + \frac{b_{j+1}}{b_j}\lambda + \dots} \\
&= O(\lambda^{k-j}) \times \frac{1}{1 + O(\lambda)} \\
&= O(\lambda^{k-j}) \times (1 + O(\lambda)) \qquad \text{(by Proposition 4.6)} \\
&= O(\lambda^{k-j}) \times O(1) \\
&\quad \text{(since, for } d \in \mathbb{R}((\lambda)),\ d = 1 + O(\lambda) \text{ implies } d = O(1)) \\
&= O(\lambda^{k-j}) \qquad \text{(by (ii) proved above).}
\end{aligned}$$

□

Given the preceding construction we now generalise our definitions of pre-ent and ent as follows:

Definition 4.8 *A λ-pre-ent over a given language L is the same as a pre-ent over L except that it is a function into $[-1,1]^{(\lambda)} = \{a \in \mathbb{R}((\lambda)) \mid -1 \leq a \leq 1\}$ instead of just $[-1,1]$ (and so it gives rise to a belief function Bel which takes values in $[0,1]^{(\lambda)} = \{a \in \mathbb{R}((\lambda)) \mid 0 \leq a \leq 1\}$). Likewise a λ-ent over L is defined to be the same as an ent over L except that it is a function into $[0,\infty)^{(\lambda)} = \{a \in \mathbb{R}((\lambda)) \mid a \geq 0\}$ rather than just $[0,\infty)$.*

It is easy to see that all the results on pre-ents and ents given in Chapters 2 and 3 remain true for λ-pre-ents and λ-ents.

Since in the rest of this chapter (in fact in the rest of this thesis!) we will mainly deal only with pre-ents and ents which conform to the above definition we will drop the prefix λ and assume, unless it is otherwise indicated, that all our pre-ents and ents are actually λ-pre-ents and λ-ents. If a particular pre-ent or ent picks up only values in $\mathbb{R}$ then we will indicate this fact by calling it a *standard* pre-ent or ent. We should emphasize that the pre-ent and ent referred to in the statement of Theorem 4.1 are both standard.

4.3 Almost-ents

In this section, purely in the interests of proving Theorem 4.1, we weaken the definition of (λ-)ent to obtain the class of almost-ents. The precise definition is as follows:

Definition 4.9 *Let L be a language. An* almost-ent over L *is a function $z : WL \to [0,\infty)^{(\lambda)}$ which satisfies the following property: For all $s \in WL$, if there exist $t_1, \ldots, t_j \in WL$ ($j \geq 0$) such that $s = t_1 \cup \ldots \cup t_j$ and $z_{t_i} > 0$ for $i = 1, \ldots, j$ then, for all $p \in L$ such that $\pm p \notin s$, there exists $t \in WL$ such that $s \cup t$ is consistent, $\pm p \in t$ and $z_t > 0$.*

An almost-ent z, then, is just like an ent (indeed we still call the value z_t the *potential* of the scenario t according to z) except that we allow that there may exist a scenario s and a propositional variable p such that $\pm p \notin s$ and, for all scenarios t such that $\pm p \in t$ and $z_t > 0$, we have $s \cup t$ is inconsistent. In other words, z has nowhere to go from s to decide p. However, the definition of almost-ent ensures that any scenarios which suffer from this "defect" are, in any case, "unimaginable" (from $\emptyset$) according to z. Note that, in the above definition, we include the possibility that $j = 0$, i.e., that $s = \emptyset$. Thus any almost-ent z satisfies the condition that, for all $p \in L$, there exists $t \in WL$ such that $\pm p \in t$ and $z_t > 0$. Also note that, by definition, any ent over L is an almost-ent over L.

As is the case with ents, and in a similar way, any almost-ent z over L yields a pre-ent G^z over L (which in turn yields a function $Bel^z : SL \to [0,1]^{(\lambda)}$ in the usual way). This can be done as follows: Suppose we are given $p \in L$, $s, t \in WL$

such that $\pm p \notin s$ and $s \subseteq t$. (Of course if $s \not\subseteq t$ then we set $G^z_p(s,t) = 0$ while if $p \in s$ ($\neg p \in s$) then we set $G^z_p(s,s) = 1$ ($G^z_p(s,s) = -1$).) If there exist $t_1, \ldots, t_j$ ($j \geq 0$) such that $s = t_1 \cup \ldots \cup t_j$ and $z_{t_i} > 0$ for $i = 1, \ldots, j$, then we define

$$G^z_p(s,t) = \begin{cases} \dfrac{\sum\{z_r \mid s \cup r = t\}}{\sum\{z_r \mid s \cup r \text{ is consistent and } \pm p \in r\}} & \text{if } p \in t \\ \dfrac{-\sum\{z_r \mid s \cup r = t\}}{\sum\{z_r \mid s \cup r \text{ is consistent and } \pm p \in r\}} & \text{if } \neg p \in t \\ 0 & \text{otherwise.} \end{cases}$$

(Note by definition of almost-ent that, here, none of the denominators will be equal to zero. Also note that so far our definition of G^z corresponds to the usual way of defining a pre-ent from a given *ent* z.) In particular, taking $j = 0$ in the above, i.e., $s = \emptyset$, we have, for all $p \in L$ and $t \in WL$,

$$G^z_p(\emptyset,t) = \begin{cases} \dfrac{z_t}{\sum\{z_r \mid \pm p \in r\}} & \text{if } p \in t \\ \dfrac{-z_t}{\sum\{z_r \mid \pm p \in r\}} & \text{if } \neg p \in t \\ 0 & \text{otherwise.} \end{cases}$$

If it is *not* the case that $s = t_1 \cup \ldots \cup t_j$ for some t_i such that $z_{t_i} > 0$ then we simply define $G^z_p(s,t)$ in any manner so as to satisfy the definition of a pre-ent. (For example defining $G^z_p(s,t)$ by

$$G^z_p(s,t) = \begin{cases} 1 & \text{if } t = s \cup \{p\} \\ 0 & \text{otherwise} \end{cases}$$

would do.) The precise details of the definition of G^z in this case are, for our purposes in this chapter, irrelevant since, in going about its business of forming beliefs (assuming it starts from $\emptyset$), the almost-ent z will never encounter a scenario s which is not formed as a union of scenarios which have non-zero potentials and so the definition of $G^z_p(s,\cdot)$ will never be called into action. To make this clear let us consider how an almost-ent z over a language L would compute its belief in a conjunction of literals $q_1 \wedge \ldots \wedge q_j$ from distinct propositional variables in L. We have, from our basic definitions in Section 2.2,

$$Bel^z(q_1 \wedge \ldots \wedge q_j) = \sum G^z_{q_1}(\emptyset, s_1) \cdot G^z_{q_2}(s_1, s_2) \cdots G^z_{q_j}(s_{j-1}, s_j) \tag{4.1}$$

where G^z is the pre-ent yielded by z in the manner described above and the sum is over all sequences of scenarios (over L) $s_1 \subseteq s_2 \subseteq \cdots \subseteq s_j$ which satisfy (taking $s_0 = \emptyset$), for each $i = 1, \ldots, j$, *(i)* $q_i \in s_i$, *(ii)* $s_i = s_{i-1}$ if $q_i \in s_{i-1}$ and *(iii)* $G^z_{q_i}(s_{i-1}, s_i) > 0$. Now, for each such sequence and for each $i = 1, \ldots, j$, we claim

$$G^z_{q_i}(s_{i-1}, s_i) = \begin{cases} \dfrac{\sum\{z_t \mid s_{i-1} \cup t = s_i\}}{\sum\{z_t \mid s_{i-1} \cup t \text{ consistent, } \pm q_i \in t\}} & \text{if } q_i \notin s_{i-1} \\ 1 & \text{if } q_i \in s_{i-1}. \end{cases}$$

(Note that, whenever q is a literal, we use $\pm q$ to denote $\pm p$ where p is the propositional variable appearing in q.) To see this we use induction on i. For $i = 1$ the above formula translates into

$$G^z_{q_1}(\emptyset, s_1) = \frac{z_{s_1}}{\sum\{z_t \mid \pm q_1 \in t\}}$$

which is true by our definition of G^z. Suppose, for inductive hypothesis that the formula is true for $i = 1, \ldots, k$. Clearly if $q_{k+1} \in s_k$ then $s_{k+1} = s_k$ and $G^z_{q_{k+1}}(s_k, s_{k+1}) = 1$ as required. So suppose also that $q_{k+1} \notin s_k$. Since we have $G^z_{q_i}(s_{i-1}, s_i) > 0$ for $i = 1, \ldots, k$, it must be the case, using the inductive hypothesis, that, for each $i = 1, \ldots, k$, either $s_i = s_{i-1}$ or there exists at least one scenario t such that $s_{i-1} \cup t = s_i$ and $z_t > 0$. Hence there must exist $t_1, \ldots, t_l$, say, such that $\bigcup_{i \leq l} t_i = s_k$ and $z_{t_i} > 0$ for $i = 1, \ldots, l$. Thus the formula also holds true for $i = k + 1$ by definition of G^z, thereby completing the inductive proof. Hence, given a sequence of scenarios $s_1 \subseteq \cdots \subseteq s_j$ which satisfies, for each $i = 1, \ldots, j$, *(i)* $q_i \in s_i$, *(ii)* $s_i = s_{i-1}$ if $q_i \in s_{i-1}$ and *(iii)* $G^z_{q_i}(s_{i-1}, s_i) > 0$, we may write

$$\prod_{i=1}^{j} G^z_{q_i}(s_{i-1}, s_i) = \sum \prod_{i=1}^{j} \Theta^z(s_{i-1} \xrightarrow{q_i} r_i) \tag{4.2}$$

where the summation here is to be taken over all sequences of scenarios $r_1, \ldots, r_j$ which satisfy, for each $i = 1, \ldots, j$, *(i)* $s_{i-1} \cup r_i = s_i$, *(ii)* if $s_i = s_{i-1}$ then $r_i = \emptyset$ and *(iii)* if $r_i \neq \emptyset$ then $z_{r_i} > 0$, and the term $\Theta^z(s_{i-1} \xrightarrow{q_i} r_i)$ which, as indicated, depends only on q_i together with the scenarios (over L) s_{i-1} and r_i is defined as follows

$$\Theta^z(s_{i-1} \xrightarrow{q_i} r_i) = \begin{cases} \dfrac{z_{r_i}}{\sum\{z_t \mid s_{i-1} \cup t \text{ consistent}, \ \pm q_i \in t\}} & \text{if } r_i \neq \emptyset \\ 1 & \text{if } r_i = \emptyset. \end{cases}$$

We now tighten up our notation even further by making the following definition.

Definition 4.10 *(a). Let $q_1 \wedge \ldots \wedge q_j$ be a conjunction of literals from distinct propositional variables in L. Then a* scenario path *(over L)* for *$q_1 \wedge \ldots \wedge q_j$ is a sequence $\vec{r} = r_1, \ldots, r_j$ of scenarios over L which satisfies (i) $q_1 \in r_1$ and (ii) for each $i \geq 1$, if $q_{i+1} \in \bigcup_{k \leq i} r_k$ then $r_{i+1} = \emptyset$, otherwise r_{i+1} is such that $q_{i+1} \in r_{i+1}$ and $\bigcup_{k \leq i} r_k \cup r_{i+1}$ is consistent.*
(b) Given an almost-ent z over L we shall say that the scenario path $\vec{r}$ for $q_1 \wedge \ldots \wedge q_j$ is non-zero for *z, to be abbreviated by $\vec{r} \neq_z 0$, iff for each $i = 1, \ldots, j$ we have $r_i \neq \emptyset$ implies $z_{r_i} > 0$.*

According to this definition, then, and combining the equations (4.1) and (4.2) we may rewrite $Bel^z(q_1 \wedge \ldots \wedge q_j)$ as

$$Bel^z(q_1 \wedge \ldots \wedge q_j) = \sum_{\vec{r} \neq_z 0} \prod_{i=1}^{j} \Theta^z(\bigcup_{k<i} r_k \xrightarrow{q_i} r_i)$$

where the sum is over all scenario paths over L for $q_1 \wedge \cdots \wedge q_j$ which are non-zero for z and the terms $\Theta^z(\bigcup_{k<i} r_k \xrightarrow{q_i} r_i)$ are as defined above. Having established the definition of almost-ents and developed a general formula for the belief an almost-ent gives to a conjunction of literals from distinct propositional variables, we now set the scene for our the proof of Theorem 4.1.

4.4 Some Preliminaries

In this section we introduce the notation which we shall need in our proof of Theorem 4.1. We shall continue to assume that we are working in the field $\mathbb{R}((\lambda))$. As we previously indicated our first aim will be to prove the following weakening of Theorem 4.1.

Theorem 4.11 *Given a language $L = \{p_1, \ldots, p_n\}$, if the function $Bel : SL \to [0,1]$ is given by a standard pre-ent over L and if, for all $\theta, \phi \in SL$, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$, then there exists an* **almost-ent** *z (over a larger language than L) such that, for all $\theta \in SL$, $Bel^z(\theta) = Bel(\theta)$. The potentials of z are elements in $[0, \infty)^{(\lambda)}$.*

Throughout the rest of this chapter we shall take $Bel : SL \to [0,1]$ to be our fixed given function which satisfies the hypotheses of Theorem 4.11 (and Theorem 4.1). As was demonstrated at the beginning of the present chapter Theorem 4.1 is certainly true in the case where we take $n = 1$, i.e., when L consists of just a single propositional variable. In view of this we shall, from now on in the rest of this chapter, assume that $n > 1$. In the course of the upcoming proofs (specifically Lemma 4.40 in Section 4.6) we shall need to rely on this assumption. Before we begin the proof, which will take us through a host of lemmas, we need to describe the various pieces of notation and abbreviation we will be using.

During the course of this chapter we shall need to talk about different languages which extend L, but when we refer to "a sequence of literals" we shall always mean a sequence of literals from distinct propositional variables **in** L. We shall denote the empty sequence of literals by $\emptyset$. (The context will always make it clear whether we are referring to an empty sequence of literals or the empty set!) We shall use σ, τ, ρ, etc, to denote sequences of literals. Given a sequence of literals σ we shall define the *length* $|\sigma|$ of σ to be the number of literals occurring in σ, i.e., $|\sigma| = j$ where $\sigma = q_1 \cdots q_j$. Thus we have that $\emptyset$ is the one and only sequence of length zero. Note that $|\sigma|$ is bounded above by $n = |L|$ since the q_i are literals drawn from *distinct* variables in L. Given two sequences of literals $\sigma = q_1 \cdots q_j$ ($j \geq 0$) and $\rho = r_1 \cdots r_s$ ($s \geq 0$) such that q_i does not appear in ρ for $i = 1, \ldots, j$ we shall sometimes denote by $\sigma\rho$ the sequence of literals $q_1 \cdots q_j r_1 \cdots r_s$ and we shall write $\sigma \subseteq \rho$ to mean that $\rho = \sigma\tau$ for some (possibly empty) sequence of literals τ. If $\sigma \subseteq \rho$ we shall sometimes say that σ is an *initial segment* of ρ. Whenever a (possibly empty) sequence of literals $\sigma = q_1 \cdots q_j$ appears as an argument of a belief function, we are simply using it as shorthand for the conjunction of literals $q_1 \wedge \ldots \wedge q_j$. So, for a pre-ent G, $Bel^G(\sigma)$ is just shorthand for $Bel^G(q_1 \wedge \ldots \wedge q_j)$ etc. Under this notation we have that $Bel^G(\emptyset)$ is the belief G has in the empty conjunction of literals which, since we here adopt the convention that any empty conjunction of sentences is a tautology, is always equal to one. Another consequence of this notation is that we have, for any two sequences σ, τ, if $\sigma \subseteq \tau$ then $Bel^G(\tau) \leq Bel^G(\sigma)$. (Since $Bel^G(\lambda \wedge \chi) \leq Bel^G(\lambda)$ for all pre-ents G and $\lambda, \chi \in SL$.) Finally, for each $j \geq 1$, given a j-tuple of objects $\vec{a} = \langle a_1, \ldots, a_j \rangle$ then, given $i \leq j$, we shall use the notation $\vec{a}{\restriction}i$ to denote the i-tuple $\langle a_1, \ldots, a_i \rangle$.

As indicated in the above statement of Theorem 4.11 the almost-ent which we produce in our proof will be defined over a language which extends the language L. This language, which we will denote by L^+, contains all the propositional variables of L together with a set of new propositional variables, one for each (non-empty) sequence of literals from L. Precisely we have

$$L^+ = L \cup \{u_{q_1 \cdots q_j} \mid q_1 \cdots q_j \text{ a sequence of literals, } j \geq 1\}.$$

Given a (non-empty) sequence $q_1 \cdots q_j$ of literals we define the scenario $s(q_1 \cdots q_j)$ over L^+ as follows:

$$s(q_1 \cdots q_j) = \{q_j\} \cup \{u_\sigma \mid \sigma \text{ is an initial segment of } q_1 \cdots q_j,\ \sigma \neq \emptyset\}$$

$$\cup \quad \{\neg u_\sigma \mid 1 \leq |\sigma| \leq j \text{ and } \sigma \text{ is not an initial segment of } q_1 \cdots q_j\}$$

So the scenario $s(q_1 \cdots q_j)$ contains just one literal from the original language L, i.e., the last literal occurring in the sequence $q_1 \cdots q_j$, together with a set of literals from $L^+ - L$ which, in effect, play the role of "markers" which show how the sequence "arrives" at q_j. Given an almost-ent z over L^+ we shall denote the potential z gives to the scenario $s(q_1 \cdots q_j)$ by $z(q_1 \cdots q_j)$. All the almost-ents over L^+ we shall encounter in our proofs in this chapter will give non-zero potential only to the scenarios of the form $s(\sigma)$ for σ a non-empty sequence of literals. In other words, all our almost-ents will be *special* according to the following definition.

Definition 4.12 *Given an almost-ent z over the language L^+, we shall say that z is* special *iff it gives non-zero potential only to scenarios of the form $s(\sigma)$ for σ a (non-empty) sequence of literals.*

The forthcoming Proposition 4.14 provides the means by which we are able to check whether or not a given function from WL^+ to $[0, \infty)^{(\lambda)}$ is in fact a legitimate special almost-ent. However, before we get to it, it will prove useful to know under what conditions two scenarios over L^+ of the form $s(\sigma)$ are jointly consistent.

Lemma 4.13 *Let σ and τ be non-empty sequences of literals. Then $s(\sigma) \cup s(\tau)$ is consistent iff either $\sigma \subseteq \tau$ or $\tau \subseteq \sigma$.*

Proof. First we show the "if" direction. By symmetry, we need look only at the case where $\sigma \subseteq \tau$. Recall that, by definition, we have

$$s(\sigma) = \{q\} \cup \{u_\rho \mid \emptyset \neq \rho \subseteq \sigma\} \cup \{\neg u_\rho \mid 1 \leq |\rho| \leq |\sigma|,\ \rho \not\subseteq \sigma\}$$

and

$$s(\tau) = \{r\} \cup \{u_\rho \mid \emptyset \neq \rho \subseteq \tau\} \cup \{\neg u_\rho \mid 1 \leq |\rho| \leq |\tau|,\ \rho \not\subseteq \tau\}$$

where q, respectively r, is the last literal in σ, respectively τ. Now, since $\sigma \subseteq \tau$, we have that, for any sequence of literals ρ, if $\rho \subseteq \sigma$ then $\rho \subseteq \tau$, Hence

$$\{u_\rho \mid \emptyset \neq \rho \subseteq \sigma\} \subseteq \{u_\rho \mid \emptyset \neq \rho \subseteq \tau\}.$$

We also have that $1 \leq |\rho| \leq |\sigma|$ implies $1 \leq |\rho| \leq |\tau|$ (since obviously $|\sigma| \leq |\tau|$) while if $\rho \not\subseteq \sigma$ and $|\rho| \leq |\sigma|$ then, clearly, also $\rho \not\subseteq \tau$. Hence

$$\{\neg u_\rho \mid 1 \leq |\rho| \leq |\sigma|,\ \rho \not\subseteq \sigma\} \subseteq \{\neg u_\rho \mid 1 \leq |\rho| \leq |\tau|,\ \rho \not\subseteq \tau\}.$$

Thus we may see that

$$\begin{aligned} s(\sigma) \cup s(\tau) &= \{q, r\} \cup \{u_\rho \mid \emptyset \neq \rho \subseteq \tau\} \cup \{\neg u_\rho \mid 1 \leq |\rho| \leq |\tau|,\ \rho \not\subseteq \tau\} \\ &= s(\tau) \cup \{q\} \end{aligned}$$

which, since q and r are either the same literal (if $\sigma = \tau$) or literals from different propositional variables in L (by definition of a sequence of literals, since both occur in τ), is clearly consistent.

Now for the "only if" direction. Suppose that both $\sigma \not\subseteq \tau$ and $\tau \not\subseteq \sigma$. Then there must exist a (possibly empty) sequence of literals ρ and literals $q \neq r$ such that $\sigma = \rho q \cdots$ and $\tau = \rho r \cdots$. But then, by the definition of $s(\sigma)$ and $s(\tau)$, we have that $u_{\rho q} \in s(\sigma)$ and $\neg u_{\rho q} \in s(\tau)$ which means $s(\sigma) \cup s(\tau)$ is inconsistent as required. □

Proposition 4.14 *Let $z : WL^+ \rightarrow [0, \infty)^{(\lambda)}$ be a function which gives non-zero values only to scenarios over L^+ of the form $s(\sigma)$ for σ a non-empty sequence of literals. Then z is an almost-ent over L^+ iff it satisfies the following conditions:*

A-E1. *For all non-empty sequences of literals σ and τ, if $z(\tau) > 0$ and $\sigma \subseteq \tau$ then $z(\sigma) > 0$.*

A-E2. *For all $p \in L$ there exists a non-empty sequence of literals σ such that σ ends with $\pm p$ and $z(\sigma) > 0$.*

A-E3. *For all sequences of literals σ such that $z(\sigma) > 0$ and all $p \in L$ such that $\pm p$ does not appear in σ there exists a sequence of literals τ such that $\sigma \subseteq \tau$, τ ends with $\pm p$ and $z(\tau) > 0$.*

Proof. Before we begin the proof let us note that, since we are assuming our function z gives non-zero values only to scenarios over L^+ of the form $s(\sigma)$ for σ a non-empty sequence of literals, the condition of Definition 4.9 reduces to:

> z is an almost-ent over L^+ iff for all $s \in WL^+$, if there exist non-empty sequences of literals $\sigma_1, \ldots, \sigma_j$ ($j \geq 0$) such that $s = s(\sigma_1) \cup \ldots \cup s(\sigma_j)$ and $z(\sigma_i) > 0$ for $i = 1, \ldots, j$ then, for all $p \in L^+$ such that $\pm p \notin s$, there exists a non-empty sequence of literals σ such that $s \cup s(\sigma)$ is consistent, $\pm p \in s(\sigma)$ and $z(\sigma) > 0$.

We first show the "only if" direction of the proposition. Let z be a special almost-ent over L^+. To show **A-E1** let σ and τ be sequences of literals such that $z(\tau) > 0$ and $\sigma \subseteq \tau$. If $\sigma = \tau$ then obviously $z(\sigma) > 0$ as required so suppose further that $\sigma \neq \tau$ and that σ ends with the literal p^ϵ. Now out of all the scenarios which are given non-zero potential by z the only ones which contain $\pm p$ are those of the form $s(\rho)$ where ρ is a sequence of literals which ends with $\pm p$ and out of these the only ones which are consistent with $s(\tau)$ are those which satisfy either $\rho \subseteq \tau$ or $\tau \subseteq \rho$ (by Lemma 4.13). Clearly we cannot have both ρ ending with $\pm p$ and $\tau \subseteq \rho$ (since this would entail $\pm p$ appearing twice in ρ), while it should also be clear that $\rho \subseteq \tau$ and ρ ends with $\pm p$ implies that $\rho = \sigma$. Hence it follows that σ is the *only* sequence of literals which satisfies both $s(\tau) \cup s(\sigma)$ is consistent and $\pm p \in s(\sigma)$ and so, applying the above reformulated condition from Definition 4.9, we are forced to conclude that $z(\sigma) > 0$ as required.

To show **A-E2** we know by the above condition that z being an almost-ent over L^+ guarantees that, for each $p \in L$, there exists a sequence of literals σ such that $\pm p \in s(\sigma)$ and $z(\sigma) > 0$. This suffices by definition of $s(\sigma)$ since $\pm p \in s(\sigma)$ iff σ ends with $\pm p$.

To show **A-E3** let σ be a non-empty sequence of literals such that $z(\sigma) > 0$ and let $p \in L$ be such that $\pm p$ does not appear in σ. Then, in particular, σ does not end with $\pm p$ and so $\pm p \notin s(\sigma)$ (by definition of $s(\sigma)$). Since z is an almost-ent there exists a non-empty sequence of literals τ such that $s(\sigma) \cup s(\tau)$ is consistent, $\pm p \in s(\tau)$ and $z(\tau) > 0$. For any such τ we have, by Lemma 4.13, that $s(\sigma) \cup s(\tau)$ is consistent iff either $\sigma \subseteq \tau$ or $\tau \subseteq \sigma$. But $\pm p \in s(\tau)$ implies τ ends $\pm p$ and hence we cannot have $\tau \subseteq \sigma$ (since otherwise $\pm p$ would appear in σ giving rise to a contradiction). Thus we have shown that there must exist some τ which satisfies $\sigma \subseteq \tau$, τ ends with $\pm p$ and $z(\tau) > 0$ as required to show **A-E3**.

To show the "if" direction suppose now that $z : WL^+ \to [0, \infty)^{(\lambda)}$ is a function which gives non-zero values only to scenarios of the form $s(\sigma)$ and that z satisfies conditions **A-E1** - **A-E3**. We must show that z is an almost-ent over L^+, i.e., that z satisfies the reformulated condition of Definition 4.9 given above. We will look at the separate cases $p \in L$ and $p \in L^+ - L$. Let ρ be a sequence of maximal length amongst $\sigma_1, \dots, \sigma_j$ (take $\rho = \emptyset$ if $j = 0$, i.e., if $s = \emptyset$). Then, since for sequences σ_i and σ_k we know, by Lemma 4.13, $s(\sigma_i) \cup s(\sigma_k)$ is consistent iff either $\sigma_i \subseteq \sigma_k$ or $\sigma_k \subseteq \sigma_i$, and since s must be consistent (by definition of scenario), we have $\sigma_i \subseteq \rho$ for all $i = 1, \dots, j$. If $p \in L$ then either $\pm p$ appears in ρ or it does not. Suppose the latter case applies. Then if $\rho = \emptyset$ then **A-E2** above tells us that there exists a sequence of literals σ such that σ ends with $\pm p$ and $z(\sigma) > 0$. By definition the scenario $s(\sigma)$ contains the last literal which appears in σ and so $\pm p \in s(\sigma)$ and the result is proved in this case. If $\rho \neq \emptyset$ then, since we have $z(\rho) > 0$, condition **A-E3** above tells us that there exists a sequence of literals

σ such that $\rho \subseteq \sigma$, σ ends with $\pm p$ and $z(\sigma) > 0$. We have that $s(\sigma)$ must be consistent with s (since $\rho \subseteq \sigma$ implies $\sigma_i \subseteq \sigma$ for all $i = 1, \ldots, j$) and again this scenario contains $\pm p$ which suffices. If $\pm p$ *does* appear in ρ, say $\rho = \tau p^\epsilon \cdots$, then we may take $s(\tau p^\epsilon)$ as our required scenario since it is clearly consistent with s and decides p, while $z(\tau p^\epsilon) > 0$ by **A-E1**. Finally we suppose that $p \in L^+ - L$. By the conditions **A-E2** (if $\rho = \emptyset$) and **A-E3** we can construct longer and longer sequences of literals δ such that $\rho \subseteq \delta$ and $z(\delta) > 0$ until we reach a δ such that $|\delta| = n$. But then, by definition of $s(\delta)$ for this δ, we have that, not only is $s \cup s(\delta)$ consistent, but $s(\delta)$ decides every propositional variable $u_\tau \in L^+ - L$ (indeed if $\tau \subseteq \delta$ then $u_\tau \in s(\delta)$ while if $\tau \not\subseteq \delta$ then $\neg u_\tau \in s(\delta)$). This completes the proof of the proposition. $\square$

Proposition 4.14 gives us an alternative way of thinking of special almost-ents over L^+. It says that we may think of them as the class of functions z, defined on the set of non-empty sequences of literals, which satisfy **A-E1**-**A-E3**. We now wish to investigate how a special almost-ent z over L^+ computes its beliefs in sequences (as conjunctions) of literals **from** L. From Section 4.3 we have, for $q_1 \cdots q_j$ a sequence of literals and *any* almost-ent z' over L^+,

$$Bel^{z'}(q_1 \ldots q_j) = \sum_{\vec{s} \neq_{z'} 0} \prod_{i=1}^{j} \Theta^{z'}(\bigcup_{k<i} s_k \xrightarrow{q_i} s_i)$$

where

$$\Theta^{z'}(\bigcup_{k<i} s_k \xrightarrow{q_i} s_i) = \begin{cases} \dfrac{z'_{s_i}}{\sum\{z'_r \mid \bigcup_{k<i} s_k \cup r \text{ consistent}, \ \pm q_i \in r\}} & \text{if } s_i \neq \emptyset \\ 1 & \text{if } s_i = \emptyset \end{cases}$$

and the above sum is taken to be over all scenario paths (over L^+) $\vec{s} = s_1, \ldots, s_j$ for $q_1 \cdots q_j$ which are non-zero for z'. In order to simplify the above formula, we would like to know what these scenario paths look like in the case when z' is taken to be special. Let z be a special almost-ent over L^+. We will now try and construct a scenario path (over L^+) $\vec{s} = s_1, \dots, s_j$ for the sequence $q_1 \cdots q_j$ which is non-zero for z. To begin with, the scenario s_1 must contain q_1. Now the only scenarios which decide q_1 one way or the other, and get non-zero potential (according to z), are those of the form $s(\rho_1)$ where ρ_1 is a sequence of literals which ends with $\pm q_1$, and $z(\rho_1) > 0$. Note that the condition **A-E2** guarantees the existence of at least one such scenario. Of these, the ones which decide q_1 positively, i.e., satisfy $q_1 \in s(\rho_1)$, are those for which ρ_1 ends q_1. Assuming such a ρ_1 exists, and given that s_1 has this form, it is clear that $q_2 \notin s(\rho_1)$ (since q_1 is the only literal from L which is in $s(\rho_1)$). Hence s_2 is required to contain q_2, be consistent with $s(\rho_1)$, and have non-zero potential. The only scenarios which decide q_2 and have non-zero potential are those of the form $s(\rho_2)$ where ρ_2 ends with a $\pm q_2$ and $z(\rho_2) > 0$. Out of these, by Lemma 4.13, the only ones which are consistent with $s(\rho_1)$ are those for which we have either $\rho_1 \subseteq \rho_2$ or $\rho_2 \subseteq \rho_1$. Hence if q_2 appears in the sequence ρ_1, say $\rho_1 = \tau_1 q_2 \cdots$, then the *only* scenario which decides q_2 and is consistent with $s(\rho_1)$ is $s(\tau_1 q_2)$, which decides q_2 positively. Thus in this case we are forced to take $s_2 = s(\tau_1 q_2)$. Note that, in this case, the condition **A-E1** guarantees that this scenario has non-zero potential. If $\overline{q}_2$ appears in ρ_1 (where, given a literal $q = p^\epsilon$, we define $\overline{q} = p^{1-\epsilon}$), say $\rho_1 = \tau_1 \overline{q}_2 \cdots$, then the only possible scenario is $s(\tau_1 \overline{q}_2)$. Thus in this case there is no scenario which is consistent with $s(\rho_1)$ and decides q_2 positively. If neither q_2 nor $\overline{q}_2$ appear in ρ_1 then the scenarios which decide q_2, are consistent with $s(\rho_1)$, and get non-zero potential are all the scenarios of the form $s(\rho_2)$ where $\rho_1 \subseteq \rho_2$, ρ_2 ends with $\pm q_2$, and $z(\rho_2) > 0$. Note that the condition **A-E3** guarantees the existence of at least one such ρ_2. Hence, provided at least one of these ρ_2's ends with q_2, we can take $s_2 = s(\rho_2)$ where $\rho_1 \subseteq \rho_2$, ρ_2 ends with q_2 and $z(\rho_2) > 0$. To complete the rest of the scenario path we may continue in this way, for suppose we have constructed $s_1, \dots, s_k$ (for some $k < j$) where $\bigcup_{i \le k} s_i$ is consistent and, for each

$i = 1, \ldots, k$, we have that $s_i = s(\rho_i)$ where ρ_i ends with q_i and $z(\rho_i) > 0$. Suppose ρ_{l_k} is the longest sequence so far constructed amongst $\{\rho_i \mid i = 1, \ldots, k\}$ and so, since $\bigcup_{i \leq k} s(\rho_i)$ is consistent, $\rho_i \subseteq \rho_{l_k}$ for each $i = 1, \ldots, k$ by Lemma 4.13. It is clear that $q_{k+1} \notin \bigcup_{i \leq k} s_i$ (since $q_1, \ldots, q_k$ are the only literals from L which are contained in this set). Hence s_{k+1} is required to contain q_{k+1}, be consistent with $\bigcup_{i \leq k} s_i$ and have non-zero potential. As above, the only scenarios which decide q_{k+1}, are consistent with $\bigcup_{i \leq k} s_i$ and get non-zero potential are those of the form $s(\rho_{k+1})$ where ρ_{k+1} ends with $\pm q_{k+1}$, $z(\rho_{k+1}) > 0$ and either $\rho_{l_k} \subseteq \rho_{k+1}$ or $\rho_{k+1} \subseteq \rho_{l_k}$, with q_{k+1} being decided positively iff ρ_{k+1} ends with q_{k+1}. Hence if q_{k+1} appears in ρ_{l_k}, say $\rho_{l_k} = \tau_k q_{k+1} \cdots$, then the only possibility is to take $\rho_{k+1} = \tau_k q_{k+1}$ while if $\overline{q}_{k+1}$ appears in ρ_{l_k} then there is no scenario consistent with $\bigcup_{i \leq k} s_i$ which decides q_{k+1} positively. If, on the other hand, $\pm q_{k+1}$ does not appear in ρ_{l_k} then we may take ρ_{k+1} to be such that $\rho_{l_k} \subseteq \rho_{k+1}$, ρ_{k+1} ends with q_{k+1} and $z(\rho_{k+1}) > 0$. Hence we may see that the scenario paths for $q_1 \cdots q_j$ which are non-zero for z have a special form, which we may express via the following definition.

Definition 4.15 *Let $q_1 \cdots q_j$ be a non-empty sequence of literals. A* non-monotonic (n-m) sequence path *for $q_1 \cdots q_j$ is a sequence of sequences of literals $\vec{\rho} = \rho_1, \ldots, \rho_j$ which satisfies (i) ρ_1 ends with q_1, and (ii) for each $i \geq 1$, if q_{i+1} appears in ρ_{l_i} (where l_i is such that $|\rho_{l_i}|$ is maximal amongst $\{|\rho_k| \mid k = 1, \ldots, i\}$, i.e., ρ_{l_i} is the longest sequence thus far constructed), say $\rho_{l_i} = \tau_i q_{i+1} \cdots$, then $\rho_{i+1} = \tau_i q_{i+1}$, otherwise $\rho_{l_i} \subseteq \rho_{i+1}$ and ρ_{i+1} ends with q_{i+1}. We shall denote the set of all n-m sequence paths for $q_1 \cdots q_j$ by $\widehat{P}(q_1 \cdots q_j)$.*

Given a special almost-ent z over L^+, we shall say the n-m sequence path $\vec{\rho}$ is non-zero for z *iff $z(\rho_{l_j}) > 0$, equivalently (by* **A-E1***) $z(\rho_i) > 0$ for all $i = 1, \ldots, j$. We shall denote the set of n-m sequence paths for $q_1 \cdots q_j$ which are non-zero for* *z by* $\widehat{N}_z(q_1 \cdots q_j)$. *Thus*

$$\widehat{N}_z(q_1 \cdots q_j) = \{\vec{\rho} \in \widehat{P}(q_1 \cdots q_j) \mid z(\rho_{l_j}) > 0\}.$$

Example 4.16 To give some examples of n-m sequence paths let us assume temporarily that $L = \{p, q, r, s\}$. Then an obvious n-m sequence path for the sequence pqr is $\vec{\rho}$ where $\rho_1 = p$, $\rho_2 = pq$ and $\rho_3 = pqr$. Another possible n-m sequence path for this particular sequence of literals can be given by setting $\rho_1 = rqp$, $\rho_2 = rq$ and $\rho_3 = r$. Yet another possibility is to take $\rho_1 = qp$, $\rho_2 = q$ and $\rho_3 = qpsr$.

Given a special almost-ent z over L^+ it should now be clear from the above discussion that the scenario paths $\vec{s}$ over L^+ for $q_1 \cdots q_j$ which are non-zero for z are precisely those paths of the form $s(\rho_1), \ldots, s(\rho_j)$ where $\vec{\rho} = \rho_1, \ldots, \rho_j$ is a n-m sequence path for $q_1 \cdots q_j$ which is non-zero for z. Hence we may write

$$Bel^z(q_1 \cdots q_j) = \sum_{\vec{\rho} \in \widehat{N}_z(q_1 \cdots q_j)} \prod_{i=1}^{j} \Theta^z(\bigcup_{k<i} s(\rho_k) \xrightarrow{q_i} s(\rho_i)) \tag{4.3}$$

where, for each $\vec{\rho} \in \widehat{N}_z(q_1 \cdots q_j)$ and for each $i = 1, \ldots, j$,

$$\begin{aligned}
\Theta^z(\bigcup_{k<i} s(\rho_k) \xrightarrow{q_i} s(\rho_i)) &= \frac{z(\rho_i)}{\sum\{z(\tau) \mid \bigcup_{k<i} s(\rho_k) \cup s(\tau) \text{ consistent, } \pm q_i \in s(\tau)\}} \\
&\qquad \text{(since we always have } s(\rho_i) \neq \emptyset) \\
&= \begin{cases} \dfrac{z(\rho_i)}{z(\rho_i)} = 1 & \text{if } \rho_i \subseteq \rho_{l_{i-1}} \\[2ex] \dfrac{z(\rho_i)}{\sum\{z(\tau) \mid \rho_{l_{i-1}} \subseteq \tau,\ \tau \text{ ends } \pm q_i\}} & \text{if } \rho_{l_{i-1}} \subseteq \rho_i \end{cases}
\end{aligned}$$

Where, as in the discussion above, $\rho_{l_{i-1}}$ is the longest sequence amongst $\{\rho_k \mid k = 1, \ldots, i-1\}$.

We shall be using the above representation in Section 4.7. For the rest of the present section, however, it will be convenient to slightly modify Definition 4.15 and work with a different type of path, one which contains essentially the same information as a non-monotonic sequence path.

Definition 4.17 *Let $q_1 \cdots q_j$ be a non-empty sequence of literals. Then a* monotonic sequence path (hereafter sequence path) for $q_1 \cdots q_j$ *is a sequence of sequences of literals $\vec{\sigma} = \sigma_1, \ldots, \sigma_{l(\vec{\sigma})}$ which satisfies (i) σ_1 ends with q_1, (ii) for each $i \geq 1$, $\sigma_i \subseteq \sigma_{i+1}$ and σ_{i+1} ends with q_t where t is minimal such that q_t does not appear in σ_i, and (iii) $l(\vec{\sigma})$ – the* length *of the sequence path $\vec{\sigma}$ – is minimal such that, for all $1 \leq i \leq j$, q_i appears in $\sigma_{l(\vec{\sigma})}$. We denote the set of all sequence paths for $q_1 \cdots q_j$ by $P(q_1 \cdots q_j)$.*

So the monotonic sequence paths for $q_1 \cdots q_j$ are just obtained from the non-monotonic sequence paths for $q_1 \cdots q_j$ by, for each $\vec{\sigma} \in \widehat{P}(q_1 \cdots q_j)$, first forming the sequence $\sigma_{l_1}, \sigma_{l_2}, \ldots, \sigma_{l_j}$ and then, reading this sequence from left to right, discarding any repeats. The canonical example of a sequence path for a sequence of literals $q_1 \cdots q_j$ is provided by the following definition.

Definition 4.18 *Let $q_1 \cdots q_j$ be a non-empty sequence of literals. The sequence path $\vec{\iota}(q_1 \cdots q_j) \in P(q_1 \cdots q_j)$ is defined to be that path $\vec{\sigma} = \sigma_1, \ldots, \sigma_j$ for which $\sigma_i = q_1 \cdots q_i$ for $i = 1, \ldots, j$.*

If the context makes it clear which sequence of literals we are talking about then we will sometimes just write $\vec{\iota}$ instead of $\vec{\iota}(q_1 \cdots q_j)$. The following proposition lists some basic characteristics of sequence paths.

Proposition 4.19 *Let $q_1 \cdots q_j$ be a non-empty sequence of literals and let $\vec{\sigma} \in P(q_1 \cdots q_j)$. Then the following are true:*

(i). $|\sigma_{i+1}| > |\sigma_i|$ for $i = 1, \ldots, l(\vec{\sigma}) - 1$.

(ii). $|\sigma_i| \geq i$ for $i = 1, \ldots, l(\vec{\sigma})$.

(iii). $0 \leq |\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) \leq n - 1$.

(iv). $|\sigma_{l(\vec{\sigma})}| = l(\vec{\sigma})$ iff $\vec{\sigma} = \vec{\iota}(q_1 \cdots q_j)$.

Proof. Property (i) is clear from the definition of sequence path. Property (ii) is provable by induction on i: it is obvious for the base case $i = 1$ while given that it is true for i we have $|\sigma_{i+1}| > |\sigma_i|$ (from (i)) $\geq i$ (from the inductive hypothesis) and so $|\sigma_{i+1}|$ must be equal to at least $i+1$ as required. To show property (iii) we have, by (ii), that $|\sigma_{l(\vec{\sigma})}| \geq l(\vec{\sigma})$ while clearly $|\sigma_{l(\vec{\sigma})}|$ is bounded above by n and $l(\vec{\sigma})$ is bounded below by 1. Combining this information gives the result. Finally to prove (iv) we have, by definition of $\vec{\iota}$, that $\vec{\sigma} = \vec{\iota}$ implies $|\sigma_{l(\vec{\sigma})}| = l(\vec{\sigma})$ while for the converse suppose $\vec{\sigma} \neq \vec{\iota}$ and let i be minimal such that $\sigma_i \neq q_1 \cdots q_i$. Then it should be clear that this implies that $|\sigma_i| \geq i+1$ and so, from (i) above, $|\sigma_{i+1}| \geq |\sigma_i| + 1 \geq (i+1)+1$ which in turn gives $|\sigma_{i+2}| \geq |\sigma_{i+1}| + 1 \geq (i+2)+1$ and so on until we reach $|\sigma_{l(\vec{\sigma})}| \geq l(\vec{\sigma}) + 1$ as required. □

Definition 4.20 *Given a non-empty sequence of literals $q_1 \cdots q_j$, a sequence path $\vec{\sigma}$ for $q_1 \cdots q_j$ and a special almost-ent z, we shall say that $\vec{\sigma}$ is* non-zero *for z iff $z(\sigma_{l(\vec{\sigma})}) > 0$, equivalently (by* **A-E1***) $z(\sigma_i) > 0$ for all $i = 1, \ldots, l(\vec{\sigma})$. We shall denote the set of sequence paths for $q_1 \cdots q_j$ which are non-zero for z by $N_z(q_1 \cdots q_j)$. Thus*

$$N_z(q_1 \cdots q_j) = \{\vec{\sigma} \in P(q_1 \cdots q_j) \mid z(\sigma_{l(\vec{\sigma})}) > 0\}.$$

It should now be clear that, taking $\sigma_0 = \emptyset$ throughout, we may rewrite (4.3) as

$$Bel^z(q_1 \cdots q_j) = \sum_{\vec{\sigma} \in N_z(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z(\sigma_i)}{\sum\{z(\tau) \mid \sigma_{i-1} \subseteq \tau,\ \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t\}}. \quad (4.4)$$

Note that all denominators in the above expression are non-zero. This is because $\vec{\sigma} \in N_z(q_1 \cdots q_j)$ implies $z(\sigma_i) > 0$, and $z(\sigma_i)$ always appears in the denominator. The term in this sum for which $\vec{\sigma} = \vec{\iota}(q_1 \cdots q_j)$ (if it occurs) we will call the *lead term*. Hence we have found a general formula, which we shall use repeatedly in what follows, for the belief given to any conjunction of literals by any special almost-ent over L^+.

The following lemma allows us to prove that all the almost-ents which we construct in the proof of Theorem 4.11 are indeed legitimate almost-ents. Its second part is a consequence (via Lemma 4.2) of the property assumed of Bel in the hypotheses of Theorem 4.11 (viz. for all $\theta, \phi \in SL$, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$).

Lemma 4.21 *Let $z : WL^+ \to [0, \infty)^{(\lambda)}$ be any function which gives non-zero values only to scenarios of the form $s(\sigma)$ for σ a non-empty sequence of literals. Then, if z satisfies*

$$z(\sigma) = 0 \textit{ iff } Bel(\sigma) = 0,$$

then z is a (special) almost-ent over L^+. Furthermore we have, for all sequences of literals $q_1 \cdots q_j$, $Bel(q_1 \cdots q_j) = 0$ implies $Bel^z(q_1 \cdots q_j) = 0$.

Proof. To show that z is an almost-ent over L^+ we simply need to check that any almost-ent z over WL^+ which satisfies the above condition also satisfies the conditions **A-E1**-**A-E3** from Proposition 4.14. Beginning with **A-E1**, let σ and τ be non-empty sequences of literals such that $z(\tau) > 0$ and $\sigma \subseteq \tau$. We must show $z(\sigma) > 0$. But if it were the case that $z(\sigma) = 0$ then we would have $Bel(\sigma) = 0$ by hypothesis and hence, since $\sigma \subseteq \tau$ implies $Bel(\tau) \leq Bel(\sigma)$, we would have $Bel(\tau) = 0$ and so $z(\tau) = 0$ giving a contradiction. Hence $z(\sigma) > 0$ as required.

To show **A-E2** let $p \in L$. We must show that there exists a sequence of literals σ such that σ ends with $\pm p$ and $z(\sigma) > 0$. But since $Bel(p) + Bel(\neg p) = 1$ we must have that either $Bel(p) > 0$ or $Bel(\neg p) > 0$ and so either $z(p) > 0$ or $z(\neg p) > 0$ which clearly suffices.

To show **A-E3** let σ be a non-empty sequence of literals such that $z(\sigma) > 0$ and let $p \in L$ be such that $\pm p$ does not appear in σ. We must prove the existence of a sequence of literals τ such that $\sigma \subseteq \tau$, τ ends with $\pm p$ and $z(\tau) > 0$. But $z(\sigma) > 0$ implies $Bel(\sigma) > 0$ and so, since $Bel(\sigma) = Bel(\sigma p) + Bel(\sigma \neg p)$, we must have either $Bel(\sigma p) > 0$ or $Bel(\sigma \neg p) > 0$, i.e., $Bel(\sigma p^\epsilon) > 0$ for some ϵ.

Hence the sequence σp^ϵ meets our required criteria since obviously $\sigma \subseteq \sigma p^\epsilon$ and σp^ϵ ends with $\pm p$, while $Bel(\sigma p^\epsilon) > 0$ implies $z(\sigma p^\epsilon) > 0$ by hypothesis.

Now let us show the last part of the lemma, i.e., that for all sequences of literals $q_1 \cdots q_j$ we have $Bel(q_1 \cdots q_j) = 0$ implies $Bel^z(q_1 \cdots q_j) = 0$. We have

$$\begin{aligned} Bel^z(q_1 \cdots q_j) &= \\ &= \sum_{\vec{\sigma} \in N_z(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z(\sigma_i)}{\sum\{z(\tau) \mid \sigma_{i-1} \subseteq \tau,\ \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t\}}. \end{aligned}$$

Hence $Bel^z(q_1 \cdots q_j) \neq 0$ iff there exists a sequence path for $q_1 \cdots q_j$ which is non-zero for z (equivalently $N_z(q_1 \cdots q_j) \neq \emptyset$). Suppose there existed such a path $\vec{\sigma}$. Then, by definition of non-zero, we would have $z(\sigma_{l(\vec{\sigma})}) > 0$ and so, by assumption, $Bel(\sigma_{l(\vec{\sigma})}) > 0$. But this would give, by Lemma 4.2, $Bel(q_1 \cdots q_j) \neq 0$ as required, since, for all $i = 1, \ldots, j$, we have that q_i appears in $\sigma_{l(\vec{\sigma})}$. □

We remark that the second part of the lemma is true if we relax the hypothesis to $Bel(\sigma) = 0$ implies $z(\sigma) = 0$ for all sequences of literals σ. Lemma 4.21 provides (via Lemma 4.2) that portion of the hypothesis of Theorem 4.11 (for all $\theta, \phi \in SL$ $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$) which will in fact be used in the proof. It will be used only in Lemmas 4.23 and 4.27.

Given that we now, hopefully, have a rigorous understanding of the workings of special almost-ents over L^+, we now make a start on stage one of our proof of Theorem 4.1, i.e., the proof of Theorem 4.11.

4.5 Stage 1 – Constructing the Almost-Ent z_∞

Having finally set up all the machinery which we shall be using in the proof of Theorem 4.11 it is now time to begin the proof proper. Our strategy is to inductively define an infinite sequence of special almost-ents over L^+ $z_0, z_1, z_2, \ldots$, showing as we go that the following are satisfied, for all $m = 0, 1, 2, \ldots$,

- S1. For all non-empty sequences of literals σ, $z_m(\sigma) = 0$ iff $Bel(\sigma) = 0$.
- S2. For all non-empty sequences of literals σ, $z_m(\sigma) \neq 0$ implies $z_m(\sigma) = O(\lambda^{|\sigma|-1})$ and $z_m(\sigma) \neq O(\lambda^{|\sigma|})$.
- S3. For all non-empty sequences of literals σ, $Bel^{z_m}(\sigma) - Bel(\sigma) = O(\lambda^{m+1})$,
- S4. For all (possibly empty) sequences of literals σ and all $p \in L$ such that $\pm p$ does not appear in σ,

$$\sum \{z_m(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma$$

  where A_σ is the term defined as follows:

$$A_\sigma = \sum \{z_0(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p'\} \tag{4.5}$$

  where p' is any propositional variable from L such that $\pm p'$ does not appear in σ. (It will soon be clear that this sum depends only on σ and not on which particular p' we choose.)

Once we have done this we will define our required special almost-ent z_∞ over L^+, i.e., that almost-ent for which we hope to show $Bel^{z_\infty}(\theta) = Bel(\theta)$ for all $\theta \in SL$, to be, in a sense to be explained later, the "limit" of these almost-ents. Note that, by Lemma 4.21, for each $m = 0, 1, 2, \ldots$, we can be sure z_m is a legitimate almost-ent over L^+ once we have shown that z_m satisfies S1. Another consequence of S1 is that, given a sequence of literals $q_1 \cdots q_j$ and a sequence path $\vec{\sigma} \in P(q_1 \cdots q_j)$, we have $\vec{\sigma} \in N_{z_m}(q_1 \cdots q_j)$ iff $z_m(\sigma_{l(\vec{\sigma})}) > 0$ iff $Bel(\sigma_{l(\vec{\sigma})}) > 0$. Hence, if we define

$$N(q_1 \cdots q_j) = \{\vec{\sigma} \in P(q_1 \cdots q_j) \mid Bel(\sigma_{l(\vec{\sigma})}) > 0\},$$

then S1 is equivalent to saying that, for all non-empty sequences of literals $q_1 \cdots q_j$,

$$N_{z_m}(q_1 \cdots q_j) = N(q_1 \cdots q_j)$$

(and so, significantly, N_{z_m} is actually independent of m).

Let us now begin our inductive process by initially defining the almost-ent z_0 by setting, for each non-empty sequence of literals σ,

$$z_0(\sigma) = \lambda^{|\sigma|-1} Bel(\sigma). \tag{4.6}$$

We may see straight away that properties S1 and S2 holds for $m = 0$ (and so, by Lemma 4.21, z_0 is a legitimate almost-ent over L^+). Before moving to the inductive stage of the process it remains to show that S3 and S4 hold for $m = 0$. S4 will hold trivially once we have shown that the term A_σ defined in (4.5) is independent of which propositional variable p' we choose amongst those which do not appear in σ. In order to do this let us re-express A_σ as follows. We have

$$A_\sigma \;=\; \sum \{ z_0(\sigma\tau p') + z_0(\sigma\tau\neg p') \}$$

where p' is a propositional variable which does not occur in σ, for definiteness let us say $p' = p_k$ where $k \leq n$ is minimal such that p_k does not occur in σ, and the sum is over all sequences of literals τ such that, for all $r \in L$ such that $\pm r$ appears in τ we have $r \neq p_k$ and $\pm r$ does not appear in σ. Splitting up these sequences τ according to length we get

$$\begin{aligned}
A_\sigma \;=\; & z_0(\sigma p_k) + z_0(\sigma\neg p_k) \\
& + \sum \Big\{ z_0(\sigma r_1 p_k) + z_0(\sigma r_1 \neg p_k) + z_0(\sigma \neg r_1 p_k) + z_0(\sigma \neg r_1 \neg p_k) \Big\} \\
& + \sum \Big\{ z_0(\sigma r_1 r_2 p_k) + z_0(\sigma r_1 r_2 \neg p_k) + z_0(\sigma r_1 \neg r_2 p_k) + z_0(\sigma r_1 \neg r_2 \neg p_k) + \\
& \quad + z_0(\sigma \neg r_1 r_2 p_k) + z_0(\sigma \neg r_1 r_2 \neg p_k) + z_0(\sigma \neg r_1 \neg r_2 p_k) + z_0(\sigma \neg r_1 \neg r_2 \neg p_k) \Big\} \\
& + \dots \\
& + \sum \Big\{ z_0(\sigma r_1^{\epsilon_1} r_2^{\epsilon_2} \cdots r_{m_\sigma}^{\epsilon_{m_\sigma}} p_k) + z_0(\sigma r_1^{\epsilon_1} r_2^{\epsilon_2} \cdots r_{m_\sigma}^{\epsilon_{m_\sigma}} \neg p_k) \Big\}
\end{aligned}$$

where the first summation here is over all $r_1 \in L$ such that $r_1 \neq p_k$ and $\pm r_1$ does not appear in σ, the second summation is over all distinct $r_1, r_2 \in L$ such that, for

$i = 1, 2$, $r_i \neq p_k$ and $\pm r_i$ does not appear in σ and so on until the final summation which is over all possible $\langle \epsilon_1, \epsilon_2, \ldots, \epsilon_{m_\sigma} \rangle \in \{0,1\}^{m_\sigma}$ (where $m_\sigma = n - (|\sigma| + 1)$) and all distinct $r_1, r_2, \ldots, r_{m_\sigma} \in L$ such that, for all $i = 1, \ldots, m_\sigma$, $r_i \neq p_k$ and $\pm r_i$ does not appear in σ. Now using the fact that, for all sequences τ , $z_0(\tau) = \lambda^{|\tau|-1} Bel(\tau)$ and repeatedly using the fact that, for all sequences τ and all $r \in L$, $Bel(\tau r) + Bel(\tau \neg r) = Bel(\tau)$ we get

$$A_\sigma = \lambda^{|\sigma|} Bel(\sigma) + a_1 \lambda^{|\sigma|+1} Bel(\sigma) + a_2 \lambda^{|\sigma|+2} Bel(\sigma) + \ldots + a_{m_\sigma} \lambda^{|\sigma|+m_\sigma} Bel(\sigma)$$

where $a_1, a_2, \ldots, a_{m_\sigma}$ are constants. Hence we may write

$$A_\sigma = \lambda^{|\sigma|} Bel(\sigma) \cdot P_\sigma(\lambda) \tag{4.7}$$

where $P_\sigma(\lambda)$ is a polynomial in λ with constant term (i.e., λ^0 term) 1. (This polynomial in fact depends only on $|\sigma|$ but this fact will not be used in any of the upcoming proofs.) We may now clearly see that A_σ is independent of p' and thus that S4 holds for $m = 0$. We also use this expression to show S3.

Lemma 4.22 *Let $q_1 \cdots q_j$ be a non-empty sequence of literals. Then $Bel^{z_0}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda)$.*

Proof. For each $\vec{\sigma} \in P(q_1 \cdots q_j)$ and for each $i = 1, \ldots, l(\vec{\sigma})$, we have

$$\sum \{ z_0(\tau) \mid \sigma_{i-1} \subseteq \tau,\ \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t \} = A_{\sigma_{i-1}}$$

so, using our general formula (4.4) applied to z_0, whilst recalling that, by S1, $N_{z_0}(q_1 \cdots q_j) = N(q_1 \cdots q_j)$, we get

$$\begin{aligned} Bel^{z_0}(q_1 \cdots q_j) &= \\ &= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_0(\sigma_i)}{\sum \{ z_0(\tau) \mid \sigma_{i-1} \subseteq \tau,\ \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t \}} \\ &= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_0(\sigma_i)}{A_{\sigma_{i-1}}} \end{aligned}$$

$$= \sum_{\vec{\sigma}\in N(q_1\cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{\lambda^{|\sigma_i|-1} Bel(\sigma_i)}{\lambda^{|\sigma_{i-1}|} Bel(\sigma_{i-1}) \cdot P_{\sigma_{i-1}}(\lambda)} \qquad \text{from (4.6) and (4.7)}$$

$$= \sum_{\vec{\sigma}\in N(q_1\cdots q_j)} \frac{Bel(\sigma_{l(\vec{\sigma})}) \cdot \lambda^{|\sigma_{l(\vec{\sigma})}|-l(\vec{\sigma})}}{Q_{\vec{\sigma}}(\lambda)}$$

where $Q_{\vec{\sigma}}(\lambda) = \prod P_{\sigma_{i-1}}(\lambda)$ is a polynomial in λ with constant term 1. The above summation is taken to be over all $\vec{\sigma} \in N(q_1 \cdots q_j)(\subseteq P(q_1 \cdots q_j))$, however we can equally take it to be over all $\vec{\sigma} \in P(q_1 \cdots q_j)$ since if $\vec{\sigma} \in P(q_1 \cdots q_j) - N(q_1 \cdots q_j)$ then $Bel(\sigma_{l(\vec{\sigma})}) = 0$ (by definition of $N(q_1 \cdots q_j)$). Hence

$$Bel^{z_0}(q_1 \cdots q_j) = \sum_{\vec{\sigma}\in P(q_1\cdots q_j)} \frac{Bel(\sigma_{l(\vec{\sigma})}) \cdot \lambda^{|\sigma_{l(\vec{\sigma})}|-l(\vec{\sigma})}}{Q_{\vec{\sigma}}(\lambda)}.$$

Now, by Proposition 4.19(iv), the only path $\vec{\sigma} \in P(q_1 \cdots q_j)$ for which $|\sigma_{l(\vec{\sigma})}| = l(\vec{\sigma})$, is the path $\vec{\iota}(q_1 \cdots q_j)$. So, given a sequence path $\vec{\sigma} \in P(q_1 \cdots q_j)$ such that $\vec{\sigma} \neq \vec{\iota}(q_1 \cdots q_j)$, we have

$$\frac{Bel(\sigma_{l(\vec{\sigma})}) \cdot \lambda^{|\sigma_{l(\vec{\sigma})}|-l(\vec{\sigma})}}{Q_{\vec{\sigma}}(\lambda)} = O(\lambda)$$

by Proposition 4.7(iii), since $|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) \geq 1$ (and so the numerator is of order $O(\lambda)$) and $Q_{\vec{\sigma}}(\lambda)$ has constant term 1 (and so the denominator is *not* of order $O(\lambda)$). Hence

$$\sum_{\vec{\iota}\neq\vec{\sigma}\in P(q_1\cdots q_j)} \frac{Bel(\sigma_{l(\vec{\sigma})}) \cdot \lambda^{|\sigma_{l(\vec{\sigma})}|-l(\vec{\sigma})}}{Q_{\vec{\sigma}}(\lambda)} = O(\lambda)$$

and so we may write

$$Bel^{z_0}(q_1 \cdots q_j) = \frac{Bel(q_1 \cdots q_j)}{Q_{\vec{\iota}}(\lambda)} + O(\lambda).$$

Hence

$$Q_{\vec{\iota}}(\lambda) \cdot Bel^{z_0}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda)$$

(by Proposition 4.7(i)) and so, remembering that $Q_{\vec{\iota}(q_1\cdots q_j)}(\lambda)$ has constant term 1, we may see that

$$Bel^{z_0}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda)$$

as required. □

Thus we have shown that S3 holds for $m = 0$, thereby completing the base stage of our inductive process. Our next step is to show how, given we have constructed a z_m which satisfies S1-4, we may construct a new special almost-ent z_{m+1} which also satisfies S1-4 but with $m + 1$ in place of m. The idea is that we will inductively define a finite sequence of intermediate special almost-ents $z_m = z_m^0, z_m^1, z_m^2, \ldots, z_m^k = z_{m+1}$. At each stage, given that z_m^l, say, is the almost-ent constructed up to that point, we focus on a pair of sequences τp and $\tau\neg p$ for some sequence of literals τ and $p \in L$ (starting with the shortest sequences and working up to the longest) and try to obtain an $O(\lambda^{m+2})$ approximation to Bel for both, i.e., we adjust z_m^l, specifically the potentials $z_m^l(\tau \pm p)$, in order to obtain a new special almost-ent z_m^{l+1} which will satisfy

$$Bel^{z_m^{l+1}}(\tau \pm p) - Bel(\tau \pm p) = O(\lambda^{m+2})$$

(although this adjustment will, in some cases, consist of doing nothing at all!) Whilst we do this we must ensure that S1-4 remain satisfied for z_m^{l+1}, and also that we maintain $O(\lambda^{m+2})$ approximations to Bel for all the sequences that have already been considered during this process. We continue this sub-process until every sequence of literals has been looked at, so after the final (k^{th}) stage we have a special almost-ent $z_{m+1} = z_m^k$ which satisfies S1-4 with m replaced by $m + 1$.

Since we begin by setting $z_m^0 = z_m$, we have, for all non-empty sequences of literals σ,

$$z_m^0(\sigma) = 0 \text{ iff } Bel(\sigma) = 0, \tag{4.8}$$

$$z_m^0(\sigma) \neq 0 \text{ implies } z_m^0(\sigma) = O(\lambda^{|\sigma|-1}) \text{ and } z_m^0(\sigma) \neq O(\lambda^{|\sigma|}), \tag{4.9}$$

and

$$Bel^{z_m^0}(\sigma) - Bel(\sigma) = O(\lambda^{m+1}) \tag{4.10}$$

while for all sequences σ and for all $p \in L$ such that $\pm p$ does not appear in σ,

$$\sum\{z_m^0(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma. \tag{4.11}$$

Let us recall that a consequence of (4.8) is that, for all non-empty sequences of literals $q_1 \cdots q_j$,

$$N_{z_m^0}(q_1 \cdots q_j) = N(q_1 \cdots q_j).$$

We shall begin by trying to improve the bound in (4.10) in the cases when $\sigma = p_1$ and $\sigma = \neg p_1$. Our actions depend on whether or not it is the case that $z_m^0(p_1) = 0$ or $z_m^0(\neg p_1) = 0$. (Note we cannot have both, since if so then, by (4.8), we would have $Bel(p_1) = 0 = Bel(\neg p_1)$, contradicting $Bel(p_1) + Bel(\neg p_1) = 1$.)

Case(i): $z_m^0(p_1) = 0$ or $z_m^0(\neg p_1) = 0$.

If one of these is true then we simply set $z_m^1 = z_m^0$, i.e., we leave all potentials unchanged, so, trivially, equations (4.8) - (4.11) are retained for z_m^1 in place of z_m^0 (and of course z_m^1 is still an almost-ent). Furthermore we may state the following.

Lemma 4.23 *If either $z_m^0(p_1) = 0$ or $z_m^0(\neg p_1) = 0$ then $Bel^{z_m^1}(p_1) = Bel(p_1)$ and $Bel^{z_m^1}(\neg p_1) = Bel(\neg p_1)$ (so certainly $Bel^{z_m^1}(\pm p_1) - Bel(\pm p_1) = O(\lambda^{m+2})$).*

Proof. Suppose $z_m^0(p_1) = 0$. Then, from (4.8), we also have $Bel(p_1) = 0$ and so

$$\begin{aligned} Bel^{z_m^1}(p_1) &= Bel^{z_m^0}(p_1) \qquad \text{since } z_m^1 = z_m^0 \\ &= 0 \qquad \text{by Lemma 4.21, since } z_m^0 \text{ satisfies (4.8)} \\ &= Bel(p_1) \end{aligned}$$

and

$$\begin{aligned} Bel^{z_m^1}(\neg p_1) &= 1 - Bel^{z_m^1}(p_1) \\ &= 1 - Bel(p_1) \qquad \text{from the above} \\ &= Bel(\neg p_1) \end{aligned}$$

as required. In the case that $z_m^0(\neg p_1) = 0$ then just repeat the above, transposing p_1 with $\neg p_1$. □

Hence in this case, we have our required $O(\lambda^{m+2})$ approximations (in fact, as it happens, an exact match) to Bel for $Bel^{z_m^1}(\pm p_1)$. Note that at this point we have used, via Lemma 4.21, the assumption on Bel that $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$ for all $\theta, \phi \in SL$. Now for our second, less straightforward, case.

Case(ii): both $z_m^0(p_1) \neq 0$ and $z_m^0(\neg p_1) \neq 0$.

Since the sequence paths for p_1 which are non-zero for z_m^0 are simply all the single element sequences $\vec{\sigma} = \sigma_1$ where σ_1 ends with p_1 and $z_m^0(\sigma_1) > 0$, we have

$$\begin{aligned} Bel^{z_m^0}(p_1) &= \sum_{\substack{\sigma_1 \text{ ends } p_1 \\ z_m^0(\sigma_1)>0}} \frac{z_m^0(\sigma_1)}{\sum\{z_m^0(\tau) \mid \tau \text{ ends } \pm p_1\}} \\ &= \sum_{\substack{\sigma_1 \text{ ends } p_1 \\ z_m^0(\sigma_1)\neq 0}} \frac{z_m^0(\sigma_1)}{A_\emptyset} \qquad \text{from (4.11)} \end{aligned}$$

Pulling out the lead term from this summation (which in this case is simply the term in the sum for which $\sigma_1 = p_1$ and which does occur in the above sum by assumption) and setting

$$T = \sum_{\substack{\sigma_1 \text{ ends } p_1 \\ z_m^0(\sigma_1)>0,\ \sigma_1 \neq p_1}} \frac{z_m^0(\sigma_1)}{A_\emptyset},$$

we get

$$Bel^{z_m^0}(p_1) = \frac{z_m^0(p_1)}{A_\emptyset} + T. \tag{4.12}$$

Now from (4.7) we have

$$A_\emptyset = P_\emptyset(\lambda)$$

where $P_\emptyset(\lambda)$ is a polynomial in λ with constant term 1. Using this together with (4.9) we have, for each sequence σ_1 which ends in p_1,

$$\frac{z_m^0(\sigma_1)}{A_\emptyset} = O(\lambda^{|\sigma_1|-1}). \tag{4.13}$$

Hence, since $\sigma_1 \neq p_1$ implies $|\sigma_1| > 1$ for all σ_1 which end p_1 , we may note that $T = O(\lambda)$. We now define the special almost-ent z^1_m from z^0_m by setting $z^1_m(\tau) = z^0_m(\tau)$ for $\tau \neq \pm p_1$, defining $z^1_m(p_1)$ via the equation

$$Bel(p_1) = \frac{z^1_m(p_1)}{A_\emptyset} + T,$$

i.e.,

$$z^1_m(p_1) = (Bel(p_1) - T) \cdot A_\emptyset, \tag{4.14}$$

and defining $z^1_m(\neg p_1)$ by the equation

$$z^1_m(\neg p_1) = (z^0_m(p_1) + z^0_m(\neg p_1)) - z^1_m(p_1). \tag{4.15}$$

The following result provides the key to enable us to show that z^1_m satisfies the properties we require of it.

Lemma 4.24 *If $z^0_m(p_1) \neq 0$ and $z^0_m(\neg p_1) \neq 0$ then the following are true*
(i). $z^1_m(p_1) = z^0_m(p_1) + O(\lambda^{m+1}) \cdot z^0_m(p_1)$.
(ii). $z^1_m(\neg p_1) = z^0_m(\neg p_1) + O(\lambda^{m+1}) \cdot z^0_m(\neg p_1)$.

Proof. *(i).* From (4.14) we get

$$z^1_m(p_1) = (Bel(p_1) - T) \cdot A_\emptyset$$

whilst from (4.12) we get

$$z^0_m(p_1) = (Bel^{z^0_m}(p_1) - T) \cdot A_\emptyset.$$

Thus we have

$$\begin{aligned} \frac{z^1_m(p_1)}{z^0_m(p_1)} &= \frac{(Bel(p_1) - T) \cdot A_\emptyset}{(Bel^{z^0_m}(p_1) - T) \cdot A_\emptyset} \\ &= \frac{Bel(p_1) - T}{Bel^{z^0_m}(p_1) - T}. \end{aligned} \tag{4.16}$$

Since we are assuming (4.10) we have

$$Bel^{z^0_m}(p_1) = Bel(p_1) + O(\lambda^{m+1})$$

which gives us, from (4.16),

$$\frac{z_m^1(p_1)}{z_m^0(p_1)} = \frac{Bel(p_1) - T}{(Bel(p_1) - T) + O(\lambda^{m+1})}. \tag{4.17}$$

Now, since we are assuming $z_m^0(p_1) \neq 0$, we have $Bel(p_1) \neq 0$ from (4.8). This, together with the fact that $T = O(\lambda)$ means that we must have both $Bel(p_1) - T = O(1)$ and $Bel(p_1) - T \neq O(\lambda)$. Hence, dividing top and bottom of (4.17) by $Bel(p_1) - T$ we get

$$\begin{aligned}
\frac{z_m^1(p_1)}{z_m^0(p_1)} &= \frac{1}{1 + \dfrac{O(\lambda^{m+1})}{Bel(p_1) - T}} \\
&= \frac{1}{1 + O(\lambda^{m+1})} \qquad \text{since } \frac{O(\lambda^{m+1})}{Bel(p_1) - T} = O(\lambda^{m+1}) \\
&= 1 + O(\lambda^{m+1}) \qquad \text{using Proposition 4.6.}
\end{aligned}$$

From this we see

$$z_m^1(p_1) = z_m^0(p_1) + O(\lambda^{m+1}) \cdot z_m^0(p_1)$$

as required.

(ii). We have, from (4.15),

$$\frac{z_m^1(\neg p_1)}{z_m^0(\neg p_1)} = \frac{z_m^0(p_1) + z_m^0(\neg p_1) - z_m^1(p_1)}{z_m^0(\neg p_1)}.$$

Hence, applying *(i)* obtained above to substitute for $z_m^1(p_1)$ we obtain

$$\begin{aligned}
\frac{z_m^1(\neg p_1)}{z_m^0(\neg p_1)} &= \frac{z_m^0(\neg p_1) + O(\lambda^{m+1}) \cdot z_m^0(p_1)}{z_m^0(\neg p_1)} \\
&= 1 + O(\lambda^{m+1}) \cdot \frac{z_m^0(p_1)}{z_m^0(\neg p_1)}.
\end{aligned}$$

Now, from (4.9), we have $z_m^0(p_1) = O(1)$, $z_m^0(\neg p_1) = O(1)$ and $z_m^0(\neg p_1) \neq O(\lambda)$. Hence

$$\frac{z_m^0(p_1)}{z_m^0(\neg p_1)} = O(1)$$

so, just as for p_1, we have

$$\frac{z_m^1(\neg p_1)}{z_m^0(\neg p_1)} = 1 + O(\lambda^{m+1})$$

which gives

$$z_m^1(\neg p_1) = z_m^0(\neg p_1) + O(\lambda^{m+1}) \cdot z_m^0(\neg p_1)$$

as required. □

Note that Lemma 4.24 is also true in the case considered earlier where $z_m^0(p_1) = 0$ or $z_m^0(\neg p_1) = 0$. This is because in this case we have $z_m^1(\pm p_1) = z_m^0(\pm p_1) = z_m^0(\pm p_1) + 0 \cdot z_m^0(\pm p_1)$, and certainly $0 = O(\lambda^{m+1})$! Back to the present case where $z_m^0(\pm p_1) \neq 0$ the above Lemma 4.24 directly tells us that, since, from (4.9), we have

$$z_m^0(\pm p_1) = O(1) \text{ and } z_m^0(\pm p_1) \neq O(\lambda),$$

the same must also be true of z_m^1. Hence, since all other potentials have remained unchanged, equation (4.9) holds for z_m^1 in place of z_m^0. In particular, Lemma 4.24 implies that $z_m^1(\pm p_1) \neq 0$, so, for all sequences σ, we have $z_m^1(\sigma) = 0$ iff $z_m^0(\sigma) = 0$ and so equation (4.8) is also preserved for z_m^1 (hence z_m^1 is a legitimate almost-ent by Lemma 4.21). To show that equation (4.11) is retained, note that, from (4.15), we have

$$z_m^1(p_1) + z_m^1(\neg p_1) = z_m^0(p_1) + z_m^0(\neg p_1)$$

and so, since $z_m^1(\sigma)$ remains unchanged from $z_m^0(\sigma)$ for all sequences of literals other than $\pm p_1$, we may write, for *all* sequences σ and *all* $p \in L$ such that $\pm p$ does not appear in σ,

$$z_m^1(\sigma p) + z_m^1(\sigma \neg p) = z_m^0(\sigma p) + z_m^0(\sigma \neg p).$$

Hence we obtain, for all such σ and p,

$$\begin{aligned} \sum\{z_m^1(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} &= \sum\{z_m^0(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} \\ &= A_\sigma \qquad \text{by (4.11)} \end{aligned} \tag{4.18}$$

as required. We now show that equation (4.10) remains true.

Lemma 4.25 *If $z^0_m(p_1) \neq 0$ and $z^0_m(\neg p_1) \neq 0$ then, for all non-empty sequences of literals $q_1 \cdots q_j$, z^1_m satisfies $Bel^{z^1_m}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda^{m+1})$.*

Proof. Let $q_1 \cdots q_j$ be a non-empty sequence of literals. Since, as we have already shown, equation (4.8) is preserved for z^1_m, i.e., $z^1_m(\tau) = 0$ iff $Bel(\tau) = 0$ for all sequences τ, we have that

$$N_{z^1_m}(q_1 \cdots q_j) = N_{z^0_m}(q_1 \cdots q_j) = N(q_1 \cdots q_j).$$

Hence we have

$$\begin{aligned} Bel^{z^1_m}(q_1 \cdots q_j) &= \\ &= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^1_m(\sigma_i)}{\sum\{z^1_m(\tau) \mid \sigma_{i-1} \subseteq \tau,\ \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t\}} \\ &= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^1_m(\sigma_i)}{A_{\sigma_{i-1}}} \qquad \text{from (4.18).} \end{aligned}$$

If $q_1 \neq \pm p_1$ then for no $\vec{\sigma} \in P(q_1 \cdots q_j)$ do we have $\sigma_i = \pm p_1$ for any $i = 1, \ldots, l(\vec{\sigma})$ (since we always have $|\sigma_i| \geq i$). Hence in this case, since $z^1_m(\tau) = z^0_m(\tau)$ for all $\tau \neq \pm p_1$, we have

$$Bel^{z^1_m}(q_1 \cdots q_j) = Bel^{z^0_m}(q_1 \cdots q_j).$$

(In particular, for all $p_i \in L$ such that $i \neq 1$, we have $Bel^{z^1_m}(\pm p_i) = Bel^{z^0_m}(\pm p_i)$. The significance of this will be explained later.)
And so, given (from (4.10)) $Bel^{z^0_m}(\tau) - Bel(\tau) = O(\lambda^{m+1})$ for all τ, we get

$$Bel^{z^1_m}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda^{m+1})$$

as required. So now suppose we do have $q_1 = \pm p_1$, say $q_1 = p_1$ (we may apply the same reasoning if $q_1 = \neg p_1$, just replace p_1 in what follows by $\neg p_1$). Then we have

$$Bel^{z^1_m}(q_1 \cdots q_j) = \sum_{\substack{\vec{\sigma} \in N(q_1 \cdots q_j) \\ \sigma_1 = p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^1_m(\sigma_i)}{A_{\sigma_{i-1}}} + \sum_{\substack{\vec{\sigma} \in N(q_1 \cdots q_j) \\ \sigma_1 \neq p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^1_m(\sigma_i)}{A_{\sigma_{i-1}}}$$

while similarly

$$Bel^{z_m^0}(q_1\cdots q_j) = \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} + \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1\neq p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}}.$$

But, for all sequence paths $\vec{\sigma}$ for $q_1\cdots q_j$, we cannot have $\sigma_i = \pm p_1$ for any $2 \le i \le l(\vec{\sigma})$ (again since $|\sigma_i| \ge i$). Thus, for all $\vec{\sigma} \in P(q_1\cdots q_j)$ such that $\sigma_1 \neq p_1$, we have $z_m^1(\sigma_i) = z_m^0(\sigma_i)$ for $1 \le i \le l(\vec{\sigma})$. Hence

$$Bel^{z_m^0}(q_1\cdots q_j) = \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} + \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1\neq p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^1(\sigma_i)}{A_{\sigma_{i-1}}}$$

and so

$$\begin{aligned}
Bel^{z_m^1}(q_1\cdots q_j) - Bel^{z_m^0}(q_1\cdots q_j) &= \\
&= \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^1(\sigma_i)}{A_{\sigma_{i-1}}} - \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} \\
&= \frac{z_m^1(p_1)}{A_\emptyset} \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=2}^{l(\vec{\sigma})} \frac{z_m^1(\sigma_i)}{A_{\sigma_{i-1}}} - \frac{z_m^0(p_1)}{A_\emptyset} \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=2}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}}.
\end{aligned}$$

Again, since for each sequence path $\vec{\sigma}$ it cannot be the case that $\sigma_i = \pm p_1$ for any $i = 2, \ldots, l(\vec{\sigma})$, we may write

$$\begin{aligned}
Bel^{z_m^1}(q_1\cdots q_j) - Bel^{z_m^0}(q_1\cdots q_j) &= \frac{z_m^1(p_1) - z_m^0(p_1)}{A_\emptyset} \cdot \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=2}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} \\
&= \frac{z_m^0(p_1)\cdot O(\lambda^{m+1})}{A_\emptyset} \cdot \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=2}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} \\
&\qquad \text{by Lemma 4.24} \\
&= O(\lambda^{m+1}) \cdot \sum_{\substack{\vec{\sigma}\in N(q_1\cdots q_j)\\ \sigma_1=p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}}.
\end{aligned}$$

For each $\vec{\sigma} \in N(q_1\cdots q_j)$, and for each $i = 1, \ldots, l(\vec{\sigma})$ we have, from (4.9),

$$z_m^0(\sigma_i) = O(\lambda^{|\sigma_i|-1})$$

and, from (4.7),

$$A_{\sigma_{i-1}} = \lambda^{|\sigma_{i-1}|} Bel(\sigma_{i-1}) \cdot P_{\sigma_{i-1}}(\lambda)$$

where $P_{\sigma_{i-1}}(\lambda)$ is a polynomial in λ with constant term 1. Hence,

$$\frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} = O(\lambda^{|\sigma_i|-1-|\sigma_{i-1}|})$$

and so, for each $\vec{\sigma} \in N(q_1 \cdots q_j)$ we have

$$\prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} = O(\lambda^{\sum_{i=1}^{l(\vec{\sigma})} |\sigma_i|-1-|\sigma_{i-1}|}) = O(\lambda^{|\sigma_{l(\vec{\sigma})}|-l(\vec{\sigma})}) = O(1).$$

Hence

$$\sum_{\substack{\vec{\sigma} \in N(q_1 \cdots q_j) \\ \sigma_1 = p_1}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^0(\sigma_i)}{A_{\sigma_{i-1}}} = O(1)$$

and so

$$Bel^{z_m^1}(q_1 \cdots q_j) - Bel^{z_m^0}(q_1 \cdots q_j) = O(\lambda^{m+1}).$$

Hence, since we are assuming (4.10) we get the required result. □

Before moving on to the next stage it remains to show that our adjustments have actually been successful in getting a closer approximation to $Bel(p_1)$ and $Bel(\neg p_1)$. In fact, as the next lemma shows, we could not have done any better.

Lemma 4.26 *If $z_m^0(p_1) \neq 0$ and $z_m^0(\neg p_1) \neq 0$ then $Bel^{z_m^1}(p_1) = Bel(p_1)$ and $Bel^{z_m^1}(\neg p_1) = Bel(\neg p_1)$. (Hence it is certainly true that $Bel^{z_m^1}(\pm p_1) - Bel(\pm p_1) = O(\lambda^{m+2})$.)*

Proof. We have

$$\begin{aligned} Bel^{z_m^1}(p_1) &= \sum_{\substack{\sigma_1 \text{ ends } p_1 \\ z_m^1(\sigma_1)>0}} \frac{z_m^1(\sigma_1)}{\sum\{z_m^1(\tau) \mid \tau \text{ ends } \pm p_1\}} \\ &= \sum_{\substack{\sigma_1 \text{ ends } p_1 \\ z_m^1(\sigma_1)>0}} \frac{z_m^1(\sigma_1)}{A_\emptyset} \\ &= \frac{z_m^1(p_1)}{A_\emptyset} + \sum_{\substack{\sigma_1 \text{ ends } p_1 \\ z_m^1(\sigma_1)>0,\ \sigma_1 \neq p_1}} \frac{z_m^1(\sigma_1)}{A_\emptyset} \end{aligned}$$

since, as we have already established, $z_m^1(p_1) \neq 0$. But $z_m^1(\sigma_1) = z_m^0(\sigma_1)$ for each σ_1 such that σ_1 ends p_1 and $\sigma_1 \neq p_1$, and so

$$Bel^{z_m^1}(p_1) = \frac{z_m^1(p_1)}{A_\emptyset} + T$$

which gives, via (4.14),

$$Bel^{z_m^1}(p_1) = Bel(p_1)$$

as required. To see that we also have $Bel^{z_m^1}(\neg p_1) = Bel(\neg p_1)$ it is enough to recall that, for *any* pre-ent (and hence any almost-ent) z,

$$Bel^z(\neg p_1) = 1 - Bel^z(p_1).$$

□

This completes case (ii) ($z_m^0(\pm p_1) \neq 0$). Unfortunately, even though in both cases (i) and (ii) described above, we have $Bel^{z_m^1}(\pm p_1) = Bel(\pm p_1)$, it is not necessarily the case that we will have $Bel^{z_m^i}(\pm p_1) = Bel^{z_m^1}(\pm p_1)$ for all $i \geq 2$ – future adjustments may slightly perturb $Bel^{z_m^i}(\pm p_1)$ from this value – and so the reader should not be lulled into thinking that that is the end of the story for $\pm p_1$.

Summarising up to this point, then, we have obtained, whether $z_m^0(\pm p_1) = 0$ or not, a new special almost-ent z_m^1 from z_m^0 which satisfies, for all non-empty sequences of literals σ,

$$z_m^1(\sigma) = 0 \text{ iff } Bel(\sigma) = 0,$$

$$z_m^1(\sigma) \neq 0 \text{ implies } z_m^1(\sigma) = O(\lambda^{|\sigma|-1}) \text{ and } z_m^1(\sigma) \neq O(\lambda^{|\sigma|})$$

and

$$Bel^{z_m^1}(\sigma) - Bel(\sigma) = O(\lambda^{m+1}),$$

while for all sequences σ and all $p \in L$ such that $\pm p$ does not appear in σ,

$$\sum\{z_m^1(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma.$$

Furthermore we now have

$$Bel^{z_m^1}(\pm p_1) - Bel(\pm p_1) = O(\lambda^{m+2}).$$

We now repeat a similar exercise for the pair $\pm p_2$ to obtain z_m^2 from z_m^1, then repeat again for $\pm p_3$ to obtain z_m^3 from z_m^2 and so on through all the propositional variables in L. Given we are at the stage i where we are focusing on the pair $\pm p_i$, and given that we are in the situation described by case *(ii)* above, i.e., $z_m^{i-1}(\pm p_i) \neq 0$, we see from the proof of Lemma 4.25 that $Bel^{z_m^i}(\pm p_j) = Bel^{z_m^{i-1}}(\pm p_j)$ for $j \neq i$. Hence if we have already established

$$Bel^{z_m^{i-1}}(\pm p_j) - Bel(\pm p_j) = O(\lambda^{m+2})$$

then this approximation will be preserved for z_m^i. Trivially this is also true if we are in the situation described by case *(i)* (since then $z_m^{i-1} = z_m^i$).

Now having obtained $O(\lambda^{m+2})$ approximations to Bel for all the sequences of literals of length one, we then go through all the sequences of length two, followed by all the sequences of length three and so on. At each stage we focus on a pair of sequences $q_1 \cdots q_{j-1} q_j$ and $q_1 \cdots q_{j-1} \neg q_j$ (where q_j is a positive literal, i.e., of the form p as opposed to $\neg p$) and try and obtain a new special almost-ent, z_m^{l+1}, from the one we have formed up to that point, z_m^l, which will give

$$Bel^{z_m^{l+1}}(q_1 \cdots q_{j-1} \pm q_j) - Bel(q_1 \cdots q_{j-1} \pm q_j) = O(\lambda^{m+2}).$$

Whilst doing this we must try and maintain

$$Bel^{z_m^{l+1}}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+2})$$

for all the sequences $r_1 \cdots r_s$ which we have already considered so at the end of this entire process we will have order $O(\lambda^{m+2})$ approximations to Bel for all sequences of literals.

Let us assume, then, that we have reached the stage where we are focusing on the pair of sequences $q_1 \cdots \pm q_j$ and that the special almost-ent z^l_m we have constructed so far satisfies, for all non-empty sequences of literals σ

$$z^l_m(\sigma) = 0 \text{ iff } Bel(\sigma) = 0, \tag{4.19}$$

$$z^l_m(\sigma) \neq 0 \text{ implies } z^l_m(\sigma) = O(\lambda^{|\sigma|-1}) \text{ and } z^l_m(\sigma) \neq O(\lambda^{|\sigma|}), \tag{4.20}$$

and, for all sequences σ and for all $p \in L$ such that $\pm p$ does not appear in σ,

$$\sum\{z^l_m(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma. \tag{4.21}$$

We also assume that, for all non-empty sequences of literals $r_1 \cdots r_s$,

$$Bel^{z^l_m}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+1}) \tag{4.22}$$

and, furthermore, that, for all non-empty sequences $r_1 \cdots r_s$ such that $s < j$,

$$Bel^{z^l_m}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+2}). \tag{4.23}$$

Let us assume that $q_1 \cdots q_j$ and $q_1 \cdots \neg q_j$ are the first sequences of length j that we are considering. Our next step depends on whether or not either $z^l_m(q_1 \cdots q_j) = 0$ or $z^l_m(q_1 \cdots \neg q_j) = 0$ (or possibly both).

Case(i): $z^l_m(q_1 \cdots q_j) = 0$ or $z^l_m(q_1 \cdots \neg q_j) = 0$ (or both).

In this case we set $z^{l+1}_m = z^l_m$, i.e., we leave all potentials unchanged. Hence equations (4.19)-(4.23) are certainly preserved for z^{l+1}_m (and of course z^{l+1}_m is still an almost-ent). Furthermore we have the following:

Lemma 4.27 *If either $z^l_m(q_1 \cdots q_j) = 0$ or $z^l_m(q_1 \cdots \neg q_j) = 0$ (or both) then*

$$Bel^{z^{l+1}_m}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda^{m+2})$$

and

$$Bel^{z^{l+1}_m}(q_1 \cdots \neg q_j) - Bel(q_1 \cdots \neg q_j) = O(\lambda^{m+2}).$$

Proof. First suppose $z_m^l(q_1 \cdots q_j) = 0$. Then, by (4.19), $Bel(q_1 \cdots q_j) = 0$. So

$$\begin{aligned} Bel^{z_m^{l+1}}(q_1 \cdots q_j) &= Bel^{z_m^l}(q_1 \cdots q_j) \qquad \text{since } z_m^{l+1} = z_m^l \\ &= 0 \qquad \text{by Lemma 4.21, since } z_m^l \text{ satsifies (4.19)} \\ &= Bel(q_1 \cdots q_j) \end{aligned}$$

so certainly $Bel^{z_m^{l+1}}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda^{m+2})$ as required.
Also

$$\begin{aligned} Bel^{z_m^{l+1}}(q_1 \cdots \neg q_j) &= Bel^{z_m^l}(q_1 \cdots \neg q_j) \qquad \text{since } z_m^{l+1} = z_m^l \\ &= Bel^{z_m^l}(q_1 \cdots q_{j-1}) - Bel^{z_m^l}(q_1 \cdots q_j) \\ &= Bel^{z_m^l}(q_1 \cdots q_{j-1}) \qquad \text{since } Bel^{z_m^l}(q_1 \cdots q_j) = 0 \end{aligned}$$

while

$$\begin{aligned} Bel(q_1 \cdots \neg q_j) &= Bel(q_1 \cdots q_{j-1}) - Bel(q_1 \cdots q_j) \\ &= Bel(q_1 \cdots q_{j-1}) \qquad \text{since } Bel(q_1 \cdots q_j) = 0. \end{aligned}$$

So

$$\begin{aligned} Bel^{z_m^{l+1}}(q_1 \cdots \neg q_j) - Bel(q_1 \cdots \neg q_j) &= Bel^{z_m^l}(q_1 \cdots q_{j-1}) - Bel(q_1 \cdots q_{j-1}) \\ &= O(\lambda^{m+2}) \qquad \text{as required from (4.23).} \end{aligned}$$

In the case when $z_m^l(q_1 \cdots \neg q_j) = 0$ we may just repeat the above proof, transposing $q_1 \cdots q_j$ everywhere with $q_1 \cdots \neg q_j$, to get the required conclusion. □

Hence in this case we have our required $O(\lambda^{m+2})$ approximations to Bel for $Bel^{z_m^{l+1}}(q_1 \cdots \pm q_j)$. Note we have once again used, via Lemma 4.21, our assumption on Bel that $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$ for all $\theta, \phi \in SL$. We now describe our second case.

Case(ii): both $z_m^l(q_1 \cdots q_j) \neq 0$ and $z_m^l(q_1 \cdots \neg q_j) \neq 0$.

In this case, noting that (4.19) gives us

$$N_{z_m^l}(q_1 \cdots q_j) = N(q_1 \cdots q_j),$$

we have

$$
\begin{aligned}
Bel^{z^l_m}(q_1 \cdots q_j) &= \\
&= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^l_m(\sigma_i)}{\sum\{z^l_m(\tau) \mid \sigma_{i-1} \subseteq \tau \text{ and } \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t\}} \\
&= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^l_m(\sigma_i)}{A_{\sigma_{i-1}}} \qquad \text{from (4.21).}
\end{aligned}
$$

Now $z^l_m(q_1 \cdots q_j) \neq 0$ implies $Bel(q_1 \cdots q_j) \neq 0$ (by (4.19)). Recall the sequence path $\vec{\iota}(q_1 \cdots q_j)$ is that sequence path $\vec{\sigma} \in P(q_1 \cdots q_j)$ for which $\sigma_i = q_1 \cdots q_i$ for $i = 1, \ldots, j$ (so $l(\vec{\iota}(q_1 \cdots q_j)) = j$). Since $\iota_{l(\vec{\iota})} = q_1 \cdots q_j$ we have that $\vec{\iota}(q_1 \cdots q_j) \in N(q_1 \cdots q_j)$ and so the lead term must appear in the above sum. Pulling out this term from the summation we get

$$
Bel^{z^l_m}(q_1 \cdots q_j) = \prod_{i=1}^{j} \frac{z^l_m(q_1 \cdots q_i)}{A_{q_1 \cdots q_{i-1}}} + \sum_{\substack{\vec{\sigma} \in N(q_1 \cdots q_j) \\ \vec{\sigma} \neq \vec{\iota}(q_1 \cdots q_j)}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^l_m(\sigma_i)}{A_{\sigma_{i-1}}}.
$$

Now setting

$$
T_1 = \prod_{i=1}^{j-1} \frac{z^l_m(q_1 \cdots q_i)}{A_{q_1 \cdots q_{i-1}}},
$$

and

$$
T_2 = \sum_{\substack{\vec{\sigma} \in N(q_1 \cdots q_j) \\ \vec{\sigma} \neq \vec{\iota}(q_1 \cdots q_j)}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z^l_m(\sigma_i)}{A_{\sigma_{i-1}}},
$$

we get

$$
Bel^{z^l_m}(q_1 \cdots q_j) = T_1 \cdot \frac{z^l_m(q_1 \cdots q_j)}{A_{q_1 \cdots q_{j-1}}} + T_2. \tag{4.24}
$$

For each $\vec{\sigma} \in N(q_1 \cdots q_j)$ and for each $i = 1, \ldots, l(\vec{\sigma})$ we have, from (4.7),

$$
A_{\sigma_{i-1}} = \lambda^{|\sigma_{i-1}|} Bel(\sigma_{i-1}) \cdot P_{\sigma_{i-1}}(\lambda)
$$

where $P_{\sigma_{i-1}}(\lambda)$ is a polynomial in λ with constant term 1. Using this together with (4.20) we have that, for each $\vec{\sigma} \in N(q_1 \cdots q_j)$

$$
\prod_{i=1}^{l(\vec{\sigma})} \frac{z^l_m(\sigma_i)}{A_{\sigma_{i-1}}} = O(\lambda^{|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma})}). \tag{4.25}
$$

Hence, since (by Proposition 4.19 (iii)) $|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) > 0$ for each $\vec{\sigma} \in N(q_1 \cdots q_j)$ such that $\vec{\sigma} \neq \vec{\iota}(q_1 \cdots q_j)$, we may note that $T_2 = O(\lambda)$. We now define the special almost-ent z_m^{l+1} by setting $z_m^{l+1}(\sigma) = z_m^l(\sigma)$ for all sequences of literals σ such that $\sigma \neq q_1 \cdots q_j$ and $\sigma \neq q_1 \cdots \neg q_j$, defining $z_m^{l+1}(q_1 \cdots q_j)$ via the equation

$$Bel(q_1 \cdots q_j) = T_1 \cdot \frac{z_m^{l+1}(q_1 \cdots q_j)}{A_{q_1 \cdots q_{j-1}}} + T_2, \tag{4.26}$$

and defining $z_m^{l+1}(q_1 \cdots \neg q_j)$ via the equation

$$z_m^{l+1}(q_1 \cdots \neg q_j) = (z_m^l(q_1 \cdots q_j) + z_m^l(q_1 \cdots \neg q_j)) - z_m^{l+1}(q_1 \cdots q_j). \tag{4.27}$$

The following lemma plays a similar role to that of Lemma 4.24 in the case examined earlier when we took $q_1 \cdots q_j = p_1$.

Lemma 4.28 *If $z_m^l(q_1 \cdots q_j) \neq 0$ and $z_m^l(q_1 \cdots \neg q_j) \neq 0$ then the following are true*

(i). $z_m^{l+1}(q_1 \cdots q_j) = z_m^l(q_1 \cdots q_j) + O(\lambda^{m+1}) \cdot z_m^l(q_1 \cdots q_j)$.

(ii). $z_m^{l+1}(q_1 \cdots \neg q_j) = z_m^l(q_1 \cdots \neg q_j) + O(\lambda^{m+1}) \cdot z_m^l(q_1 \cdots \neg q_j)$.

Proof. *(i).* From equations (4.24) and (4.26) we may see that

$$\frac{z_m^{l+1}(q_1 \cdots q_j)}{z_m^l(q_1 \cdots q_j)} = \frac{Bel(q_1 \cdots q_j) - T_2}{Bel^{z_m^l}(q_1 \cdots q_j) - T_2}$$

and hence, since we are assuming (4.22), we obtain

$$\frac{z_m^{l+1}(q_1 \cdots q_j)}{z_m^l(q_1 \cdots q_j)} = \frac{Bel(q_1 \cdots q_j) - T_2}{(Bel(q_1 \cdots q_j) - T_2) + O(\lambda^{m+1})} = 1 + O(\lambda^{m+1}). \tag{4.28}$$

From this we get

$$z_m^{l+1}(q_1 \cdots q_j) = z_m^l(q_1 \cdots q_j) + O(\lambda^{m+1}) \cdot z_m^l(q_1 \cdots q_j)$$

as required.

(ii). We have, from (4.27),

$$\frac{z_m^{l+1}(q_1 \cdots \neg q_j)}{z_m^l(q_1 \cdots \neg q_j)} = \frac{z_m^l(q_1 \cdots q_j) + z_m^l(q_1 \cdots \neg q_j) - z_m^{l+1}(q_1 \cdots q_j)}{z_m^l(q_1 \cdots \neg q_j)}.$$

Hence, applying part *(i)* proved above to substitute for $z_m^{l+1}(q_1 \cdots q_j)$ we obtain

$$\begin{aligned}\frac{z_m^{l+1}(q_1 \cdots \neg q_j)}{z_m^l(q_1 \cdots \neg q_j)} &= \frac{z_m^l(q_1 \cdots \neg q_j) + O(\lambda^{m+1}) \cdot z_m^l(q_1 \cdots q_j)}{z_m^l(q_1 \cdots \neg q_j)} \\ &= 1 + O(\lambda^{m+1}) \cdot \frac{z_m^l(q_1 \cdots q_j)}{z_m^l(q_1 \cdots \neg q_j)}.\end{aligned}$$

Now, from (4.20), we have $z_m^l(q_1 \cdots q_j) = O(\lambda^{j-1}), z_m^l(q_1 \cdots \neg q_j) = O(\lambda^{j-1})$ and $z_m^l(q_1 \cdots \neg q_j) \neq O(\lambda^j)$. Hence

$$\frac{z_m^l(q_1 \cdots q_j)}{z_m^l(q_1 \cdots \neg q_j)} = O(1)$$

so, just as for $q_1 \cdots q_j$, we have

$$\frac{z_m^{l+1}(q_1 \cdots \neg q_j)}{z_m^l(q_1 \cdots \neg q_j)} = 1 + O(\lambda^{m+1})$$

which gives the result. □

Note again that Lemma 4.28 holds also if $z_m^l(q_1 \cdots q_j) = 0$ or $z_m^l(q_1 \cdots \neg q_j) = 0$. Back to the present case of $z_m^l(q_1 \cdots \pm q_j) \neq 0$ we see that, since, by (4.20),

$$z_m^l(q_1 \cdots \pm q_j) = O(\lambda^{j-1}) \text{ and } z_m^l(q_1 \cdots \pm q_j) \neq O(\lambda^j),$$

Lemma 4.28 tells us that the same must be true of z_m^{l+1}. Hence, since all other potentials have remained unchanged, equation (4.20) holds for z_m^{l+1} in place of z_m^l. In particular, Lemma 4.28 implies that $z_m^{l+1}(q_1 \cdots \pm q_j) \neq 0$ so, for all sequences σ, we have $z_m^{l+1}(\sigma) = 0$ iff $z_m^l(\sigma) = 0$ and so equation (4.19) is also preserved for z_m^{l+1} (hence z_m^{l+1} is a legitimate almost-ent by Lemma 4.21). To show (4.21) is retained, note that, from (4.27), we have

$$z_m^{l+1}(q_1 \cdots q_j) + z_m^{l+1}(q_1 \cdots \neg q_j) = z_m^l(q_1 \cdots q_j) + z_m^l(q_1 \cdots \neg q_j)$$

which ensures that, for all sequences of literals σ and all $p \in L$ such that $\pm p$ does not appear in σ,

$$z_m^{l+1}(\sigma p) + z_m^{l+1}(\sigma \neg p) = z_m^l(\sigma p) + z_m^l(\sigma \neg p)$$

(since the potential given to scenarios other than $s(q_1 \cdots \pm q_j)$ has remained unchanged). Hence, for all sequences σ and all $p \in L$ such that $\pm p$ does not appear in σ,

$$\begin{aligned} \sum\{z_m^{l+1}(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} &= \sum\{z_m^{l}(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} \\ &= A_\sigma \qquad \text{by (4.21).} \qquad (4.29) \end{aligned}$$

We now show that equations (4.22) and (4.23) remain true for our new almost-ent.

Lemma 4.29 *If $z_m^l(q_1 \cdots q_j) \neq 0$ and $z_m^l(q_1 \cdots \neg q_j) \neq 0$ then, for all non-empty sequences of literals sequences $r_1 \cdots r_s$,*

$$Bel^{z_m^{l+1}}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+1}).$$

In addition, for all sequences $r_1 \cdots r_s$ such that $s < j$,

$$Bel^{z_m^{l+1}}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+2}).$$

Proof. Let $r_1 \cdots r_s$ be a non-empty sequence of literals. The fact that equation (4.19) is satisfied for z_m^{l+1} means that we have

$$N_{z_m^{l+1}}(r_1 \cdots r_s) = N_{z_m^l}(r_1 \cdots r_s) = N(r_1 \cdots r_s).$$

Hence we may write

$$\begin{aligned} & Bel^{z_m^{l+1}}(r_1 \cdots r_s) = \\ & = \sum_{\vec{\sigma} \in N(r_1 \cdots r_s)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{\sum\{z_m^{l+1}(\tau) \mid \sigma_{i-1} \subseteq \tau,\ \tau \text{ ends } \pm r_t \text{ where } \sigma_i \text{ ends } r_t\}} \\ & = \sum_{\vec{\sigma} \in N(r_1 \cdots r_s)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{A_{\sigma_{i-1}}} \qquad \text{from (4.29).} \end{aligned}$$

Hence if for no $\vec{\sigma} \in N(r_1 \cdots r_s)$ do we have $\sigma_i = q_1 \cdots \pm q_j$ for any $i = 1, \ldots, l(\vec{\sigma})$ then we must have

$$Bel^{z_m^{l+1}}(r_1 \cdots r_s) = Bel^{z_m^l}(r_1 \cdots r_s).$$

In this case we derive the required conclusion from either equation (4.22) or equation (4.23), according to whether $s < j$ or not. Suppose, on the other hand, that either the sequence $q_1 \cdots q_j$ or the sequence $q_1 \cdots \neg q_j$ appears in at least one sequence path $\vec{\sigma} \in N(r_1 \cdots r_s)$. It cannot be the case that both occur, since if $q_1 \cdots q_j$ appears then it must be that $q_j = r_{i_1}$ for some $1 \leq i_1 \leq s$, while if $q_1 \cdots \neg q_j$ also appears then we have $\neg q_j = r_{i_2}$ for some $1 \leq i_2 \leq s$. Hence both q_j and $\neg q_j$ appear in $r_1 \cdots r_s$ giving a contradiction. Let us assume that it is $q_1 \cdots q_j$ which appears (the same reasoning will apply if we assume it is $q_1 \cdots \neg q_j$ which appears — just replace $q_1 \cdots q_j$ everywhere in what follows by $q_1 \cdots \neg q_j$). Then we may write

$$\begin{aligned} Bel^{z_m^{l+1}}(r_1 \cdots r_s) &= \sum_{\substack{\vec{\sigma} \in N(r_1 \cdots r_s) \\ q_1 \cdots q_j \text{ appears in } \vec{\sigma}}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{A_{\sigma_{i-1}}} + \\ &+ \sum_{\substack{\vec{\sigma} \in N(r_1 \cdots r_s) \\ q_1 \cdots q_j \text{ doesn't appear in } \vec{\sigma}}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{A_{\sigma_{i-1}}} \end{aligned}$$

while similarly

$$\begin{aligned} Bel^{z_m^{l}}(r_1 \cdots r_s) &= \sum_{\substack{\vec{\sigma} \in N(r_1 \cdots r_s) \\ q_1 \cdots q_j \text{ appears in } \vec{\sigma}}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l}(\sigma_i)}{A_{\sigma_{i-1}}} + \\ &+ \sum_{\substack{\vec{\sigma} \in N(r_1 \cdots r_s) \\ q_1 \cdots q_j \text{ doesn't appear in } \vec{\sigma}}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l}(\sigma_i)}{A_{\sigma_{i-1}}}. \end{aligned}$$

For each $\vec{\sigma} \in N(r_1 \cdots r_s)$ for which $q_1 \cdots q_j$ does not appear in $\vec{\sigma}$ we have $z_m^{l+1}(\sigma_i) = z_m^l(\sigma_i)$ for all $i = 1, \ldots, l(\vec{\sigma})$. Hence we have

$$\begin{aligned} & Bel^{z_m^{l+1}}(r_1 \cdots r_s) - Bel^{z_m^{l}}(r_1 \cdots r_s) = \\ &= \sum_{\substack{\vec{\sigma} \in N(r_1 \cdots r_s) \\ q_1 \cdots q_j \text{ appears in } \vec{\sigma}}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{A_{\sigma_{i-1}}} - \sum_{\substack{\vec{\sigma} \in N(r_1 \cdots r_s) \\ q_1 \cdots q_j \text{ appears in } \vec{\sigma}}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l}(\sigma_i)}{A_{\sigma_{i-1}}} \end{aligned}$$

$$= \sum_{\substack{\vec{\sigma}\in N(r_1\cdots r_s)\\ q_1\cdots q_j \text{ appears in } \vec{\sigma}}} \left\{ \frac{\prod_{i=1}^{l(\vec{\sigma})} z_m^{l+1}(\sigma_i) - \prod_{i=1}^{l(\vec{\sigma})} z_m^{l}(\sigma_i)}{\prod_{i=1}^{l(\vec{\sigma})} A_{\sigma_{i-1}}} \right\}.$$

Let $\vec{\sigma} \in N(r_1 \cdots r_s)$ be such that $q_1 \cdots q_j$ appears in $\vec{\sigma}$, say $q_1 \cdots q_j = \sigma_t$ where $1 \le t \le l(\vec{\sigma})$. Then for such a $\vec{\sigma}$ we have

$$\frac{\prod_{i=1}^{l(\vec{\sigma})} z_m^{l+1}(\sigma_i) - \prod_{i=1}^{l(\vec{\sigma})} z_m^{l}(\sigma_i)}{\prod_{i=1}^{l(\vec{\sigma})} A_{\sigma_{i-1}}} =$$

$$= \frac{(z_m^{l+1}(q_1\cdots q_j) - z_m^{l}(q_1\cdots q_j)) \cdot \prod_{i=1}^{t-1} z_m^l(\sigma_i) \cdot \prod_{i=t+1}^{l(\vec{\sigma})} z_m^l(\sigma_i)}{\prod_{i=1}^{l(\vec{\sigma})} A_{\sigma_{i-1}}}$$

since $z_m^{l+1}(\sigma) = z_m^l(\sigma)$ for all sequences $\sigma \neq q_1 \cdots \pm q_j$ and clearly $\sigma_i \neq q_1 \cdots \pm q_j$ for $i = 1, \ldots, t-1, t+1, \ldots, l(\vec{\sigma})$. Hence, since $z_m^{l+1}(q_1 \cdots q_j) - z_m^l(q_1 \cdots q_j) = O(\lambda^{m+1}) \cdot z_m^l(q_1 \cdots q_j)$ from Lemma 4.28, we get

$$Bel^{z_m^{l+1}}(r_1\cdots r_s) - Bel^{z_m^{l}}(r_1\cdots r_s) = \sum_{\substack{\vec{\sigma}\in N(r_1\cdots r_s)\\ q_1\cdots q_j \text{ appears in } \vec{\sigma}}} O(\lambda^{m+1}) \cdot \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^l(\sigma_i)}{A_{\sigma_{i-1}}}. \quad (4.30)$$

Now for each $\vec{\sigma} \in N(r_1 \cdots r_s)$ we have, from (4.25),

$$\prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^l(\sigma_i)}{A_{\sigma_{i-1}}} = O(\lambda^{|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma})}).$$

Hence we may see from (4.30) that we certainly have

$$Bel^{z_m^{l+1}}(r_1\cdots r_s) - Bel^{z_m^{l}}(r_1\cdots r_s) = O(\lambda^{m+1}). \quad (4.31)$$

Indeed if $q_1 \cdots q_j$ does not appear in $\vec{\iota}(r_1 \cdots r_s)$ (the only sequence path $\vec{\sigma} \in P(r_1 \cdots r_s)$ for which $|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) = 0$) then for each $\vec{\sigma} \in N(r_1 \cdots r_s)$ such that

$q_1 \cdots q_j$ appears in $\vec{\sigma}$ we have $\vec{\sigma} \neq \vec{\iota}(r_1 \cdots r_s)$ and consequently for each such $\vec{\sigma}$ we must have

$$\prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^l(\sigma_i)}{A_{\sigma_{i-1}}} = O(\lambda^y)$$

for some $y \geq 1$ (which depends on $\vec{\sigma}$) Hence in this case we may strengthen (4.31) to

$$Bel^{z_m^{l+1}}(r_1 \cdots r_s) - Bel^{z_m^l}(r_1 \cdots r_s) = O(\lambda^{m+2}). \tag{4.32}$$

If $r_1 \cdots r_s$ is a sequence of literals such that $s < j$ then it cannot be the case that $q_1 \cdots q_j$ appears in $\vec{\iota}(r_1 \cdots r_s)$ (since $|\iota_i| \leq s$ for all ι_i in $\vec{\iota}(r_1 \cdots r_s)$). Hence the above equation (4.32) holds in particular for such sequences. Combining all this with the equations (4.22) and (4.23) we may now see that, for all sequences of literals $r_1 \cdots r_s$,

$$Bel^{z_m^{l+1}}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+1}),$$

while for all sequences $r_1 \cdots r_s$ such that $s < j$ we maintain

$$Bel^{z_m^{l+1}}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+2})$$

as required. □

Hence we have now shown that, in the case where $z_m^l(q_1 \cdots \pm q_j) \neq 0$, all the equations (4.19)-(4.23) remain true for our newly defined almost-ent z_m^{l+1}. It remains to show that the changes we have made to $z_m^l(q_1 \cdots \pm q_j)$ have had the desired effect of bringing $Bel^{z_m^{l+1}}(q_1 \cdots \pm q_j)$ closer to $Bel(q_1 \cdots \pm q_j)$.

Lemma 4.30 *If* $z_m^l(q_1 \cdots q_j) \neq 0$ *and* $z_m^l(q_1 \cdots \neg q_j) \neq 0$ *then*

$$Bel^{z_m^{l+1}}(q_1 \cdots \pm q_j) - Bel(q_1 \cdots \pm q_j) = O(\lambda^{m+2}).$$

Proof. We will in fact show that $Bel^{z_m^{l+1}}(q_1 \cdots q_j) = Bel(q_1 \cdots q_j)$. As in the earlier situation it should not be assumed that this automatically means

$Bel^{z_m^{l+i}}(q_1 \cdots q_j) = Bel(q_1 \cdots q_j)$ for all $i \geq 2$.

We have

$$
\begin{aligned}
Bel^{z_m^{l+1}}(q_1 \cdots q_j) &= \\
&= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{\sum\{z_m^{l+1}(\tau) \mid \sigma_{i-1} \subseteq \tau \text{ and } \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t\}} \\
&= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{A_{\sigma_{i-1}}} \\
&= \prod_{i=1}^{j} \frac{z_m^{l+1}(q_1 \cdots q_i)}{A_{q_1 \cdots q_{i-1}}} + \sum_{\substack{\vec{\sigma} \in N(q_1 \cdots q_j) \\ \vec{\sigma} \neq \vec{\iota}(q_1 \cdots q_j)}} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m^{l+1}(\sigma_i)}{A_{\sigma_{i-1}}}
\end{aligned}
$$

and so, since $z_m^{l+1}(\sigma) = z_m^l(\sigma)$ for all sequences $\sigma \neq q_1 \cdots \pm q_j$ and since, as is easily seen, $q_1 \cdots q_j$ appears only in the lead term (while $q_1 \cdots \neg q_j$ does not appear at all), we have

$$Bel^{z_m^{l+1}}(q_1 \cdots q_j) \;=\; T_1 \cdot \frac{z_m^{l+1}(q_1 \cdots q_j)}{A_{q_1 \cdots q_{j-1}}} + T_2.$$

Hence, from (4.26),

$$Bel^{z_m^{l+1}}(q_1 \cdots q_j) = Bel(q_1 \cdots q_j).$$

as required.

To show the other part of the lemma recall that for any pre-ent (and hence any almost-ent) z we have

$$Bel^z(q_1 \cdots \neg q_j) = Bel^z(q_1 \cdots q_{j-1}) - Bel^z(q_1 \cdots q_j).$$

Hence we have

$$
\begin{aligned}
Bel^{z_m^{l+1}}(q_1 \cdots \neg q_j) - Bel(q_1 \cdots \neg q_j) &= (Bel^{z_m^{l+1}}(q_1 \cdots q_{j-1}) - Bel^{z_m^{l+1}}(q_1 \cdots q_j)) - \\
&\quad -(Bel(q_1 \cdots q_{j-1}) - Bel(q_1 \cdots q_j)) \\
&= Bel^{z_m^{l+1}}(q_1 \cdots q_{j-1}) - Bel(q_1 \cdots q_{j-1}) \\
&\quad \text{by the part proved above} \\
&= O(\lambda^{m+2}) \qquad \text{by Lemma 4.29}
\end{aligned}
$$

as required. □

This completes our discussion of case (ii) ($z^l_m(q_1 \cdots \pm q_j) \neq 0$).

To summarise up to this point, then, we have created a new special almost-ent z^{l+1}_m from z^l_m by changing (or not, as the case may be) the potentials $z^l_m(q_1 \cdots \pm q_j)$. Our new almost-ent satisfies, for all non-empty sequences of literals σ,

$$z^{l+1}_m(\sigma) = 0 \text{ iff } Bel(\sigma) = 0,$$

$$z^{l+1}_m(\sigma) \neq 0 \text{ implies } z^{l+1}_m(\sigma) = O(\lambda^{|\sigma|-1}) \text{ and } z^{l+1}_m(\sigma) \neq O(\lambda^{|\sigma|})$$

and

$$Bel^{z^{l+1}_m}(\sigma) - Bel(\sigma) = O(\lambda^{m+1}),$$

while, for all sequences σ and all $p \in L$ such that p does not appear in σ,

$$\sum \{z^{l+1}_m(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma.$$

Furthermore for all non-empty sequences σ such that $|\sigma| < j$ we have

$$Bel^{z^{l+1}_m}(\sigma) - Bel(\sigma) = O(\lambda^{m+2})$$

and now also

$$Bel^{z^{l+1}_m}(q_1 \cdots \pm q_j) - Bel(q_1 \cdots \pm q_j) = O(\lambda^{m+2}).$$

We assumed in the above that $q_1 \cdots \pm q_j$ were the first sequences of literals of length j to be considered. We may now go through each of the other pairs of sequences $r_1 \cdots \pm r_j$ in similar fashion. Note that, as well as retaining $O(\lambda^{m+2})$ approximations to Bel for all sequences of length less than j, we will retain $O(\lambda^{m+2})$ approximations to Bel for all the sequences of length j which we have already considered. The reason for this can be found in the proof of Lemma 4.29 where we showed that if, for any sequence $r_1 \cdots r_s$, we had

$$Bel^{z^l_m}(r_1 \cdots r_s) - Bel(r_1 \cdots r_s) = O(\lambda^{m+2})$$

then this will remain true when we change l to $l+1$, provided that neither $q_1 \cdots q_j$ nor $q_1 \cdots \neg q_j$ appear in the path $\vec{\iota}(r_1 \cdots r_s)$, and, given a sequence $r_1 \cdots r_j$, the only sequence of length j which appears in $\vec{\iota}(r_1 \cdots r_j)$ is $r_1 \cdots r_j$ itself.

In this way, then, we obtain $O(\lambda^{m+2})$ approximations to Bel for *every* sequence of literals to finally create the special almost-ent z_{m+1} from z_m in such a way that S1-4 hold when m is replaced by $m+1$. Hence, by induction, S1-4 are true for all $m \geq 0$. In other words, by way of a reminder, the following are satisfied for all $m = 0, 1, \ldots$

- S1. For all non-empty sequences of literals σ, $z_m(\sigma) = 0$ iff $Bel(\sigma) = 0$.
- S2. For all non-empty sequences of literals σ, $z_m(\sigma) \neq 0$ implies $z_m(\sigma) = O(\lambda^{|\sigma|-1})$ and $z_m(\sigma) \neq O(\lambda^{|\sigma|})$.
- S3. For all non-empty sequences of literals σ, $Bel^{z_m}(\sigma) - Bel(\sigma) = O(\lambda^{m+1})$,
- S4. For all (possibly empty) sequences of literals σ and all $p \in L$ such that $\pm p$ does not appear in σ,

$$\sum \{z_m(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma.$$

Note that S1 gives us, for all non-empty sequences of literals σ and all $m = 0, 1, \ldots,$

$$N_{z_m}(\sigma) = N(\sigma).$$

The preceding lemmas have also shown that, for each $m = 0, 1, \ldots$ and each sequence of literals σ,

$$z_{m+1}(\sigma) = z_m(\sigma) + O(\lambda^{m+1}) \cdot z_m(\sigma) \tag{4.33}$$

The condition S3 says that, for each sequence of literals σ, the values $Bel^{z_m}(\sigma)$ are getting closer and closer to $Bel(\sigma)$ as m gets bigger. Indeed the almost-ent

z_∞ we now construct to prove Theorem 4.11 may be thought of, in a strong sense, to be the limit of this sequence. In view of S1 and S2 we may write

$$z_m(\sigma) = \lambda^{|\sigma|-1} \sum_{i=0}^{\infty} z_m^{(i)}(\sigma)\lambda^i$$

where we use $z_m^{(i)}(\sigma)$ simply to denote the i'th coefficient in the above series and where $z_m^{(0)}(\sigma) = 0$ implies $z_m^{(i)}(\sigma) = 0$ for all $i \geq 0$. For example, for $m = 0$ we defined $z_0(\sigma) = \lambda^{|\sigma|-1} Bel(\sigma)$. Hence $z_0^{(0)}(\sigma) = Bel(\sigma)$ and $z_0^{(i)}(\sigma) = 0$ for all $i = 1, 2, \ldots$

Lemma 4.31 *Let σ be a non-empty sequence of literals. Then, for each $m = 0, 1, \ldots$, given that we may expand $z_m(\sigma)$ as above, we have $z_m^{(i)}(\sigma) = z_i^{(i)}(\sigma)$ for $i = 0, 1, \ldots, m$.*

Proof. We use induction on m. For $m = 0$ the result is clear. Let $k \geq 0$ and suppose for inductive hypothesis that the result is true for $m = k$, i.e., that

$$z_k^{(i)}(\sigma) = z_i^{(i)}(\sigma) \text{ for } i = 0, 1, \ldots, k.$$

We need to show that the result holds for $m = k + 1$. From (4.33) we have

$$z_{k+1}(\sigma) = z_k(\sigma) + O(\lambda^{k+1}) \cdot z_k(\sigma)$$

and so

$$\begin{aligned} z_{k+1}(\sigma) &= \lambda^{|\sigma|-1} \left\{ \sum_{i=0}^{\infty} z_k^{(i)}(\sigma)\lambda^i + O(\lambda^{k+1}) \sum_{i=0}^{\infty} z_k^{(i)}(\sigma)\lambda^i \right\} \\ &= \lambda^{|\sigma|-1} \left\{ \sum_{i=0}^{k} z_k^{(i)}(\sigma)\lambda^i + O(\lambda^{k+1}) \right\}. \end{aligned}$$

Hence we may see that $z_{k+1}^{(i)}(\sigma) = z_k^{(i)}(\sigma)$ for $i = 0, 1, \ldots, k$ and so, by inductive hypothesis, that $z_{k+1}^{(i)}(\sigma) = z_i^{(i)}(\sigma)$ for $i = 0, 1, \ldots, k$. Since this is clearly also true for $i = k + 1$ we may see that the lemma is true for $m = k + 1$, thereby completing the proof. □

The above lemma says, then, that for each non-empty sequence of literals σ and for each $m = 0, 1, \ldots$,

$$z_m(\sigma) = \lambda^{|\sigma|-1} \left\{ \sum_{i=0}^{m} z_i^{(i)}(\sigma)\lambda^i + \sum_{i=m+1}^{\infty} z_m^{(i)}(\sigma)\lambda^i \right\}.$$

We now define the special almost-ent z_∞ by setting, for each σ a non-empty sequence of literals,

$$z_\infty(\sigma) = \lambda^{|\sigma|-1} \sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i.$$

We now set out to show that z_∞ is the almost-ent required to prove Theorem 4.11. The next three lemmas describe a series of important properties of z_∞. The first of these suffices (by Lemma 4.21) to assure us that z_∞ is indeed an almost-ent.

Lemma 4.32 *Let σ be a non-empty sequence of literals. Then we have $z_\infty(\sigma) = 0$ iff $Bel(\sigma) = 0$.*

Proof. Suppose $z_\infty(\sigma) = 0$. Then $z_i^{(i)}(\sigma) = 0$ for all $i \geq 0$, in particular $z_0^{(0)}(\sigma) = 0$. As we remarked above, $z_0^{(0)}(\sigma) = Bel(\sigma)$ and so $Bel(\sigma) = 0$. Conversely suppose $z_\infty(\sigma) \neq 0$. Then it must be that $z_i^{(i)}(\sigma) \neq 0$ for some $i \geq 0$. Hence $z_i(\sigma) \neq 0$ and so $Bel(\sigma) \neq 0$ from S1 as required. □

Corollary 4.33 *Let σ be a non-empty sequence of literals. Then, for each $m = 0, 1, \ldots$, $z_\infty(\sigma) = 0$ iff $z_m(\sigma) = 0$ (and so $N_{z_\infty}(\sigma) = N_{z_m}(\sigma) = N(\sigma)$).*

Proof. Immediate from Lemma 4.32 and S1. □

Lemma 4.34 *For each $m = 0, 1, \ldots$ and for each non-empty sequence of literals σ,*

$$z_\infty(\sigma) = z_m(\sigma) + O(\lambda^{m+1}) \cdot z_m(\sigma).$$

Proof. First note that if $z_m(\sigma) = 0$ then, by Corollary 4.33, we also have $z_\infty(\sigma) = 0$ and so the result is true. So let us assume $z_m(\sigma) \neq 0$. By Lemma 4.31, we have $z_m^{(i)}(\sigma) = z_i^{(i)}(\sigma)$ for $i \leq m$, hence

$$\begin{aligned}
z_m(\sigma) &= \lambda^{|\sigma|-1} \sum_{i=0}^{\infty} z_m^{(i)}(\sigma)\lambda^i \\
&= \lambda^{|\sigma|-1} \left\{ \sum_{i=0}^{m} z_i^{(i)}(\sigma)\lambda^i + \sum_{i=m+1}^{\infty} z_m^{(i)}(\sigma)\lambda^i \right\} \\
&= \lambda^{|\sigma|-1} \left\{ \sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i + \sum_{i=m+1}^{\infty} z_m^{(i)}(\sigma)\lambda^i - \sum_{i=m+1}^{\infty} z_i^{(i)}(\sigma)\lambda^i \right\} \\
&= \lambda^{|\sigma|-1} \left\{ \sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i + \sum_{i=m+1}^{\infty} (z_m^{(i)}(\sigma) - z_i^{(i)}(\sigma))\lambda^i \right\} \\
&= \lambda^{|\sigma|-1} \left\{ \left(\sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i \right) + O(\lambda^{m+1}) \right\}
\end{aligned}$$

and so

$$\begin{aligned}
\frac{z_\infty(\sigma)}{z_m(\sigma)} &= \frac{\lambda^{|\sigma|-1} \sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i}{\lambda^{|\sigma|-1} \left\{ \left(\sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i \right) + O(\lambda^{m+1}) \right\}} \\
&= \frac{\sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i}{\left(\sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i \right) + O(\lambda^{m+1})}.
\end{aligned}$$

Now, since we are assuming $z_m(\sigma) \neq 0$, we have $z_0^{(0)}(\sigma) = Bel(\sigma) \neq 0$ from S1 and so

$$\sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i = O(1) \text{ and } \sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i \neq O(\lambda).$$

Hence we may write

$$\begin{aligned}
\frac{z_\infty(\sigma)}{z_m(\sigma)} &= \frac{1}{1+B} \qquad \text{where } B = \frac{O(\lambda^{m+1})}{\sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i}. \\
&= \frac{1}{1+O(\lambda^{m+1})} \\
&= 1 + O(\lambda^{m+1})
\end{aligned}$$

which gives the result. □

Lemma 4.35 *For all sequences of literals σ and all $p \in L$ such that $\pm p$ does not appear in σ,*

$$\sum\{z_\infty(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma.$$

Proof. We will show that, for each $m = 0, 1, \ldots$,

$$\sum\{z_\infty(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} = A_\sigma + O(\lambda^{m+1})$$

which will suffice by Corollary 4.5. So let m be fixed. From Lemma 4.34 we have, for each sequence τ such that $\sigma \subseteq \tau$ and τ ends $\pm p$,

$$z_\infty(\tau) = z_m(\tau) + O(\lambda^{m+1}) \cdot z_m(\tau).$$

We have $z_m(\tau) = O(\lambda^{|\tau|-1})$ and so, since obviously $|\tau| \geq 1$ we have have $z_m(\tau) = O(1)$. Hence we certainly have

$$z_\infty(\tau) = z_m(\tau) + O(\lambda^{m+1}).$$

To be more accurate we have

$$z_\infty(\tau) = z_m(\tau) + P^\tau(\lambda)$$

where $P^\tau(\lambda)$ is a power series in λ such that $P^\tau(\lambda) = O(\lambda^{m+1})$. Hence

$$\begin{aligned}
\sum\{z_\infty(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} &= \\
&= \sum\{(z_m(\tau) + P^\tau(\lambda)) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} \\
&= \sum\{z_m(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} + \\
&\quad + \sum\{P^\tau(\lambda) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} \\
&= \sum\{z_m(\tau) \mid \sigma \subseteq \tau \text{ and } \tau \text{ ends } \pm p\} + \\
&\quad + O(\lambda^{m+1}) \\
&= A_\sigma + O(\lambda^{m+1})
\end{aligned}$$

from S4 as required. □

We are now finally in a position to prove Theorem 4.11.

Theorem 4.11 Given a language $L = \{p_1, \ldots, p_n\}$ $(n > 1)$, if the function $Bel : SL \to [0,1]$ is given by a standard pre-ent over L and if, for all $\theta, \phi \in SL$, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$, then there exists an **almost-ent** z (over a larger language than L) such that, for all $\theta \in SL$, $Bel^z(\theta) = Bel(\theta)$. The potentials of z are elements in $[0, \infty)^{(\lambda)}$.

Proof. Given Bel we define the almost-ent z_∞ as in the preceding construction. We will show that, for all non-empty sequences of literals $q_1 \cdots q_j$,

$$Bel^{z_\infty}(q_1 \cdots q_j) = Bel(q_1 \cdots q_j).$$

This suffices to show the conclusion of Theorem 4.11, namely that, for all $\theta \in L$, $Bel^{z_\infty}(\theta) = Bel(\theta)$. Before that we will show that, for each $m = 0, 1, 2, \ldots$,

$$Bel^{z_\infty}(q_1 \cdots q_j) - Bel^{z_m}(q_1 \cdots q_j) = O(\lambda^{m+1}).$$

This will suffice since we already have, by S3, that

$$Bel^{z_m}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda^{m+1})$$

and so we will have that, for each $m = 0, 1, 2, \ldots$,

$$Bel^{z_\infty}(q_1 \cdots q_j) - Bel(q_1 \cdots q_j) = O(\lambda^{m+1})$$

and so we must have (using Corollary 4.5) $Bel^{z_\infty}(q_1 \cdots q_j) = Bel(q_1 \cdots q_j)$ as required.

From Corollary 4.33 we have that $N_{z_\infty}(q_1 \cdots q_j) = N(q_1 \cdots q_j)$. Hence we may write

$$Bel^{z_\infty}(q_1 \cdots q_j) =$$

$$= \sum_{\vec{\sigma}\in N(q_1\cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_\infty(\sigma_i)}{\sum\{z_\infty(\tau) \mid \sigma_{i-1}\subseteq\tau \text{ and } \tau \text{ ends } \pm q_t \text{ where } \sigma_i \text{ ends } q_t\}}$$

$$= \sum_{\vec{\sigma}\in N(q_1\cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_\infty(\sigma_i)}{A_{\sigma_{i-1}}} \qquad \text{by Lemma 4.35}$$

while similarly (since also $N_{z_m}(q_1\cdots q_j) = N(q_1\cdots q_j)$)

$$Bel^{z_m}(q_1\cdots q_j) = \sum_{\vec{\sigma}\in N(q_1\cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m(\sigma_i)}{A_{\sigma_{i-1}}}.$$

Hence we have

$$Bel^{z_\infty}(q_1\cdots q_j) - Bel^{z_m}(q_1\cdots q_j) = \sum_{\vec{\sigma}\in N(q_1\cdots q_j)} \left\{ \frac{\prod_{i=1}^{l(\vec{\sigma})} z_\infty(\sigma_i) - \prod_{i=1}^{l(\vec{\sigma})} z_m(\sigma_i)}{\prod_{i=1}^{l(\vec{\sigma})} A_{\sigma_{i-1}}} \right\}.$$

Now for each path $\vec{\sigma}\in N(q_1\cdots q_j)$ we have

$$\begin{aligned}
\prod_{i=1}^{l(\vec{\sigma})} z_\infty(\sigma_i) &= \prod_{i=1}^{l(\vec{\sigma})} \{z_m(\sigma_i) + z_m(\sigma_i)\cdot O(\lambda^{m+1})\} \qquad \text{by Lemma 4.34} \\
&= \prod_{i=1}^{l(\vec{\sigma})} z_m(\sigma_i) + O(\lambda^{m+1})\cdot \prod_{i=1}^{l(\vec{\sigma})} z_m(\sigma_i).
\end{aligned}$$

Hence

$$\begin{aligned}
Bel^{z_\infty}(q_1\cdots q_j) - Bel^{z_m}(q_1\cdots q_j) &= \sum_{\vec{\sigma}\in N(q_1\cdots q_j)} O(\lambda^{m+1})\cdot \prod_{i=1}^{l(\vec{\sigma})} \frac{z_m(\sigma_i)}{A_{\sigma_i - 1}} \\
&= O(\lambda^{m+1})
\end{aligned}$$

as required. □

Hence we have proved Theorem 4.11, i.e., that given a function $Bel : SL \to [0,1]$ which is given by a standard pre-ent over L and which satisfies, for all $\theta, \phi \in SL$, $Bel(\theta\wedge\phi) = 0$ implies $Bel(\phi\wedge\theta) = 0$, there exists an almost-ent, which we have denoted by z_∞ in the above proof and which was defined over a larger language than L (namely L^+), such that $Bel^{z_\infty}(\theta) = Bel(\theta)$ for all $\theta \in SL$.

Unfortunately we are not *yet* able to say that there exists an *ent* which gives the same belief values to sentences in SL, for, as we indicated earlier, the almost-ent z_∞ is *not* an ent over L^+. In fact no almost-ent over L^+ which is special (according to Definition 4.12) can be an ent over L^+. To see this let $s_1 \in WL^+$ be given by $s_1 = \{u_{p_1}, u_{p_2}\}$ (recall we assume $n > 1$). Then, given a special almost-ent z, there can be no scenario over L^+ which is both awarded non-zero potential by z *and* is consistent with s_1. This is because, for any non-empty sequence of literals σ, the scenario $s(\sigma)$ contains at least one of $\neg u_{p_1}$ or $\neg u_{p_2}$. Indeed if $\sigma = p_1 \cdots$ then $\neg u_{p_2} \in s(\sigma)$; if $\sigma = p_2 \cdots$ then $\neg u_{p_1} \in s(\sigma)$; while if $\sigma = q \cdots$ for some literal $q \neq p_i$ $(i = 1, 2)$ then $\neg u_{p_i} \in s(\sigma)$ $(i = 1, 2)$. We will show in Section 4.7 how to convert z_∞ into an ent which gives equivalent belief values to sentences in SL. Before that, however, we show how we can dispense with using the infinitesimal λ.

4.6 Stage 2 – The Potentials of z_∞

So far we have established the existence of an almost-ent z_∞ which, given a belief function Bel which is given by a standard pre-ent over L and which satisfies, for all $\theta, \phi \in SL$, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$, gives the same belief values to sentences in SL as Bel (although z_∞ is defined over a language L^+ that extends L). However, although the function Bel^{z_∞} is real-valued on the interval $[0, 1]$ (at least when Bel^{z_∞} is regarded as a function on SL), the potentials of the almost-ent z_∞ are non-standard real numbers. Indeed, as we saw in the last section, given a scenario $s(\sigma)$ over L^+ where σ is a non-empty sequence of literals, the potential z_∞ assigns to $s(\sigma)$ is given by

$$z_\infty(\sigma) = \lambda^{|\sigma|-1} \sum_{i=0}^{\infty} z_i^{(i)}(\sigma)\lambda^i$$

where λ is an indeterminate which may be thought of as a positive infinitesimal and $z_0^{(0)}(\sigma) = 0$ implies $z_\infty(\sigma) = 0$. Now ideally we would like to be able to take the potentials of z_∞ to be standard real numbers and this section is devoted to showing how we may do just that. The problem amounts to showing how, by taking λ to be a small enough real number, all the power series $z_\infty(\sigma)$ for all non-empty sequences σ will converge. Our strategy for showing this will be to find a sequence of real numbers η_k for $k = 0, 1, \dots$ such that, for all non-empty sequences σ, $|z_k^{(k)}(\sigma)| \leq \eta_k$, and then showing that $\sum_{k=0}^{\infty} \eta_k \lambda^k$ converges. This will then suffice by appealing to the following two propositions:

Proposition 4.36 (Comparison Test) *Suppose $\sum_{i=0}^{\infty} a_i \lambda^i$ and $\sum_{i=0}^{\infty} b_i \lambda^i$ are two power series in λ such that $0 \leq a_i \leq b_i$ for all $i \geq M$ (some $M \geq 0$). If $\sum_{i=0}^{\infty} b_i \lambda^i$ converges for all $|\lambda| < R$ (for some real number $R > 0$) then so too does $\sum_{i=0}^{\infty} a_i \lambda^i$.* □

Proposition 4.37 *Let $\sum_{i=0}^{\infty} a_i \lambda^i$ be a power series in λ. If $\sum_{i=0}^{\infty} |a_i| \lambda^i$ converges for all $|\lambda| < R$ then so too does $\sum_{i=0}^{\infty} a_i \lambda^i$.* □

Henceforth, to ease clutter on notation, we shall use $a_i(\sigma)$ to denote $z_i^{(i)}(\sigma)$. We shall define the numbers η_k inductively. Note that, for $k = 0$, we may take

$$\eta_0 = 1$$

since, for all non-empty sequences σ, we know $a_0(\sigma) = Bel(\sigma)$. Now suppose $k > 0$ and that we have found $\eta_0, \eta_1, \dots, \eta_{k-1}$ such that, for each $i = 0, 1, \dots, k-1$ and for all non-empty sequences σ, $|a_i(\sigma)| \leq \eta_i$. Our task now is to find a suitable η_k. We will do this by firstly, for each $l = 1, \dots, n$, finding a separate bound η_k^l for all the $|a_k(\sigma)|$ where $|\sigma| = l$. The overall bound η_k will then be gleaned from these bounds, which will be found using a sub-inductive process. Let $1 \leq j \leq n$, then, and let us assume for the moment that we have already found numbers $\eta_k^1, \dots, \eta_k^{j-1}$ such that

$$|a_k(\sigma)| \leq \eta_k^1 \text{ for all } \sigma \text{ such that } |\sigma| = 1.$$
$$\vdots$$
$$|a_k(\sigma)| \leq \eta_k^{j-1} \text{ for all } \sigma \text{ such that } |\sigma| = j-1.$$

Given these bounds, we will now try to find a number η_k^j which satisfies, for all sequences σ such that $|\sigma| = j$,

$$|a_k(\sigma)| \leq \eta_k^j.$$

Let $q_1 \cdots q_j$ be a sequence of literals of length j. By results in the previous section we have

$$Bel(q_1 \cdots q_j) = Bel^{z_\infty}(q_1 \cdots q_j) = \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{z_\infty(\sigma_i)}{A_{\sigma_{i-1}}}.$$

For each non-empty sequence of literals σ we define $z_*(\sigma) \in \mathbb{R}((\lambda))$ as follows:

$$z_*(\sigma) = \frac{z_\infty(\sigma)}{\lambda^{|\sigma|-1}} = \sum_{i=0}^{\infty} a_i(\sigma)\lambda^i.$$

Recall equation (4.7) for any sequence of literals σ:

$$A_\sigma = \lambda^{|\sigma|} Bel(\sigma) \cdot P_\sigma(\lambda)$$

where $P_\sigma(\lambda)$ is a polynomial in λ with constant term 1. Using these two identities to substitute in the above expression for $Bel(q_1 \cdots q_j)$ we get

$$\begin{aligned} Bel(q_1 \cdots q_j) &= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \prod_{i=1}^{l(\vec{\sigma})} \frac{\lambda^{|\sigma_i|-1} z_*(\sigma_i)}{\lambda^{|\sigma_{i-1}|} \cdot Bel(\sigma_{i-1}) \cdot P_{\sigma_{i-1}}(\lambda)} \\ &= \sum_{\vec{\sigma} \in N(q_1 \cdots q_j)} \left\{ \frac{\lambda^{|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma})}}{Q_{\vec{\sigma}}(\lambda) \cdot \prod Bel(\sigma_{i-1})} \cdot \prod_{i=1}^{l(\vec{\sigma})} z_*(\sigma_i) \right\} \end{aligned}$$

where, recall, for each sequence path $\vec{\sigma}$ we define $Q_{\vec{\sigma}}(\lambda) = \prod P_{\sigma_{i-1}}(\lambda)$. Let us assume that $Bel(q_1 \cdots q_j) \neq 0$. Then we must have $\vec{\iota}(q_1 \cdots q_j) \in N(q_1 \cdots q_j)$ (by definition of $\vec{\iota}$ and $N(q_1 \cdots q_j)$) and so we may pull out the lead term from the above sum and write

$$Bel(q_1 \cdots q_j) \quad = \quad \frac{z_*(q_1) \cdot z_*(q_1 q_2) \cdots z_*(q_1 \cdots q_j)}{Q_{\vec{\iota}}(\lambda) \cdot \prod Bel(q_1 \cdots q_{i-1})} +$$

$$+ \sum_{\vec{\iota} \neq \vec{\sigma} \in N(q_1 \cdots q_j)} \left\{ \frac{\lambda^{|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma})}}{Q_{\vec{\sigma}}(\lambda) \cdot \prod Bel(\sigma_{i-1})} \cdot \prod_{i=1}^{l(\vec{\sigma})} z_*(\sigma_i) \right\}. \tag{4.34}$$

Now for each non-empty sequence of literals $r_1 \cdots r_s$ let us define a polynomial $S_{r_1 \cdots r_s}(\lambda)$ as follows:

$$S_{r_1 \cdots r_s}(\lambda) = \prod_{\vec{\sigma} \in N(r_1 \cdots r_s)} \left\{ Q_{\vec{\sigma}}(\lambda) \cdot \prod_{i=1}^{l(\vec{\sigma})} Bel(\sigma_{i-1}) \right\}$$

and, for each $\vec{\sigma} \in N(r_1 \cdots r_s)$, define the polynomial $R^{\vec{\sigma}}_{r_1 \cdots r_s}(\lambda)$ by

$$R^{\vec{\sigma}}_{r_1 \cdots r_s}(\lambda) = \prod_{\vec{\sigma} \neq \vec{\rho} \in N(r_1 \cdots r_s)} \left\{ Q_{\vec{\rho}}(\lambda) \cdot \prod_{i=1}^{l(\vec{\rho})} Bel(\rho_{i-1}) \right\}.$$

For each i we shall denote the i^{th} coefficients of $S_{r_1 \cdots r_s}(\lambda)$ and $R^{\vec{\sigma}}_{r_1 \cdots r_s}(\lambda)$ by $S_{r_1 \cdots r_s}(\lambda)^{(i)}$ and $R^{\vec{\sigma}}_{r_1 \cdots r_s}(\lambda)^{(i)}$. Multiplying both sides of equation (4.34) by $S_{q_1 \cdots q_j}(\lambda)$ we get

$$\begin{aligned} S_{q_1 \cdots q_j}(\lambda) \cdot Bel(q_1 \cdots q_j) \;=\; & R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda) \cdot z_*(q_1) \cdot z_*(q_1 q_2) \cdots z_*(q_1 \cdots q_j) + \\ & + \sum_{\vec{\iota} \neq \vec{\sigma} \in N(q_1 \cdots q_j)} R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda) \cdot \lambda^{|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma})} \cdot \prod_{i=1}^{l(\vec{\sigma})} z_*(\sigma_i). \end{aligned} \tag{4.35}$$

We now equate the k^{th} coefficient of each side of the above formula (treated as power series in λ). The k^{th} coefficient of the left hand side is simply equal to

$$S_{q_1 \cdots q_j}(\lambda)^{(k)} \cdot Bel(q_1 \cdots q_j).$$

To find the k^{th} coefficient of the right hand side let us begin by considering its first term, namely,

$$R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda) \cdot z_*(q_1) \cdot z_*(q_1 q_2) \cdots z_*(q_1 \cdots q_j).$$

The k^{th} coefficient of this term is given by

$$\begin{aligned}
&\sum_{i_0+i_1+\cdots+i_j=k} R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)} \cdot a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \;= \\
&= R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot \sum_{i_1+\cdots+i_j=k} a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \;+ \\
&\quad + \sum_{\substack{i_0+i_1+\cdots+i_j=k\\ i_0\neq 0}} R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)} \cdot a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \\
&= R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot a_0(q_1)\cdot a_0(q_1q_2)\cdots a_0(q_1\cdots q_{j-1})\cdot a_k(q_1\cdots q_j) \;+ \\
&\quad + R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot a_k(q_1)\cdot a_0(q_1q_2)\cdots a_0(q_1\cdots q_{j-1})\cdot a_0(q_1\cdots q_j) \;+ \\
&\quad + R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot a_0(q_1)\cdot a_k(q_1q_2)\cdots a_0(q_1\cdots q_{j-1})\cdot a_0(q_1\cdots q_j) \;+ \\
&\quad + \;\cdots\; + \\
&\quad + R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot a_0(q_1)\cdot a_0(q_1q_2)\cdots a_k(q_1\cdots q_{j-1})\cdot a_0(q_1\cdots q_j) \;+ \\
&\quad + R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot \sum_{\substack{i_1+\cdots+i_j=k\\ \forall l\;\; i_l\neq k}} a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \;+ \\
&\quad + \sum_{\substack{i_0+i_1+\cdots+i_j=k\\ i_0\neq 0}} R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)} \cdot a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j).
\end{aligned}$$

Now let us consider the second term of the right hand side of (4.35), namely,

$$\sum_{\vec{\iota}\neq\vec{\sigma}\in N(q_1\cdots q_j)} R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)\cdot \lambda^{|\sigma_{l(\vec{\sigma})}|-l(\vec{\sigma})}\cdot \prod_{i=1}^{l(\vec{\sigma})} z_*(\sigma_i).$$

Let us denote this term by V for the moment. For each $i \geq 0$ let us define

$$N^i(q_1\cdots q_j) = \{\vec{\sigma}\in N(q_1\cdots q_j) \mid |\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) = i\}.$$

Note that under this definition we have

$$N(q_1\cdots q_j) = \bigcup_{i=0}^{n-1} N^i(q_1\cdots q_j)$$

since $|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) \leq n-1$ for all $\vec{\sigma}\in P(q_1\cdots q_j)$ (by Proposition 4.19(iii)). Let $\vec{\sigma}\in N(q_1\cdots q_j)$ be such that $\vec{\sigma}\neq\vec{\iota}$. Then we have that $|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) > 0$ (by

Proposition 4.19(iv)) and so $\vec{\sigma} \in N^i(q_1 \cdots q_j)$ for some $i > 0$. Hence, since this means that $|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) = i$, the k^{th} coefficient of

$$R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda) \cdot \lambda^{|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma})} \cdot \prod_{i=1}^{l(\vec{\sigma})} z_*(\sigma_i)$$

will be equal to the $(k-i)^{\text{th}}$ coefficient of

$$R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda) \cdot \prod_{i=1}^{l(\vec{\sigma})} z_*(\sigma_i)$$

which in turn is equal to

$$\sum_{i_0 + i_1 + \cdots + i_{l(\vec{\sigma})} = k-i} R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda)^{(i_0)} \cdot a_{i_1}(\sigma_1) \cdots a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})}).$$

Note that this sum becomes zero when $i > k$. Hence the k^{th} coefficient of the term V is given by

$$\sum_{i=1}^{k} \sum_{\vec{\sigma} \in N^i(q_1 \cdots q_j)} \sum_{i_0 + i_1 + \cdots + i_{l(\vec{\sigma})} = k-i} R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda)^{(i_0)} \cdot a_{i_1}(\sigma_1) \cdots a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})}).$$

However, since we have $i > n-1$ implies $N^i(q_1 \cdots q_j) = \emptyset$, we may just as well replace the upper limit k in the first summation by $n-1$ and thus express the k^{th} coefficient of V as

$$\sum_{i=1}^{n-1} \sum_{\vec{\sigma} \in N^i(q_1 \cdots q_j)} \sum_{i_0 + i_1 + \cdots + i_{l(\vec{\sigma})} = k-i} R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda)^{(i_0)} \cdot a_{i_1}(\sigma_1) \cdots a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})}).$$

Putting all this together we can expand (4.35) as

$$\begin{aligned}
S_{q_1 \cdots q_j}(\lambda)^{(k)} \cdot Bel(q_1 \cdots q_j) \ &= \\
= \ & R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)} \cdot a_0(q_1) \cdot a_0(q_1 q_2) \cdots a_0(q_1 \cdots q_{j-1}) \cdot a_k(q_1 \cdots q_j) \ + \\
& + R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)} \cdot a_k(q_1) \cdot a_0(q_1 q_2) \cdots a_0(q_1 \cdots q_{j-1}) \cdot a_0(q_1 \cdots q_j) \ + \\
& + R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)} \cdot a_0(q_1) \cdot a_k(q_1 q_2) \cdots a_0(q_1 \cdots q_{j-1}) \cdot a_0(q_1 \cdots q_j) \ + \\
& + \ \cdots \ + \\
& + R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)} \cdot a_0(q_1) \cdot a_0(q_1 q_2) \cdots a_k(q_1 \cdots q_{j-1}) \cdot a_0(q_1 \cdots q_j) \ +
\end{aligned}$$

$$+ R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l\neq k}} a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \;+$$
$$+ \sum_{\substack{i_0+i_1+\cdots+i_j=k \\ i_0\neq 0}} R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}\cdot a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \;+$$
$$+ \sum_{i=1}^{n-1} \sum_{\vec{\sigma}\in N^i(q_1\cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}\cdot a_{i_1}(\sigma_1)\cdots a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})$$

and so

$$\begin{aligned}
R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} &\cdot a_0(q_1)\cdot a_0(q_1q_2)\cdots a_0(q_1\cdots q_{j-1})\cdot a_k(q_1\cdots q_j) \;= \\
= \;& S_{q_1\cdots q_j}(\lambda)^{(k)}\cdot Bel(q_1\cdots q_j) \;- \\
& - R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}\cdot a_k(q_1)\cdot a_0(q_1q_2)\cdots a_0(q_1\cdots q_{j-1})\cdot a_0(q_1\cdots q_j) \;- \\
& - R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}\cdot a_0(q_1)\cdot a_k(q_1q_2)\cdots a_0(q_1\cdots q_{j-1})\cdot a_0(q_1\cdots q_j) \;- \\
& - \;\cdots\; - \\
& - R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}\cdot a_0(q_1)\cdot a_0(q_1q_2)\cdots a_k(q_1\cdots q_{j-1})\cdot a_0(q_1\cdots q_j) \;- \\
& - R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)} \cdot \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l\neq k}} a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \;- \\
& - \sum_{\substack{i_0+i_1+\cdots+i_j=k \\ i_0\neq 0}} R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}\cdot a_{i_1}(q_1)\cdot a_{i_2}(q_1q_2)\cdots a_{i_j}(q_1\cdots q_j) \;- \\
& - \sum_{i=1}^{n-1} \sum_{\vec{\sigma}\in N^i(q_1\cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}\cdot a_{i_1}(\sigma_1)\cdots a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})}).
\end{aligned}$$

Taking the modulus of each side and using the triangle inequality gives us

$$\begin{aligned}
|R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| &\cdot |a_0(q_1)|\cdot |a_0(q_1q_2)|\cdots |a_0(q_1\cdots q_{j-1})|\cdot |a_k(q_1\cdots q_j)| \;\leq \\
\leq \;& |S_{q_1\cdots q_j}(\lambda)^{(k)}|\cdot Bel(q_1\cdots q_j) \;+ \\
& + |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}|\cdot |a_k(q_1)|\cdot |a_0(q_1q_2)|\cdots |a_0(q_1\cdots q_{j-1})|\cdot |a_0(q_1\cdots q_j)| \;+ \\
& + |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}|\cdot |a_0(q_1)|\cdot |a_k(q_1q_2)|\cdots |a_0(q_1\cdots q_{j-1})|\cdot |a_0(q_1\cdots q_j)| \;+ \\
& + \;\cdots\; + \\
& + |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}|\cdot |a_0(q_1)|\cdot |a_0(q_1q_2)|\cdots |a_k(q_1\cdots q_{j-1})|\cdot |a_0(q_1\cdots q_j)| \;+
\end{aligned}$$

$$+ |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l\neq k}} |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| \; +$$
$$+ \sum_{\substack{i_0+i_1+\cdots+i_j=k \\ i_0\neq 0}} |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| \; +$$
$$+ \sum_{i=1}^{n-1} \sum_{\vec{\sigma}\in N^i(q_1\cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(\sigma_1)| \cdots |a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})|. \tag{4.36}$$

Let us abbreviate this inequality by defining the following:

$$G = |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot |a_0(q_1)| \cdot |a_0(q_1q_2)| \cdots |a_0(q_1\cdots q_{j-1})|,$$

$$H_1 = |S_{q_1\cdots q_j}(\lambda)^{(k)}| \cdot Bel(q_1\cdots q_j),$$

$$\begin{aligned} H_2 \;\; = \;\; & |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot |a_k(q_1)| \cdot |a_0(q_1q_2)| \cdots |a_0(q_1\cdots q_{j-1})| \cdot |a_0(q_1\cdots q_j)| \; + \\ & + |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot |a_0(q_1)| \cdot |a_k(q_1q_2)| \cdots |a_0(q_1\cdots q_{j-1})| \cdot |a_0(q_1\cdots q_j)| \; + \\ & + \;\cdots\; + \\ & + |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot |a_0(q_1)| \cdot |a_0(q_1q_2)| \cdots |a_k(q_1\cdots q_{j-1})| \cdot |a_0(q_1\cdots q_j)|, \end{aligned}$$

$$H_3 = |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l\neq k}} |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)|,$$

$$H_4 = \sum_{\substack{i_0+i_1+\cdots+i_j=k \\ i_0\neq 0}} |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)|$$

and

$$H_5 = \sum_{i=1}^{n-1} \sum_{\vec{\sigma}\in N^i(q_1\cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(\sigma_1)| \cdots |a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})|.$$

So, with these abbreviations, (4.36) becomes

$$G \cdot |a_k(q_1\cdots q_j)| \leq H_1 + H_2 + H_3 + H_4 + H_5 \tag{4.37}$$

where, remember, we are assuming $Bel(q_1\cdots q_j) \neq 0$. At this point let us remind ourselves that we are seeking an upper bound for the set

$$\{|a_k(\sigma)| \mid \sigma \text{ a sequence of literals of length } j\}.$$

Thus our task now is to find an upper bound for $|a_k(q_1 \cdots q_j)|$ which is *independent* of the particular sequence $q_1 \cdots q_j$ of length j which we are considering. Our next step in this direction is to find an upper bound independent of $q_1 \cdots q_j$ for each of the H_i's. First of all let m be such that, for all sequences of literals $r_1 \cdots r_s$ and for all $\vec{\sigma} \in N(r_1 \cdots r_s)$, the degree of $S_{r_1 \cdots r_s}(\lambda)$ and the degree of each $R^{\vec{\sigma}}_{r_1 \cdots r_s}(\lambda)$ is less than or equal to m. Such an m exists since there are only finitely many sequences $r_1 \cdots r_s$ and, given $r_1 \cdots r_s$, only finitely many $\vec{\sigma} \in N(r_1 \cdots r_s)$, hence there are only finitely many polynomials of the form $S_{r_1 \cdots r_s}(\lambda)$ and $R^{\vec{\sigma}}_{r_1 \cdots r_s}(\lambda)$. For the same reason there exists some number B such that, for all sequences $r_1 \cdots r_s$ and all $\vec{\sigma} \in N(r_1 \cdots r_s)$, we have $|S_{r_1 \cdots r_s}(\lambda)^{(i)}| \leq B$ and $|R^{\vec{\sigma}}_{r_1 \cdots r_s}(\lambda)^{(i)}| \leq B$ for all $i = 0, \ldots, m$. Hence we have

$$H_1 = |S_{q_1 \cdots q_j}(\lambda)^{(k)}| \cdot Bel(q_1 \cdots q_j) \leq |S_{q_1 \cdots q_j}(\lambda)^{(k)}| \leq B \tag{4.38}$$

which takes care of H_1. Now let us look at H_2. From the definition of $R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)$, we have

$$R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)} = \prod_{\vec{\iota} \neq \vec{\sigma} \in N(q_1 \cdots q_j)} \left\{ \prod_{i=1}^{l(\vec{\sigma})} Bel(\sigma_{i-1}) \right\}$$

(since, for each path $\vec{\sigma}$, $Q_{\vec{\sigma}}(\lambda)$ has constant term 1) and so clearly $|R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)}| \leq 1$. Hence, recalling that $a_0(\sigma) = Bel(\sigma)$ for each sequence σ and also that we are assuming $|a_k(\sigma)| \leq \eta^1_k$ for all σ of length 1,

$$\begin{aligned}
|R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)}| \cdot |a_k(q_1)| \cdot |a_0(q_1 q_2)| \cdots |a_0(q_1 \cdots q_{j-1})| \cdot |a_0(q_1 \cdots q_j)| &= \\
= \ |R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)}| \cdot |a_k(q_1)| \cdot Bel(q_1 q_2) \cdots Bel(q_1 \cdots q_{j-1}) \cdot Bel(q_1 \cdots q_j) & \\
\leq \ |a_k(q_1)| & \\
\leq \ \eta^1_k. &
\end{aligned}$$

Similarly we get

$$\begin{array}{rcl}
|R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)}| \cdot |a_0(q_1)| \cdot |a_k(q_1 q_2)| \cdots |a_0(q_1 \cdots q_{j-1})| \cdot |a_0(q_1 \cdots q_j)| & \leq & \eta^2_k \\
\vdots & \vdots & \vdots \\
|R^{\vec{\iota}}_{q_1 \cdots q_j}(\lambda)^{(0)}| \cdot |a_0(q_1)| \cdot |a_0(q_1 q_2)| \cdots |a_k(q_1 \cdots q_{j-1})| \cdot |a_0(q_1 \cdots q_j)| & \leq & \eta^{j-1}_k
\end{array}$$

and so we may see that

$$H_2 \leq \eta_k^1 + \eta_k^2 + \cdots + \eta_k^{j-1}. \tag{4.39}$$

Now, for H_3, since $|R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \leq 1$, we have

$$\begin{aligned}
|R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l \neq k}} |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1 q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| \;\; &\leq \\
\leq \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l \neq k}} |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1 q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| & \\
\leq \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l \neq k}} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_j}. & \qquad (4.40)
\end{aligned}$$

We now make use of a slight variant of the following result:

Lemma 4.38 *For each $i \geq 0$ and for any $l \leq n$,*

$$\sum_{i_1+\cdots+i_l=k-i} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_l} \leq \sum_{i_1+\cdots+i_n=k-i} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_n}.$$

Proof. Since $\eta_0 = 1$ we have

$$\sum_{i_1+\cdots+i_l=k-i} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_l} = \sum_{i_1+\cdots+i_l=k-i} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_l} \cdot \overbrace{\eta_0 \cdots \eta_0}^{n-l \text{ copies}}.$$

We also have

$$\sum_{i_1+\cdots+i_l=k-i} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_l} \cdot \overbrace{\eta_0 \cdots \eta_0}^{n-l \text{ copies}} \leq \sum_{i_1+\cdots+i_n=k-i} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_n}$$

since every term which appears in the left-hand sum also appears in the right-hand sum. This gives the result. □

By similar reasoning to that in the proof of the above lemma we may see that

$$\begin{aligned}
\sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l \neq k}} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_j} &= \sum_{\substack{i_1+\cdots+i_j=k \\ \forall l \;\; i_l \neq k}} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_j} \cdot \overbrace{\eta_0 \cdots \eta_0}^{n-j \text{ copies}} \\
&\leq \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \;\; i_l \neq k}} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_n}
\end{aligned}$$

and so, from this and (4.40), we get our bound for H_3:

$$H_3 \leq \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \;\; i_l \neq k}} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_n}. \tag{4.41}$$

For H_4 we have

$$\begin{aligned} \sum_{\substack{i_0+i_1+\cdots+i_j=k \\ i_0 \neq 0}} |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| &= \\ = \sum_{i_0=1}^{k} |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot \sum_{i_1+\cdots+i_j=k-i_0} |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| \end{aligned}$$

Note that, in the first summation in the above line, we may replace k as the upper limit of that summation by any number bigger than k. This is because, for any $i_0 > k$, there can be no $i_1, \ldots, i_j$ which sum to $k - i_0$ and so the second summation will be empty. In particular if $m \geq k$ (where, recall, m is our upper bound on the degrees of all the polynomials $S(\lambda)$ and $R(\lambda)$) then we may replace k by m. Indeed we may also replace k by m in the case where $m < k$ since we know that $R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)} = 0$ for $i_0 > m$ (by the definition of m). Hence

$$\begin{aligned} \sum_{\substack{i_0+i_1+\cdots+i_j=k \\ i_0 \neq 0}} |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| &= \\ = \sum_{i_0=1}^{m} |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \cdot \sum_{i_1+\cdots+i_j=k-i_0} |a_{i_1}(q_1)| \cdot |a_{i_2}(q_1q_2)| \cdots |a_{i_j}(q_1\cdots q_j)| & \\ \leq B \sum_{i_0=1}^{m} \sum_{i_1+\cdots+i_j=k-i_0} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_j} \qquad \text{since } |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(i_0)}| \leq B \text{ for all } i_0 & \\ \leq B \sum_{i_0=1}^{m} \sum_{i_1+\cdots+i_n=k-i_0} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_n} \qquad \text{by Lemma 4.38.} & \qquad (4.42) \end{aligned}$$

It will be more convenient, for when we come to combine the bounds for the H_i's later on, to enlarge this bound even further by increasing the upper limit of this last summation from m to $m + n - 1$. Thus we now have

$$H_4 \;\; \leq \;\; B \sum_{i_0=1}^{m} \sum_{i_1+\cdots+i_n=k-i_0} \eta_{i_1} \cdot \eta_{i_2} \cdots \eta_{i_n} \qquad \text{from (4.42)}$$

$$\begin{aligned} &\leq B \sum_{i_0=1}^{m+n-1} \sum_{i_1+\cdots+i_n=k-i_0} \eta_{i_1}\cdot\eta_{i_2}\cdots\eta_{i_n} \\ &= B \sum_{k-t\leq i_1+\cdots+i_n\leq k-1} \eta_{i_1}\cdot\eta_{i_2}\cdots\eta_{i_n} \end{aligned} \tag{4.43}$$

where $t = m+n-1$. Finally let us consider H_5, which, we recall, is equal to

$$\sum_{i=1}^{n-1} \sum_{\vec{\sigma}\in N^i(q_1\cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}|\cdot|a_{i_1}(\sigma_1)|\cdots|a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})|.$$

Now, for each $i = 1,\ldots,n-1$ and for each $\vec{\sigma}\in N^i(q_1\cdots q_j)$, we have

$$\begin{aligned} &\sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}|\cdot|a_{i_1}(\sigma_1)|\cdots|a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})| = \\ &= \sum_{i_0=0}^{k-i} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}|\cdot \sum_{i_1+\cdots+i_{l(\vec{\sigma})}=k-i-i_0} |a_{i_1}(\sigma_1)|\cdots|a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})| \\ &= \sum_{i_0=0}^{m} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}|\cdot \sum_{i_1+\cdots+i_{l(\vec{\sigma})}=k-i-i_0} |a_{i_1}(\sigma_1)|\cdots|a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})| \end{aligned} \tag{4.44}$$

since we may replace the upper limit $k-i$ in the above by m for exactly the same reasons as we replaced k by m in a similar situation when looking at H_4. Now, since $|R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}|\leq B$ for all i_0 and $|a_{i_l}(\sigma)|\leq \eta_{i_l}$ for any $i_l<k$ and any non-empty sequence of literals σ, we then have

$$\begin{aligned} &\sum_{i_0=0}^{m} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}|\cdot \sum_{i_1+\cdots+i_{l(\vec{\sigma})}=k-i-i_0} |a_{i_1}(\sigma_1)|\cdots|a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})| \leq \\ &\leq B\sum_{i_0=0}^{m} \sum_{i_1+\cdots+i_{l(\vec{\sigma})}=k-i-i_0} \eta_{i_1}\cdots\eta_{i_{l(\vec{\sigma})}} \\ &\leq B\sum_{i_0=0}^{m} \sum_{i_1+\cdots+i_n=k-i-i_0} \eta_{i_1}\cdots\eta_{i_n} \qquad \text{by Lemma 4.38.} \end{aligned} \tag{4.45}$$

Hence, combining (4.44) and (4.45) gives us, for each $i=1,\ldots,n-1$,

$$\sum_{\vec{\sigma}\in N^i(q_1\cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} |R^{\vec{\sigma}}_{q_1\cdots q_j}(\lambda)^{(i_0)}|\cdot|a_{i_1}(\sigma_1)|\cdots|a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})|$$

$$\leq B \sum_{\vec{\sigma} \in N^i(q_1 \cdots q_j)} \sum_{i_0=0}^{m} \sum_{i_1+\cdots+i_n=k-i-i_0} \eta_{i_1} \cdots \eta_{i_n}$$
$$= B \cdot |N^i(q_1 \cdots q_j)| \sum_{i_0=0}^{m} \sum_{i_1+\cdots+i_n=k-i-i_0} \eta_{i_1} \cdots \eta_{i_n}$$

where $|N^i(q_1 \cdots q_j)|$ denotes the cardinality of the set $N^i(q_1 \cdots q_j)$. Now $N^i(q_1 \cdots q_j)$ is defined as the set of all paths $\vec{\sigma}$ for $q_1 \cdots q_j$ for which $Bel(\sigma_{l(\vec{\sigma})}) \neq 0$ and $|\sigma_{l(\vec{\sigma})}| - l(\vec{\sigma}) = i$, and as such may be identified with a subset of the set of *all* sequences $\vec{\sigma} = \sigma_1, \ldots, \sigma_{l(\vec{\sigma})}$ where $l(\vec{\sigma})$ satisfies $1 \leq l(\vec{\sigma}) \leq n$ and, for each $l = 1, \ldots, l(\vec{\sigma})$, σ_l may be *any* sequence of literals, i.e., there are no constraints on the choices of the σ_l's. Clearly, since there are only finitely many sequences of literals and since $l(\vec{\sigma})$ is bounded above by n, this latter set is finite, say it has C elements. Hence $|N^i(q_1 \cdots q_j)| \leq C$ (and note that this argument works independently of i and $q_1 \cdots q_j$). Hence, for each $i = 1, \ldots, n-1$,

$$\sum_{\vec{\sigma} \in N^i(q_1 \cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} |R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(\sigma_1)| \cdots |a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})|$$
$$\leq BC \sum_{i_0=0}^{m} \sum_{i_1+\cdots+i_n=k-i-i_0} \eta_{i_1} \cdots \eta_{i_n}$$

and so

$$\sum_{i=1}^{n-1} \sum_{\vec{\sigma} \in N^i(q_1 \cdots q_j)} \sum_{i_0+i_1+\cdots+i_{l(\vec{\sigma})}=k-i} |R^{\vec{\sigma}}_{q_1 \cdots q_j}(\lambda)^{(i_0)}| \cdot |a_{i_1}(\sigma_1)| \cdots |a_{i_{l(\vec{\sigma})}}(\sigma_{l(\vec{\sigma})})|$$
$$\leq BC \sum_{i=1}^{n-1} \sum_{i_0=0}^{m} \sum_{i_1+\cdots+i_n=k-i-i_0} \eta_{i_1} \cdots \eta_{i_n}. \qquad (4.46)$$

Now, for each $i = 1, \ldots, n-1$, we have

$$\sum_{i_0=0}^{m} \sum_{i_1+\cdots+i_n=k-i-i_0} \eta_{i_1} \cdots \eta_{i_n} \leq \sum_{i_0=1}^{n-1+m} \sum_{i_1+\cdots+i_n=k-i_0} \eta_{i_1} \cdots \eta_{i_n}$$

since every term in the first summation on the left hand side, i.e., each term of the form

$$\sum_{i_1+\cdots+i_n=k-i-i_0} \eta_{i_1} \cdots \eta_{i_n},$$

appears as a term in the first summation on the right hand side. Hence

$$\begin{aligned} BC\sum_{i=1}^{n-1}\sum_{i_0=0}^{m}\sum_{i_1+\cdots+i_n=k-i-i_0}\eta_{i_1}\cdots\eta_{i_n} &\leq \\ \leq\ BC\sum_{i=1}^{n-1}\sum_{i_0=1}^{n-1+m}\sum_{i_1+\cdots+i_n=k-i_0}\eta_{i_1}\cdots\eta_{i_n} & \\ =\ BC(n-1)\sum_{i_0=1}^{n-1+m}\sum_{i_1+\cdots+i_n=k-i_0}\eta_{i_1}\cdots\eta_{i_n} & \\ =\ BC(n-1)\sum_{k-t\leq i_1+\cdots+i_n\leq k-1}\eta_{i_1}\cdots\eta_{i_n} & \end{aligned} \tag{4.47}$$

And so (4.46) together with (4.47) gives us

$$H_5 \leq BC(n-1)\sum_{k-t\leq i_1+\cdots+i_n\leq k-1}\eta_{i_1}\cdots\eta_{i_n}. \tag{4.48}$$

And so we now have a further upper bound (which depends at the most on k) for each of the H_i's. Using these bounds (equations (4.38), (4.39), (4.41), (4.43), and (4.48)) together with (4.37) we may write

$$\begin{aligned} G\cdot|a_k(q_1\cdots q_j)| &\leq H_1+H_2+H_3+H_4+H_5 \qquad \text{(from (4.37))} \\ &\leq B+\eta_k^1+\eta_k^2+\cdots+\eta_k^{j-1}+\sum_{\substack{i_1+\cdots+i_n=k \\ \forall l\ i_l\neq k}}\eta_{i_1}\cdots\eta_{i_n}+ \\ &\quad +B\sum_{k-t\leq i_1+\cdots+i_n\leq k-1}\eta_{i_1}\cdots\eta_{i_n}+ \\ &\quad +BC(n-1)\sum_{k-t\leq i_1+\cdots+i_n\leq k-1}\eta_{i_1}\cdots\eta_{i_n} \\ &= B+\eta_k^1+\eta_k^2+\cdots+\eta_k^{j-1}+\sum_{\substack{i_1+\cdots+i_n=k \\ \forall l\ i_l\neq k}}\eta_{i_1}\cdots\eta_{i_n}+ \\ &\quad +D\sum_{k-t\leq i_1+\cdots+i_n\leq k-1}\eta_{i_1}\cdots\eta_{i_n}. \end{aligned} \tag{4.49}$$

where $D = B(1+C(n-1))$, a constant. Now we turn our attention to finding a suitable lower bound for G. Recall that

$$G = |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}|\cdot|a_0(q_1)|\cdots|a_0(q_1\cdots q_{j-1})|.$$

Now, since for all sequences of literals σ, $a_0(\sigma) = Bel(\sigma)$, we have

$$G = |R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| \cdot \left\{ \prod_{i=1}^{j-1} Bel(q_1 \cdots q_i) \right\}$$

We also know that

$$|R^{\vec{\iota}}_{q_1\cdots q_j}(\lambda)^{(0)}| = \prod_{\vec{\iota}\neq\vec{\sigma}\in N(q_1\cdots q_j)} \left\{ \prod_{i=1}^{l(\vec{\sigma})} Bel(\sigma_{i-1}) \right\}$$

while for the sequence path $\vec{\iota}(q_1 \cdots q_j)$ we have $l(\vec{\iota}) = j$ and $\iota_i = q_1 \cdots q_i$ for each $i = 1, \ldots, j$ which gives

$$\prod_{i=1}^{l(\vec{\iota})} Bel(\iota_{i-1}) = \prod_{i=1}^{j} Bel(q_1 \cdots q_{i-1}) = \prod_{i=1}^{j-1} Bel(q_1 \cdots q_j)$$

(remembering that $Bel(\emptyset) = 1$). Therefore

$$\begin{aligned} G &= \prod_{\vec{\iota}\neq\vec{\sigma}\in N(q_1\cdots q_j)} \left\{ \prod_{i=1}^{l(\vec{\sigma})} Bel(\sigma_{i-1}) \right\} \cdot \prod_{i=1}^{l(\vec{\iota})} Bel(\iota_{i-1}) \\ &= \prod_{\vec{\sigma}\in N(q_1\cdots q_j)} \left\{ \prod_{i=1}^{l(\vec{\sigma})} Bel(\sigma_{i-1}) \right\} \\ &= S_{q_1\cdots q_j}(\lambda)^{(0)} \qquad \text{by definition of } S_{q_1\cdots q_j}(\lambda). \end{aligned}$$

Hence, substituting this expression for G in (4.49),

$$\begin{aligned} S_{q_1\cdots q_j}(\lambda)^{(0)} \cdot |a_k(q_1 \cdots q_j)| \;\leq\; & B + \eta_k^1 + \eta_k^2 + \cdots + \eta_k^{j-1} + \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \;\; i_l \neq k}} \eta_{i_1} \cdots \eta_{i_n} + \\ & + D \sum_{k-t \leq i_1+\cdots+i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n}. \end{aligned} \tag{4.50}$$

Let b be the minimum of the finite set

$$\{S_{r_1\cdots r_s}(\lambda)^{(0)} \mid r_1 \cdots r_s \text{ a sequence of literals}\}.$$

Then, obviously,

$$b \cdot |a_k(q_1 \cdots q_j)| \;\leq\; S_{q_1\cdots q_j}(\lambda)^{(0)} \cdot |a_k(q_1 \cdots q_j)|$$

so from (4.50)

$$b \cdot |a_k(q_1 \cdots q_j)| \leq B + \eta_k^1 + \eta_k^2 + \cdots + \eta_k^{j-1} + \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \ i_l \neq k}} \eta_{i_1} \cdots \eta_{i_n} + \\ +D \sum_{k-t \leq i_1+\cdots+i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n}. \quad (4.51)$$

Note that, for all sequences of literals $r_1 \cdots r_s$, we have

$$S_{r_1 \cdots r_s}(\lambda)^{(0)} = \prod_{\vec{\sigma} \in N(r_1 \cdots r_s)} \left\{ \prod_{i=1}^{l(\vec{\sigma})} Bel(\sigma_{i-1}) \right\} \\ > 0$$

since if $N(r_1 \cdots r_s) \neq \emptyset$ then for all $\vec{\sigma} \in N(r_1 \cdots r_s)$ we have $Bel(\sigma_i) > 0$ for all $i = 1, \ldots, l(\vec{\sigma})$ (by definition of $N(r_1 \cdots r_s)$) while if $N(r_1 \cdots r_s) = \emptyset$ then $S_{r_1 \cdots r_s}(\lambda)^{(0)} = 1$ since we are adopting the convention that the empty product is equal to 1. Hence $b > 0$ and so we may divide throughout the inequality (4.51) by b to get

$$|a_k(q_1 \cdots q_j)| \leq X + Y(\eta_k^1 + \eta_k^2 + \cdots + \eta_k^{j-1}) + Y \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \ i_l \neq k}} \eta_{i_1} \cdots \eta_{i_n} + \\ +Z \sum_{k-t \leq i_1+\cdots+i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n}$$

where X, Y and Z are constants. ($X = \frac{B}{b}$, $Y = \frac{1}{b}$, $Z = \frac{D}{b}$.) Recall that we made the assumption that $Bel(q_1 \cdots q_j) > 0$ but this inequality clearly must still hold even if $Bel(q_1 \cdots q_j) = 0$ (since, in this case, $a_k(q_1 \cdots q_j) = 0$ also). Hence we have finally found a suitable definition for η_k^j, namely

$$\eta_k^j = X + Y(\eta_k^1 + \eta_k^2 + \cdots + \eta_k^{j-1}) + Y \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \ i_l \neq k}} \eta_{i_1} \cdots \eta_{i_n} + \\ +Z \sum_{k-t \leq i_1+\cdots+i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n}. \quad (4.52)$$

Recall, though, that our aim is to find a single number η_k for which, for *all* sequences of literals σ (regardless of length) we have

$$|a_k(\sigma)| \leq \eta_k.$$

We may do this as follows. First let us define the abbreviations

$$\Sigma_k^1 = \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \ i_l \neq k}} \eta_{i_1} \cdots \eta_{i_n}$$

and

$$\Sigma_k^2 = \sum_{k-t \leq i_1+\cdots+i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n}.$$

Putting $j = 1$ in (4.52) we see that

$$\eta_k^1 = X + Y\Sigma_k^1 + Z\Sigma_k^2.$$

Putting $j = 2$ in (4.52) gives us

$$\begin{aligned} \eta_k^2 &= X + Y\eta_k^1 + Y\Sigma_k^1 + Z\Sigma_k^2 \\ &= X + Y(X + Y\Sigma_k^1 + Z\Sigma_k^2) + Y\Sigma_k^1 + Z\Sigma_k^2 \\ &= (X + YX) + (Y^2 + Y)\Sigma_k^1 + (YZ + Z)\Sigma_k^2. \end{aligned}$$

In general, by a simple inductive argument, we may assert that, for each $j = 1, \ldots, n$, there exist constants X_j, Y_j and Z_j (which do not depend on k) such that

$$\eta_k^j = X_j + Y_j\Sigma_k^1 + Z_j\Sigma_k^2.$$

Let

$$K = \max_{1 \leq j \leq n} X_j, \ L = \max_{1 \leq j \leq n} Y_j,$$

and let

$$M = \max(\max_{1 \leq j \leq n} Z_j, \ 1).$$

(We choose M in this way to make absolutely sure that the η_k are increasing – a property which will be needed later.) Then, for any non-empty sequence of literals σ, we have

$$\begin{aligned} |a_k(\sigma)| &\leq \eta_k^j \qquad \text{where } j = |\sigma|. \\ &= X_j + Y_j\Sigma_k^1 + Z_j\Sigma_k^2 \\ &\leq K + L\Sigma_k^1 + M\Sigma_k^2. \end{aligned}$$

Hence this shows that we may take, for $k > 1$,

$$\eta_k = K + L\Sigma_k^1 + M\Sigma_k^2$$

(with, recall, $\eta_0 = 1$). Note that the actual values of K, L and M are not relevant in the rest of this section. The important thing is to note that they are constants which do not vary with k. And so, in our efforts to establish the convergence of (for all non-empty sequences of literals σ) the series $\sum_{k=0}^{\infty} a_k(\sigma)\lambda^k$, we now turn our attention to establishing the convergence of the series $\sum_{k=0}^{\infty} \eta_k \lambda^k$. By Propositions 4.36 and 4.37 the convergence of this latter series will then imply the convergence of the former series. However, the convergence of $\sum_{k=0}^{\infty} \eta_k \lambda^k$ itself is not immediately provable and we will, in fact, need recourse to Proposition 4.36 in several more places in the rest of this section in order to show it. The next thing we shall do is to find a bound for η_k (for $k > 1$) in terms of Σ_k^1 only. To do this we need to know that the sequence $\eta_0, \eta_1, \eta_2, \ldots$ is increasing.

Lemma 4.39 *For each $k \geq 1$ we have $\eta_k \geq \eta_{k-1}$.*

Proof. First note that $K \geq 0$ (since $X_1 = X = \frac{B}{b} \geq 0$ and $K \geq X_1$ by definition of K) and $L \geq 0$ (since $Y_1 = Y = \frac{1}{b} \geq 0$ and $L \geq Y_1$ by definition of L) while clearly $M \geq 1 \geq 0$ (by definition of M). Hence, for $k > 0$,

$$\begin{aligned}
\eta_k &= K + L\Sigma_k^1 + M\Sigma_k^2 \\
&\geq M\Sigma_k^2 \qquad \text{since } \eta_i \geq 0 \text{ for all } i \text{ (recall } |a_i(\sigma)| \leq \eta_i \text{ for all } \sigma) \\
&\qquad\qquad\quad \text{and so } L\Sigma_k^1 \geq 0. \\
&= M \sum_{k-t \leq i_1 + \cdots + i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n}.
\end{aligned}$$

Now one of the terms in the above sum will be $\eta_{k-1} \cdot \overbrace{\eta_0 \cdots \eta_0}^{n-1 \text{ times}} = \eta_{k-1}$ (since $\eta_0 = 1$). Hence

$$\eta_k \geq M \sum_{k-t \leq i_1 + \cdots + i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n}$$

$$\begin{aligned} &\geq M\eta_{k-1} \\ &\geq \eta_{k-1} \qquad \text{since we have forced } M \geq 1. \end{aligned}$$

Hence the η_k are increasing as required. □

The next lemma will help to give us a bound in terms of Σ_k^1 for Σ_k^2.

Lemma 4.40 *For each $k > 1$ and for each $s > 0$*

$$\sum_{i_1+\cdots+i_n=k-s} \eta_{i_1}\cdots\eta_{i_n} \leq 2\Sigma_k^1.$$

Proof. We look at the three separate cases $k < s$, $k = s$ and $k > s$.

Case (i): $k < s$

In this case we have $k - s < 0$ and so, since for no $i_1, \ldots, i_n$ can we have $i_1 + \cdots + i_n < 0$, we have

$$\sum_{i_1+\cdots+i_n=k-s} \eta_{i_1}\cdots\eta_{i_n} = 0.$$

Hence the result is proved since we certainly have $2\Sigma_k^1 \geq 0$.

Case (ii): $k = s$

If $k = s$ then

$$\sum_{i_1+\cdots+i_n=k-s} \eta_{i_1}\cdots\eta_{i_n} = \sum_{i_1+\cdots+i_n=0} \eta_{i_1}\cdots\eta_{i_n} = \eta_0^n = 1$$

and so the proof is reduced to showing that

$$1 \leq 2\Sigma_k^1.$$

But

$$\begin{aligned} \Sigma_k^1 &= \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l\ i_l\neq k}} \eta_{i_1}\cdots\eta_{i_n} \\ &\geq \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l\ i_l\neq k}} \eta_0\cdots\eta_0 \qquad \text{since the } \eta_i \text{ are increasing by Lemma 4.39.} \\ &= \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l\ i_l\neq k}} 1 \qquad \text{since } \eta_0 = 1. \\ &= N \end{aligned}$$

where N here is the number of possible ways of choosing $i_1, \ldots, i_n$ such that $i_1 + \cdots + i_n = k$ and, for all l, $i_l \neq k$. Clearly, since we are assuming $k > 1$, N must be at least 1 (take $i_1 = k - 1$, $i_2 = 1$ and $i_l = 0$ for $l = 2, \ldots, n$ – note we also need here the assumption made at the beginning of Section 4.4 (just after the statement of Theorem 4.11) that $n > 1$) and so the result is proved. (Note that if $k = 1$ then there is no possible way of choosing $i_1, \ldots, i_n$ such that $i_1 + \cdots + i_n = 1$ and $i_l \neq 1$ for all l so $\Sigma^1_k = 0$ in this case and the result does not hold.)

Case (iii): $k > s$

In this case let us firstly set

$$U_1 = \{\langle i_1, \ldots, i_n \rangle \mid i_1 + \cdots + i_n = k - s\}$$

and

$$U_2 = \{\langle i_1, \ldots, i_n \rangle \mid i_1 + \cdots + i_n = k \text{ and } i_l \neq k \text{ for all } l\}.$$

We define a function $f : U_1 \to U_2$ by setting, for each $\langle i_1, i_2, \ldots, i_{n-1}, i_n \rangle \in U_1$,

$$f(\langle i_1, i_2, \ldots, i_{n-1}, i_n \rangle) = \begin{cases} \langle i_1 + s, i_2, \ldots, i_{n-1}, i_n \rangle & \text{if } i_1 < k - s \\ \langle i_1, i_2, \ldots, i_{n-1}, i_n + s \rangle & \text{if } i_1 = k - s. \end{cases}$$

Note that, in the second case in the above definition of f, if $i_1 = k - s$ then $i_l = 0$ for $l = 2, \ldots, n$. In particular $i_n = 0$, so, since we are assuming $s < k$, we have $i_n + s = s < k$. This shows that we do indeed have $f(\langle i_1, \ldots, i_n \rangle) \in U_2$. Also note that we are again employing the assumption that $n > 1$. Now let us extend f to a function $f^+ : U_1 \to \mathbb{R}$ by setting, for each $\langle i_1, \ldots, i_n \rangle \in U_1$,

$$f^+(\langle i_1, \ldots, i_n \rangle) = \eta_{j_1} \cdots \eta_{j_n}, \text{ where } \langle j_1, \ldots, j_n \rangle = f(\langle i_1, \ldots, i_n \rangle).$$

Note that, for any $\langle i_1, \ldots, i_n \rangle \in U_1$, if $\eta_{j_1} \cdots \eta_{j_n} = f^+(\langle i_1, \ldots, i_n \rangle)$ then

$$\eta_{i_1} \cdots \eta_{i_n} \leq \eta_{j_1} \cdots \eta_{j_n}$$

since $i_l \leq j_l$ for each $l = 1, \ldots, n$ (by definition of f) and the η_l's are increasing (by Lemma 4.39). Hence we have

$$\sum_{i_1+\cdots+i_n=k-s} \eta_{i_1} \cdots \eta_{i_n} \leq \sum_{i_1+\cdots+i_n=k-s} f^+(\langle i_1, \ldots, i_n \rangle)$$

and so the result will be proved if we can show

$$\sum_{i_1+\cdots+i_n=k-s} f^+(\langle i_1, \ldots, i_n \rangle) \leq 2\Sigma_k^1.$$

To see that this is true, note that the function f, although it is not injective, has the property that each $\langle j_1, \ldots, j_n \rangle \in U_2$ has at most two pre-images under f. In fact, given $\langle j_1, \ldots, j_n \rangle \in U_2$, we will have that $\langle j_1, \ldots, j_n \rangle = f(\langle j_1 - s, \ldots, j_n \rangle)$ (as long as $j_1 - s \geq 0$), but, additionally, if $j_1 = k - s$ and $j_n = s$ then we will also have $\langle j_1, \ldots, j_n \rangle = f(\langle j_1, \ldots, j_n - s \rangle)$. Clearly, looking at the definition of f, this exhausts the possibilities for constructing a pre-image for $\langle j_1, \ldots, j_n \rangle$. Because f has this property we may see that each term $\eta_{i_1} \cdots \eta_{i_n}$ appearing in Σ_k^1 appears at most twice in the sum

$$\sum_{i_1+\cdots+i_n=k-s} f^+(\langle i_1, \ldots, i_n \rangle)$$

which entails that

$$\sum_{i_1+\cdots+i_n=k-s} f^+(\langle i_1, \ldots, i_n \rangle) \leq 2\Sigma_k^1$$

as required. □

Thus we have, for $k > 1$,

$$\begin{aligned}
M\Sigma_k^2 &= M \sum_{k-t \leq i_1+\cdots+i_n \leq k-1} \eta_{i_1} \cdots \eta_{i_n} \\
&= M \sum_{s=1}^{t} \sum_{i_1+\cdots+i_n=k-s} \eta_{i_1} \cdots \eta_{i_n} \\
&\leq M \sum_{s=1}^{t} 2\Sigma_k^1 \qquad \text{by Lemma 4.40} \\
&= 2Mt\Sigma_k^1.
\end{aligned}$$

We also know, from case (ii) in the proof of Lemma 4.40, that $\Sigma_k^1 \geq 1$. Hence, for $k > 1$, we have

$$\begin{aligned} \eta_k &= K + L\Sigma_k^1 + M\Sigma_k^2 \\ &\leq (K+L)\Sigma_k^1 + M\Sigma_k^2 \\ &\leq W\Sigma_k^1 \end{aligned}$$

where $W = K + L + 2Mt$ — another constant which is independent of k. Thus we now have a bound on the η_k given in terms of Σ_k^1. For technical reasons we now define a new sequence $\eta'_0, \eta'_1, \eta'_2, \ldots$ from the sequence $\eta_0, \eta_1, \eta_2, \ldots$ as follows:

$$\eta'_k = \frac{\eta_k}{\eta_1^k} \text{ for } k \geq 0.$$

Please note that in the above definition η_1^k means "η_1 to the power k". The usage of the superscript k here should not be confused with our earlier notation where we used η_1^k to denote an upper bound for the set $\{|a_1(\sigma)| \mid |\sigma| = k\}$. Thus we have $\eta'_0 = \eta_0 = 1 = \eta'_1$, while, for $k > 1$,

$$\begin{aligned} \eta'_k &= \frac{\eta_k}{\eta_1^k} \\ &\leq \frac{W}{\eta_1^k} \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \;\; i_l \neq k}} \eta_{i_1} \cdots \eta_{i_n} \qquad \text{since } \eta_k \leq W\Sigma_k^1 \\ &= W \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \;\; i_l \neq k}} \frac{\eta_{i_1}}{\eta_1^{i_1}} \cdots \frac{\eta_{i_n}}{\eta_1^{i_n}} \\ &= W \sum_{\substack{i_1+\cdots+i_n=k \\ \forall l \;\; i_l \neq k}} \eta'_{i_1} \cdots \eta'_{i_n}. \end{aligned} \tag{4.53}$$

(The purpose of this "normalisation" of the series η_k is to ensure that $\eta'_1 = 1$ since this is required for the following development.) Recall that our current aim is to prove the convergence, for some real $\lambda > 0$, of the series $\sum_{k=0}^{\infty} \eta_k \lambda^k$. However to do this it will suffice to show that the series $\sum_{k=0}^{\infty} \eta'_k \lambda^k$ has a non-trivial radius of convergence. This is because

$$\sum_{k=0}^{\infty} \eta_k \lambda^k = \sum_{k=0}^{\infty} \eta'_k (\eta_1 \lambda)^k,$$

and so if $\sum_{k=0}^{\infty}\eta'_k\lambda^k$ converges for all $|\lambda| < R$ for some radius of convergence $R > 0$ then $\sum_{k=0}^{\infty}\eta_k\lambda^k$ will converge for all $|\lambda| < \frac{R}{\eta_1}$.

To show the convergence of this new series we shall compare it (for the purpose of applying Proposition 4.36) with the solution to the following formal equations:

$$(Wn+1)\sum_{i=0}^{\infty}\mu_i\lambda^i = W(\sum_{i=0}^{\infty}\mu_i\lambda^i)^n + Wn + 1 - W + \lambda \tag{4.54}$$

with W the constant as defined above. Equating the constant terms on each side gives us

$$(Wn+1)\mu_0 = W\mu_0^n + Wn + 1 - W$$

which in turn shows that

$$\mu_0 = 1. \tag{4.55}$$

Equating the λ coefficients on each side gives us

$$(Wn+1)\mu_1 = W\sum_{i_1+\cdots+i_n=1}\mu_{i_1}\cdots\mu_{i_n} + 1. \tag{4.56}$$

Now

$$\begin{aligned}\sum_{i_1+\cdots+i_n=1}\mu_{i_1}\cdots\mu_{i_n} &= (\mu_1\cdot\mu_0\cdots\mu_0) + (\mu_0\cdot\mu_1\cdots\mu_0) + \cdots + (\mu_0\cdot\mu_0\cdots\mu_1) \\ &= \overbrace{\mu_1 + \mu_1 + \cdots + \mu_1}^{n \text{ times}} \qquad \text{since } \mu_0 = 1.\end{aligned}$$

Hence (4.56) gives us

$$(Wn+1)\mu_1 = Wn\mu_1 + 1$$

and so we must have

$$\mu_1 = 1. \tag{4.57}$$

Now let us take $k > 1$ and equate the k^{th} coefficients of each side of (4.54). We obtain

$$(Wn+1)\mu_k = W\sum_{i_1+\cdots+i_n=k}\mu_{i_1}\cdots\mu_{i_n}. \tag{4.58}$$

We may break down the sum on the right hand side here as follows:

$$\begin{aligned}\sum_{i_1+\cdots+i_n=k} \mu_{i_1}\cdots\mu_{i_n} &= (\mu_k\cdot\mu_0\cdots\mu_0)+(\mu_0\cdot\mu_k\cdots\mu_0)+\cdots \\ &\qquad \cdots+(\mu_0\cdot\mu_0\cdots\mu_k)+\sum_{\substack{i_1+\cdots+i_n=k\\ \forall l\ i_l\neq k}} \mu_{i_1}\cdots\mu_{i_n} \\ &= n\mu_k+\sum_{\substack{i_1+\cdots+i_n=k\\ \forall l\ i_l\neq k}} \mu_{i_1}\cdots\mu_{i_n} \qquad \text{since } \mu_0=1.\end{aligned}$$

Substituting this into (4.58) gives

$$(Wn+1)\mu_k = Wn\mu_k + W\sum_{\substack{i_1+\cdots+i_n=k\\ \forall l\ i_l\neq k}} \mu_{i_1}\cdots\mu_{i_n}$$

whereby we can see that

$$\mu_k = W\sum_{\substack{i_1+\cdots+i_n=k\\ \forall l\ i_l\neq k}} \mu_{i_1}\cdots\mu_{i_n}. \tag{4.59}$$

Lemma 4.41 *For all $k=0,1,2,\ldots$, we have $\eta'_k\leq\mu_k$.*

Proof. The proof is by induction on k. For $k=0$ and $k=1$ we have $\eta'_0=\eta'_1=1$ (by definition of the sequence η'_k) and also $\mu_0=\mu_1=1$ (from (4.55) and (4.57)) and so the result holds in these cases. Now suppose that $k>1$ and that, for $l=0,1,\ldots,k-1$, we have $\eta'_l\leq\mu_l$. Then

$$\begin{aligned}\eta'_k &\leq W\sum_{\substack{i_1+\cdots+i_n=k\\ \forall l\ i_l\neq k}} \eta'_{i_1}\cdots\eta'_{i_n} && \text{from (4.53)} \\ &\leq W\sum_{\substack{i_1+\cdots+i_n=k\\ \forall l\ i_l\neq k}} \mu_{i_1}\cdots\mu_{i_n} && \text{from the inductive hypothesis} \\ &= \mu_k \qquad \text{from (4.59).}\end{aligned}$$

Hence $\eta'_k\leq\mu_k$ as required. □

Hence, in view of Proposition 4.36, to show that the series $\sum_{i=0}^{\infty}\eta'_i\lambda^i$ converges we may now switch our attention to showing that the series $\sum_{i=0}^{\infty}\mu_i\lambda^i$ has a non-trivial radius of convergence. Let us define, for $i=0,1,\ldots$, $\nu_i=\mu_{i+1}$ and set

$T(\lambda) = \sum_{i=1}^{\infty} \mu_i \lambda^{i-1} = \sum_{i=0}^{\infty} \nu_i \lambda^i$. So we have

$$\sum_{i=0}^{\infty} \mu_i \lambda^i = 1 + \lambda T(\lambda).$$

Clearly $\sum_{i=0}^{\infty} \mu_i \lambda^i$ will have a non-trivial radius of convergence iff $T(\lambda)$ has a non-trivial radius of convergence. Substituting $T(\lambda)$ into (4.54) we get

$$(Wn+1)(1+\lambda T(\lambda)) = W(1+\lambda T(\lambda))^n + Wn + 1 - W + \lambda$$

and so

$$(Wn+1)\lambda T(\lambda) = W\left(\sum_{i=1}^{n} \binom{n}{i} \lambda^i T(\lambda)^i\right) + \lambda.$$

Therefore we have

$$(Wn+1)T(\lambda) = W\left(\sum_{i=2}^{n} \binom{n}{i} \lambda^{i-1} T(\lambda)^i\right) + WnT(\lambda) + 1$$

and from here we may see that

$$T(\lambda) = W\left(\sum_{i=2}^{n} \binom{n}{i} \lambda^{i-1} T(\lambda)^i\right) + 1. \tag{4.60}$$

Equating the constant coefficients of this identity gives

$$\mu_1 = \nu_0 = 1 \tag{4.61}$$

while for $k > 0$ we have that, for each $i \geq 2$, the k^{th} coefficient of $\lambda^{i-1}T(\lambda)^i$ is equal to the $(k-i+1)^{\text{th}}$ coefficient of $T(\lambda)^i$ which in turn is equal to

$$\sum_{j_1+\cdots+j_i=k-i+1} \nu_{j_1}\cdots\nu_{j_i}.$$

Hence equating the k^{th} coefficients of (4.60) for $k > 0$ gives us

$$\mu_{k+1} = \nu_k = W\sum_{i=2}^{n} \binom{n}{i} \sum_{j_1+\cdots+j_i=k-i+1} \nu_{j_1}\cdots\nu_{j_i}. \tag{4.62}$$

We next show that the ν_k are increasing for $k \geq 0$. First note that all the ν_k are non-zero, since

$$\begin{aligned} \nu_k = \mu_{k+1} &\geq \eta'_{k+1} && \text{by Lemma 4.41} \\ &= \frac{\eta_{k+1}}{\eta_1^{k+1}} && \text{by definition of } \eta'_{k+1} \end{aligned}$$

and this last term is clearly strictly positive. Hence, for each $i = 2, \ldots, n$, we have

$$\sum_{j_1+\cdots+j_i=k-i+1} \nu_{j_1} \cdots \nu_{j_i} \geq 0.$$

Hence we have

$$\begin{aligned}
\nu_k &= W \sum_{i=2}^{n} \binom{n}{i} \sum_{j_1+\cdots+j_i=k-i+1} \nu_{j_1} \cdots \nu_{j_i} \\
&\geq W \binom{n}{2} \sum_{j_1+j_2=k-1} \nu_{j_1} \nu_{j_2} \\
&= W \binom{n}{2} \left(\nu_{k-1}\nu_0 + \sum_{\substack{j_1+j_2=k-1 \\ j_1 \neq k-1,\ j_2 \neq 0}} \nu_{j_1} \nu_{j_2} \right) \\
&\geq W \binom{n}{2} \nu_{k-1}\nu_0 \\
&= W \binom{n}{2} \nu_{k-1} \qquad \text{since } \nu_0 = 1 \text{ from (4.61)}
\end{aligned}$$

Now it should be clear (by considering the various other constants of which it is composed) that $W \geq 1$ and so $W\binom{n}{2} \geq 1$. Hence

$$\nu_k \geq W \binom{n}{2} \nu_{k-1} \geq \nu_{k-1}$$

which shows the ν_k are increasing as required. Now, since $\nu_0 = 1$, we have, for each $i = 2, \ldots, n$,

$$\begin{aligned}
\sum_{j_1+\cdots+j_i=k-i+1} \nu_{j_1} \cdots \nu_{j_i} &= \sum_{j_1+\cdots+j_i=k-i+1} \nu_{j_1} \cdots \nu_{j_i} \cdot \overbrace{\nu_0 \cdots \nu_0}^{n-i \text{ copies}} \\
&\leq \sum_{j_1+\cdots+j_n=k-i+1} \nu_{j_1} \cdots \nu_{j_n} \\
&\leq \sum_{j_1+\cdots+j_n=k-i+1} \nu_{j_1+i-2} \nu_{j_2} \cdots \nu_{j_n} \\
&\qquad\qquad \text{(since the } \nu_k \text{ are increasing)} \\
&\leq \sum_{j_1+\cdots+j_n=k-1} \nu_{j_1} \cdots \nu_{j_n}
\end{aligned}$$

since each term in the summation in the preceding line appears in the summation in this last line. And so from this together with (4.62) we get, for $k > 0$,

$$\begin{aligned} \nu_k &\leq W \sum_{i=2}^{n} \binom{n}{i} \sum_{j_1+\cdots+j_n=k-1} \nu_{j_1} \cdots \nu_{j_n} \\ &= F \sum_{j_1+\cdots+j_n=k-1} \nu_{j_1} \cdots \nu_{j_n} \end{aligned} \tag{4.63}$$

where

$$F = W \sum_{i=2}^{n} \binom{n}{i} = W(2^n - (n+1))$$

— another constant. Now in order to establish the convergence of the series $\sum_{i=0}^{\infty} \nu_i \lambda^i$ we shall again rely on Proposition 4.36. This time, though, the series we shall compare with will arise from a very different source.

Let $\mathcal{A}$ be an alphabet which consists of F distinct letters $*^1, \ldots, *^F$ together with an additional letter p, i.e.,

$$\mathcal{A} = \{p, *^1, \ldots, *^F\}.$$

We form the language $\mathcal{L}$ over $\mathcal{A}$, inductively, as follows: starting with $l = 0$ we set

$$\mathcal{L}_0 = \{p\}.$$

For $l \geq 0$, having defined $\mathcal{L}_l$, we set

$$\mathcal{L}_{l+1} = \mathcal{L}_l \cup \{*^i\theta_1 \cdots \theta_n \mid i \in \{1, \ldots, F\},\ \theta_j \in \mathcal{L}_l \text{ for } j = 1, \ldots, n\}.$$

Then, finally,

$$\mathcal{L} = \bigcup_{l=0}^{\infty} \mathcal{L}_l.$$

The reader may think of the language $\mathcal{L}$ as the set of propositional sentences built up using a single propositional variable p and a stock of F distinct n-ary connectives $*^1, \ldots, *^F$, but written in a "Polish notation style", i.e., without the use of parentheses or commas and with a connective always placed to the left of its arguments whenever it is applied. For this reason we shall, from now on, refer to any letter of the form $*^i$ as "a connective". We now define a sequence of numbers $\tau_0, \tau_1, \tau_2, \ldots$ by setting, for each $k = 0, 1, 2, \ldots$,

$$\tau_k = \text{the number of distinct strings in } \mathcal{L} \text{ in which the total number of connectives is equal to } k.$$

For example, for $k = 0$, the only sentence which may be formed from p and using no connectives whatsoever is the sentence "p", i.e., the sentence consisting of just p itself. Hence

$$\tau_0 = 1. \tag{4.64}$$

For $k = 1$ the strings of interest will be all the strings which have the form

$$*^{i_1} \overbrace{p \cdots p}^{n}$$

where $*^{i_1}$ is a connective chosen from our stock of n-ary connectives. Since there are F possible choices for the letter $*^{i_1}$ it follows that

$$\tau_1 = F.$$

For $k = 2$ the strings will be those of the form

$$*^{i_2} \phi_1 \cdots \phi_n$$

where $*^{i_2}$ is a connective and there is precisely one occurrence of a single connective, $*^{i_1}$ say, in the whole of $\phi_1, \ldots, \phi_n$. Suppose first of all that this connective occurs in ϕ_1. Then our string would look like

$$*^{i_2} *^{i_1} \overbrace{p \cdots p}^{n} \overbrace{p \cdots p}^{n-1}.$$

Given that there are F possible choices for both $*^{i_1}$ and $*^{i_2}$, there are F^2 possibilities for the above string. Similarly if the connective $*^{i_1}$ appears in ϕ_2 we get another F^2 possibilities, if it appears in ϕ_3 we get another F^2 possibilities, and so on up to $*^{i_1}$ appearing in ϕ_n. Hence the total number of strings that may be formed just from p and two connectives is

$$\tau_2 = nF^2.$$

In fact it turns out that the sequence $\tau_0, \tau_1, \tau_2, \ldots$ may be defined inductively and that, for a general $k \geq 1$,

$$\tau_k = F \sum_{j_1+\cdots+j_n=k-1} \tau_{j_1} \cdots \tau_{j_n}. \tag{4.65}$$

(Note that this does indeed hold for the cases $k = 1$ and $k = 2$ already considered above.) To see this, suppose we have established $\tau_0, \tau_1, \ldots, \tau_{k-1}$. Now note that the strings which may be formed in $\mathcal{L}$ using k n-ary connectives must take the form

$$*^{i_k}\theta_1 \cdots \theta_n$$

where $*^{i_k}$ is a connective, each θ_i is a string in $\mathcal{L}$, and the total number of connectives appearing in the whole of $\theta_1, \ldots, \theta_n$ is equal to $k-1$, say there are j_1 connectives appearing in θ_1, j_2 connectives in θ_2 and so on up to j_n connectives appearing in θ_n ($j_1 + \cdots + j_n = k-1$). Since we have already found $\tau_0, \ldots, \tau_{k-1}$ we know that there are τ_{j_i} possibilities for θ_i ($i = 1, \ldots, n$) and so, remembering also that there are F choices for $*^{i_k}$, there must be a total of $F \cdot \tau_{j_1} \cdots \tau_{j_n}$ choices for our original string. Summing over all the possible distributions $\langle j_1, \ldots, j_n \rangle$ of connectives gives the result.

Lemma 4.42 *For all $k = 0, 1, 2, \ldots$, we have $\nu_k \leq \tau_k$.*

Proof. The proof is by induction on k. For $k = 0$ we have $\nu_0 = 1 = \tau_0$ from (4.61) and (4.64) and so the result certainly holds in this case. Now suppose $k > 0$ and assume that for $l = 0, 1, \ldots, k-1$ we have $\nu_l \leq \tau_l$. Then

$$\nu_k \;\leq\; F \sum_{j_1+\cdots+j_n=k-1} \nu_{j_1} \cdots \nu_{j_n} \qquad \text{from (4.63)}$$

$$\begin{aligned} &\leq F \sum_{j_1+\cdots+j_n=k-1} \tau_{j_1}\cdots\tau_{j_n} && \text{from the inductive hypothesis} \\ &= \tau_k && \text{from (4.65).} \end{aligned}$$

Hence $\nu_k \leq \tau_k$ as required. □

By Lemma 4.42 and Proposition 4.36, to show that $\sum_{k=0}^{\infty} \nu_k \lambda^k$ has a non-trivial radius of convergence it is enough to show that $\sum_{k=0}^{\infty} \tau_k \lambda^k$ has a non-trivial radius of convergence, and it is this last series whose convergence we will now establish directly (albeit with one more use of Proposition 4.36). We begin by finding yet another upper bound, this time for τ_k. We need the following result.

Lemma 4.43 *Let $k \geq 0$ and let θ be a string in $\mathcal{L}$ in which the total number of occurrences of all connectives is equal to k. Then the number of occurrences of p in θ is equal to $k(n-1)+1$.*

Proof. The proof is by induction on k. For $k = 0$ the only possible choice for θ is p itself and so the number of occurrences of p in any string in $\mathcal{L}$ formed only from p and no connectives is equal to $1 = k(n-1)+1$ as required. So now, for our induction hypothesis, let us assume that $k > 0$ and that, for all $l < k$, the number of times p occurs in any sentence formed from p and l n-ary connectives chosen from our stock of F connectives is equal to $l(n-1)+1$. Suppose now that θ is a string in $\mathcal{L}$ formed with k connectives. Then θ must be of the form

$$*^{i_k}\phi_1\cdots\phi_n$$

where $*^{i_k}$ is a connective, each ϕ_i is a string in $\mathcal{L}$, and the total number of connectives appearing in the whole of $\phi_1,\ldots,\phi_n$ is equal to $k-1$, say there are j_1 connectives appearing in ϕ_1, j_2 connectives appearing in ϕ_2, and so on up to j_n connectives appearing in ϕ_n ($j_1+\cdots+j_n = k-1$). Clearly the total number of occurrences of p in θ is equal to the sum of the number of p's occurring in each of the ϕ_i's. But, by inductive hypothesis, this number is equal to

$$\begin{aligned}\sum_{i=1}^{n} j_i(n-1)+1 &= \left((n-1)\sum_{i=1}^{n} j_i\right)+n \\ &= (n-1)(k-1)+n \qquad \text{since } j_1+\cdots+j_n = k-1 \\ &= k(n-1)+1\end{aligned}$$

as required. □

A corollary of this result is that, given k and θ as in the statement, the length of θ is equal to $kn+1$, since it equals the number of occurrences of p (which is $k(n-1)+1$, by the above lemma) plus the number of connectives (which is k). Hence the set of those strings in $\mathcal{L}$ which contain a total number of k connectives taken from our stock of F distinct n-ary connectives is a subset of the set of *all* strings γ of length $kn+1$ from the alphabet $\mathcal{A}$ in which the total number of connective is equal to k, but where those connectives occurring in γ may be distributed freely throughout γ without strict adherence to the "formation rules" of $\mathcal{L}$. Hence the cardinality of this latter set, which we shall call Γ, serves as an upper bound for τ_k. The question is, what is the cardinality $|\Gamma|$ of Γ? To help us find out, let us consider how we might construct a string γ in order for it to be admitted to Γ. We start off by imagining γ in its embryonic state as a sequence of spaces numbered from 1 to $kn+1$ which are each to be filled with either a connective or the letter p. We then suppose we are given a k-tuple $\langle *^{i_1}, \ldots, *^{i_k}\rangle$ of connectives, taken from our stock of F n-ary connectives, which represents the k connectives which are to appear in γ. To this k-tuple of connectives we assign a k-tuple $\langle j_1, \ldots, j_k\rangle$ of numbers which satisfy $1 \le j_1 < \cdots < j_n \le kn+1$ with the intention that, for each $l = 1, \ldots, k$, the connective $*^{i_l}$ should be placed in that space in γ which is numbered j_l. The string γ is then completed by filling up the remaining spaces, i.e., all those spaces that are not numbered j_l for any $l = 1, \ldots, k$, with p. Now, for each tuple of connectives $\langle *^{i_1}, \ldots, *^{i_k}\rangle$, the number of strings formable in this way is equal to the number of different ways we may choose the tuple $\langle j_1, \ldots, j_k \rangle$, i.e., the number of different ways that these connectives may be distributed throughout γ. But this is simply equal to the number of different ways of choosing a set of k numbers from the set $\{1, \ldots, kn+1\}$ which is, of course, equal to

$$\binom{kn+1}{k}.$$

And so, given also that there are F^k different choices for $\langle *^{i_1}, \ldots, *^{i_k} \rangle$, we get

$$|\Gamma| = F^k \binom{kn+1}{k}.$$

Hence

$$\begin{aligned} \tau_k &\leq F^k \binom{kn+1}{k} \\ &= F^k \frac{(kn+1) \cdot kn \cdots (kn+2-k)}{k!} \\ &\leq F^k \frac{(kn+1)^k}{k!}. \end{aligned} \tag{4.66}$$

Now according to Stirling's formula we have

$$k! \sim \sqrt{2\pi k}\ k^k e^{-k},$$

i.e.,

$$k! = \sqrt{2\pi k}\ k^k e^{-k} c_k$$

where $c_k \to 1$ as $k \to \infty$. Substituting this into (4.66) gives us

$$\tau_k \leq F^k \frac{(kn+1)^k e^k}{\sqrt{2\pi k}\ k^k c_k}. \tag{4.67}$$

Now, since $c_k \to 1$ and $\sqrt{2\pi k} \to \infty$ as $k \to \infty$, we have $\sqrt{2\pi k}\ c_k \to \infty$ as $k \to \infty$. Hence there exists $K \geq 0$ such that

$$k \geq K \text{ implies } \sqrt{2\pi k}\ c_k \geq 1.$$

Thus, from (4.67), for $k \geq K$,

$$\begin{aligned} \tau_k &\leq \frac{F^k(kn+1)^k e^k}{k^k} \\ &= (Fe)^k \left(\frac{kn+1}{k}\right)^k \\ &= (Fen)^k \left(1+\frac{1}{kn}\right)^k. \end{aligned} \tag{4.68}$$

Now we know that

$$\lim_{k\to\infty}\left(1+\frac{1}{kn}\right)^{kn} = e$$

and, furthermore, since the sequence $\left(1+\frac{1}{kn}\right)^{kn}$ is increasing for $k = 0, 1, 2, \ldots$, that

$$e = \sup\{\left(1+\frac{1}{kn}\right)^{kn} \mid k = 0, 1, 2, \ldots\}.$$

Hence, for all $k = 0, 1, 2, \ldots$,

$$\left(1+\frac{1}{kn}\right)^{kn} \leq e,$$

so

$$\begin{aligned} \left(1+\frac{1}{kn}\right)^k &= \left(\left(1+\frac{1}{kn}\right)^{kn}\right)^{\frac{1}{n}} \\ &\leq e^{\frac{1}{n}}. \end{aligned}$$

Therefore, from (4.68), for $k \geq K$ we get

$$\tau_k \leq e^{\frac{1}{n}}(Fen)^k \tag{4.69}$$

Lemma 4.44 *The series $\sum_{i=0}^{\infty}\tau_k\lambda^k$ has a non-trivial radius of convergence.*

Proof. The geometric series

$$e^{\frac{1}{n}}\sum_{k=0}^{\infty}(Fen\lambda)^k$$

converges for all λ such that $|\lambda| < \frac{1}{Fen}$. Hence, from (4.69) and from Proposition 4.36, so does the series $\sum_{k=0}^{\infty}\tau_k\lambda^k$. Hence it has a non-trivial radius of convergence as required. □

We are now finally in a position to be able to prove the convergence of all the series $z_\infty(\sigma)$.

Lemma 4.45 *There exists a real number $R > 0$ such that, simultaneously, for all sequences of literals σ, the series $\sum_{i=0}^{\infty} a_i(\sigma)\lambda^i$ (and thus also the series $z_\infty(\sigma) = \lambda^{|\sigma|-1}\sum_{i=0}^{\infty} a_i(\sigma)\lambda^i$) converges for any $|\lambda| < R$.*

Proof. By Lemma 4.44 the series $\sum_{i=0}^{\infty} \tau_i\lambda^i$ has a non-trivial radius of convergence. Hence, using Lemma 4.42 and Proposition 4.36, so too does the series $\sum_{i=0}^{\infty} \nu_i\lambda^i$. Recall that $\nu_i = \mu_{i+1}$ for each $i = 0, 1, 2, \ldots$ Hence, since

$$\sum_{i=0}^{\infty} \mu_i\lambda^i = \mu_0 + \lambda\sum_{i=0}^{\infty} \tau_i\lambda^i,$$

we have that $\sum_{i=0}^{\infty} \mu_i\lambda^i$ also has a non-trivial radius of convergence. Then, by Lemma 4.41 and Proposition 4.36, the series $\sum_{i=0}^{\infty} \eta_i'\lambda^i$ must have a non-trivial radius of convergence. As indicated earlier, this implies that the series $\sum_{i=0}^{\infty} \eta_i\lambda^i$ has a non-trivial radius of convergence and so, since we have $|a_i(\sigma)| \le \eta_i$ for all sequences of literals σ, the series $\sum_{i=0}^{\infty} |a_i(\sigma)|\lambda^i$ also converges non-trivially. We conclude by Proposition 4.37. □

The above Lemma 4.45 then completes the second stage of our proof of Theorem 4.1. It allows us to drop all reference to infinitesimals and λ from the statement of Theorem 4.11 to arrive at

Theorem 4.46 *Given a language $L = \{p_1, \ldots, p_n\}$, if the function $Bel : SL \to [0, 1]$ is given by a standard pre-ent over L and if, for all $\theta, \phi \in SL$, $Bel(\theta\wedge\phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$, then there exists a standard (i.e., whose potentials are standard non-negative real numbers) almost-ent z (over a larger language than L) such that, for all $\theta \in SL$, $Bel^z(\theta) = Bel(\theta)$.*

Proof. Now immediate from Lemma 4.45. We can ensure that $z_\infty(\sigma) \in [0, \infty)$ by choosing λ such that $0 < \lambda < R$ where R was found in Lemma 4.45. □

All that needs to be done from here to achieve a proof of Theorem 4.1 is to show how we can replace "almost-ent" in the above by "ent". It is to this – the final stage in our proof of Theorem 4.1 – that we turn to next.

4.7 Stage 3 – Converting z_∞ into an Ent

The work in the previous two sections, which together comprised the first two stages of our proof of Theorem 4.1, has shown that, given a belief-function Bel which was given by a standard pre-ent over the language L and which satisfied, for all $\theta, \phi \in SL$, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$, there exists a standard almost-ent $z_\infty : SL^+ \to [0, \infty)$ for which $Bel^{z_\infty}(\theta) = Bel(\theta)$ for all $\theta \in SL$. This almost-ent was defined over a language L^+ which extended L. However, the purpose in defining almost-ents was merely to provide a stepping stone for showing how any such function Bel can be given by a standard *ent*. We showed at the end of Section 4.5 that the almost-ent z_∞ failed to be an ent over L^+. In this section we complete the proof of Theorem 4.1 by showing how z_∞ can be converted into an ent which will give the same beliefs as z_∞ to all sentences in our original language L. This ent, which will be denoted by y_∞, will, like z_∞, be defined over a language which contains L, though this language will be different from L^+.

To begin with, we choose a number $d \in \mathbb{N}$ such that $2^d \geq 2n$. For each $0 \leq j < n$, given distinct $p_{i_1}, \ldots, p_{i_j} \in L$ and some $\epsilon_1, \ldots, \epsilon_j \in \{0, 1\}$, if $Bel(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j}) \neq 0$ then let the sets $\bar{S}(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} p^\epsilon)$, as the pair $\langle p, \epsilon \rangle$ ranges over the set $(L - \{p_{i_1}, \ldots, p_{i_j}\}) \times \{0, 1\}$, form a partition of the set $\{0, 1\}^d$ (so $\bigcup_{\langle p, \epsilon \rangle} \bar{S}(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} p^\epsilon) = \{0, 1\}^d$ and $\bar{S}(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} p^\epsilon) \cap \bar{S}(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} q^\delta) \neq \emptyset$ implies $p^\epsilon = q^\delta$) such that

$$\bar{S}(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} p^\epsilon) = \emptyset \text{ iff } Bel(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} p^\epsilon) = 0. \tag{4.70}$$

If $Bel(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j}) = 0$ then just define $\bar{S}(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} p^{\epsilon}) = \emptyset$ for all $\langle p, \epsilon \rangle$, so (4.70) still holds since in this case $Bel(p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j} p^{\epsilon}) = 0$ for all $\langle p, \epsilon \rangle$. Let us now define the language L^* over which our ent y_∞ will be defined.

$$L^* = L \cup \{u_{k,r} \mid k = 1, \ldots, n;\ r = 1, \ldots, d\}.$$

So L^* consists of all the propositional variables in L together with a set of new propositional variables consisting of one variable for each pair $\langle k, r \rangle$ such that $k = 1, \ldots, n$ and $r = 1, \ldots, d$. For each non-empty sequence of literals $q_1 \cdots q_j$ we define a set of scenarios (over L^*) $S(q_1 \cdots q_j)$ as follows:

$$S(q_1 \cdots q_j) = \Big\{ \{u_{j,r}^{\epsilon_r} \mid r = 1, \ldots, d\} \mid \langle \epsilon_1, \ldots, \epsilon_d \rangle \in \bar{S}(q_1 \cdots q_j) \Big\}.$$

Note that we have $S(q_1 \cdots q_j) = \emptyset$ iff $\bar{S}(q_1 \cdots q_j) = \emptyset$. We then define the set $T(q_1 \cdots q_j)$ for all (possibly empty) sequences of literals $q_1 \cdots q_j$ by

$$T(q_1 \cdots q_j) = S(q_1) \times S(q_1 q_2) \times \cdots \times S(q_1 \cdots q_j),$$

(so following this definition $T(\emptyset) = \{\emptyset\}$). Note that we have $T(q_1 \cdots q_j) \neq \emptyset$ iff $S(q_1 \cdots q_i) \neq \emptyset$ for each $i = 1, \ldots, j$ iff $\bar{S}(q_1 \cdots q_i) \neq \emptyset$ for each $i = 1, \ldots, j$. Hence, from (4.70) we may see that, for each sequence of literals $q_1 \cdots q_j$,

$$T(q_1 \cdots q_j) = \emptyset \text{ iff } Bel(q_1 \cdots q_j) = 0. \tag{4.71}$$

Our next step is to describe what sorts of scenarios over L^* will be assigned non-zero potential by y_∞. First of all, all singleton scenarios of the form $\{u_{k,r}\}$ will be assigned non-zero potential (which may be any fixed arbitrary non-zero real number). This is to ensure that, for any scenario $s \in WL^*$ not deciding a variable $u_{k,r} \in L^* - L$, there exists a scenario with non-zero potential (namely $\{u_{k,r}\}$) which does decide that variable and which is consistent with s. The only other scenarios to get non-zero potential will be those of the following form, for each non-empty sequence of literals $q_1 \cdots q_j$ such that $Bel(q_1 \cdots q_j) \neq 0$ and for each $\vec{\mathcal{U}} = \langle \mathcal{U}_1, \ldots, \mathcal{U}_j \rangle \in T(q_1 \cdots q_j)$,

$$s(q_1 \cdots q_j; \vec{\mathcal{U}}) = \mathcal{U}_1 \cup \mathcal{U}_2 \cup \ldots \cup \mathcal{U}_j \cup \{q_j\}.$$

Note in the above that, for each $q_1 \cdots q_j$, there will in general be several choices for $\vec{\mathcal{U}}$ in $s(q_1 \cdots q_j; \vec{\mathcal{U}})$ (since it may be that $|\bar{S}(q_1 \cdots q_i)| > 1$ for some $i = 1, \ldots, j$ and so therefore $|S(q_1 \cdots q_i)| > 1$). For each such scenario the potential given to it by y_∞, which we shall denote by $y_\infty(q_1 \cdots q_j; \vec{\mathcal{U}})$, is set as follows:

$$y_\infty(q_1 \cdots q_j; \vec{\mathcal{U}}) = \frac{z_\infty(q_1 \cdots q_j)}{|T(q_1 \cdots q_j)|} = \frac{z_\infty(q_1 \cdots q_j)}{\prod_{i=1}^{j} |S(q_1 \cdots q_i)|}$$

The reader may notice immediately that $y_\infty(q_1 \cdots q_j; \vec{\mathcal{U}})$ is actually independent of $\vec{\mathcal{U}}$. We also note that $Bel(q_1 \cdots q_j) \neq 0$ ensures that both the numerator and the denominator in the above expression are non-zero. The numerator is non-zero by Lemma 4.32 while the denominator is non-zero by equation (4.71). We now show that y_∞ is, unlike z_∞, a standard ent.

Lemma 4.47 *y_∞ is a standard ent (over L^*).*

Proof. We must check that, for each scenario $s \in WL^*$ and for each $p \in L^*$ such that $\pm p \notin s$, there exists a scenario $t \in WL^*$ which is consistent with s and is such that $\pm p \in t$ and t has non-zero potential according to y_∞. As we remarked above, if $p \in L^* - L$ then such a scenario always exists. Thus let us assume that $p \in L$, i.e., that $p = p_k$ where $k \in \{1, \ldots, n\}$. The scenario s will, generally speaking, contain a mixture of, on the one hand, literals from propositional variables in $L^* - L$ and, on the other, literals from variables in L. Let us denote the set of literals in s from this latter group by s_L. Now for each $i = 1, \ldots, n$ we define a set of literals (over $L^* - L$) $\mathcal{U}_i$ as follows:

$$\mathcal{U}_i = \{u_{i,r}^{\delta} \mid u_{i,r}^{\delta} \in s\} \cup \{u_{i,r} \mid \pm u_{i,r} \notin s\}.$$

So $\mathcal{U}_i$ is obtained by firstly including all the literals in s which are of the form $\pm u_{i,r}$ for some r. Then, for each $r = 1, \ldots, d$, if the propositional variable $u_{i,r}$ is not decided by s then we simply add it to the set. In this way the set $\mathcal{U}_i$ decides all the variables of the form $u_{i,r}$. We define the scenario $s' \supseteq s$ by

$$s' = \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_n \cup s_L.$$

(So s' decides all the propositional variables in $L^* - L$.) We shall now check for the existence of a scenario which decides p_k, has non-zero potential, and is consistent with this larger scenario s'. This will clearly suffice since any such scenario will also be consistent with s. Let us assume that, for each $i = 1, \ldots, n$, the d-tuple $\langle \epsilon_1^i, \ldots, \epsilon_d^i \rangle$ is such that

$$\mathcal{U}_i = \{u_{i,r}^{\epsilon_r^i} \mid r = 1, \ldots, d\}.$$

Now, by the construction, there is a literal q_1 from L such that $Bel(q_1) \neq 0$ and $\langle \epsilon_1^1, \ldots, \epsilon_d^1 \rangle \in \bar{S}(q_1)$, equivalently $\mathcal{U}_1 \in S(q_1)$ (in fact this literal will be unique, though this does not matter for the present proof). In turn there exists a unique literal q_2 from L such that $Bel(q_1 q_2) \neq 0$ and $\langle \epsilon_1^2, \ldots, \epsilon_d^2 \rangle \in \bar{S}(q_1 q_2)$, equivalently $\mathcal{U}_2 \in S(q_1 q_2)$. Continuing in this way we will, eventually, arrive at some q_r such that $q_r = \pm p_k$, $\mathcal{U}_r \in S(q_1 \cdots q_r)$ and $Bel(q_1 \cdots q_r) \neq 0$. But then the scenario $s(q_1 \cdots q_r; \langle \mathcal{U}_1, \ldots, \mathcal{U}_r \rangle)$ is consistent with s', decides p_k, and, since $Bel(q_1 \cdots q_r) \neq 0$, has non-zero potential as required. Finally y_∞ is clearly standard since z_∞ is. □

Having established the ent-hood of y_∞ we would now like to show that y_∞ gives the same belief values to sentences from SL as z_∞, and hence that we can replace *almost-ent* in the statement of Theorem 4.46 by *ent*. To help us do this, we now examine how y_∞ computes its belief in sentences consisting of a sequence of literals $q_1 \cdots q_j$ from L. From Section 4.3 we know that, for any almost-ent (and hence for any ent) z (over L^*),

$$Bel^z(q_1 \cdots q_j) = \sum_{\vec{r} \neq_z 0} \prod_{i=1}^{j} \Theta^z(\bigcup_{k<i} r_k \xrightarrow{q_i} r_i) \tag{4.72}$$

where the sum is over all scenario paths (over L^*) $\vec{r} = r_1, \ldots, r_j$ which are non-zero for z (see Definition 4.10), and the terms $\Theta^z(\bigcup_{k<i} r_k \xrightarrow{q_i} r_i)$ are given by

$$\Theta^z(\bigcup_{k<i} r_k \xrightarrow{q_i} r_i) = \begin{cases} \dfrac{z_{r_i}}{\sum\{z_t \mid \bigcup_{k<i} r_k \cup t \text{ consistent, } \pm q_i \in t\}} & \text{if } r_i \neq \emptyset \\ 1 & \text{if } r_i = \emptyset. \end{cases}$$

Our first step in determining $Bel^{y_\infty}(q_1 \cdots q_j)$ is to identify the form of the scenario paths over L^* for $q_1 \cdots q_j$ which are non-zero for y_∞. The next lemma, which should be compared to Lemma 4.13, will make it easier to identify which scenarios with non-zero potential are consistent with a given scenario $s(\sigma;\vec{\mathcal{U}})$ with non-zero potential. In all what follows, given $\vec{\mathcal{U}} = \langle \mathcal{U}_1, \ldots, \mathcal{U}_{|\sigma|} \rangle \in T(\sigma)$ and $\vec{\mathcal{V}} = \langle \mathcal{V}_1, \ldots, \mathcal{V}_{|\tau|} \rangle \in T(\tau)$ for some, possibly empty, sequences of literals σ, τ such that $|\sigma| \leq |\tau|$, we shall write $\vec{\mathcal{U}} \subseteq \vec{\mathcal{V}}$ if $\mathcal{U}_i = \mathcal{V}_i$ for all $i \leq |\sigma|$. We shall write $\langle \sigma, \vec{\mathcal{U}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle$ to mean that $\sigma \subseteq \tau$ and $\vec{\mathcal{U}} \subseteq \vec{\mathcal{V}}$.

Lemma 4.48 *Let σ, τ be non-empty sequences of literals such that $Bel(\sigma) \neq 0 \neq Bel(\tau)$, and let $\vec{\mathcal{U}} \in T(\sigma)$ and $\vec{\mathcal{V}} \in T(\tau)$. Then $s(\sigma;\vec{\mathcal{U}}) \cup s(\tau;\vec{\mathcal{V}})$ is consistent iff either $\langle \sigma, \vec{\mathcal{U}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle$ or $\langle \tau, \vec{\mathcal{V}} \rangle \subseteq \langle \sigma, \vec{\mathcal{U}} \rangle$.*

Proof. We suppose that $\sigma = q_1 \cdots q_j$, $\tau = r_1 \cdots r_k$, $\vec{\mathcal{U}} = \langle \mathcal{U}_1, \ldots, \mathcal{U}_j \rangle$ and $\vec{\mathcal{V}} = \langle \mathcal{V}_1, \ldots, \mathcal{V}_k \rangle$. So we have

$$s(\sigma;\vec{\mathcal{U}}) = \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_j \cup \{q_j\} \text{ and } s(\tau;\vec{\mathcal{V}}) = \mathcal{V}_1 \cup \ldots \cup \mathcal{V}_k \cup \{r_k\}.$$

We first show the "only if" direction. Without loss of generality we assume that $j \leq k$. We will show that (under the assumption just made) $s(\sigma;\vec{\mathcal{U}}) \cup s(\tau;\vec{\mathcal{V}})$ is consistent implies $\langle \sigma, \vec{\mathcal{U}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle$. For each $i = 1, \ldots, j$, since $\mathcal{U}_i \in S(q_1 \cdots q_i)$ we have

$$\mathcal{U}_i = \{u_{i,r}^{\epsilon_{i,r}} \mid r = 1, \ldots, d\}$$

for some $\langle \epsilon_{i,1}, \ldots, \epsilon_{i,d} \rangle \in \bar{S}(q_1 \cdots q_i)$, while similarly, since $\mathcal{V}_i \in S(r_1 \cdots r_i)$,

$$\mathcal{V}_i = \{u_{i,r}^{\delta_{i,r}} \mid r = 1, \ldots, d\}$$

for some $\langle \delta_{i,1}, \ldots, \delta_{i,d} \rangle \in \bar{S}(r_1 \cdots r_i)$. Hence straight away we see that, for consistency, we must have $\mathcal{U}_i = \mathcal{V}_i$ for $i = 1, \ldots, j$, i.e., $\vec{\mathcal{U}} \subseteq \vec{\mathcal{V}}$. It remains to show that $\sigma \subseteq \tau$. We will do this by showing, by induction on i, that $q_i = r_i$ for $i = 1, \ldots, j$. For $i = 1$ we have

$$\mathcal{U}_1 = \{u_{1,r}^{\epsilon_{1,r}} \mid r = 1, \ldots, d\}$$

for some $\langle \epsilon_{1,1}, \ldots, \epsilon_{1,d} \rangle \in \bar{S}(q_1)$ and

$$\mathcal{V}_1 = \{u_{1,r}^{\delta_{1,r}} \mid r = 1, \ldots, d\}$$

for some $\langle \delta_{1,1}, \ldots, \delta_{1,d} \rangle \in \bar{S}(r_1)$. But, since, as we have already said that to keep consistency we must have $\mathcal{U}_1 = \mathcal{V}_1$, we must have $\langle \epsilon_{1,1}, \ldots, \epsilon_{1,d} \rangle = \langle \delta_{1,1}, \ldots, \delta_{1,d} \rangle$ and so $\langle \epsilon_{1,1}, \ldots, \epsilon_{1,d} \rangle \in \bar{S}(q_1) \cap \bar{S}(r_1)$. Hence, since the sets $\bar{S}(p^\epsilon)$ form a partition of $\{0,1\}^d$, it must be the case that $q_1 = r_1$ as required. Now suppose $1 < l \leq j$ and that, for inductive hypothesis, $q_i = r_i$ for all $i < l$. We must show that $q_l = r_l$. Again we have

$$\mathcal{U}_l = \{u_{l,r}^{\epsilon_{l,r}} \mid r = 1, \ldots, d\}$$

for some $\langle \epsilon_{l,1}, \ldots, \epsilon_{l,d} \rangle \in \bar{S}(q_1 \cdots q_{l-1} q_l)$ and

$$\mathcal{V}_l = \{u_{l,r}^{\delta_{l,r}} \mid r = 1, \ldots, d\}$$

for some $\langle \delta_{l,1}, \ldots, \delta_{l,d} \rangle \in \bar{S}(r_1 \cdots r_{l-1} r_l) = \bar{S}(q_1 \cdots q_{l-1} r_l)$, and again, since $\mathcal{U}_l = \mathcal{V}_l$, we have $\langle \epsilon_{l,1}, \ldots, \epsilon_{l,d} \rangle = \langle \delta_{l,1}, \ldots, \delta_{l,d} \rangle$ which gives $\langle \epsilon_{l,1}, \ldots, \epsilon_{l,d} \rangle \in \bar{S}(q_1 \cdots q_{l-1} q_l)$ $\cap\, \bar{S}(q_1 \cdots q_{l-1} r_l)$. Now $Bel(q_1 \cdots q_{l-1}) \neq 0$, since otherwise we would have $Bel(q_1 \cdots q_j) = 0$ which contradicts one of the original assumptions of the lemma. Hence the sets $\bar{S}(q_1 \cdots q_{l-1} p^\epsilon)$, as p^ϵ varies, form a partition of $\{0,1\}^d$ and hence we must have $q_l = r_l$. This completes the inductive step. Thus we have shown that $\langle \sigma, \vec{\mathcal{U}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle$ which proves the "only if" direction.

To show the "if" direction, by symmetry, we need only look at the case where $\langle \sigma, \vec{\mathcal{U}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle$. Then $\sigma \subseteq \tau$ and $\vec{\mathcal{U}} \subseteq \vec{\mathcal{V}}$ which gives $\mathcal{U}_i = \mathcal{V}_i$ for $i = 1, \ldots, j$ and so

$$s(\tau; \vec{\mathcal{V}}) = \mathcal{V}_1 \cup \ldots \cup \mathcal{V}_k \cup \{r_k\} = \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_j \cup \mathcal{V}_{j+1} \cup \ldots \cup \mathcal{V}_k \cup \{r_k\}.$$

Hence

$$s(\sigma; \vec{\mathcal{U}}) \cup s(\tau; \vec{\mathcal{V}}) = \mathcal{U}_1 \cup \ldots \cup \mathcal{U}_j \cup \mathcal{V}_{j+1} \cup \ldots \cup \mathcal{V}_k \cup \{q_j, r_k\}.$$

Now clearly

$$\mathcal{U}_1 \cup \ldots \cup \mathcal{U}_j \cup \mathcal{V}_{j+1} \cup \ldots \cup \mathcal{V}_k$$

is consistent, while $\sigma \subseteq \tau$ implies either that $q_j = r_k$ (if $j = k$) or that q_j and r_k are literals from distinct propositional variables in L (since otherwise we would have a propositional variable occurring twice during τ). Either way the set $\{q_j, r_k\}$ is consistent. Hence $s(\sigma; \vec{\mathcal{U}}) \cup s(\tau; \vec{\mathcal{V}})$ is consistent as required. $\square$

Recall (Definition 4.10(a)) that a scenario path (over L^*) for a non-empty sequence of literals $q_1 \cdots q_j$ is a sequence of scenarios (over L^*) $\vec{s} = s_1, \ldots, s_j$ such that *(i)* $q_1 \in s_1$, and *(ii)* for each $i \geq 1$, if $q_{i+1} \in \bigcup_{k \leq i} s_k$ then $s_{i+1} = \emptyset$, otherwise s_{i+1} is such that $q_{i+1} \in s_{i+1}$ and $\bigcup_{k \leq i} s_k \cup s_{i+1}$ is consistent. Also recall (Definition 4.10(b)) that such a scenario path for $q_1 \cdots q_j$ is labelled non-zero for y_∞ iff y_∞ assigns non-zero potential to each of the non-empty scenarios amongst $s_1, \ldots, s_j$. To give us an idea of what the scenario paths for $q_1 \cdots q_j$ which are non-zero for y_∞ look like let us now try and construct one. The following explanation should be compared closely with the one in Section 4.5 just after Lemma 4.13, in which we constructed a scenario path for a special almost-ent over the language L^+ such as z_∞.

Firstly the only scenarios which decide q_1 one way or the other and which are given non-zero potential by y_∞ are those of the form $s(\sigma_1; \vec{\mathcal{U}}_1)$ where σ_1 is a sequence of literals which ends with $\pm q_1$, $Bel(\sigma_1) \neq 0$ and $\vec{\mathcal{U}}_1 \in T(\sigma_1)$. Out of these the ones which decide q_1 positively, i.e., include q_1, are those such that σ_1 ends q_1. Given that such a sequence exists and that s_1 is of this form, it is clear that $q_2 \notin s_1$ (since q_1 is the only literal from L which is contained in s_1 by definition of $s(\sigma_1;\vec{\mathcal{U}}_1))$. Hence s_2 is required to contain q_2 and be consistent with s_1. Again the only scenarios which decide q_2 one way or the other and which are given non-zero potential by y_∞ are those of the form $s(\sigma_2;\vec{\mathcal{U}}_2)$ where σ_2 ends $\pm q_2$, $Bel(\sigma_2) \neq 0$ and $\vec{\mathcal{U}}_2 \in T(\sigma_2)$. Of these, by Lemma 4.48, the only ones which are consistent with s_1 are those such that either $\langle \sigma_1,\vec{\mathcal{U}}_1\rangle \subseteq \langle \sigma_2,\vec{\mathcal{U}}_2\rangle$ or $\langle \sigma_2,\vec{\mathcal{U}}_2\rangle \subseteq \langle \sigma_1,\vec{\mathcal{U}}_1\rangle$, with q_2 being decided positively iff σ_2 ends q_2. Hence if q_2 appears in σ_1, say $\sigma_1 = \tau_1 q_2 \cdots$, then the only possible choice for s_2 (indeed the only scenario with non-zero potential which even *decides* q_2 and is consistent with s_1) is $s(\tau_1 q_2;\vec{\mathcal{U}}_1\lceil(|\tau_1|+1))$, i.e., take $\sigma_2 = \tau_1 q_2$ and $\vec{\mathcal{U}}_2$ to be that tuple which consists of just the first $|\tau_1|+1$ entries of $\vec{\mathcal{U}}_1$. If $\overline{q}_2$ appears in σ_1, say $\sigma_1 = \rho_1 \overline{q}_2 \cdots$ then the only possible scenario is $s(\rho_1\overline{q}_2;\vec{\mathcal{U}}_1\lceil(|\rho_1|+1))$ thus in this case there is no scenario consistent with s_1 which decides q_2 positively. If neither q_2 nor $\overline{q}_2$ appear in σ_1 then we must have $s_2 = s(\sigma_2;\vec{\mathcal{U}}_2)$ where $\langle \sigma_1,\vec{\mathcal{U}}_1\rangle \subseteq \langle \sigma_2,\vec{\mathcal{U}}_2\rangle$, σ_2 ends q_2, $\vec{\mathcal{U}}_2 \in T(\sigma_2)$ and $Bel(\sigma_2) \neq 0$ (provided such a σ_2 exists). Now suppose we have found scenarios $s_1, \ldots, s_i$ $(i < j)$ such that $\bigcup_{k\leq i} s_k$ is consistent and which satisfy, for each $k = 1, \ldots, i$, $s_k = s(\sigma_k;\vec{\mathcal{U}}_k)$ where σ_k ends with q_k, $\vec{\mathcal{U}}_k \in T(\sigma_k)$ and $Bel(\sigma_k) \neq 0$ (so $q_k \in s_k$). If we choose l_i such that $|\sigma_{l_i}|$ is maximal amongst $\{|\sigma_k| \mid k = 1, \ldots, i\}$ then, by Lemma 4.48, since $\bigcup_{k\leq i} s_k$ is consistent, we must have $\langle \sigma_k,\vec{\mathcal{U}}_k\rangle \subseteq \langle \sigma_{l_i},\vec{\mathcal{U}}_{l_i}\rangle$ for all $k = 1, \ldots, i$. Clearly $q_{i+1} \notin \bigcup_{k\leq i} s_k$ so we would like to find a scenario s_{i+1} which is consistent with $\bigcup_{k\leq i} s_k$ and which includes q_{i+1}. The only scenarios which decide q_{i+1}, are consistent with $\bigcup_{k\leq i} s_k$ and are given non-zero potential by y_∞ are those of the form $s(\sigma_{i+1};\vec{\mathcal{U}}_{i+1})$ where σ_{i+1} ends $\pm q_{i+1}$, $\vec{\mathcal{U}}_{i+1} \in T(\sigma_{i+1})$, $Bel(\sigma_{i+1}) \neq 0$ and either $\langle \sigma_{i+1},\vec{\mathcal{U}}_{i+1}\rangle \subseteq \langle \sigma_{l_i},\vec{\mathcal{U}}_{l_i}\rangle$ or

$\langle \sigma_{l_i}, \vec{\mathcal{U}}_{l_i} \rangle \subseteq \langle \sigma_{i+1}, \vec{\mathcal{U}}_{i+1} \rangle$; q_{i+1} being decided positively iff σ_{i+1} ends with q_{i+1}. As above, if q_{i+1} appears in σ_{l_i}, say $\sigma_{l_i} = \tau_i q_{i+1} \cdots$, then the only choice for s_{i+1} is to take $\sigma_{i+1} = \tau_i q_{i+1}$ and $\vec{\mathcal{U}}_{i+1} = \vec{\mathcal{U}}_{l_i} \restriction (|\tau_i| + 1)$, while if $\overline{q}_{i+1}$ appears in σ_{l_i} then there is no scenario consistent with $\bigcup_{k \leq i} s_k$ which decides q_{i+1} positively. If $\pm q_{i+1}$ does not appear in σ_{l_i} then we take $\langle \sigma_{i+1}, \vec{\mathcal{U}}_{i+1} \rangle$ to be such that $\langle \sigma_{l_i}, \vec{\mathcal{U}}_{l_i} \rangle \subseteq \langle \sigma_{i+1}, \vec{\mathcal{U}}_{i+1} \rangle$.

In summary, then, it should be clear from the above that the scenario paths for $q_1 \cdots q_j$ which are non-zero for y_∞ are all those scenario paths of the form

$$s(\sigma_1; \vec{\mathcal{U}}_1), s(\sigma_2; \vec{\mathcal{U}}_2), \ldots, s(\sigma_j; \vec{\mathcal{U}}_j)$$

where $\langle \sigma_1, \vec{\mathcal{U}}_1 \rangle, \langle \sigma_2, \vec{\mathcal{U}}_2 \rangle, \ldots, \langle \sigma_j, \vec{\mathcal{U}}_j \rangle$ conform to the following behaviour:

- (i). σ_1 ends with q_1 and $\vec{\mathcal{U}}_1 \in T(\sigma_1)$.
- (ii). For $i \geq 1$, given that l_i is such that $|\sigma_{l_i}|$ is maximal amongst $\{|\sigma_k| \mid k = 1, \ldots, i\}$, if q_{i+1} appears in σ_{l_i}, say $\sigma_{l_i} = \tau_i q_{i+1} \cdots$, then $\sigma_{i+1} = \tau_i q_{i+1}$ and $\vec{\mathcal{U}}_{i+1} = \vec{\mathcal{U}}_{l_i} \restriction (|\tau_i| + 1)$. Otherwise $\langle \sigma_{l_i}, \vec{\mathcal{U}}_{l_i} \rangle \subseteq \langle \sigma_{i+1}, \vec{\mathcal{U}}_{i+1} \rangle$, σ_{i+1} ends q_{i+1} and $\vec{\mathcal{U}}_{i+1} \in T(\sigma_{i+1})$.
- (iii). $Bel(\sigma_i) \neq 0$ for $i = 1, \ldots, j$.

It should be evident that, in fact, in such a sequence $\langle \sigma_1, \vec{\mathcal{U}}_1 \rangle, \ldots, \langle \sigma_j, \vec{\mathcal{U}}_j \rangle$, the sequence $\sigma_1, \ldots, \sigma_j$ forms a n-m sequence path for $q_1 \cdots q_j$ (see Definition 4.15) which, using (iii) above with Lemma 4.32, is non-zero for z_∞. Hence we may rephrase the above and say that the scenario paths for $q_1 \cdots q_j$ which are non-zero for y_∞ are all those scenario paths of the form

$$s(\sigma_1; \vec{\mathcal{U}}_1), s(\sigma_2; \vec{\mathcal{U}}_2), \ldots, s(\sigma_j; \vec{\mathcal{U}}_j)$$

where $\vec{\sigma} = \sigma_1, \ldots, \sigma_j$ is a n-m sequence path for $q_1 \cdots q_j$ which is non-zero for z_∞, i.e., $\vec{\sigma} \in \widehat{N}_{z_\infty}(q_1 \cdots q_j)$, and $\vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_j$ satisfy the following:

- (a). $\vec{\mathcal{U}}_1 \in T(\sigma_1)$.

- (b). For $1 \leq i \leq j-1$, if $\sigma_{i+1} \subseteq \sigma_{l_i}$ then $\vec{\mathcal{U}}_{i+1} = \vec{\mathcal{U}}_{l_i} \restriction |\sigma_{i+1}|$. Otherwise $\vec{\mathcal{U}}_{i+1}$ satisfies $\vec{\mathcal{U}}_{l_i} \subseteq \vec{\mathcal{U}}_{i+1}$ and $\vec{\mathcal{U}}_{i+1} \in T(\sigma_{i+1})$.

Let us denote by $X(\vec{\sigma})$ the set of j-tuples $\langle \vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_j \rangle$ which satisfy conditions (a) and (b) above.

Hence, given a non-empty sequence of literals $q_1 \cdots q_j$, we now know what the n-m sequence paths (over L^*) for $q_1 \cdots q_j$ which are non-zero for y_∞ look like, and we have established the close link between these paths and the set of n-m sequence paths (over L^+) for $q_1 \cdots q_j$ which are non-zero for z_∞. The next lemma further indicates the connection between y_∞ and z_∞.

Lemma 4.49 *Let σ and τ be sequences of literals such that $\sigma \subseteq \tau$ and $\sigma \neq \tau$ and $Bel(\sigma) \neq 0 \neq Bel(\tau)$, and let $\vec{\mathcal{U}} \in T(\sigma)$. Then*

$$\sum_{\substack{\vec{\mathcal{V}} \text{ such that} \\ \vec{\mathcal{U}} \subseteq \vec{\mathcal{V}},\ \vec{\mathcal{V}} \in T(\tau)}} y_\infty(\tau; \vec{\mathcal{V}}) = z_\infty(\tau) \cdot |T(\sigma)|^{-1}.$$

Proof. By definition of $y_\infty(\tau; \vec{\mathcal{V}})$ we have

$$\begin{aligned} \sum_{\substack{\vec{\mathcal{V}} \text{ such that} \\ \vec{\mathcal{U}} \subseteq \vec{\mathcal{V}},\ \vec{\mathcal{V}} \in T(\tau)}} y_\infty(\tau; \vec{\mathcal{V}}) &= \sum_{\substack{\vec{\mathcal{V}} \text{ such that} \\ \vec{\mathcal{U}} \subseteq \vec{\mathcal{V}},\ \vec{\mathcal{V}} \in T(\tau)}} \frac{z_\infty(\tau)}{|T(\tau)|} \\ &= N \times \frac{z_\infty(\tau)}{|T(\tau)|} \end{aligned}$$

where N is equal to the number of ways of choosing $\vec{\mathcal{V}}$ such that $\vec{\mathcal{U}} \subseteq \vec{\mathcal{V}}$ and $\vec{\mathcal{V}} \in T(\tau)$. Let us assume that $\tau = \sigma r_1 \cdots r_s$ for some $s \geq 1$ and literals $r_1, \ldots, r_s$. Then

$$N = \prod_{i=1}^{s} |S(\sigma r_1 \cdots r_i)|$$

and so

$$\begin{aligned} N \times \frac{z_\infty(\tau)}{|T(\tau)|} &= \prod_{i=1}^{s} |S(\sigma r_1 \cdots r_i)| \times \frac{z_\infty(\tau)}{|T(\tau)|} \\ &= \prod_{i=1}^{s} |S(\sigma r_1 \cdots r_i)| \times \frac{z_\infty(\tau)}{|T(\sigma)| \cdot \prod_{i=1}^{s} |S(\sigma r_1 \cdots r_i)|} \\ &= z_\infty(\tau) \cdot |T(\sigma)|^{-1} \end{aligned}$$

as required. □

We will now use the preceding lemma to show that y_∞ gives the same belief to all conjunctions of literals from over L as z_∞. Thus y_∞ gives the same belief to all sentences in SL as z_∞. This will allow us to prove Theorem 4.1.

Lemma 4.50 *For all non-empty sequences of literals $q_1 \cdots q_j$ from L we have*

$$Bel^{y_\infty}(q_1 \cdots q_j) = Bel^{z_\infty}(q_1 \cdots q_j).$$

Proof. Tailoring equation (4.72) to fit our current situation, we may write $Bel^{y_\infty}(q_1 \cdots q_j)$ as

$$\begin{aligned} &Bel^{y_\infty}(q_1 \cdots q_j) = \\ &= \sum_{\vec{\sigma} \in \widehat{N}_{z_\infty}(q_1 \cdots q_j)} \; \sum_{\langle \vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_j \rangle \in X(\vec{\sigma})} \prod_{i=1}^{j} \Theta^{y_\infty}\Big(\bigcup_{k<i} s(\sigma_k; \vec{\mathcal{U}}_k) \xrightarrow{q_i} s(\sigma_i; \vec{\mathcal{U}}_i)\Big) \end{aligned} \tag{4.73}$$

where, for each $\vec{\sigma} \in \widehat{N}_{z_\infty}(q_1 \cdots q_j)$, $\langle \vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_j \rangle \in X(\vec{\sigma})$ and $i = 1, \ldots, j$,

$$\begin{aligned} &\Theta^{y_\infty}\Big(\bigcup_{k<i} s(\sigma_k; \vec{\mathcal{U}}_k) \xrightarrow{q_i} s(\sigma_i; \vec{\mathcal{U}}_i)\Big) = \\ &= \frac{y_\infty(\sigma_i; \vec{\mathcal{U}}_i)}{\sum\{y_\infty(\tau; \vec{\mathcal{V}}) \mid \bigcup_{k<i} s(\sigma_k; \vec{\mathcal{U}}_k) \cup s(\tau; \vec{\mathcal{V}}) \text{ is consistent and } \pm q_i \in s(\sigma_i; \vec{\mathcal{U}}_i)\}} \\ &\qquad \text{(since } s(\sigma_i; \vec{\mathcal{U}}_i) \neq \emptyset) \\ &= \begin{cases} \dfrac{y_\infty(\sigma_i; \vec{\mathcal{U}}_i)}{y_\infty(\sigma_i; \vec{\mathcal{U}}_i)} = 1 & \text{if } \sigma_i \subseteq \sigma_{l_{i-1}} \\[2ex] \dfrac{y_\infty(\sigma_i; \vec{\mathcal{U}}_i)}{\sum\{y_\infty(\tau; \vec{\mathcal{V}}) \mid \langle \sigma_{l_{i-1}}, \vec{\mathcal{U}}_{l_{i-1}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle,\ \tau \text{ ends } \pm q_i,\ \vec{\mathcal{V}} \in T(\tau)\}} & \text{if } \sigma_{l_{i-1}} \subseteq \sigma_i \end{cases} \end{aligned} \tag{4.74}$$

where, as usual, l_i is such that $|\sigma_{l_i}|$ is maximal amongst $\{|\sigma_k| \mid k = 1, \ldots, i\}$. Compare identity (4.74) with the formula obtained in Section 4.5 (just after equation 4.3) for the term $\Theta^z(\bigcup_{k<i} s(\rho_k) \xrightarrow{q_i} s(\rho_i))$ for z a special almost-ent over L^+ (such as z_∞) and $\vec{\rho} \in \widehat{N}_z(q_1 \cdots q_j)$:

$$\Theta^z(\bigcup_{k<i} s(\rho_k) \xrightarrow{q_i} s(\rho_i)) = \begin{cases} 1 & \text{if } \rho_i \subseteq \rho_{l_{i-1}} \\ \dfrac{z(\rho_i)}{\sum\{z(\tau) \mid \rho_{l_{i-1}} \subseteq \tau,\ \tau \text{ ends } \pm q_i\}} & \text{if } \rho_{l_{i-1}} \subseteq \rho_i \end{cases} \tag{4.75}$$

We will show that, for each $\vec{\sigma} \in \widehat{N}_{z_\infty}(q_1 \cdots q_j)$,

$$\sum_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_j\rangle \in X(\vec{\sigma})} \prod_{i=1}^{j} \Theta^{y_\infty}(\bigcup_{k<i} s(\sigma_k;\vec{\mathcal{U}}_k) \xrightarrow{q_i} s(\sigma_i;\vec{\mathcal{U}}_i)) \ = \ \prod_{i=1}^{j} \Theta^{z_\infty}(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)).$$

This will suffice to prove the lemma since substituting this into (4.73) will give us

$$\begin{aligned} Bel^{y_\infty}(q_1 \cdots q_j) &= \sum_{\vec{\sigma} \in \widehat{N}_{z_\infty}(q_1 \cdots q_j)} \prod_{i=1}^{j} \Theta^{z_\infty}(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)) \\ &= Bel^{z_\infty}(q_1 \cdots q_j) \end{aligned}$$

as required. To prove this identity let us start by, first of all, given $\vec{\sigma} \in \widehat{N}_{z_\infty}(q_1 \cdots q_j)$ and $\langle \vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_l\rangle \in X(\vec{\sigma}\lceil l)$ where $1 \leq l \leq j$, defining the following terms:

$$\mathcal{A}_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_l\rangle} = \prod_{i=1}^{l} \Theta^{y_\infty}(\bigcup_{k<i} s(\sigma_k;\vec{\mathcal{U}}_k) \xrightarrow{q_i} s(\sigma_i;\vec{\mathcal{U}}_i)). \tag{4.76}$$

So our task is to show

$$\sum_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_j\rangle \in X(\vec{\sigma})} \mathcal{A}_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_j\rangle} = \prod_{i=1}^{j} \Theta^{z_\infty}(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)).$$

We will do this by proving, by induction on $m = 1, \ldots, j$, that

$$\sum_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_m\rangle \in X(\vec{\sigma}\lceil m)} \mathcal{A}_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_m\rangle} = \prod_{i=1}^{m} \Theta^{z_\infty}(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)).$$

Starting off with the base case $m = 1$ we must show

$$\sum_{\langle \vec{\mathcal{U}}_1\rangle \in X(\vec{\sigma}\lceil 1)} \mathcal{A}_{\langle \vec{\mathcal{U}}_1\rangle} = \Theta^{z_\infty}(\emptyset \xrightarrow{q_1} s(\sigma_1)).$$

The definition of $\mathcal{A}_{\langle\vec{\mathcal{U}}_1\rangle}$ gives us

$$\mathcal{A}_{\langle\vec{\mathcal{U}}_1\rangle} = \Theta^{y_\infty}(\emptyset \xrightarrow{q_1} s(\sigma_1;\vec{\mathcal{U}}_1)),$$

which in turn, by (4.74) (since certainly $\sigma_1 \supseteq \sigma_0 = \emptyset$), gives

$$\mathcal{A}_{\langle\vec{\mathcal{U}}_1\rangle} = \frac{y_\infty(\sigma_1;\vec{\mathcal{U}}_1)}{\sum\{y_\infty(\tau;\vec{\mathcal{V}}) \mid \tau \text{ ends } \pm q_1,\ \vec{\mathcal{V}} \in T(\tau)\}}.$$

Hence

$$\begin{aligned}\sum_{\langle\vec{\mathcal{U}}_1\rangle \in X(\vec{\sigma}\restriction 1)} \mathcal{A}_{\langle\vec{\mathcal{U}}_1\rangle} &= \sum_{\vec{\mathcal{U}}_1 \in T(\sigma_1)} \mathcal{A}_{\langle\vec{\mathcal{U}}_1\rangle} \qquad \text{by definition of } X(\vec{\sigma}\restriction 1) \\ &= \sum_{\vec{\mathcal{U}}_1 \in T(\sigma_1)} \frac{y_\infty(\sigma_1;\vec{\mathcal{U}}_1)}{\sum\{y_\infty(\tau;\vec{\mathcal{V}}) \mid \tau \text{ ends } \pm q_1,\ \vec{\mathcal{V}} \in T(\tau)\}}\end{aligned} \tag{4.77}$$

Consider the denominator in the above expression for $\mathcal{A}_{\langle\vec{\mathcal{U}}_1\rangle}$. We may write it as

$$\begin{aligned}\sum\{y_\infty(\tau;\vec{\mathcal{V}}) \mid \tau \text{ ends } \pm q_1,\ \vec{\mathcal{V}} \in T(\tau)\} &= \sum_{\tau \text{ ends } \pm q_1} \sum_{\vec{\mathcal{V}} \in T(\tau)} y_\infty(\tau;\vec{\mathcal{V}}) \\ &= \sum_{\tau \text{ ends } \pm q_1} z_\infty(\tau) \\ &\quad \text{(by Lemma 4.49 applied to } \sigma = \emptyset) \\ &= \sum\{z_\infty(\tau) \mid \tau \text{ ends } \pm q_1\}.\end{aligned}$$

Lemma 4.49 also gives us

$$\sum_{\vec{\mathcal{U}}_1 \in T(\sigma_1)} y_\infty(\sigma_1;\vec{\mathcal{U}}_1) = z_\infty(\sigma_1),$$

Hence, plugging this information into (4.77) yields

$$\begin{aligned}\sum_{\langle\vec{\mathcal{U}}_1\rangle \in X(\vec{\sigma}\restriction 1)} \mathcal{A}_{\langle\vec{\mathcal{U}}_1\rangle} &= \frac{z_\infty(\sigma_1)}{\sum\{z_\infty(\tau) \mid \tau \text{ ends } \pm q_1\}} \\ &= \Theta^{z_\infty}(\emptyset \xrightarrow{q_1} s(\sigma_1)) \qquad \text{from (4.75) as required.}\end{aligned}$$

Now let $1 < m \leq j$ and suppose for inductive hypothesis that

$$\sum_{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1}\rangle \in X(\vec{\sigma}\restriction(m-1))} \mathcal{A}_{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1}\rangle} = \prod_{i=1}^{m-1} \Theta^{z_\infty}(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)).$$

We must show that this equation remains true when we substitute $m-1$ everywhere by m. To do this let us first notice that

$$\begin{aligned}
\sum_{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_m\rangle\in X(\vec{\sigma}\restriction m)} \mathcal{A}_{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_m\rangle} &= \\
&= \sum_{\substack{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1}\rangle \\ \in X(\vec{\sigma}\restriction(m-1))}} \sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1},\vec{\mathcal{U}}_m\rangle\in X(\vec{\sigma}\restriction m)}} \mathcal{A}_{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_m\rangle} \\
&= \sum_{\substack{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1}\rangle \\ \in X(\vec{\sigma}\restriction(m-1))}} \mathcal{A}_{\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1}\rangle} \sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1},\vec{\mathcal{U}}_m\rangle\in X(\vec{\sigma}\restriction m)}} \Theta^{y\infty}(\bigcup_{k<m} s(\sigma_k;\vec{\mathcal{U}}_k) \stackrel{q_m}{\to} s(\sigma_m;\vec{\mathcal{U}}_m)) \\
&= \prod_{i=1}^{m-1} \Theta^{z\infty}(\bigcup_{k<i} s(\sigma_k) \stackrel{q_i}{\to} s(\sigma_i)) \sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1},\vec{\mathcal{U}}_m\rangle\in X(\vec{\sigma}\restriction m)}} \Theta^{y\infty}(\bigcup_{k<m} s(\sigma_k;\vec{\mathcal{U}}_k) \stackrel{q_m}{\to} s(\sigma_m;\vec{\mathcal{U}}_m))
\end{aligned}$$

by inductive hypothesis.

Hence our inductive step will be completed if we can show

$$\sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1},\vec{\mathcal{U}}_m\rangle\in X(\vec{\sigma}\restriction m)}} \Theta^{y\infty}(\bigcup_{k<m} s(\sigma_k;\vec{\mathcal{U}}_k) \stackrel{q_m}{\to} s(\sigma_m;\vec{\mathcal{U}}_m)) = \Theta^{z\infty}(\bigcup_{k<m} s(\sigma_k) \stackrel{q_m}{\to} s(\sigma_m)).$$

We must have either $\sigma_m \subseteq \sigma_{l_{m-1}}$ or $\sigma_{l_{m-1}} \subseteq \sigma_m$. If the former case obtains then, by definition of $X(\vec{\sigma}\restriction m)$, we know that $\vec{\mathcal{U}}_m$ satisfies $\langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1},\vec{\mathcal{U}}_m\rangle \in X(\vec{\sigma}\restriction m)$ iff $\vec{\mathcal{U}}_m = \vec{\mathcal{U}}_{l_{m-1}}\restriction|\sigma_m|$. Also, for this particular choice of $\vec{\mathcal{U}}_m$ we have, from (4.74),

$$\Theta^{y\infty}(\bigcup_{k<m} s(\sigma_k;\vec{\mathcal{U}}_k) \stackrel{q_m}{\to} s(\sigma_m;\vec{\mathcal{U}}_m)) = 1.$$

Hence, in this case,

$$\begin{aligned}
\sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_m\rangle\in X(\vec{\sigma})}} \Theta^{y\infty}(\bigcup_{k<m} s(\sigma_k;\vec{\mathcal{U}}_k) \stackrel{q_m}{\to} s(\sigma_m;\vec{\mathcal{U}}_m)) &= 1 \\
&= \Theta^{z\infty}(\bigcup_{k<m} s(\sigma_k) \stackrel{q_m}{\to} s(\sigma_m))
\end{aligned}$$

using the identity 4.75.

If, on the other hand, $\sigma_{l_{m-1}} \subseteq \sigma_m$ then we must have

$$\sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \langle\vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_{m-1},\vec{\mathcal{U}}_m\rangle\in X(\vec{\sigma}\restriction m)}} \Theta^{y\infty}(\bigcup_{k<m} s(\sigma_k;\vec{\mathcal{U}}_k) \stackrel{q_m}{\to} s(\sigma_m;\vec{\mathcal{U}}_m)) =$$

$$= \sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \vec{\mathcal{U}}_{l_{m-1}} \subseteq \vec{\mathcal{U}}_m,\ \vec{\mathcal{U}}_m \in T(\sigma_m)}} \frac{y_\infty(\sigma_m; \vec{\mathcal{U}}_m)}{\sum\{y_\infty(\tau; \vec{\mathcal{V}}) \mid \langle \sigma_{l_{m-1}}, \vec{\mathcal{U}}_{l_{m-1}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle,\ \tau \text{ ends } \pm q_m,\ \vec{\mathcal{V}} \in T(\tau)\}}. \tag{4.78}$$

Consider the denominator in the above expression. We have

$$\begin{aligned}
&\sum\{y_\infty(\tau; \vec{\mathcal{V}}) \mid \langle \sigma_{l_{m-1}}, \vec{\mathcal{U}}_{l_{m-1}} \rangle \subseteq \langle \tau, \vec{\mathcal{V}} \rangle,\ \tau \text{ ends } \pm q_m,\ \vec{\mathcal{V}} \in T(\tau)\} = \\
&= \sum_{\substack{\tau \text{ such that} \\ \sigma_{l_{m-1}} \subseteq \tau,\ \tau \text{ ends } \pm q_m}} \sum_{\substack{\vec{\mathcal{V}} \text{ such that} \\ \vec{\mathcal{U}}_{l_{m-1}} \subseteq \vec{\mathcal{V}},\ \vec{\mathcal{V}} \in T(\tau)}} y_\infty(\tau; \vec{\mathcal{V}}) \\
&= \sum_{\substack{\tau \text{ such that} \\ \sigma_{l_{m-1}} \subseteq \tau,\ \tau \text{ ends } \pm q_m}} z_\infty(\tau) \cdot |T(\sigma_{l_{m-1}})|^{-1} \\
&\qquad \text{(by Lemma 4.49 (since clearly } \sigma_{l_{m-1}} \neq \emptyset)) \\
&= |T(\sigma_{l_{m-1}})|^{-1} \sum\{z_\infty(\tau) \mid \sigma_{l_{m-1}} \subseteq \tau,\ \tau \text{ ends } \pm q_m\}.
\end{aligned}$$

Similarly by Lemma 4.49 we have

$$\sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \vec{\mathcal{U}}_{l_{m-1}} \subseteq \vec{\mathcal{U}}_m,\ \vec{\mathcal{U}}_m \in T(\sigma_m)}} y_\infty(\sigma_m; \vec{\mathcal{U}}_m) = z_\infty(\sigma_m) \cdot |T(\sigma_{l_{m-1}})|^{-1}.$$

Hence, plugging all this into equation (4.78), we obtain

$$\begin{aligned}
&\sum_{\substack{\vec{\mathcal{U}}_m \text{ such that} \\ \langle \vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_{m-1}, \vec{\mathcal{U}}_m \rangle \in X(\vec{\sigma} \restriction m)}} \Theta^{y_\infty}\Big(\bigcup_{k<m} s(\sigma_k; \vec{\mathcal{U}}_k) \xrightarrow{q_m} s(\sigma_m; \vec{\mathcal{U}}_m)\Big) = \\
&= \frac{z_\infty(\sigma_m) \cdot |T(\sigma_{l_{m-1}})|^{-1}}{|T(\sigma_{l_{m-1}})|^{-1} \sum\{z_\infty(\tau) \mid \sigma_{l_{m-1}} \subseteq \tau,\ \tau \text{ ends } \pm q_m\}} \\
&= \frac{z_\infty(\sigma_m)}{\sum\{z_\infty(\tau) \mid \sigma_{l_{m-1}} \subseteq \tau,\ \tau \text{ ends } \pm q_m\}} \\
&= \Theta^{z_\infty}\Big(\bigcup_{k<m} s(\sigma_k) \xrightarrow{q_m} s(\sigma_m)\Big) \\
&\qquad \text{from 4.75, as required.}
\end{aligned}$$

This completes the inductive proof that, for all $m = 1, \ldots, j$,

$$\sum_{\langle \vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_m \rangle \in X(\vec{\sigma} \restriction m)} \mathcal{A}_{\langle \vec{\mathcal{U}}_1, \ldots, \vec{\mathcal{U}}_m \rangle} = \prod_{i=1}^{m} \Theta^{z_\infty}\Big(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)\Big).$$

In particular, putting $m = j$ in the above, we have shown

$$\sum_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_j\rangle \in X(\vec{\sigma})} \mathcal{A}_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_j\rangle} = \prod_{i=1}^{j} \Theta^{z_\infty}(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)).$$

Thus

$$\begin{aligned} Bel^{y_\infty}(q_1 \cdots q_j) &= \sum_{\vec{\sigma} \in N_{z_\infty}(q_1 \cdots q_j)} \sum_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_j\rangle \in X(\vec{\sigma})} \mathcal{A}_{\langle \vec{\mathcal{U}}_1,\ldots,\vec{\mathcal{U}}_j\rangle} \\ &\qquad \text{(from (4.73) and (4.76))} \\ &= \sum_{\vec{\sigma} \in N_{z_\infty}(q_1 \cdots q_j)} \prod_{i=1}^{j} \Theta^{z_\infty}(\bigcup_{k<i} s(\sigma_k) \xrightarrow{q_i} s(\sigma_i)) \\ &= Bel^{z_\infty}(q_1 \cdots q_j) \end{aligned}$$

as required. □

Given the preceding lemma, we are now finally in a position to prove Theorem 4.1.

Theorem 4.1 Given a language $L = \{p_1, \ldots, p_n\}$, if the function $Bel : SL \to [0, 1]$ is given by a standard pre-ent over L and if, for all $\theta, \phi \in SL$, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$, then there exists a standard ent z (over a larger language than L) such that, for all $\theta \in SL$, $Bel^z(\theta) = Bel(\theta)$.

Proof. Let $L = \{p_1, \ldots, p_n\}$. As pointed out at the beginning of the present chapter (just after the statement of Theorem 4.1) if $n = 1$ then *any* function $Bel : SL \to [0, 1]$ given by a pre-ent may be given by an ent over L so assume $n > 1$. Then by Theorems 4.11 and 4.46 there exists an *almost-ent* z_∞ over a language L^+ which extends L for which we have, for all $\theta \in SL$, $Bel^{z_\infty}(\theta) = Bel(\theta)$. By Lemma 4.50 there exists an *ent* y_∞ over a language L^* which extends L for which we have, for all $\theta \in SL$, $Bel^{y_\infty}(\theta) = Bel^{z_\infty}(\theta)$. This gives the result. □

Thus we have shown in this chapter that, assuming an open-ended underlying propositional language, if we start from the class of pre-ents and then force the belief functions of those pre-ents to satisfy the desirable property that, for all sentences θ and ϕ, $Bel(\theta \wedge \phi) = 0$ implies $Bel(\phi \wedge \theta) = 0$ then we are led automatically to the class of ents. This result can be interpreted as showing that ents are perhaps more general than they might first appear.

Chapter 5

Pre-Ents and Consequence Relations

5.1 Introduction

The work comprised in the rest of this thesis is motivated by considering the question of what possible notions of *entailment* between sentences of a propositional language can be captured using ents or pre-ents, or what form of *consequence relation* can a pre-ent or ent give rise to. In its most abstract terms a consequence relation (over a given language L) is just a binary relation on SL. Indeed we have already seen one example of such a consequence relation yielded by pre-ents, namely the relation $\mid\!\ddot{\sim}$ studied in Section 3.4. Another example, which we shall give (for ents only) in Section 5.2, was examined by Gladstone in [4], and used there to characterise the class of *monotonic* consequence relations (i.e., the class to which the relation $\vdash \subseteq SL \times SL$ of classical logical consequence belongs). The consequence relation which forms the main object of study for the next two chapters was borne out of an attempt to characterise the more general class of *rational* consequence relations (which were defined in [16]) in terms of pre-ents.

As we shall see in Section 5.2 (where we shall also give a brief review of such relations), our attempt, via the relation which we shall denote by $\mid\!\sim_G$, fails and so we are led in Section 5.3 to generalise even further by defining what we call *natural* consequence relations. We define this new class of consequence relation by a set of rules which are intended to be weakened versions of the rules for rational consequence. In Section 5.3 we show that $\mid\!\sim_G$ satisfies all these rules and give some further rules which follow from this set. We also confirm that every rational consequence relation is a natural consequence relation. In the rest of the chapter we give another family of natural consequence relations which is different from, though, as it turns out, closely related to, the family $\mid\!\sim_G$. This family is described in the framework of *permatoms* which is the subject of Section 5.4. We give some basic results about this framework which include a characterisation of the relation $\mid\!\ddot{\sim}$ from Section 3.4 in terms of permatoms. We get some good practice in using this new framework by showing in Section 5.5 how the class of rational consequence relations can be characterised in terms of those sequences $\vec{\mathcal{U}} = \mathcal{U}_1, \dots, \mathcal{U}_k$ of sets of permatoms which satisfy a certain condition of *admissibility.* This characterisation is essentially the same as that given in [7] although the proof is different. Finally in Section 5.6 we weaken the admissibility condition in the hope that we may be able to characterise the more general class of natural consequence relations in terms of a more general family of sequences $\vec{\mathcal{U}}$. We show that the consequence relations $\mid\!\sim_{\vec{\mathcal{U}}}$ arising from such *weakly admissible* sequences $\vec{\mathcal{U}}$ do at least satisfy the rules for natural consequence by showing how each such sequence gives rise to a pre-ent G (over a language which extends the language of $\vec{\mathcal{U}}$) such that (on its restriction to the language of $\vec{\mathcal{U}}$) $\mid\!\sim_G = \mid\!\sim_{\vec{\mathcal{U}}}$. As we shall see in Chapter 6, we shall encounter problems in showing that, conversely, every natural consequence relation is given by a weakly admissible $\vec{\mathcal{U}}$.

We remind the reader that, unless specified otherwise, we assume that $L =$

$\{p_1, \ldots, p_n\}$.

5.2 Pre-Ents and Rational Consequence

Rational consequence relations were defined (along with other families of consequence relation) by Kraus, Lehmann and Magidor in [16] and studied extensively by Lehmann and Magidor in [7]. The motivation for their definition came from considering the following question: Given a propositional language L and a binary relation $\mid\sim$ on SL with $\theta \mid\sim \phi$ having the intended interpretation "if θ is true then, typically, ϕ is also true", what closure properties should $\mid\sim$ have given that it corresponds to the set of beliefs (of this form) of an intelligent, rational agent (such as ourselves!)? The list of properties that Kraus et al. arrived at will now be given.

Definition 5.1 *A* rational consequence relation *on L is a binary relation $\mid\sim$ on SL which satisfies the following rules for all $\theta, \phi, \psi \in SL$. (Rules of the same form as rules (2)–(7) should be read as: from the truth of the numerator, deduce the truth of the denominator.)*

1. $\theta \mid\sim \theta$ *(Reflexivity (REF))*
2. $\dfrac{\theta \mid\sim \phi,\ \theta \equiv \psi}{\psi \mid\sim \phi}$ *(Left Logical Equivalence(LLE))*
3. $\dfrac{\theta \mid\sim \phi,\ \phi \vdash \psi}{\theta \mid\sim \psi}$ *(Right Weakening(RWE))*
4. $\dfrac{\theta \mid\sim \phi,\ \theta \mid\sim \psi}{\theta \mid\sim \phi \wedge \psi}$ *(AND)*
5. $\dfrac{\theta \mid\sim \phi,\ \psi \mid\sim \phi}{\theta \vee \psi \mid\sim \phi}$ *(OR)*
6. $\dfrac{\theta \mid\sim \phi,\ \theta \mid\sim \psi}{\theta \wedge \phi \mid\sim \psi}$ *(Cautious Monotonicity(CMO))*

7. $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \theta \mathrel{|\!\not\sim} \neg\psi}{\theta \wedge \psi \mathrel{|\!\sim} \phi}$ *(Rational Monotonicity(RMO))*

The full justifications for these rules may be found in [16]. Here are some more rules which can be derived from the above rules (see [16] and [7] for the (simple) proofs), and hence are satisfied by any rational consequence relation.

Proposition 5.2 *Every rational consequence relation* $\mathrel{|\!\sim}$ *satisfies the following rules:*

1. $\dfrac{\theta \vdash \phi}{\theta \mathrel{|\!\sim} \phi}$ *(Supraclassicality(SCL))*
2. $\dfrac{\theta \wedge \phi \mathrel{|\!\sim} \psi,\ \theta \mathrel{|\!\sim} \phi}{\theta \mathrel{|\!\sim} \psi}$ *(Cautious Cut(CC))*
3. $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \phi \mathrel{|\!\sim} \theta,\ \theta \mathrel{|\!\sim} \psi}{\phi \mathrel{|\!\sim} \psi}$ *(Equivalence)*
4. $\dfrac{\theta \vee \phi \mathrel{|\!\sim} \psi,\ \theta \mathrel{|\!\not\sim} \psi}{\phi \mathrel{|\!\sim} \psi}$ *(Disjunctive Rationality(DR))* □

The definition of rational consequence relations is given in a syntactic form (i.e., as a set of rules). However, several semantic characterisations of rational consequence relations have been provided in the literature. One of these will be expounded (in our more general framework) in Section 5.5. Another one (see [7], [14], though the original idea can be traced back to [1]) is given in terms of non-standard probability functions, or, as we may call them here, λ-probability functions, where we define a λ-probability function to be just like a probability function (see Definition 2.12) except we interpret it as a function into $[0,1]^{(\lambda)}$ rather than just $[0,1]$. According to this characterisation, a binary relation $\mathrel{|\!\sim} \subseteq SL \times SL$ is a non-trivial rational consequence relation iff there exists a λ-probability function F on L such that, for all $\theta, \phi \in SL$,

$$\begin{aligned}\theta \mathrel{|\!\sim} \phi \quad \text{iff} \quad & \textit{either } F(\theta) = 0 \\ & \textit{or } F(\theta) \neq 0 \text{ and } F(\neg\phi \mid \theta) = O(\lambda).\end{aligned}$$

(By "non-trivial rational consequence relation" we mean a rational consequence relation for which it is not the case that $\theta \mid\!\sim \phi$ for *all* $\theta, \phi \in SL$.) This representation of rational consequence relations seems very intuitive. It says that $\mid\!\sim$ is a (non-trivial) rational consequence relation just in case there exists some λ-probability function F such that $\theta \mid\!\sim \phi$ holds iff either θ is considered totally unbelievable by F or the conditional probability, according to F, of $\neg\phi$ given θ is infinitesimally small (or zero).

Note that one rule which a rational consequence relation is *not* required to satisfy, at least in its full generality (since we *do* require the special cases of it LLE, CMO and RMO), is the following

$$\frac{\phi \mid\!\sim \psi, \;\; \theta \vdash \phi}{\theta \mid\!\sim \psi} \textit{(Monotonicity)}$$

and indeed this is how it should be, for suppose we held the belief "if the cake contains butter then, typically, it will taste delicious". If we then strengthened the "if" clause here to "the cake contains butter **and** drawing pins" would we still be willing to take a bite? Thus rational consequence relations provide examples of *non-monotonic* inference relations. If we add the rule of Monotonicity to the rules for rational consequence then we arrive (following Lemma 7.3 of [16]) at a subclass of rational consequence relations – the class of *monotonic* consequence relations. From [4] (Theorems 9 and 10) we already have a result which completely characterises monotonic consequence relations in terms of standard ents, namely:

Theorem 5.3 *Let $\mid\!\sim$ be a binary relation on SL. Then $\mid\!\sim$ is a monotonic consequence relation (on L) iff there exists a (standard) ent z over L such that for all $\theta, \phi \in SL$, $\theta \mid\!\sim \phi$ iff $Bel^z(\theta) = Bel^z(\theta \wedge \phi)$.* □

In view of this result it would be hoped that a similar characterisation in terms of ents or pre-ents would be possible for (at least the non-trivial) rational consequence relations (initially, we shall actually seek a characterisation in terms of pre-ents rather than just ents), especially if we widen our attention to λ-pre-ents and λ-ents (see Section 4.2). Indeed, as a first guess, it might be expected that the following definition, by analogy with conditional probability, would give us what we want.

Definition 5.4 *Given a (λ-)pre-ent G over L, we define the consequence relation $\mid\!\sim_G \subseteq SL \times SL$ by, for all $\theta, \phi \in SL$,*

$$\theta \mid\!\sim_G \phi \quad \textit{iff} \quad \textit{either } Bel^G(\theta) = 0$$
$$\textit{or } Bel^G(\theta) \neq 0 \text{ and } \frac{Bel^G(\theta \wedge \neg\phi)}{Bel^G(\theta)} = O(\lambda)$$

(We shall again assume that, from now on, all pre-ents and ents are in fact λ-pre-ents and λ-ents.) We would now like to able to show that a given non-trivial binary relation $\mid\!\sim$ on SL is a rational consequence relation on L iff $\mid\!\sim \, = \, \mid\!\sim_G$ for some pre-ent G over L. Note that, for arbitrary $\psi, \chi \in SL$, if $Bel^G(\psi) > 0$ then $0 \leq \frac{Bel^G(\psi \wedge \chi)}{Bel^G(\psi)} \leq 1$ while

$$\frac{Bel^G(\psi \wedge \neg\chi)}{Bel^G(\psi)} = 1 - \frac{Bel^G(\psi \wedge \chi)}{Bel^G(\psi)}.$$

The relation $\mid\!\sim_G$ defined above has an equivalent formulation which we shall find useful to keep in mind.

Proposition 5.5 *For all pre-ents G over L and for all $\theta, \phi \in SL$,*

$$\theta \mid\!\sim_G \phi \quad \textit{iff} \quad \textit{either } Bel^G(\theta) = 0$$
$$\textit{or } Bel^G_s(\neg\phi) = O(\lambda) \text{ for all } s \in WL \text{ such that } s \vdash \theta \text{ and } \frac{G_\theta(\emptyset, s)}{Bel^G(\theta)} \neq O(\lambda)$$

Proof. The result is clear if $Bel^G(\theta) = 0$, so assume $Bel^G(\theta) > 0$. We have

$$\frac{Bel^G(\theta \wedge \neg\phi)}{Bel^G(\theta)} = \frac{1}{Bel^G(\theta)} \sum_{r \vdash \theta \wedge \neg\phi} G_{\theta \wedge \neg\phi}(\emptyset, r)$$

$$
\begin{aligned}
&= \frac{1}{Bel^G(\theta)} \sum_{s \vdash \theta} G_\theta(\emptyset, s) \cdot \sum_{t \vdash \neg\phi} G_{\neg\phi}(s,t) \\
&= \frac{1}{Bel^G(\theta)} \sum_{s \vdash \theta} G_\theta(\emptyset, s) \cdot Bel_s^G(\neg\phi) \\
&= \sum_{s \vdash \theta} (\frac{G_\theta(\emptyset, s)}{Bel^G(\theta)}) \cdot Bel_s^G(\neg\phi) \\
&= \overbrace{\sum_{\substack{s \vdash \theta \\ (\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)}) = O(\lambda)}} (\frac{G_\theta(\emptyset, s)}{Bel^G(\theta)}) \cdot Bel_s^G(\neg\phi)}^{\gamma} \; + \\
&\quad + \overbrace{\sum_{\substack{s \vdash \theta \\ (\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)}) \neq O(\lambda)}} (\frac{G_\theta(\emptyset, s)}{Bel^G(\theta)}) \cdot Bel_s^G(\neg\phi)}^{\delta} \\
&= \gamma \; + \; \delta \qquad\qquad \text{where } \gamma = O(\lambda)
\end{aligned}
$$

If $\frac{Bel^G(\theta \wedge \neg\phi)}{Bel^G(\theta)} = O(\lambda)$ then we must have $\delta = O(\lambda)$ and so for all s such that $s \vdash \theta$ and $\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)} \neq O(\lambda)$ we must have $\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)} \cdot Bel_s^G(\neg\phi) = O(\lambda)$ and so we must have $Bel_s^G(\neg\phi) = O(\lambda)$. If $\frac{Bel^G(\theta \wedge \neg\phi)}{Bel^G(\theta)} \neq O(\lambda)$ then $\delta \neq O(\lambda)$ and so for some s such that $s \vdash \theta$ and $\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)} \neq O(\lambda)$ we must have $Bel_s^G(\neg\phi) \neq O(\lambda)$ as required. □

Given $s \vdash \theta$, the identity $\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)} = O(\lambda)$ says that the probability that G imagines s when called upon to imagine a scenario which decides θ is, relative to the probability of imagining *any* scenario which decides θ positively, infinitesimally small or zero. In other words, s is exceptional amongst those scenarios which decide θ positively. Thus the interpretation of $\mid\!\sim_G$ provided by the above proposition is that ϕ should be considered a consequence of θ iff either θ is totally unbelievable according to G or $Bel_s^G(\neg\phi)$ is infinitesimally small or zero for all those s which are "unexceptional" amongst the scenarios which decide θ positively.

So does $\mid\!\sim_G$ form a rational consequence relation? Well to start with we can easily show the following result which shows that $\mid\!\sim_G$ satisfies the first property from Definition 5.1.

Theorem 5.6 *Given a pre-ent G over L, the consequence relation $\mid\!\sim_G$ satisfies the rule REF.*

Proof. Let $\theta \in SL$. We must show $\theta \mid\!\sim_G \theta$. If $Bel^G(\theta) = 0$ then trivially $\theta \mid\!\sim_G \theta$ so suppose $Bel^G(\theta) \neq 0$. Then since $\vdash \neg(\theta \wedge \neg\theta)$ we have $Bel^G(\neg(\theta \wedge \neg\theta)) = 1$ and so $Bel^G(\theta \wedge \neg\theta) = 0$. Hence $\frac{Bel^G(\theta \wedge \neg\theta)}{Bel^G(\theta)} = 0$ giving the result. □

Unfortunately, as the following examples show, REF is the *only* rule for rational consequence which $\mid\!\sim_G$ will satisfy in general.

Example 5.7 In each of the following examples we take $L = \{p, q, r\}$.
(i). Let z^1 be the ent defined as follows:

s	$\{r\}$	$\{p\}$	$\{\neg p, \neg q\}$	$\{q\}$
z^1_s	1	1	λ	λ

Then we have $p \vee \neg p \mid\!\sim_{z^1} p$, $p \vee \neg p \equiv q \vee \neg q$, but $q \vee \neg q \not\mid\!\sim_{z^1} p$. Hence $\mid\!\sim_{z^1}$ fails to satisfy LLE. Also we have $r \mid\!\sim_{z^1} p \wedge q$, $p \wedge q \vdash q$, but $r \not\mid\!\sim_{z^1} q$. Hence $\mid\!\sim_{z^1}$ also fails to satisfy RWE.

(ii). Let z^2 be the ent defined as follows:

s	$\{p\}$	$\{\neg p, q\}$	$\{\neg q\}$	$\{r\}$
z^2_s	1	λ	λ^2	1

Then we have $r \mid\!\sim_{z^2} p$, $r \mid\!\sim_{z^2} q$ but $r \not\mid\!\sim_{z^2} p \wedge q$. Hence $\mid\!\sim_{z^2}$ fails to satisfy AND.

(iii). let z^3 be the ent defined as follows:

s	$\{r\}$	$\{\neg p, r\}$	$\{p, q\}$	$\{\neg q\}$	$\{p\}$
z^3_s	1	λ	λ	λ^2	λ

Then we have $p \mathrel{|\!\sim}_{z^3} q,\ r \mathrel{|\!\sim}_{z^3} q$, but $p \vee r \mathrel{|\!\not\sim}_{z^3} q$. Hence $\mathrel{|\!\sim}_{z^3}$ fails to satisfy OR.

(iv). let z^4 be the ent defined as follows:

s	$\{p\}$	$\{q\}$	$\{r, \neg q\}$	$\{\neg r\}$
z^4_s	1	1	λ	λ^2

Then we have $p \mathrel{|\!\sim}_{z^4} q,\ p \mathrel{|\!\sim}_{z^4} r$, but $p \wedge q \mathrel{|\!\not\sim}_{z^4} r$. Hence $\mathrel{|\!\sim}_{z^4}$ fails to satisfy CMO.

(v). let z^5 be the ent defined as follows:

s	$\{p\}$	$\{\neg q\}$	$\{q, \neg r\}$	$\{r\}$
z^5_s	1	λ	1	1

Then we have $p \mathrel{|\!\sim}_{z^5} q,\ p \mathrel{|\!\not\sim}_{z^5} \neg r$, but $p \wedge r \mathrel{|\!\not\sim}_{z^5} q$. Hence $\mathrel{|\!\sim}_{z^5}$ fails to satisfy RMO.

Hence $\mathrel{|\!\sim}_G$ fails, at least in general, to satisfy most of the rules for rational consequence. However, by considering the characterisation of rational consequence relations described earlier in terms of λ-probability functions, we may see that if Bel^G turns out to be a λ-probability function, equivalently (by Theorem 2.6, which clearly remains true in our more general λ-framework) if it is the case that

$$\forall \theta, \phi,\ Bel^G(\theta \wedge \phi) = Bel^G(\phi \wedge \theta),$$

then $\mathrel{|\!\sim}_G$ *will* turn out to be rational.

In the face of the counter-examples of Example 5.7 we are met with two possibilities. We can either search for a different consequence relation arising from pre-ents which hopefully *does* satisfy all the rules for rational consequence, or we can persevere with $\mathrel{|\!\sim}_G$ and try to establish what properties it *does* satisfy. We could then maybe define a new class of consequence relation to contain those relations which satisfy these new properties and then characterise *this* class in terms of pre-ents instead. Since the relation $\mathrel{|\!\sim}_G$ seems, after all, to be a very natural relation to consider, we choose here the second option.

5.3 Natural Consequence Relations

We now present a set of rules with which, we hope, we will be able to completely characterise the family of relations $\mathrel{|\!\sim}_G$. Each rule is intended to be a suitably weakened counterpart of a rule for rational consequence.

1. $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \theta \mathrel{\ddot{\sim}} \psi,\ \theta \wedge \phi \mathrel{\ddot{\sim}} \psi \wedge \phi}{\psi \mathrel{|\!\sim} \phi}$ (P-LLE)

2. $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \theta \wedge \phi \mathrel{|\!\ddot{\sim}} \theta \wedge \psi}{\theta \mathrel{|\!\sim} \psi}$ (P-RWE)

3. $\dfrac{\theta \mathrel{|\!\sim} \phi,\ \theta \mathrel{|\!\sim} \neg\phi \vee \psi}{\theta \mathrel{|\!\sim} \phi \wedge \psi}$ (P-AND)

4. $\dfrac{\theta \mathrel{|\!\sim} \psi,\ \neg\theta \wedge \phi \mathrel{|\!\sim} \psi}{\theta \vee \phi \mathrel{|\!\sim} \psi}$ (P-OR)

5. $\dfrac{\theta \mathrel{|\!\sim} \phi \wedge \psi}{\theta \wedge \phi \mathrel{|\!\sim} \psi}$ (P-CMO)

6. $\dfrac{\theta \mathrel{|\!\not\sim} \neg\phi,\ \theta \mathrel{|\!\sim} \neg\phi \vee \psi}{\theta \wedge \phi \mathrel{|\!\sim} \psi}$ (P-RMO)

Note that, in the rule P-LLE above, we do need both $\theta \mathrel{\ddot{\sim}} \psi$ **and** $\theta \wedge \phi \mathrel{\ddot{\sim}} \psi \wedge \phi$ since, as we have already remarked, it is *not* necessarily the case that $\theta \mathrel{\ddot{\sim}} \psi$ implies $\theta \wedge \phi \mathrel{\ddot{\sim}} \psi \wedge \phi$. Also note that we could equally well have replaced "$\theta \wedge \phi \mathrel{\ddot{\sim}} \psi \wedge \phi$" in the numerator of P-LLE by $\theta \wedge \neg\phi \mathrel{\ddot{\sim}} \psi \wedge \neg\phi$" since these two are equivalent in the presence of $\theta \mathrel{\ddot{\sim}} \phi$. We now show that these rules are satisfied by $\mathrel{|\!\sim}_G$ for G a pre-ent over L. Note for the proof that, for all $a, b \in \mathbb{R}((\lambda))$, if $a \in [0,1]^{(\lambda)}$ then $a = O(\lambda^m)$ for some m and so if $a, b \in [0,1]^{(\lambda)}$ and $b = O(\lambda^k)$ then $a \leq b$ implies also $a = O(\lambda^k)$. The simple proof of this is left to the reader.

Theorem 5.8 *Let G be a pre-ent over L. Then the relation $\mathrel{|\!\sim}_G$ satisfies the above rules.*

Proof. P-LLE: Suppose $\theta \mathrel{|}\!\sim_G \phi$, $\theta \mathrel{\ddot{\sim}} \psi$, and $\theta \wedge \phi \mathrel{\ddot{\sim}} \psi \wedge \phi$. Then $Bel^G(\theta) = Bel^G(\psi)$ and $Bel^G(\theta \wedge \phi) = Bel^G(\psi \wedge \phi)$. If $Bel^G(\theta) = 0$ then $Bel^G(\psi) = 0$ giving $\psi \mathrel{|}\!\sim_G \phi$ as required. So suppose $Bel^G(\theta) \neq 0 \neq Bel^G(\psi)$. Then $\theta \mathrel{|}\!\sim_G \phi$ gives

$$1 - \frac{Bel^G(\theta \wedge \phi)}{Bel^G(\theta)} = 1 - \frac{Bel^G(\psi \wedge \phi)}{Bel^G(\psi)} = O(\lambda)$$

again giving $\psi \mathrel{|}\!\sim_G \phi$.

P-RWE: Suppose $\theta \mathrel{|}\!\sim_G \phi$ and $\theta \wedge \phi \mathrel{|}\!\ddot{\sim} \theta \wedge \psi$. Then $Bel^G(\theta \wedge \phi) \leq Bel^G(\theta \wedge \psi)$. If $Bel^G(\theta) = 0$ then $\theta \mathrel{|}\!\sim_G \psi$ trivially, so suppose $Bel^G(\theta) \neq 0$. We have

$$1 - \frac{Bel^G(\theta \wedge \psi)}{Bel^G(\theta)} \leq 1 - \frac{Bel^G(\theta \wedge \phi)}{Bel^G(\theta)}$$

Hence $\theta \mathrel{|}\!\sim_G \phi$ implies $1 - \frac{Bel^G(\theta \wedge \phi)}{Bel^G(\theta)} = O(\lambda)$ which, in turn, implies $1 - \frac{Bel^G(\theta \wedge \psi)}{Bel^G(\theta)} = O(\lambda)$ as required to show $\theta \mathrel{|}\!\sim_G \psi$.

P-AND: Suppose $\theta \mathrel{|}\!\sim_G \phi$ and $\theta \mathrel{|}\!\sim_G \neg\phi \vee \psi$. If $Bel^G(\theta) = 0$ then $\theta \mathrel{|}\!\sim_G \phi \wedge \psi$ so suppose $Bel^G(\theta) \neq 0$. Let $s \in WL$ be such that $s \vdash \theta$ and $\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)} \neq O(\lambda)$. We must show $Bel^G_s(\neg(\phi \wedge \psi)) = O(\lambda)$. Now, for $\chi, \rho \in SL$, we have, by definition of $\dot{\sim}$, $\chi \mathrel{\dot{\sim}} \rho$ iff $G'_\chi = G'_\rho$ for all pre-ents G' over L. Hence clearly we have that $\chi \mathrel{\dot{\sim}} \rho$ implies $Bel^G_s(\chi) = Bel^G_s(\rho)$ and so, since $\neg(\phi \wedge \psi) \mathrel{\dot{\sim}} \neg\phi \vee \neg\psi$, we may write

$$\begin{aligned}
Bel^G_s(\neg(\phi \wedge \psi)) &= Bel^G_s(\neg\phi \vee \neg\psi) \\
&= Bel^G_s(\neg\phi) + Bel^G_s(\phi \wedge \neg\psi) \qquad \text{(using Theorem 2.5(d))} \\
&= Bel^G_s(\neg\phi) + Bel^G_s(\neg(\neg\phi \vee \psi))
\end{aligned}$$

Now $Bel^G_s(\neg\phi) = O(\lambda)$ since $\theta \mathrel{|}\!\sim_G \phi$, and $Bel^G_s(\neg(\neg\phi \vee \psi)) = O(\lambda)$ since $\theta \mathrel{|}\!\sim_G \neg\phi \vee \psi$. Hence $Bel^G_s(\neg(\phi \wedge \psi)) = O(\lambda)$ as required.

P-OR: Suppose $\theta \mathrel{|}\!\sim_G \psi$ and $\neg\theta \wedge \phi \mathrel{|}\!\sim_G \psi$. If $Bel^G(\theta \vee \phi) = 0$ then $\theta \vee \phi \mathrel{|}\!\sim_G \psi$, so suppose $Bel^G(\theta \vee \phi) \neq 0$. Let $s \in WL$ be such that $s \vdash \theta \vee \phi$ and $\frac{G_{\theta\vee\phi}(\emptyset,s)}{Bel^G(\theta\vee\phi)} \neq O(\lambda)$. We must show that $Bel^G_s(\neg\psi) = O(\lambda)$. Now $G_{\theta\vee\phi}(\emptyset, s) > 0$ (by (2.1) from Section

2.2) and so, by the inductive definition of $G_{\theta\vee\phi}$ either $s \vdash \theta$ or $s \vdash \neg\theta \wedge \phi$. If $s \vdash \theta$ then $G_{\theta\vee\phi}(\emptyset, s) = G_\theta(\emptyset, s) > 0$ so $Bel^G(\theta) \neq 0$ and

$$\frac{G_\theta(\emptyset, s)}{Bel^G(\theta)} \geq \frac{G_\theta(\emptyset, s)}{Bel^G(\theta) + Bel^G(\neg\theta \wedge \phi)} = \frac{G_\theta(\emptyset, s)}{Bel^G(\theta \vee \phi)} = \frac{G_{\theta\vee\phi}(\emptyset, s)}{Bel^G(\theta \vee \phi)}$$

Hence $\frac{G_\theta(\emptyset,s)}{Bel^G(\theta)} \neq O(\lambda)$ and so, since $\theta \mid\!\sim_G \psi$, $Bel^G_s(\neg\psi) = O(\lambda)$.

If $s \vdash \neg\theta \wedge \phi$ then $G_{\theta\vee\phi}(\emptyset, s) = G_{\neg\theta\wedge\phi}(\emptyset, s) > 0$ so $Bel^G(\neg\theta \wedge \phi) \neq 0$ and, similarly to the above,

$$\frac{G_{\neg\theta\wedge\phi}(\emptyset, s)}{Bel^G(\neg\theta \wedge \phi)} \geq \frac{G_{\theta\vee\phi}(\emptyset, s)}{Bel^G(\theta \vee \phi)}$$

giving $\frac{G_{\neg\theta\wedge\phi}(\emptyset,s)}{Bel^G(\neg\theta\wedge\phi)} \neq O(\lambda)$ and so, since $\neg\theta \wedge \phi \mid\!\sim_G \psi$, $Bel^G(\neg\psi) = O(\lambda)$ as required.

<u>P-CMO:</u> Suppose $\theta \mid\!\sim_G \phi \wedge \psi$. If $Bel^G(\theta \wedge \phi) = 0$ then $\theta \wedge \phi \mid\!\sim_G \psi$ so suppose $Bel^G(\theta \wedge \phi) \neq 0$. Then $Bel^G(\theta) \geq Bel^G(\theta \wedge \phi) > 0$ and $1 - \frac{Bel^G(\theta\wedge(\phi\wedge\psi))}{Bel^G(\theta)} = O(\lambda)$.

$$\begin{aligned} 1 - \frac{Bel^G((\theta \wedge \phi) \wedge \psi)}{Bel^G(\theta \wedge \phi)} &= 1 - \frac{Bel^G(\theta \wedge (\phi \wedge \psi))}{Bel^G(\theta \wedge \phi)} \\ &\leq 1 - \frac{Bel^G(\theta \wedge (\phi \wedge \psi))}{Bel^G(\theta)} \end{aligned}$$

Hence $1 - \frac{Bel^G((\theta\wedge\phi)\wedge\psi)}{Bel^G(\theta\wedge\phi)} = O(\lambda)$ giving $\theta \wedge \phi \mid\!\sim_G \psi$.

<u>P-RMO:</u> Suppose $\theta \not\mid\!\sim_G \neg\phi$ and $\theta \mid\!\sim_G \neg\phi \vee \psi$. Since $\theta \not\mid\!\sim_G \neg\phi$ we have

$$Bel^G(\theta) > 0 \text{ and } \frac{Bel^G(\theta \wedge \neg\neg\phi)}{Bel^G(\theta)} = \frac{Bel^G(\theta \wedge \phi)}{Bel^G(\theta)} \neq O(\lambda).$$

Hence $Bel^G(\theta \wedge \phi) > 0$. Let $a \in \mathbb{R}$ be such that $a > 0$ and $\frac{Bel^G(\theta\wedge\phi)}{Bel^G(\theta)} \geq a$ (such an a must exist since otherwise we would have $\frac{Bel^G(\theta\wedge\phi)}{Bel^G(\theta)} = O(\lambda)$). Then

$$\begin{aligned} \frac{Bel^G((\theta \wedge \phi) \wedge \neg\psi)}{Bel^G(\theta \wedge \phi)} = \frac{Bel^G(\theta \wedge (\phi \wedge \neg\psi))}{Bel^G(\theta \wedge \phi)} &= \frac{Bel^G(\theta \wedge \neg(\neg\phi \vee \psi))}{Bel^G(\theta \wedge \phi)} \\ &\leq \frac{Bel^G(\theta \wedge \neg(\neg\phi \vee \psi))}{a\, Bel^G(\theta)} \end{aligned}$$

which is of order $O(\lambda)$ since $\theta \mathrel{|\!\sim}_G \neg\phi \vee \psi$. Hence $\frac{Bel^G((\theta\wedge\phi)\wedge\neg\psi)}{Bel^G(\theta\wedge\phi)} = O(\lambda)$ giving $\theta \wedge \phi \mathrel{|\!\sim}_G \psi$. □

We now make the following definition.

Definition 5.9 *A* natural consequence relation on L *is a binary relation on* SL *which satisfies the six properties from Theorem 5.8 together with REF.*

Theorem 5.10 *Given a pre-ent* G *over* L*, the consequence relation* $\mathrel{|\!\sim}_G$ *forms a natural consequence relation on* L*.*

Proof. This is simply Theorems 5.6 and 5.8. □

The following lemma gives some rules which follow from the rules of natural consequence. They will be useful in some of the upcoming proofs.

Lemma 5.11 *Let* $\mathrel{|\!\sim}$ *be a natural consequence relation on* L*. Then* $\mathrel{|\!\sim}$ *satisfies the rule SCL from Proposition 5.2 and also following rules:*

1. $\dfrac{\theta \mathrel{|\!\sim} \psi,\ \theta \mathrel{\dot{\sim}} \phi}{\phi \mathrel{|\!\sim} \psi}$ *(Left G-Equivalence (LGE))*
2. $\dfrac{\theta \wedge \phi \mathrel{|\!\sim} \psi}{\theta \mathrel{|\!\sim} \neg\phi \vee \psi}$ *(Conditionalisation (CON))*
3. $\dfrac{\theta \vee \phi \mathrel{|\!\sim} \theta}{\theta \vee (\phi \wedge \psi) \mathrel{|\!\sim} \theta}$ *(A)*
4. $\dfrac{\theta \vee (\phi \wedge \psi) \mathrel{|\!\sim} \neg\theta}{\theta \vee \phi \mathrel{|\!\sim} \neg\theta}$ *(B)*

Proof. To show SCL suppose $\theta \vdash \phi$. Then $\theta \mathrel{\ddot{\sim}} \theta \wedge \phi$ (see Theorem 3.7). Hence

$$\theta \wedge \theta \mathrel{\dot{\sim}} \theta \mathrel{\ddot{\sim}} \theta \wedge \phi$$

and so $\theta \wedge \theta \mathrel{|\!\ddot{\sim}} \theta \wedge \phi$. Also $\theta \mathrel{|\!\sim} \theta$ by REF and so by P-RWE, $\theta \mathrel{|\!\sim} \phi$ as required. To show $\mathrel{|\!\sim}$ satisfies LGE suppose $\theta \mathrel{|\!\sim} \psi$ and $\theta \mathrel{\dot{\sim}} \phi$. Then $\theta \mathrel{\ddot{\sim}} \phi$. Also $\theta \wedge \psi \mathrel{\dot{\sim}} \phi \wedge \psi$

and so $\theta \wedge \psi \ddot{\sim} \phi \wedge \psi$ giving $\phi \mid\!\sim \psi$ by P-LLE.

To prove CON suppose $\theta \wedge \phi \mid\!\sim \psi$. Since $\phi \wedge \psi \dot{\sim} \phi \wedge (\neg\phi \vee \psi)$ (by Proposition 3.2(s)) we have

$$\theta \wedge \phi \wedge \psi \dot{\sim} \theta \wedge \phi \wedge (\neg\phi \vee \psi)$$

which gives $\theta \wedge \phi \wedge \psi \mid\!\ddot{\sim} \theta \wedge \phi \wedge (\neg\phi \vee \psi)$. And so, by P-RWE, $\theta \wedge \phi \mid\!\sim \neg\phi \vee \psi$. Using this together with P-OR and $\neg(\theta \wedge \phi) \wedge (\theta \wedge \neg\phi) \mid\!\sim \neg\phi \vee \psi$ (an instance of SCL proved above) gives $(\theta \wedge \phi) \vee (\theta \wedge \neg\phi) \mid\!\sim \neg\phi \vee \psi$ and the conclusion follows from P-LLE, since

$$(\theta \wedge \phi) \vee (\theta \wedge \neg\phi) \dot{\sim} \theta \wedge (\phi \vee \neg\phi) \ddot{\sim} \theta$$

and

$$\begin{aligned}((\theta \wedge \phi) \vee (\theta \wedge \neg\phi)) \wedge (\neg\phi \vee \psi) \quad &\dot{\sim} \quad \theta \wedge (\phi \vee \neg\phi) \wedge (\neg\phi \vee \psi)\\ &\dot{\sim} \quad \theta \wedge (\neg\phi \vee \phi) \wedge (\neg\phi \vee \psi)\\ &\dot{\sim} \quad \theta \wedge (\neg\phi \vee (\phi \wedge \psi))\\ &\dot{\sim} \quad \theta \wedge (\neg\phi \vee \psi).\end{aligned}$$

To prove (A) suppose $\theta \vee \phi \mid\!\sim \theta$. We have

$$\begin{aligned}\theta \quad &\ddot{\sim} \quad \theta \vee (\neg\theta \wedge \psi \wedge \theta)\\ &\dot{\sim} \quad \theta \vee (\psi \wedge \theta)\\ &\dot{\sim} \quad (\theta \vee \psi) \wedge (\theta \vee \theta)\\ &\dot{\sim} \quad (\theta \vee \psi) \wedge \theta.\end{aligned}$$

Hence, by Proposition 3.6, $(\theta \vee \phi) \wedge \theta \ddot{\sim} (\theta \vee \phi) \wedge (\theta \vee \psi) \wedge \theta$ and so $(\theta \vee \phi) \wedge \theta \mid\!\ddot{\sim} (\theta \vee \phi) \wedge (\theta \vee \psi) \wedge \theta$ which gives $\theta \vee \phi \mid\!\sim (\theta \vee \psi) \wedge \theta$ by P-RWE with $\theta \vee \phi \mid\!\sim \theta$. From this and P-CMO we get $(\theta \vee \phi) \wedge (\theta \vee \psi) \mid\!\sim \theta$ and the conclusion follows from this and LGE proved above.

To prove (B) suppose $\theta \vee (\phi \wedge \psi) \mathrel{|}\!\sim \neg\theta$. Then, by LGE, $(\theta \vee \phi) \wedge (\theta \vee \psi) \mathrel{|}\!\sim \neg\theta$. From this and CON proved above we obtain $\theta \vee \phi \mathrel{|}\!\sim \neg(\theta \vee \psi) \vee \neg\theta$ and the conclusion then follows from P-RWE, since

$$\begin{aligned} \neg(\theta \vee \psi) \vee \neg\theta \quad & \dot{\sim} \quad (\neg\theta \wedge \neg\psi) \vee \neg\theta \\ & \dot{\sim} \quad (\neg\theta \vee \neg\theta) \wedge (\theta \vee \neg\psi \vee \neg\theta) \qquad \text{from Proposition 3.2(q)} \\ & \ddot{\sim} \quad \neg\theta \vee \neg\theta \\ & \dot{\sim} \quad \neg\theta \end{aligned}$$

and so, via Proposition 3.6, $(\theta \vee \phi) \wedge (\neg(\theta \vee \psi) \vee \neg\theta) \mathrel{|}\!\ddot{\sim} (\theta \vee \phi) \wedge \neg\theta$. □

Note that we also have $\theta \mathrel{|}\!\ddot{\sim} \phi$ implies $\theta \mathrel{|}\!\sim \phi$ for any natural consequence relation, i.e., any natural consequence relation extends $\mathrel{|}\!\ddot{\sim}$. This follows from SCL and the fact that $\vdash$ extends $\mathrel{|}\!\ddot{\sim}$.

As expected, natural consequence relations are more general than rational consequence relations as we shall now show.

Theorem 5.12 *Every rational consequence relation on L is a natural consequence relation on L.*

Proof. Let $\mathrel{|}\!\sim$ be a rational consequence relation on L. We must check that $\mathrel{|}\!\sim$ satisfies each rule for natural consequence. $\mathrel{|}\!\sim$ satisfies REF by definition.

<u>P-LLE:</u> Suppose $\theta \mathrel{|}\!\sim \phi$, $\theta \ddot{\sim} \psi$ and $\theta \wedge \phi \ddot{\sim} \psi \wedge \phi$. Since $\equiv$ extends $\ddot{\sim}$, from $\theta \ddot{\sim} \phi$ we get $\theta \equiv \psi$ and so by LLE for $\mathrel{|}\!\sim$ we get $\psi \mathrel{|}\!\sim \phi$ as required.

<u>P-RWE:</u> Suppose $\theta \mathrel{|}\!\sim \phi$ and $\theta \wedge \phi \mathrel{|}\!\ddot{\sim} \theta \wedge \psi$. Since $\vdash$ extends $\mathrel{|}\!\ddot{\sim}$ we have that $\theta \wedge \phi \mathrel{|}\!\ddot{\sim} \theta \wedge \psi$ gives us $\theta \wedge \phi \vdash \theta \wedge \psi$, equivalently $\theta \wedge \phi \vdash \psi$. Hence $\theta \wedge \phi \mathrel{|}\!\sim \psi$ by SCL and so we conclude $\theta \mathrel{|}\!\sim \psi$ by CC (see Proposition 5.2) with $\theta \mathrel{|}\!\sim \phi$.

<u>P-AND:</u> Suppose $\theta \mathrel{|}\!\sim \phi$ and $\theta \mathrel{|}\!\sim \neg\phi \vee \psi$. Then, by AND, $\theta \mathrel{|}\!\sim \phi \wedge (\neg\phi \vee \psi)$ so $\theta \mathrel{|}\!\sim \phi \wedge \psi$ by RWE as required.

P-OR: Suppose $\theta \mid\!\sim \psi$ and $\neg\theta \wedge \phi \mid\!\sim \psi$. Then, by OR, $\theta \vee (\neg\theta \wedge \phi) \mid\!\sim \psi$ and so, by LLE, $\theta \vee \phi \mid\!\sim \psi$ as required.

P-CMO: Suppose $\theta \mid\!\sim \phi \wedge \psi$. Then $\theta \mid\!\sim \phi$ and $\theta \mid\!\sim \psi$ by RWE. So $\theta \wedge \phi \mid\!\sim \psi$ by CMO as required.

P-RMO: Suppose $\theta \not\mid\!\sim \neg\phi$ and $\theta \mid\!\sim \neg\phi \vee \psi$. Then $\theta \wedge \phi \mid\!\sim \neg\phi \vee \psi$ by RMO. Also $\theta \wedge \phi \mid\!\sim \phi$ by SCL so $\theta \wedge \phi \mid\!\sim \phi \wedge (\neg\phi \vee \psi)$ by AND and then $\theta \wedge \phi \mid\!\sim \psi$ by RWE as required. □

As can be seen from Examples 5.7 the converse to Theorem 5.12 is false. The fact that all the rules for natural consequence are sound for rational consequence relations can at least be seen as an indication of their reasonableness for an intelligent agent.

Recall that our aim now is to try and characterise our newly-defined class of natural consequence relations in terms of the family of relations $\mid\!\sim_G$. By Theorem 5.10 we know that the rules for natural consequence are sound for $\mid\!\sim_G$. Hence it remains to show that those rules are also *complete* for $\mid\!\sim_G$, i.e., that *any* binary relation on SL which satisfies those rules is given by $\mid\!\sim_G$ for some pre-ent G over L. This representation theorem we seek for natural consequence relations in terms of λ-pre-ents may be thought of as the analogue of the one given for rational consequence relations in terms of λ-probability functions just after Proposition 5.2. But what of the other representation theorems in the rational case? Might not a consideration of *them* yield a characterisation of natural consequence relations? For example, we know (from [7], see also Section 1 of [2]) that for every finite sequence $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At^L$, if we define the relation $\mid\!\sim_{\vec{\mathcal{U}}}$ on SL by, for $\theta, \phi \in SL$,

$$\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi \quad \text{iff} \quad \textit{either } \mathcal{U}_i \cap S_\theta = \emptyset \text{ for all } 1 \leq i \leq k$$

or there exists an i such that $\mathcal{U}_i \cap S_\theta \neq \emptyset$ and for the least such i we have $\mathcal{U}_i \cap S_\theta \subseteq S_\phi$

then $\mid\!\sim_{\vec{\mathcal{U}}}$ forms a rational consequence relation on L and, conversely for every rational consequence relation $\mid\!\sim$ on L there corresponds a $\vec{\mathcal{U}}$ such that $\mid\!\sim = \mid\!\sim_{\vec{\mathcal{U}}}$. Can a representation analogous to this one exist for natural consequence relations? The next section makes a start on finding an answer to this question.

5.4 Permatoms and T_θ

We begin this section with some new notation. Recall the definition of the set At^L of atoms over L:

$$At^L = \{p_1^{\epsilon_1} \wedge p_2^{\epsilon_2} \wedge \ldots \wedge p_n^{\epsilon_n} \mid \epsilon_i \in \{0,1\} \text{ for } i = 1, 2, \ldots, n\}.$$

Given that, so far, our examples of natural consequence relations have given the impression of natural consequence relations as being "like rational consequence relations for which the order matters", we shall make the following definition.

Definition 5.13 *We define the set of* permatoms*, At_*^L, of L as follows*

$$\begin{aligned} At_*^L \;=\; & \{p_{\sigma(1)}^{\epsilon_1} \wedge \ldots \wedge p_{\sigma(n)}^{\epsilon_n} \mid \sigma \text{ is a permutation on } \{1, \ldots, n\} \text{ and} \\ & \epsilon_i \in \{0,1\} \text{ for } i = 1, \ldots, n\}. \end{aligned}$$

Clearly we have $At^L \subseteq At_*^L$ and so permatoms are a generalisation of atoms. Recall the definition of S_θ for $\theta \in SL$:

$$S_\theta = \{\alpha \in At^L \mid \alpha \vdash \theta\}.$$

By analogy, we define the set $T_\theta \subseteq At_*^L$ as follows:

Definition 5.14 *for each $\theta \in SL$ we define the set $T_\theta \subseteq At_*^L$ by*

$$T_\theta = \begin{cases} At_*^L & \text{if} \vdash \theta \\ \emptyset & \text{if} \vdash \neg\theta \\ \{\delta \in At_*^L \mid \delta \text{ has an initial segment which} \\ \qquad\qquad \text{is an element of } rT(\theta)^+\} & \text{if } \theta \text{ is contingent} \end{cases}$$

Thus, according to the above definition, if θ is contingent then T_θ contains all those permatoms which have an initial segment which is a positive clause, minus the last repeat, of $rT(\theta)$. In particular, for τ a conjunction of literals from distinct propositional variables in L, T_τ consists of all those permatoms of L which contain τ as an initial segment. Note that, for $\delta \in At_*^L$, the only element in $rT(\delta)^+$ is δ itself and so, in light of Proposition 3.23, for contingent $\theta \in SL$ we may write

$$T_\theta = \{\delta \in At_*^L \mid \delta \mid\!\overset{\cdot\cdot}{\sim} \theta\}.$$

In fact we may also do this even if θ is not contingent, since if $\vdash \theta$ then for all $\delta \in At_*^L$ we have $\delta \mid\!\overset{\cdot\cdot}{\sim} \theta$, giving the required $T_\theta = At_*^L$, while if $\vdash \neg\theta$ then, given $\delta \in At_*^L$, we have that $\delta \mid\!\overset{\cdot\cdot}{\sim} \theta$ iff $Bel^G(\delta) \leq Bel^G(\theta) = 0$ for all G iff $Bel^G(\delta) = 0$ for all G iff $\vdash \neg\delta$ which is false. Hence in this case we also get the required answer, i.e., $T_\theta = \emptyset$.

The above Definition 5.14 states that if θ is not a contingent sentence then T_θ is one of the two "extreme" subsets of At_*^L. The following proposition says that *only* if θ is a non-contingent sentence can T_θ be one of the extreme subsets of At_*^L.

Proposition 5.15 *Let $\theta, \phi \in SL$. Then*
(i). $T_\theta = At_*^L$ iff $\vdash \theta$.
(ii). $T_\theta = \emptyset$ iff $\vdash \neg\theta$.

Proof. *(i).* If $\vdash \theta$ then $T_\theta = At_*^L$ by definition of T_θ. Suppose $\nvdash \theta$. Then if $\vdash \neg\theta$ we have $T_\theta = \emptyset \neq At_*^L$ while if θ is contingent then $rT(\theta)$ is well-defined and, by

Proposition 3.17, there exists at least one negative clause τ_1, say, in $rT(\theta)$. Choose any $\delta \in At_*^L$ such that δ has τ_1 (minus its last repeated propositional variable) as an initial segment. By property (4) of Lemma 3.12, no initial segment of δ can be an element of $rT(\theta)^+$. Hence $\delta \notin T_\theta$ so $T_\theta \neq At_*^L$ as required.

(ii). If $\vdash \neg\theta$ then $T_\theta = \emptyset$ by definition of T_θ. Suppose $\nvdash \neg\theta$. Then if $\vdash \theta$ we have $T_\theta = At_*^L \neq \emptyset$ while if θ is contingent then $rT(\theta)$ is well defined and, again by Lemma 3.17, must have at least one positive clause τ_2, say. Choose any $\delta \in At_*^L$ such that δ has τ_2 (minus its last repeat) as an initial segment. Then clearly $\delta \in T_\theta$ so $T_\theta \neq \emptyset$ as required. □

We also have the following:

Proposition 5.16 *Let $\theta \in SL$. Then, for any $\delta \in At_*^L$, $\delta \in T_\theta$ implies $\delta \vdash \theta$.*

Proof. We have $\delta \in T_\theta$ iff $\delta \mid\!\ddot{\sim} \theta$. Hence the result follows since $\vdash$ extends $\mid\!\ddot{\sim}$. □

The following lemma and proposition will help to characterise the binary relation R defined on SL by, for $\theta, \phi \in SL$, $\theta R \phi$ iff $T_\theta = T_\phi$ and will, in fact show that this relation coincides with the relation $\ddot{\sim}$ on SL.

Lemma 5.17 *Let $\theta \in SL$ be contingent and let τ be a conjunction of literals from distinct propositional variables in L. Then $\tau \wedge \nu \in T_\theta$ for all conjunctions ν of literals from the remaining propositional variables in L iff τ has an initial segment which is an element $rT(\theta)^+$.*

Proof. The "if" direction is obvious. For the "only if" direction suppose that, for all ν we have $\tau \wedge \nu \in T_\theta$. Let $q_1, \ldots, q_l$ be all the propositional variables not appearing in τ (we may assume $l > 0$ since otherwise the result holds trivially). Then $\tau \wedge q_1 \wedge \ldots \wedge q_l \in T_\theta$. We will show that if no initial segment of τ is an element of $rT(\theta)^+$ then there must exist some conjunction ν such that $\tau \wedge \nu \notin T_\theta$. But $\tau \wedge q_1 \wedge \ldots \wedge q_l \in T_\theta$ implies that if no initial segment of τ is an element of

$rT(\theta)^+$ then τ itself must be a proper initial segment of an element of $rT(\theta)^+$. So, by Proposition 3.17, τ must be an initial segment of a negative clause of $rT(\theta)$. Let this negative clause, minus its last repeated propositional variable, be $\tau \wedge \rho_1$ where ρ_1 is a, possibly empty, conjunction of literals. Let ρ_2 be a conjunction of all the propositional variables not appearing in $\tau \wedge \rho_1$. Then, using property (4) of Lemma 3.12, no initial segment of $\tau \wedge \rho_1 \wedge \rho_2$ can be an element of $rT(\theta)^+$ and so $\tau \wedge \rho_1 \wedge \rho_2 \notin T_\theta$ as required. □

Proposition 5.18 *Let $\theta, \phi \in SL$ be contingent sentences. Then the following are equivalent.*
(i). $T_\theta \subseteq T_\phi$.
(ii). Every $\tau \in rT(\theta)^+$ has an initial segment which is an element of $rT(\phi)^+$.

Proof. To show that (i) implies (ii), let τ be an element of $rT(\theta)^+$ for which no initial segment is an element of $rT(\phi)^+$. Then, for all conjunctions ν of literals from all the propositional variables not appearing in τ, $\tau \wedge \nu \in T_\theta$, but by Lemma 5.17, there exists ν such that $\tau \wedge \nu \notin T_\phi$. Hence $T_\theta \not\subseteq T_\phi$.
To show that (ii) implies (i), suppose every element of $rT(\theta)^+$ has an initial segment which is an element of $rT(\phi)^+$. Let $\delta \in T_\theta$. Then $\delta = \tau \wedge \ldots$ for some τ an element of $rT(\theta)^+$. Hence δ must have an initial segment which is an element of $rT(\phi)^+$ (since τ does), so $\delta \in T_\phi$ as required. □

Corollary 5.19 *Let θ, $\phi \in SL$. Then $T_\theta \subseteq T_\phi$ iff $\theta \mathrel{|\!\ddot{\sim}} \phi$ (and so $T_\theta = T_\phi$ iff $\theta \mathrel{\ddot{\sim}} \phi$).*

Proof. The case where θ and ϕ are both contingent sentences is handled by Propositions 3.23 and 5.18.
Let us suppose, then, that it is not the case that both θ and ϕ are contingent. First let us assume that θ is non-contingent. If $\vdash \neg\theta$ then $T_\theta = \emptyset$ while $Bel(\theta) = 0$ for all pre-ents over L. Hence, for *any* ϕ, we automatically have both $T_\theta \subseteq T_\phi$

and $\theta \mathrel{|\!\ddot{\sim}} \phi$ which suffices. If $\vdash \theta$ then $T_\theta = At^L_*$ while $Bel(\theta) = 1$ for all pre-ents over L. Hence $T_\theta \subseteq T_\phi$ iff $T_\phi = At^L_*$ iff (Proposition 5.15) $\vdash \phi$ while $\theta \mathrel{|\!\ddot{\sim}} \phi$ iff $Bel(\phi) = 1$ for all pre-ents iff (Theorem 2.7) $\vdash \phi$. Hence $T_\theta \subseteq T_\phi$ iff $\vdash \phi$ iff $\theta \mathrel{|\!\ddot{\sim}} \phi$ as required.

Let us suppose now that θ is contingent and that it is ϕ which is non-contingent. Suppose $T_\theta \subseteq T_\phi$. Then we cannot have $\vdash \neg\phi$ since if we did we would have $T_\phi = \emptyset$ and so $T_\theta = \emptyset$ which implies, by Proposition 5.15, $\vdash \neg\theta$, contradicting θ being contingent. Hence we must have $\vdash \phi$ which gives $\theta \mathrel{|\!\ddot{\sim}} \phi$ as required since $Bel(\phi) = 1$ for all pre-ents over L. Conversely suppose $\theta \mathrel{|\!\ddot{\sim}} \phi$. Then again we cannot have $\vdash \neg\phi$ since if so then $Bel(\phi) = 0$ for all pre-ents over L which gives $Bel(\theta) = 0$ for all pre-ents over L and so, by Theorem 2.7, $\vdash \neg\theta$, contradicting θ being contingent. Hence again we must have $\vdash \phi$ and so $T_\theta \subseteq T_\phi$ as required since in this case $T_\phi = At^L_*$. □

With our new notation in place, we now provide a first opportunity to see it in action by using it to prove a characterisation of rational consequence relations.

5.5 Characterising Rational Consequence

In this section we shall provide a representation result for rational consequence relations. This characterisation is essentially the same as the one mentioned at the end of Section 5.3 in terms of atoms which was given originally (for arbitrary, possibly infinite propositional languages) by Lehmann and Magidor in [7]. Our proof, however, uses slightly alternative methods. We set it in the framework of the previous section to allow it to fit more easily with the results obtained for *natural* consequence later in this thesis. The techniques used in proving this result will, to an extent, be transferable to our proof of those results.

We begin by defining the binary relation $\mathrel{|\!\sim}_{\vec{\mathcal{U}}} \subseteq SL \times SL$ for $\vec{\mathcal{U}}$ a completely general finite sequence of subsets of permatoms.

Definition 5.20 *Let* $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ *(*$k \geq 0$*) be a finite sequence of sets of permatoms over* L*. We define the consequence relation* $\mathrel{|\!\sim}_{\vec{\mathcal{U}}} \subseteq SL \times SL$ *as follows, for* $\theta, \phi \in SL$*:*

$$\theta \mathrel{|\!\sim}_{\vec{\mathcal{U}}} \phi \quad \textit{iff} \quad \textit{either} \quad \mathcal{U}_i \cap T_\theta = \emptyset \textit{ for all } i$$
$$\textit{or} \quad \mathcal{U}_i \cap T_{\theta \wedge \neg\phi} = \emptyset \textit{ for the least } i \textit{ such that } \mathcal{U}_i \cap T_\theta \neq \emptyset.$$

We shall characterise rational consequence by concentrating on those sequences $\vec{\mathcal{U}}$ which are *admissible* according to the following definition.

Definition 5.21 *Let* $\mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ *be a finite sequence of sets of permatoms. Then this sequence is* admissible *iff it satisfies the following condition:*

(AD) *For all* $i = 1, \ldots, k$ *and all* $\delta \in At_*^L$*, if* $\delta \in \mathcal{U}_i$ *then* $\delta' \in \mathcal{U}_i$ *for all* $\delta' \in At_*^L$ *such that* $\delta' \equiv \delta$*.*

Hence, if a sequence $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k$ is admissible, the position of any permatom δ in that sequence (i.e., the set of i such that $\delta \in \mathcal{U}_i$) is independent of the order we take the literals in δ to be in. Thus we really only need to look at the position of each *atom*. In the next chapter we shall weaken this condition for our attempts to characterise natural consequence. Note that if $\mathcal{U}_i = \emptyset$ for all $i = 1, \ldots, k$ then the sequence $\mathcal{U}_1, \ldots, \mathcal{U}_k$ is vacuously admissible.

The condition **(AD)** was given in that form to make it easy to check. However it has an equivalent formulation which we will find useful. To show it we need the following lemma, which should be obvious.

Lemma 5.22 *For all* $\delta \in At_*^L$ *and all* $\theta \in SL$*, if* $\delta \vdash \theta$ *then there exists some* $\delta' \in At_*^L$ *such that* $\delta' \equiv \delta$ *and* $\delta' \in T_\theta$*.* □

Lemma 5.23 *Let* $\mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ *be a finite sequence of sets of permatoms. Then the following are equivalent:*

(i). $\mathcal{U}_1, \ldots, \mathcal{U}_k$ *is admissible.*

(ii). For all $i = 1, \ldots, k$ *and all* $\theta \in SL$ *we have* $\mathcal{U}_i \cap T_\theta \neq \emptyset$ *iff* $\mathcal{U}_i \cap S_\theta \neq \emptyset$*.*

Proof. To show that (i) implies (ii), let $i \in \{1, \ldots, k\}$ and let $\theta \in SL$. Suppose $\mathcal{U}_i \cap T_\theta \neq \emptyset$ and $\delta \in \mathcal{U}_i \cap T_\theta$. Then, by Proposition 5.16, $\delta \vdash \theta$ and so there exists $\alpha \in S_\theta$ such that $\alpha \equiv \delta$ (take α to be such that $\{\alpha\} = S_\delta$). By admissibility we have $\alpha \in \mathcal{U}_i$ which gives $\mathcal{U}_i \cap S_\theta \neq \emptyset$ as required. Conversely let $\alpha \in \mathcal{U}_i \cap S_\theta$. Then $\alpha \vdash \theta$ so, by Lemma 5.22, there exists $\delta \in T_\theta$ such that $\delta \equiv \alpha$. By admissibility we get $\delta \in \mathcal{U}_i$ and so $\mathcal{U}_i \cap T_\theta \neq \emptyset$ as required.
To show that (ii) implies (i) suppose that $\mathcal{U}_1, \ldots, \mathcal{U}_k$ is *not* admissible. Then there exists $i \in \{1, \ldots, k\}$ and $\delta, \delta' \in At_*^L$ such that $\delta \in \mathcal{U}_i$, $\delta' \notin \mathcal{U}_i$ and $\delta \equiv \delta'$. Suppose $S_\delta = \{\alpha\}$. Then if $\alpha \in \mathcal{U}_i$ we have $\mathcal{U}_i \cap S_{\delta'} = \mathcal{U}_i \cap S_\delta \neq \emptyset$ and $\mathcal{U}_i \cap T_{\delta'} = \mathcal{U}_i \cap \{\delta'\} = \emptyset$ as required. If $\alpha \notin \mathcal{U}_i$ then $\mathcal{U}_i \cap T_\delta = \mathcal{U}_i \cap \{\delta\} \neq \emptyset$ and $\mathcal{U}_i \cap S_\delta = \emptyset$ as required. □

This reformulation of **(AD)** makes the following easier to show.

Theorem 5.24 *For* $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ *an admissible sequence of sets of permatoms, the consequence relation* $\mid\!\sim_{\vec{\mathcal{U}}}$ *is a rational consequence relation on* L*.*

Proof. By Lemma 5.23 we have

$$\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi \quad \text{iff} \quad \textit{either} \quad \mathcal{U}_i \cap S_\theta = \emptyset \text{ for all } i$$
$$\textit{or} \quad \mathcal{U}_i \cap S_{\theta \wedge \neg\phi} = \emptyset \text{ for the least } i \text{ such that } \mathcal{U}_i \cap S_\theta \neq \emptyset$$

From here it is straightforward to check that $\mid\!\sim_{\vec{\mathcal{U}}}$ given in this way satisfies the properties given in Definition 5.1 and so $\mid\!\sim_{\vec{\mathcal{U}}}$ is a rational consequence relation as required. □

Thus we have that, for an admissible sequence $\vec{\mathcal{U}}$, the relation $\mid\!\sim_{\vec{\mathcal{U}}}$ forms a rational consequence relation. We now show that *every* rational consequence relation on L is given by an admissible sequence of sets of permatoms, i.e., that for every rational consequence relation $\mid\!\sim$ on L there exists an admissible sequence

$\vec{\mathcal{U}}$ such that $\mid\!\sim=\mid\!\sim_{\vec{\mathcal{U}}}$. Before we describe how to construct $\vec{\mathcal{U}}$ from $\mid\!\sim$ we give some derived rules for rational consequence which we shall find useful. Properties (1) (without proof) and (4) below were given in [7] (in Lemmas 3.10 and 3.11 there). We include both their proofs below for completeness.

Lemma 5.25 *The following are derived rules for rational consequence relations:*

1. $\dfrac{\theta \vee \phi \mid\!\sim \neg\phi}{\theta \vee \phi \vee \psi \mid\!\sim \neg\phi}$

2. $\dfrac{\theta \mid\!\sim \phi,\ \theta \vee \psi \mid\!\sim \neg\psi}{\theta \vee \psi \mid\!\sim \phi}$

3. $\dfrac{\theta \vee \phi \mid\!\sim \bot}{\theta \mid\!\sim \bot}$

4. $\dfrac{\theta \vee \phi \mid\!\sim \neg\phi,\ \theta \vee \psi \not\mid\!\sim \neg\psi}{\psi \vee \phi \mid\!\sim \neg\phi}$

5. $\dfrac{\theta_1 \vee \ldots \vee \theta_m \mid\!\sim \phi,\ \theta_i \vee \psi \mid\!\sim \neg\psi \text{ for some } i = 1, \ldots, m}{\theta_1 \vee \ldots \vee \theta_m \vee \psi \mid\!\sim \phi}$

6. $\dfrac{\theta \mid\!\sim \bot}{\phi \mid\!\sim \neg\theta}$

Proof. (1). From $\theta \vee \phi \mid\!\sim \neg\phi$ and $\theta \vee \phi \mid\!\sim \theta \vee \phi \vee \psi$ (SCL) using CMO we get $(\theta \vee \phi) \wedge (\theta \vee \phi \vee \psi) \mid\!\sim \neg\phi$ which, in turn, gives $(\theta \vee \phi) \vee (\psi \wedge \phi) \mid\!\sim \neg\phi$ by LLE. Combining this with $\psi \wedge \neg\phi \mid\!\sim \neg\phi$ (SCL again) using OR yields $(\theta \vee \phi) \vee (\psi \wedge \phi) \vee (\psi \wedge \neg\phi) \mid\!\sim \neg\phi$ and we conclude by LLE.

(2). From $\theta \vee \psi \mid\!\sim \neg\psi$ together with $\theta \vee \psi \mid\!\sim \theta \vee \psi$ (REF) we obtain, using AND and RWE, $\theta \vee \psi \mid\!\sim \theta$. We also have, by SCL, $\theta \mid\!\sim \theta \vee \psi$ and so we may conclude using Equivalence (see Proposition 5.2) with $\theta \mid\!\sim \phi$.

(3). From $\theta \vee \phi \mid\!\sim \bot$ we get $\theta \vee \phi \mid\!\sim \theta$ by RWE. We also have $\theta \mid\!\sim \theta \vee \phi$ by SCL. Hence, from these two, we may apply Equivalence with $\theta \vee \phi \mid\!\sim \bot$ to obtain $\theta \mid\!\sim \bot$ as required.

(4). From $\theta \vee \phi \mid\!\sim \neg\phi$ we obtain $\theta \vee \phi \vee \psi \mid\!\sim \neg\phi$ from (1) proved above. If it were the case that $\theta \vee \psi \not\mid\!\sim \neg\psi$ implies $\theta \vee \phi \vee \psi \not\mid\!\sim \neg(\phi \vee \psi)$ then we could use

this with $\theta \vee \phi \vee \psi \mathrel{|\!\sim} \neg\phi$ and RMO to obtain $(\theta \vee \phi \vee \psi) \wedge (\phi \vee \psi) \mathrel{|\!\sim} \neg\phi$ and then conclude by LLE. Hence our result will be proved if we can show $\theta \vee \psi \mathrel{|\!\not\sim} \neg\psi$ implies $\theta \vee \phi \vee \psi \mathrel{|\!\not\sim} \neg(\phi \vee \psi)$, equivalently $\theta \vee \phi \vee \psi \mathrel{|\!\sim} \neg(\phi \vee \psi)$ implies $\theta \vee \psi \mathrel{|\!\sim} \neg\psi$. But from $\theta \vee \phi \vee \psi \mathrel{|\!\sim} \neg(\phi \vee \psi)$ we get both $\theta \vee \phi \vee \psi \mathrel{|\!\sim} \theta \vee \psi$ (using REF, AND, RWE) and $\theta \vee \phi \vee \psi \mathrel{|\!\sim} \neg\psi$ (using RWE). Hence, applying CMO to these two, we obtain $(\theta \vee \phi \vee \psi) \wedge (\theta \vee \psi) \mathrel{|\!\sim} \neg\psi$ and so $\theta \vee \psi \mathrel{|\!\sim} \neg\psi$, from LLE, as required.
(5). From $\theta_i \vee \psi \mathrel{|\!\sim} \neg\psi$ and property (1) of this lemma we obtain $\theta_1 \vee \ldots \vee \theta_m \vee \psi \mathrel{|\!\sim} \neg\psi$. Using this together with $\theta_1 \vee \ldots \vee \theta_m \mathrel{|\!\sim} \phi$ and property (2) of this lemma gives the required conclusion.
(6). From $\theta \mathrel{|\!\sim} \bot$ we get $\theta \mathrel{|\!\sim} \phi$ from RWE. Hence using these two with CMO gives $\theta \wedge \phi \mathrel{|\!\sim} \bot$ and so, by RWE, $\theta \wedge \phi \mathrel{|\!\sim} \neg\theta$. Meanwhile, by SCL, we have $\neg\theta \wedge \phi \mathrel{|\!\sim} \neg\theta$ and so applying OR to this and $\theta \wedge \phi \mathrel{|\!\sim} \neg\theta$ yields $(\neg\theta \wedge \phi) \vee (\theta \wedge \phi) \mathrel{|\!\sim} \neg\theta$ and so the conclusion follows from LLE. □

We will now show how to construct, from a given rational consequence relation $\mathrel{|\!\sim}$, an admissible sequence $\vec{\mathcal{U}}$ such that $\mathrel{|\!\sim} = \mathrel{|\!\sim}_{\vec{\mathcal{U}}}$. So let $\mathrel{|\!\sim}$ be our given rational consequence relation on L. We begin our construction process by setting

$$\mathcal{U} = \mathcal{U}(\mathrel{|\!\sim}) = \{\delta \in At^L_* \mid \delta \mathrel{|\!\not\sim} \bot\}.$$

For a permatom $\delta \in At^L_*$, $\delta \mathrel{|\!\sim} \bot$ has the intuitive meaning: "if δ is true then, typically, $\bot$ is also true". In effect what this is saying is that, according to $\mathrel{|\!\sim}$, the possibility that δ is true should not be entertained at all. In this sense, then, we may think of $\mathcal{U}$ as the set consisting of those permatoms which are "consistent" for $\mathrel{|\!\sim}$. As a next step we define, from $\mathrel{|\!\sim}$, a binary relation $\prec_\sim$ on At^L_* by setting, for $\delta_1, \delta_2 \in At^L_*$,

$$\delta_1 \prec_\sim \delta_2 \text{ iff } \delta_1 \vee \delta_2 \mathrel{|\!\sim} \neg\delta_2.$$

(Note that we could define $\prec_\sim$ in this way from *any* consequence relation on SL.) For $\delta_1, \delta_2 \in At^L_*$, the intuitive meaning behind $\delta_1 \vee \delta_2 \mathrel{|\!\sim} \neg\delta_2$ is: "if either δ_1 or δ_2

is true then, typically, δ_2 will be false". According to $\mid\!\sim$ then, $\delta_1 \prec_\sim \delta_2$ says that δ_1 is more natural than, or "preferred" to, δ_2. We need the following properties of $\prec_\sim$.

Lemma 5.26 *The relation $\prec_\sim$ defined from $\mid\!\sim$ above is irreflexive on $\mathcal{U}$ and transitive on At_*^L.*

Proof. Let $\delta_i \in \mathcal{U}$ for $i = 1, 2, 3$. To show that $\prec_\sim$ is irreflexive on $\mathcal{U}$ suppose for contradiction that we had $\delta_1 \prec_\sim \delta_1$, equivalently $\delta_1 \vee \delta_1 \mid\!\sim \neg\delta_1$. Then, by LLE, we would get $\delta_1 \mid\!\sim \neg\delta_1$. We also have $\delta_1 \mid\!\sim \delta_1$ from REF and so using these two with AND gives us $\delta_1 \mid\!\sim \delta_1 \wedge \neg\delta_1$ which gives $\delta_1 \mid\!\sim \bot$ by RWE. Hence $\delta_1 \notin \mathcal{U}$ giving the required contradiction. Note that the restriction to $\mathcal{U}$ is necessary here – $\prec_\sim$ is not irreflexive on the whole set At_*^L.
To show transitivity suppose $\delta_1 \prec_\sim \delta_2$ and $\delta_2 \prec_\sim \delta_3$, i.e., $\delta_1 \vee \delta_2 \mid\!\sim \neg\delta_2$ and $\delta_2 \vee \delta_3 \mid\!\sim \neg\delta_3$. We must show $\delta_1 \vee \delta_3 \mid\!\sim \neg\delta_3$. But from $\delta_1 \vee \delta_2 \mid\!\sim \neg\delta_2$ we get $\delta_1 \vee \delta_2 \vee \delta_3 \mid\!\sim \neg\delta_2$ from Lemma 5.25(1) (and LLE). Using this together with $\delta_1 \vee \delta_2 \vee \delta_3 \mid\!\sim \delta_1 \vee \delta_2 \vee \delta_3$ (an instance of REF), AND and RWE yields $\delta_1 \vee \delta_2 \vee \delta_3 \mid\!\sim \delta_1 \vee \delta_3$. Meanwhile we can also apply Lemma 5.25(1) (and LLE) to $\delta_2 \vee \delta_3 \mid\!\sim \neg\delta_3$ to obtain $\delta_1 \vee \delta_2 \vee \delta_3 \mid\!\sim \neg\delta_3$ which, when combined with the just obtained $\delta_1 \vee \delta_2 \vee \delta_3 \mid\!\sim \delta_1 \vee \delta_3$ and CMO, gives us $(\delta_1 \vee \delta_2 \vee \delta_3) \wedge (\delta_1 \vee \delta_3) \mid\!\sim \neg\delta_3$. Hence $\delta_1 \vee \delta_3 \mid\!\sim \neg\delta_3$ by LLE as required. □

By Lemma 5.26 the relation $\prec_\sim$ forms a strict partial order on the set $\mathcal{U}$. Hence it makes sense to talk about the minimal elements in subsets of $\mathcal{U}$ under $\prec_\sim$. Let us inductively define a sequence of sets of permatoms $\mathcal{U}_1, \mathcal{U}_2, \ldots$ by, for each $i = 1, 2, \ldots$, setting

$$\mathcal{U}_i = \{\delta \in At_*^L \mid \delta \in \mathcal{U} \text{ and } \delta \text{ is minimal in } \mathcal{U} - \textstyle\bigcup_{j<i} \mathcal{U}_j \text{ under } \prec_\sim\}.$$

So $\mathcal{U}_1$ contains all the most natural (according to $\prec_\sim$) permatoms in $\mathcal{U}$, $\mathcal{U}_2$ contains all the most natural permatoms in $\mathcal{U} - \mathcal{U}_1$ etc. Clearly the $\mathcal{U}_i$'s so constructed are pairwise disjoint so, by the finiteness of At^L_*, there exists k such that $\mathcal{U}_i = \emptyset$ for all $i > k$. Hence we arrive, from our given rational consequence relation $\mid\!\sim$, at a finite sequence $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mid\!\sim) = \mathcal{U}_1, \ldots, \mathcal{U}_k$ of pairwise disjoint sets of permatoms with $\mathcal{U} = \bigcup_{i=1}^k \mathcal{U}_i$. We next show that this sequence is admissible.

Lemma 5.27 *The sequence $\vec{\mathcal{U}}$ defined above is admissible.*

Proof. We will show by induction on i that, for each $i = 1, \ldots, k$, for each $\delta \in At^L_*$, if $\delta \in \mathcal{U}_i$ then $\delta' \in \mathcal{U}_i$ for all $\delta' \in At^L_*$ such that $\delta' \equiv \delta$. To begin with, for $i = 1$, let $\delta \in \mathcal{U}_1$. Then $\delta \in \mathcal{U}$ and δ is minimal in $\mathcal{U}$ under $\prec_\sim$. Let $\delta' \in At^L_*$ be such that $\delta' \equiv \delta$. We must firstly make sure that $\delta' \in \mathcal{U}$. But $\delta \in \mathcal{U}$ implies $\delta \mid\!\not\sim \bot$ which, in turn, implies $\delta' \mid\!\not\sim \bot$ by LLE, and so $\delta' \in \mathcal{U}$ as required. It remains to show that δ' is minimal in $\mathcal{U}$ under $\prec_\sim$. But suppose for contradiction that δ' was not minimal in $\mathcal{U}$ under $\prec_\sim$. Then there must exist $\gamma \in \mathcal{U}$ such that $\gamma \prec_\sim \delta'$, i.e., $\gamma \vee \delta' \mid\!\sim \neg\delta'$ which gives $\gamma \vee \delta \mid\!\sim \neg\delta$ by LLE and RWE. Hence also $\gamma \prec_\sim \delta$ and so δ is not minimal in $\mathcal{U}$ under $\prec_\sim$, giving the required contradiction. This completes the base stage of the induction. Now suppose, for inductive hypothesis, that $1 < i \leq k$ and that for all $j < i$ we have that, for all $\delta \in At^L_*$, if $\delta \in \mathcal{U}_j$ then $\delta' \in \mathcal{U}_j$ for all $\delta' \in At^L_*$ such that $\delta' \equiv \delta$. Fix $\delta \in \mathcal{U}_i$. So $\delta \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ and δ is minimal in this set under $\prec_\sim$. Fix $\delta' \in At^L_*$ such that $\delta' \equiv \delta$. We must show that $\delta' \in \mathcal{U}_i$. As was proved above, we know that, since $\delta \in \mathcal{U}$, we have $\delta' \in \mathcal{U}$. Also if it were the case that $\delta' \in \bigcup_{j<i} \mathcal{U}_j$ then, by our inductive hypothesis, we would also have $\delta \in \bigcup_{j<i} \mathcal{U}_j$ giving a contradiction. Hence $\delta' \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$. If δ' were not minimal in this set under $\prec_\sim$ then, similarly to the case proved above, neither would δ be – contradiction. Hence $\delta' \in \mathcal{U}_i$ as required. □

We are now in a position to prove our required characterisation.

Theorem 5.28 *Let $\mid\!\sim$ be a rational consequence relation on L. Then there exists an admissible sequence $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At^L_*$ of sets of permatoms such that*

$\mid\!\sim = \mid\!\sim_{\vec{\mathcal{U}}}$.

Proof. From our given $\mid\!\sim$ we define the sequence $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mid\!\sim) = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At^L_*$ via the preference relation $\prec_\sim$ as in the construction process outlined above. This sequence is admissible by Lemma 5.27. We will now show that

$$\theta \mid\!\sim \phi \quad \text{iff} \quad \textit{either} \quad \mathcal{U}_i \cap S_\theta = \emptyset \text{ for all } i$$
$$\textit{or} \quad \mathcal{U}_i \cap S_{\theta \wedge \neg\phi} = \emptyset \text{ for the least } i \text{ such that } \mathcal{U}_i \cap S_\theta \neq \emptyset$$

which suffices by Lemma 5.23. To show the "only if" direction here suppose that $\theta \mid\!\sim \phi$. If $\mathcal{U}_i \cap S_\theta = \emptyset$ for all i then we are done so suppose i is such that $\mathcal{U}_i \cap S_\theta \neq \emptyset$ and furthermore that i is minimal such that this is true. Let $\alpha \in S_{\theta\wedge\neg\phi}$. We will show that $\alpha \notin \mathcal{U}_i$ which will prove $\mathcal{U}_i \cap S_{\theta\wedge\neg\phi} = \emptyset$ as required. Firstly if $\alpha \notin \mathcal{U}$ then $\alpha \notin \mathcal{U}_i$ as required so let us assume $\alpha \in \mathcal{U}$. If $\alpha \in \mathcal{U}_j$ for some $j < i$ then we would have $\mathcal{U}_j \cap S_\theta \neq \emptyset$ (since $\alpha \in S_{\theta\wedge\neg\phi} \subseteq S_\theta$) which would contradict the minimality of i. Hence $\alpha \notin \bigcup_{j<i} \mathcal{U}_j$ so it remains to show that α is not minimal in $\mathcal{U} - \bigcup_{j<i} \mathcal{U}_i$ under $\prec_\sim$. But $\theta \mid\!\sim \phi$ implies, by LLE, $\bigvee S_\theta \mid\!\sim \phi$ which, in turn, implies $\bigvee S_\theta \mid\!\sim \neg\alpha$ by RWE since $\alpha \in S_{\neg\phi}$. If $S_\theta = \{\alpha\}$ then we have $\alpha \mid\!\sim \neg\alpha$ and so $\alpha \mid\!\sim \bot$ (using REF, AND, RWE) which contradicts $\alpha \in \mathcal{U}$. Hence we must have $S_\theta = \{\alpha, \gamma_1, \ldots, \gamma_l\}$ where $l > 0$, and so $\alpha \vee \gamma_1 \vee \ldots \vee \gamma_l \mid\!\sim \neg\alpha$. By LLE this gives $(\gamma_1 \vee \alpha) \vee (\gamma_2 \vee \alpha) \vee \ldots \vee (\gamma_l \vee \alpha) \mid\!\sim \neg\alpha$ and so, from the derived rule DR (see Proposition 5.2), we have $\gamma_r \vee \alpha \mid\!\sim \neg\alpha$ for some $r \in \{1, \ldots, l\}$, i.e., $\gamma_r \prec_\sim \alpha$. Now by the minimality of i we know $\gamma_r \notin \bigcup_{j<i} \mathcal{U}_j$ and so it remains to show $\gamma_r \in \mathcal{U}$ which will show that α cannot be minimal in $\mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ as required. But suppose $\gamma_r \notin \mathcal{U}$, i.e., $\gamma_r \mid\!\sim \bot$. Then, by RWE, we also have $\gamma_r \mid\!\sim \neg\gamma_r$. Also $\alpha \mid\!\sim \neg\gamma_r$ by SCL so $\gamma_r \vee \alpha \mid\!\sim \neg\gamma_r$ by OR. This together with $\gamma_r \vee \alpha \mid\!\sim \neg\alpha$ gives $\gamma_r \vee \alpha \mid\!\sim \neg(\gamma_r \vee \alpha)$ by AND, RWE which in turn (using REF, AND, RWE) gives $\gamma_r \vee \alpha \mid\!\sim \bot$, and so $\alpha \mid\!\sim \bot$ by property (3) from Lemma 5.25 (and LLE). Hence $\alpha \notin \mathcal{U}$ – contradiction. Hence we must have $\gamma_r \in \mathcal{U}$ as required.

Let us now turn to the "if" direction. Firstly if $\mathcal{U}_i \cap S_\theta = \emptyset$ for all $i = 1, \ldots, k$

then, since $\mathcal{U} = \bigcup_{i=1}^{k} \mathcal{U}_i$, this is saying that $\mathcal{U} \cap S_\theta = \emptyset$. Hence, in this case, $\alpha \mathrel{|}\!\sim \bot$ for all $\alpha \in S_\theta$ and so, by OR repeatedly, $\bigvee S_\theta \mathrel{|}\!\sim \bot$ which gives $\theta \mathrel{|}\!\sim \bot$ by LLE. Hence $\theta \mathrel{|}\!\sim \phi$ as required by RWE. Now let us suppose i is such that $\mathcal{U}_i \cap S_\theta \neq \emptyset$ and furthermore that i is minimal with this property. We will show that if $\theta \not\mathrel{|}\!\sim \phi$ then $\mathcal{U}_i \cap S_{\theta\wedge\neg\phi} \neq \emptyset$. We have that $\theta \not\mathrel{|}\!\sim \phi$ implies (by LLE) $\bigvee S_\theta \not\mathrel{|}\!\sim \phi$, i.e., $\delta_1 \vee \ldots \vee \delta_l \vee \gamma_1 \vee \ldots \vee \gamma_r \not\mathrel{|}\!\sim \phi$ where $S_\theta = \{\delta_1, \ldots, \delta_l, \gamma_1, \ldots, \gamma_r\}$ and $\delta_1, \ldots, \delta_l$ $(l > 0)$ are those elements of S_θ which are minimal in $\mathcal{U} \cap S_\theta$ under $\prec_\sim$. For each $j = 1, \ldots, r$ we have $\delta_s \vee \gamma_j \mathrel{|}\!\sim \neg\gamma_j$ for some $s \in \{1, \ldots, l\}$. This is clear if $\gamma_j \in \mathcal{U}$ (since otherwise γ_j would be one of the minimal elements) while if $\gamma_j \notin \mathcal{U}$, i.e., $\gamma_j \mathrel{|}\!\sim \bot$, then $\delta_1 \vee \gamma_j \mathrel{|}\!\sim \neg\gamma_j$ by Lemma 5.25(6). Hence we may repeatedly apply property (5) from Lemma 5.25 to obtain $\delta_1 \vee \ldots \vee \delta_l \not\mathrel{|}\!\sim \phi$ and so $\delta_y \not\mathrel{|}\!\sim \phi$ for some $y \in \{1, \ldots, l\}$ (otherwise we would have $\delta_1 \vee \ldots \vee \delta_l \mathrel{|}\!\sim \phi$ by OR repeatedly). Now if $\delta_y \vdash \phi$ then $\delta_y \mathrel{|}\!\sim \phi$ by SCL giving a contradiction. Hence $\delta_y \nvdash \phi$, equivalently $\delta_y \vdash \neg\phi$, so $\delta_y \in S_{\theta\wedge\neg\phi}$. Our result will then be proved if we can show $\delta_y \in \mathcal{U}_i$. We know $\mathcal{U}_i \cap S_\theta \neq \emptyset$ so let $\delta \in \mathcal{U}_i \cap S_\theta$. If $\delta = \gamma_j$ for some $j \in \{1, \ldots, r\}$ then δ is not minimal in S_θ under $\prec_\sim$ so, since i is minimal such that $\mathcal{U}_i \cap S_\theta \neq \emptyset$, δ is not minimal in $\mathcal{U} - \bigcup_{j<i} \mathcal{U}_i$ under $\prec_\sim$ contradicting $\delta \in \mathcal{U}_i$. Hence we must have $\delta = \delta_m$ for some $m \in \{1, \ldots, l\}$. If $m = y$ then $\delta_y = \delta \in \mathcal{U}_i$ as required so suppose $y \neq m$. Then if $\delta_y \notin \mathcal{U}_i$ we would have $\tau \vee \delta_y \mathrel{|}\!\sim \neg\delta_y$ for some $\tau \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$. By the minimality of δ_y in S_θ we also have $\delta \vee \delta_y \not\mathrel{|}\!\sim \neg\delta_y$ and therefore (using property (4) from Lemma 5.25) $\tau \vee \delta \mathrel{|}\!\sim \neg\delta$ which contradicts $\delta \in \mathcal{U}_i$. Hence $\delta_y \in \mathcal{U}_i$ and so $\mathcal{U}_i \cap S_{\theta\wedge\neg\phi} \neq \emptyset$ as required. □

Thus we have proved a representation theorem for rational consequence relations. We have shown that such relations correspond to admissible sequences of sets of permatoms.

5.6 Weakly Admissible Sequences

Our aim now is to find a characterisation of natural consequence relations analogous to that given in the previous section for rational consequence relations. Since natural consequence relations are more general than rational consequence relations, we need first of all to weaken the condition **(AD)**. We do this in the following way.

Definition 5.29 *Let* $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ $(k \geq 0)$ *be a finite sequence of sets of permatoms. We shall say* $\vec{\mathcal{U}}$ *is* weakly admissible *(over L) iff it satisfies the following condition:*

(WA) *For each* τ *a (possibly empty) conjunction of literals from distinct propositional variables from* L*, and for each* $p \in L$ *such that* $\pm p$ *does not appear in* τ*, either* $\mathcal{U}_i \cap T_\tau = \emptyset$ *for all* $i = 1, \ldots, k$*, or* $\mathcal{U}_i \cap (T_{\tau\wedge p} \cup T_{\tau\wedge\neg p}) \neq \emptyset$ *for the minimal* i *such that* $\mathcal{U}_i \cap T_\tau \neq \emptyset$.

Note we are adopting the convention that $\vdash \bigwedge \emptyset$ so if $\tau = \emptyset$ then $T_\tau = At_*^L$. In this case the condition **(WA)** reduces to: For each $p \in L$ if i is minimal such that $\mathcal{U}_i \neq \emptyset$ then $\mathcal{U}_i \cap (T_p \cup T_{\neg p}) \neq \emptyset$. Note that if a sequence $\vec{\mathcal{U}}$ is admissible then it is indeed weakly admissible. Our objective now is to characterise natural consequence relations in terms of weakly admissible sequences $\vec{\mathcal{U}}$. The results of this section will show that, for such a sequence $\vec{\mathcal{U}}$, $\mid\!\sim_{\vec{\mathcal{U}}}$ forms a natural consequence relation. However rather than show this directly by proving $\mid\!\sim_{\vec{\mathcal{U}}}$ satisfies all the rules for natural consequence we will instead tie together weakly admissible sequences with pre-ents and show how each weakly admissible sequence $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k$ such that $\mathcal{U}_i \neq \emptyset$ for some $i = 1, \ldots, k$ gives rise to a (λ-)pre-ent G (over a larger language L' than L) such that, for all $\theta, \phi \in SL$, $\theta \mid\!\sim_G \phi$ iff $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$. This will suffice to show that $\mid\!\sim_{\vec{\mathcal{U}}}$ is a natural consequence relation since, by results in Section 5.3, $\mid\!\sim_G$ forms a natural consequence relation on the language L' and hence clearly also on L. By proceeding in this way we ensure that any characterisation of natural consequence in terms of weakly admissible sequences will automatically give us a kind of characterisation of natural consequence in terms of pre-ents. Note that we do need the assumption here on $\vec{\mathcal{U}}$ that $\mathcal{U}_i \neq \emptyset$ for some i since if $\mathcal{U}_i = \emptyset$ for all i then we have $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$ for *all* $\theta, \phi \in SL$. In particular we have $\eta \mid\!\sim_{\vec{\mathcal{U}}} \neg\eta$ for any $\eta \in SL$ such that $\vdash \eta$. However for G a pre-ent we can never have $\eta \mid\!\sim_G \neg\eta$ since we always have $Bel^G(\eta) = 1 \neq 0$ and

$$\frac{Bel^G(\eta \wedge \neg\neg\eta)}{Bel^G(\eta)} = \frac{Bel^G(\eta)}{Bel^G(\eta)} = 1 \neq O(\lambda).$$

Hence if $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k$ is such that $\mathcal{U}_i = \emptyset$ for all i then there can be no pre-ent G such that $\mid\!\sim_G = \mid\!\sim_{\vec{\mathcal{U}}}$. However, for such a $\vec{\mathcal{U}}$, the relation $\mid\!\sim_{\vec{\mathcal{U}}}$ trivially satisfies all the rules from Definition 5.9 and so still forms a natural consequence relation.

So let $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k$ be a weakly admissible sequence of subsets of At^L_* such that $\mathcal{U}_i \neq \emptyset$ for some $i = 1, \ldots, k$. It should be clear that, for each permatom δ, only the least i such that $\delta \in \mathcal{U}_i$ is relevant to the relation $\mid\!\sim_{\vec{\mathcal{U}}}$ – any appearances of δ in any later $\mathcal{U}_i$ have no effect on $\mid\!\sim_{\vec{\mathcal{U}}}$. Also, since it is only the overall ordering of these least i's which is important, we may insert or delete appearances of the empty set $\emptyset$ in the sequence $\mathcal{U}_1, \ldots, \mathcal{U}_k$ without changing $\mid\!\sim_{\vec{\mathcal{U}}}$. Hence we may as well make the assumption that it is $\mathcal{U}_1$ which is non-empty and that $\mathcal{U}_i \subseteq \mathcal{U}_{i+1}$ for $i = 1, \ldots, k-1$. The pre-ent we define from $\vec{\mathcal{U}}$ will be defined over the language $L' \supseteq L$ which we define as follows:

$$L' = L \cup \{x_{p_{i_1}^{\epsilon_1} \cdots p_{i_j}^{\epsilon_j}} \mid j > 0,\ i_k \neq i_l \text{ for } k \neq l,\ \epsilon_k \in \{0,1\} \text{ for } 1 \leq k \leq j\}.$$

So L' consists of the propositional variables in L together with a new set of variables consisting of one for each non-empty sequence of literals taken from distinct propositional variables in L. For each τ a (possibly empty) sequence of literals from distinct propositional variables **in** L, we define the scenario $s^\tau \in WL'$ by

$$s^\tau = \{p_i^{\nu_i} \mid p_i^{\nu_i} \text{ appears in } \tau\} \cup \{x_\mu \mid \mu \subseteq \tau,\ \mu \neq \emptyset\}.$$

So s^τ contains all the literals which appear in τ together with those variables $x_\mu \in L' - L$ for which μ is a non-empty initial segment of τ. We remark that the reason for the x_μ's is to ensure $\tau_1 \neq \tau_2$ implies $s^{\tau_1} \neq s^{\tau_2}$. Note that $s^\emptyset = \emptyset$.

We use $\vec{\mathcal{U}}$ to define a pre-ent $G = G(\vec{\mathcal{U}})$ over L' as follows:

Let $p \in L'$, $s, t \in WL'$. Firstly, if $p \in L' - L$ then just define $G_p(s,t)$ in any correct way (since this case will not be needed in the proof). So suppose $p \in L$. Of course if $s \not\subseteq t$ then we set $G_p(s,t) = 0$ while if $p \in s$ ($\neg p \in s$) then we set $G_p(s,t) = 1$ ($G_p(s,t) = -1$), so suppose also that $s \subseteq t$ and $\pm p \notin s$. The only case here we are interested in is if $s = s^\tau$ for some sequence of literals τ from L. (If it is *not* the case that $s = s^\tau$ for any such sequence τ then again we just define $G_p(s,t)$ in any correct way.) Note that (for $p \in L$) $\pm p \notin s^\tau$ iff $\pm p$ does not appear anywhere in the sequence τ. So let s be of this form. First of all let us suppose that τ is such that $\mathcal{U}_i \cap T_\tau \neq \emptyset$ for some $i \in \{1, \ldots, k\}$, equivalently (since we assume the $\mathcal{U}_i$'s are increasing) $\mathcal{U}_k \cap T_\tau \neq \emptyset$ (note that whenever a sequence of literals τ appears as a subscript as in T_τ we are using it as shorthand for the conjunction, in sequence order, of those literals). In this case we define

$$G_p(s,t) = \begin{cases} \dfrac{\overline{G}_p(s^\tau, s^{\tau p})}{\overline{G}_p(s^\tau, s^{\tau p}) + \overline{G}_{\neg p}(s^\tau, s^{\tau \neg p})} & \text{if } t = s^{\tau p} \\[2ex] \dfrac{-\overline{G}_{\neg p}(s^\tau, s^{\tau \neg p})}{\overline{G}_p(s^\tau, s^{\tau p}) + \overline{G}_{\neg p}(s^\tau, s^{\tau \neg p})} & \text{if } t = s^{\tau \neg p} \\[2ex] 0 & \text{otherwise} \end{cases}$$

where, for $\nu \in \{0,1\}$,

$$\overline{G}_{p^\nu}(s^\tau, s^{\tau p^\nu}) = \begin{cases} 0 & \text{if } \mathcal{U}_k \cap T_{\tau \wedge p^\nu} = \emptyset \\ \lambda^{j-i} & \text{otherwise} \end{cases}$$

where i is minimal such that $\mathcal{U}_i \cap T_\tau \neq \emptyset$ and $j \geq i$ is minimal such that $\mathcal{U}_j \cap T_{\tau \wedge p^\nu} \neq \emptyset$. (Note that, since $\vec{\mathcal{U}}$ is weakly admissible we know that if $\mathcal{U}_i \cap T_\tau \neq \emptyset$ then either $\mathcal{U}_i \cap T_{\tau \wedge p} \neq \emptyset$ or $\mathcal{U}_i \cap T_{\tau \wedge \neg p} \neq \emptyset$. Hence $\overline{G}_{p^\nu}(s^\tau, s^{\tau p^\nu}) = 1$ for some $\nu \in \{0,1\}$ which ensures that the denominator in the above definition of $G_p(s,t)$ is non-zero.) If $s = s^\tau$ for some τ but $\mathcal{U}_k \cap T_\tau = \emptyset$ then we define

$$G_p(s,t) = \begin{cases} \frac{1}{2} & \text{if } t = s^{\tau p} \\ -\frac{1}{2} & \text{if } t = s^{\tau \neg p} \\ 0 & \text{otherwise.} \end{cases}$$

We will now show that, for all $\theta, \phi \in SL$, $\theta \mathrel{|\!\sim}_{\vec{\mathcal{U}}} \phi$ iff $\theta \mathrel{|\!\sim}_G \phi$. The main work required to prove this lies in showing that, for all $\theta \in SL$ and each $i = 1, \ldots, k$, $Bel^G(\theta) = O(\lambda^i)$ iff $\mathcal{U}_i \cap T_\theta = \emptyset$. Our next couple of results help us work towards that aim.

Lemma 5.30 *Let $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At^L_*$ be a weakly admissible sequence of sets of permatoms such that $\mathcal{U}_1 \neq \emptyset$ and $\mathcal{U}_i \subseteq \mathcal{U}_{i+1}$ for $i = 1, \ldots, k$ and let $G = G(\vec{\mathcal{U}})$ be the pre-ent over L' defined from $\vec{\mathcal{U}}$ as above. Then, for all $\delta \in At^L_*$, if $\delta \notin \mathcal{U}_k$ then $Bel^G(\delta) = 0$, otherwise*

$$Bel^G(\delta) = \frac{\lambda^{i-1}}{a}$$

where $i \in \{1, \ldots, k\}$ is minimal such that $\delta \in \mathcal{U}_i$ and $a \in \mathbb{R}((\lambda))$ is such that $a = O(1)$ and $a \neq O(\lambda)$.

Proof. Let $\delta \in At^L_*$ and suppose that $\delta = q_1 \wedge \ldots \wedge q_n$. Then we have, from the definitions of Section 2.2,

$$Bel^G(\delta) = \sum_{\substack{s_1 \subseteq s_2 \subseteq \cdots \subseteq s_n \\ q_i \in s_i \text{ for } i=1,\ldots,n}} G_{q_1}(\emptyset, s_1) \cdot G_{q_2}(s_1, s_2) \cdots G_{q_n}(s_{n-1}, s_n).$$

We now claim that $G_{q_1}(\emptyset, s_1) \cdot G_{q_2}(s_1, s_2) \cdots G_{q_n}(s_{n-1}, s_n) \neq 0$ only if $s_i = s^{q_1 q_2 \cdots q_i}$ for $i = 1, \ldots, n$. We show this using induction on i. For $i = 1$ we have $G_{q_1}(\emptyset, s_1) \cdot G_{q_2}(s_1, s_2) \cdots G_{q_n}(s_{n-1}, s_n) \neq 0$ implies $G_{q_1}(\emptyset, s_1) \neq 0$. Now, since $\emptyset = s^\emptyset$, our definition of G in this case gives us that $G_{q_1}(\emptyset, s_1) \neq 0$ only if either $s_1 = s^{q_1}$ or $s_1 = s^{\overline{q}_1}$ (where, recall, given a literal $q = p^\epsilon$, we define $\overline{q} = p^{1-\epsilon}$). Hence, since we must have $q_1 \in s_1$, we conclude that $s_1 = s^{q_1}$ as required. Now suppose $1 < i \leq n$ and that, for induction hypothesis, $s_{i-1} = s^{q_1 \cdots q_{i-1}}$. We must show $s_i = s^{q_1 \cdots q_i}$. But $G_{q_1}(\emptyset, s_1) \cdot G_{q_2}(s_1, s_2) \cdots G_{q_n}(s_{n-1}, s_n) \neq 0$ implies $G_{q_i}(s_{i-1}, s_i) = G_{q_i}(s^{q_1 \cdots q_{i-1}}, s_i) \neq 0$. Again, our definition of G now forces either $s_i = s^{q_1 \cdots q_{i-1} q_i}$ or $s_i = s^{q_1 \cdots q_{i-1} \overline{q}_i}$ and again the requirement that $q_i \in s_i$ forces us to conclude $s_i = s^{q_1 \cdots q_{i-1} q_i}$ as required. Thus our claim is proved and we have

$$Bel^G(\delta) = \prod_{j=1}^{n} G_{q_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} q_j}).$$

Let us now suppose $\delta \notin \mathcal{U}_k$. We must show that, in this case, $Bel^G(\delta) = 0$. But $\delta \notin \mathcal{U}_k$ is equivalent to saying $\mathcal{U}_k \cap T_{q_1 \wedge \ldots \wedge q_n} = \emptyset$ (since $\{\delta\} = T_{q_1 \wedge \ldots \wedge q_n}$) and so we may talk of the least $l \in \{0, 1, \ldots, n\}$ such that $\mathcal{U}_k \cap T_{q_1 \wedge \ldots \wedge q_l} = \emptyset$. Note that $l > 0$ since $\mathcal{U}_k \cap T_{\bigwedge \emptyset} = \mathcal{U}_k \supseteq \mathcal{U}_1 \neq \emptyset$. For this choice of l we have $\mathcal{U}_k \cap T_{q_1 \wedge \ldots \wedge q_{l-1}} \neq \emptyset$ and so, by definition of G,

$$G_{q_l}(s^{q_1 \cdots q_{l-1}}, s^{q_1 \cdots q_{l-1} q_l}) = \frac{\overline{G}_{q_l}(s^{q_1 \cdots q_{l-1}}, s^{q_1 \cdots q_{l-1} q_l})}{\overline{G}_{q_l}(s^{q_1 \cdots q_{l-1}}, s^{q_1 \cdots q_{l-1} q_l}) + \overline{G}_{\overline{q}_l}(s^{q_1 \cdots q_{l-1}}, s^{q_1 \cdots q_{l-1} \overline{q}_l})}.$$

But $\mathcal{U}_k \cap T_{q_1 \wedge \ldots \wedge q_l} = \emptyset$ implies, from the definition of the function $\overline{G}$, that

$$\overline{G}_{q_l}(s^{q_1 \cdots q_{l-1}}, s^{q_1 \cdots q_{l-1} q_l}) = 0.$$

Hence

$$G_{q_l}(s^{q_1 \cdots q_{l-1}}, s^{q_1 \cdots q_{l-1} q_l}) = 0$$

and so

$$Bel^G(\delta) = \prod_{j=1}^{n} G_{q_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} q_j}) = 0$$

as required.

Now suppose that we do have $\delta \in \mathcal{U}_k$. Then let i be minimal such that $\delta \in \mathcal{U}_i$. We must show that, in this case, we have

$$Bel^G(\delta) = \frac{\lambda^{i-1}}{a}$$

for some $a \in \mathbb{R}((\lambda))$ such that $a = O(1)$ and $a \neq O(\lambda)$. Since $\delta \in \mathcal{U}_k$ we have that $\mathcal{U}_k \cap T_{q_1 \wedge \ldots \wedge q_n} \neq \emptyset$ and so, since $T_{q_1 \wedge \ldots \wedge q_n} \subseteq T_{q_1 \wedge \ldots \wedge q_j}$ (by Corollary 5.19), $\mathcal{U}_k \cap T_{q_1 \wedge \ldots \wedge q_j} \neq \emptyset$ for all $j = 1, \ldots, n$. For each $j = 0, 1, \ldots, n$ let i_j be minimal such that $\mathcal{U}_{i_j} \cap T_{q_1 \wedge \ldots \wedge q_j} \neq \emptyset$ (so $i_0 = 1$ by assumption and $i_n = i$). For each $j = 1, \ldots, n$ our definition of G gives us

$$\begin{aligned} G_{q_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} q_j}) &= \frac{\overline{G}_{q_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} q_j})}{\overline{G}_{q_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} q_j}) + \overline{G}_{\overline{q}_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} \overline{q}_j})} \\ &= \frac{\lambda^{i_j - i_{j-1}}}{a_j} \end{aligned}$$

where

$$a_j = \overline{G}_{q_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} q_j}) + \overline{G}_{\overline{q}_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} \overline{q}_j}).$$

By the weak admissibility of $\vec{\mathcal{U}}$ we know that at least one of the two terms in a_j is equal to 1, and the other is of order $O(\lambda^y)$ for some $y \geq 0$. Hence we know that $a_j = O(1)$ and $a_j \neq O(\lambda)$. Hence

$$\begin{aligned} Bel^G(\delta) &= \prod_{j=1}^{n} G_{q_j}(s^{q_1 \cdots q_{j-1}}, s^{q_1 \cdots q_{j-1} q_j}) \\ &= \prod_{j=1}^{n} \frac{\lambda^{i_j - i_{j-1}}}{a_j} \\ &= \frac{\lambda^{i_n - i_0}}{a} \end{aligned}$$

where $a = \prod a_j$. Now clearly $a = O(1)$ while if it were the case that $a = O(\lambda)$ then (see Proposition 4.7(ii)) we would have to have $a_j = O(\lambda)$ for some j – contradiction. Hence our result is proved since $i_n = i$ and $i_0 = 1$. □

Corollary 5.31 *Let $\vec{\mathcal{U}}$ and $G = G(\vec{\mathcal{U}})$ be as in Lemma 5.30 and let $\delta \in At^L_*$. Then, for all $i = 1, \ldots, k$,*

$$\delta \in \mathcal{U}_i \text{ iff } Bel^G(\delta) \neq O(\lambda^i).$$

Proof. For the "only if" direction suppose $\delta \in \mathcal{U}_i$. By Lemma 5.30,

$$Bel^G(\delta) = \frac{\lambda^{j-1}}{a}$$

where j is minimal such that $\delta \in \mathcal{U}_j$ and $a \in \mathbb{R}((\lambda))$ is such that $a = O(1)$ and $a \neq O(\lambda)$. Now if it were the case that $Bel^G(\delta) = O(\lambda^j)$ then we would have

$$\frac{\lambda^{j-1}}{a} = b$$

where $b = O(\lambda^j)$, and so

$$\lambda^{j-1} = a \cdot b = O(1) \times O(\lambda^j) = O(\lambda^j)$$

which is a contradiction. Hence $Bel^G(\delta) \neq O(\lambda^j)$ which means, since $j \leq i$, $Bel^G(\delta) \neq O(\lambda^i)$ as required.

Conversely suppose $\delta \notin \mathcal{U}_i$. If in fact $\delta \notin \mathcal{U}_k$ then Lemma 5.30 gives us $Bel^G(\delta) = 0$ and so certainly $Bel^G(\delta) = O(\lambda^i)$ as required. So suppose $\delta \in \mathcal{U}_k$ and again suppose j is minimal such that $\delta \in \mathcal{U}_j$ (so, since the $\mathcal{U}_l$ are increasing, $i \leq j-1$). Then, by Lemma 5.30,

$$Bel^G(\delta) = \frac{\lambda^{j-1}}{a} = \frac{O(\lambda^{j-1})}{O(1)} = O(\lambda^{j-1})$$

using Proposition 4.7(iii), since $a \neq O(\lambda)$. Hence, since $i \leq j-1$, we also have $Bel^G(\delta) = O(\lambda^i)$ as required. □

We now give the key result which allows us to prove the equivalence (on L) of $\mid\!\sim_{\vec{\mathcal{U}}}$ and $\mid\!\sim_{G(\vec{\mathcal{U}})}$.

Lemma 5.32 *Let $\vec{\mathcal{U}}$ and $G = G(\vec{\mathcal{U}})$ be as in Lemma 5.30. Then, for all $\theta \in SL$ and $i \in \{1, \ldots, k\}$,*

$$Bel^G(\theta) = O(\lambda^i) \text{ iff } \mathcal{U}_i \cap T_\theta = \emptyset.$$

Proof. We firstly consider the case when θ is non-contingent.

If $\vdash \neg\theta$ then $Bel^G(\theta) = 0$ and so $Bel^G(\theta) = O(\lambda^i)$ for all i. Hence we must show that $\mathcal{U}_i \cap T_\theta = \emptyset$ for all i. But this is true since $\vdash \neg\theta$ implies $T_\theta = \emptyset$ by definition of T_θ.

If, on the other hand, $\vdash \theta$ then $Bel^G(\theta) = 1$ and so $Bel^G(\theta) \neq O(\lambda^i)$ for all $i \in \{1, \ldots, k\}$. Hence in this case we must show $\mathcal{U}_i \cap T_\theta \neq \emptyset$ for all i, equivalently (since the $\mathcal{U}_i$ are increasing) $\mathcal{U}_1 \cap T_\theta \neq \emptyset$. But $\vdash \theta$ implies $T_\theta = At_*^L$ by definition of T_θ and so $\mathcal{U}_1 \cap T_\theta \neq \emptyset$ iff $\mathcal{U}_1 \neq \emptyset$, which we have by assumption as required. Now suppose that θ is contingent. Then, from Corollary 3.16, we have

$$Bel^G(\theta) = \sum_{\tau \in rT(\theta)^+} Bel^G(\tau).$$

Suppose that $Bel^G(\theta) \neq O(\lambda^i)$. Then $Bel^G(\tau) \neq O(\lambda^i)$ for some $\tau \in rT(\theta)^+$. Let $q_1, \ldots, q_l \in L$ $(l \geq 0)$ be all those propositional variables which do not appear in τ. Then we have

$$Bel^G(\tau) = \sum_{\langle \nu_1, \ldots, \nu_l \rangle \in \{0,1\}^l} Bel^G(\tau \wedge q_1^{\nu_1} \wedge \ldots \wedge q_l^{\nu_l}).$$

Hence there exists $\nu_1, \ldots, \nu_l \in \{0, 1\}$ such that

$$Bel^G(\tau \wedge q_1^{\nu_1} \wedge \ldots \wedge q_l^{\nu_l}) \neq O(\lambda^i).$$

Hence, putting $\delta = \tau \wedge q_1^{\nu_1} \wedge \ldots \wedge q_l^{\nu_l}$ gives us $\delta \in \mathcal{U}_i$ by Corollary 5.31 while obviously $\delta \in T_\theta$ so $\mathcal{U}_i \cap T_\theta \neq \emptyset$. Conversely suppose $\mathcal{U}_i \cap T_\theta \neq \emptyset$ and let $\delta \in \mathcal{U}_i \cap T_\theta$. Then δ is of the form $\tau' \wedge \rho$ for some $\tau' \in rT(\theta)^+$ and we have

$$\begin{aligned} Bel^G(\theta) &= \sum_{\tau \in rT(\theta)^+} Bel^G(\tau) \\ &\geq Bel^G(\tau') \\ &\geq Bel^G(\tau' \wedge \rho) \\ &\neq O(\lambda^i) \qquad \text{form Corollary 5.31.} \end{aligned}$$

Hence if $Bel^G(\theta) = O(\lambda^i)$ then, since $Bel^G(\theta) \geq Bel^G(\tau' \wedge \rho) \geq 0$, we must also have $Bel^G(\tau' \wedge \rho) = O(\lambda^i)$ – contradiction. Hence we get $Bel^G(\theta) \neq O(\lambda^i)$ as required. □

We are now in a position to prove the main result of this section.

Theorem 5.33 *For each weakly admissible sequence* $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ *of sets of permatoms of* L *such that* $\mathcal{U}_i \neq \emptyset$ *for some* $i = 1, \ldots, k$ *there exists a pre-ent* G *(possibly over a larger language than* L*) such that, for all* $\theta, \phi \in SL$, $\theta \mid\!\sim_G \phi$ *iff* $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$.

Proof. Given $\vec{\mathcal{U}}$ satisfying the conditions of the theorem and given that, as we have already said, we may assume the $\mathcal{U}_i$ are increasing and $\mathcal{U}_1 \neq \emptyset$, let $G = G(\vec{\mathcal{U}})$ be the pre-ent over the language L' defined as in the above construction process. By Lemma 5.32 we have, for all $\theta \in SL$ and $i \in \{1, \ldots, k\}$, $Bel^G(\theta) = O(\lambda^i)$ iff $\mathcal{U}_i \cap T_\theta = \emptyset$. We claim that this G fulfills the requirements of the theorem, i.e., that, for all $\theta, \phi \in SL$, $\theta \mid\!\sim_G \phi$ iff $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$.

To show the "if" direction suppose $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$. If $\mathcal{U}_i \cap T_\theta = \emptyset$ for all i then we must have $Bel^G(\theta) = O(\lambda^i)$ for all i and so, by Proposition 4.4, we must have $Bel^G(\theta) = 0$ and so $\theta \mid\!\sim_G \phi$ as required. So suppose i is minimal such that $\mathcal{U}_i \cap T_\theta \neq \emptyset$. Then i is minimal such that $Bel^G(\theta) \neq O(\lambda^i)$ (and so $Bel^G(\theta) \neq 0$ and $Bel^G(\theta) = O(\lambda^{i-1})$) while $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$ gives us $\mathcal{U}_i \cap T_{\theta\wedge\neg\phi} = \emptyset$, i.e., $Bel^G(\theta \wedge \neg\phi) = O(\lambda^i)$. Hence, using Proposition 4.7 (iii),

$$\frac{Bel^G(\theta \wedge \neg\phi)}{Bel^G(\theta)} = O(\lambda^{i-(i-1)}) = O(\lambda)$$

which gives $\theta \mid\!\sim_G \phi$ as required.

To show the "only if" direction suppose $\theta \mid\!\sim_G \phi$. If $Bel^G(\theta) = 0$ then clearly $Bel^G(\theta) = O(\lambda^i)$ for all i and so $\mathcal{U}_i \cap T_\theta = \emptyset$ for all i giving $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$ as required. So suppose $Bel^G(\theta) \neq 0$. Then we must have $\mathcal{U}_i \cap T_\theta \neq \emptyset$ for some i (since otherwise $Bel^G(\theta) = 0$ by Proposition 4.4). Let us assume that i is minimal such that this occurs (so i is also minimal such that $Bel^G(\theta) \neq O(\lambda^i)$). We must show that $\mathcal{U}_i \cap T_{\theta\wedge\neg\phi} = \emptyset$. Let j be minimal such that $\mathcal{U}_j \cap T_{\theta\wedge\neg\phi} \neq \emptyset$, so $Bel^G(\theta \wedge \neg\phi) = O(\lambda^{j-1})$. Then we know $i \leq j$ since $T_{\theta\wedge\neg\phi} \subseteq T_\theta$ (using Corollary 5.19), which gives $\mathcal{U}_j \cap T_\theta \neq \emptyset$ and so $j < i$ would contradict the minimality of i.

But if $j = i$ then, by Proposition 4.7 (iii),

$$\frac{Bel^G(\theta \wedge \neg\phi)}{Bel^G(\theta)} = O(\lambda^{(j-1)-(i-1)}) = O(1)$$

which contradicts $\theta \mid\!\sim_G \phi$. Hence $i < j$, i.e., $\mathcal{U}_i \cap T_{\theta \wedge \neg\phi} = \emptyset$ as required. □

With Theorem 5.33 in place it is now easy to show that $\mid\!\sim_{\vec{\mathcal{U}}}$ for $\vec{\mathcal{U}}$ a weakly admissible sequence forms a natural consequence relation.

Corollary 5.34 *Let $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ be a weakly admissible sequence. Then the relation $\mid\!\sim_{\vec{\mathcal{U}}} \subseteq SL \times SL$ forms a natural consequence relation on L.*

Proof. If $\mathcal{U}_i = \emptyset$ for all $i = 1, \ldots, k$ then we have $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$ for all $\theta, \phi \in SL$. Hence, in this case, $\mid\!\sim_{\vec{\mathcal{U}}}$ trivially satisfies all the rules for natural consequence. If it is not the case that $\mathcal{U}_i = \emptyset$ for all i then the result is still true since, by Theorem 5.33, we may now assert the existence of a pre-ent G such that (on SL) $\mid\!\sim_G = \mid\!\sim_{\vec{\mathcal{U}}}$ and we know $\mid\!\sim_G$ forms a natural consequence relation by Theorem 5.10. □

The results of this section have thus provided us with an example, other than (though closely linked to) $\mid\!\sim_G$ for G a pre-ent on L, of a family of natural consequence relations on L, namely the family $\mid\!\sim_{\vec{\mathcal{U}}}$ for $\vec{\mathcal{U}}$ a weakly admissible sequence of sets of permatoms of L. In the next chapter we will attempt to show that *every* natural consequence relation on L arises from such a weakly admissible sequence $\vec{\mathcal{U}}$. By Theorem 5.33 any such result would show that every (non-trivial) natural consequence relation on L is given by the restriction to $SL \times SL$ of $\mid\!\sim_G$ for a pre-ent over a larger language $L' \supseteq L$. Thus giving us a kind of characterisation of natural consequence of the type originally called for in terms of pre-ents. We end the present section by giving a reformulation of the condition **(WA)** which will prove useful in some of the proofs in the next chapter.

Corollary 5.35 *Let $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ be a sequence of sets of permatoms of L. Then $\vec{\mathcal{U}}$ is weakly admissible iff the following holds:*

- *For all* $\theta, \phi \in SL$, *either* $\mathcal{U}_i \cap T_\theta = \emptyset$ *for all* $i = 1, \ldots, k$, *or* $\mathcal{U}_i \cap (T_{\theta\wedge\phi} \cup T_{\theta\wedge\neg\phi}) \neq \emptyset$ *for* i *minimal such that* $\mathcal{U}_i \cap T_\theta \neq \emptyset$.

Proof. We must show that this condition is equivalent to the condition **(WA)**. That it implies **(WA)** is clear. To show the converse implication let $\vec{\mathcal{U}}$ be a weakly admissible (i.e., satisfies **(WA)**) sequence of sets of permatoms over L. If $\mathcal{U}_j = \emptyset$ for all j then the condition holds trivially, so suppose $\mathcal{U}_j \neq \emptyset$ for some j. Then, by Lemma 5.32 there exists a pre-ent G over a language $L' \supseteq L$ such that, for all $\psi \in SL$ and $i = 1, \ldots, k$, $\mathcal{U}_i \cap T_\psi = \emptyset$ iff $Bel^G(\psi) = O(\lambda^i)$. Let $\theta, \phi \in SL$. If $\mathcal{U}_j \cap T_\theta = \emptyset$ for all j then again the condition is true, so suppose otherwise and let i be minimal such that $\mathcal{U}_i \cap T_\theta \neq \emptyset$. Then i is also minimal such that $Bel^G(\theta) \neq O(\lambda^i)$. If both $\mathcal{U}_i \cap T_{\theta\wedge\phi} = \emptyset$ and $\mathcal{U}_i \cap T_{\theta\wedge\neg\phi} = \emptyset$ then $Bel^G(\theta\wedge\phi) = O(\lambda^i) = Bel^G(\theta\wedge\neg\phi)$ which would give

$$\begin{aligned} Bel^G(\theta) &= Bel^G(\theta\wedge\phi) + Bel^G(\theta\wedge\neg\phi) \\ &= O(\lambda^i) + O(\lambda^i) \\ &= O(\lambda^i) \end{aligned}$$

which is a contradiction. Hence we have either $\mathcal{U}_i \cap T_{\theta\wedge\phi} \neq \emptyset$ or $\mathcal{U}_i \cap T_{\theta\wedge\neg\phi} \neq \emptyset$ as required. □

Chapter 6

Characterising F.T. Natural Consequence

6.1 Introduction

Given a weakly admissible sequence $\vec{\mathcal{U}}$ of sets of permatoms of L, we have seen (Corollary 5.34) that if we define a consequence relation $\mid\sim_{\vec{\mathcal{U}}}$ on SL by setting, for all $\theta, \phi \in SL$,

$$\theta \mid\sim_{\vec{\mathcal{U}}} \phi \quad \text{iff} \quad \textit{either} \quad \mathcal{U}_i \cap T_\theta = \emptyset \text{ for all } i$$
$$\textit{or} \quad \mathcal{U}_i \cap T_{\theta\wedge\neg\phi} = \emptyset \text{ for the least } i \text{ such that } \mathcal{U}_i \cap T_\theta \neq \emptyset$$

then $\mid\sim_{\vec{\mathcal{U}}}$ forms a natural consequence relation on L. The results in this chapter are motivated by a desire to completely characterise the class of natural consequence relations on L in terms of weakly admissible sequences of sets of permatoms of L. In other words, given a natural consequence relation $\mid\sim$ on L, we would like to show that there exists a weakly admissible sequence $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mid\sim) = \mathcal{U}_1, \ldots, \mathcal{U}_k$ such that $\mid\sim = \mid\sim_{\vec{\mathcal{U}}}$. We shall show that we *almost* arrive at such a result. We begin with the idea that we would like to prove this characterisation analogously to the way we proved the corresponding result for *rational* consequence relations in Section 5.5, where we showed that every rational consequence relation arises as $\mid\!\sim_{\vec{\mathcal{U}}}$ for some *admissible* sequence $\vec{\mathcal{U}}$. Section 6.2 is devoted to making the link between our current case and the rational case more transparent by studying what we call *permutation trees*, which are subsets of permatoms in which all the elements are syntactically "comparable" to each other and on which a natural consequence relation behaves like a rational consequence relation. Any proof of a characterisation for natural consequence must utilise all the rules of Definition 5.9 and any additional property of $\mid\!\sim$ that we use must be shown to be derivable from this our basic set of properties. Unfortunately the proof we shall give employs a rule which, although sound for $\mid\!\sim_{\vec{\mathcal{U}}}$, does not obviously appear to follow from this set. We shall describe this rule, which we call **(FT)** (standing for Full Transitivity) in Section 6.3. Any binary relation on SL which satisfies all the rules for natural consequence together with **(FT)** we shall call a *fully transitive (f. t.) natural consequence relation*. In Section 6.3 we shall show that the relation $\mid\!\sim_G$ for G a pre-ent, and hence (applying Theorem 5.33) the relation $\mid\!\sim_{\vec{\mathcal{U}}}$ for $\vec{\mathcal{U}}$ a weakly admissible sequence, satisfies the extra rule **(FT)** and so is a f. t. natural consequence relation. We shall also show that every rational consequence relation satisfies **(FT)** and so the class of rational consequence relations forms a subclass of the class of f. t. natural consequence relations. Then in Section 6.4 we give the main result of this chapter – we show that each f. t. natural consequence relation is given by a weakly admissible sequence $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k$. Thus what we end up with at the end of this chapter is not a characterisation of natural consequence but a characterisation of f. t. natural consequence. If our given f. t. natural consequence relation $\mid\!\sim$ is non-trivial, i.e., it is not the case that $\theta \mid\!\sim \phi$ for all $\theta, \phi \in SL$, then it will be apparent that $\mathcal{U}_i \neq \emptyset$ for some i. Hence in that case we may then apply Theorem 5.33 to show that there exists a pre-ent G (over a language which extends L) such that, on its restriction to SL, $\mid\!\sim_G = \mid\!\sim$.

Thus the results in this chapter will extend to give us a kind of characterisation of (non-trivial) f. t. natural consequence relations in terms of pre-ents. Finally in Section 6.5 we use the main result of Chapter 4 (Theorem 4.1) to show how the family of consequence relations $\mid\!\sim_z$ for z an *ent* may be said to correspond to those (non-trivial) f. t. natural consequence relations which satisfy a further, natural, property. We end the thesis with some concluding remarks in Section 6.6.

6.2 Permutation Trees

Recall from Section 5.5 that, given a rational consequence relation $\mid\!\sim$, we defined the (admissible) sequence of permatoms $\vec{\mathcal{U}}$ such that $\mid\!\sim = \mid\!\sim_{\vec{\mathcal{U}}}$ by, firstly, forming the set $\mathcal{U} = \mathcal{U}(\mid\!\sim)$ of those permatoms which were consistent for $\mid\!\sim$ (i.e., those permatoms δ for which we had $\delta \not\mid\!\sim \bot$) then, secondly, defining a preference order on $\mathcal{U}$ that was irreflexive and transitive and then, finally, taking $\mathcal{U}_1$ to be those permatoms which were minimal in $\mathcal{U}$ under this ordering, $\mathcal{U}_2$ to be those permatoms which were minimal in $\mathcal{U} - \mathcal{U}_1$ under this ordering, etc. We would like to be able to carry this process over to the natural consequence situation to find, from a given natural consequence relation $\mid\!\sim$, a weakly admissible sequence $\vec{\mathcal{U}}$ such that $\mid\!\sim = \mid\!\sim_{\vec{\mathcal{U}}}$. Translating the first stage across presents no problem – we can still begin by setting $\mathcal{U} = \mathcal{U}(\mid\!\sim)$ to be the set of permatoms consistent for $\mid\!\sim$. However when we try to replicate the second stage, i.e., defining a preference order that is both irreflexive and transitive on $\mathcal{U}$, we do run into trouble. Recall that in the rational case we just defined our preference order $\prec_\sim$ from $\mid\!\sim$ by, for $\delta_1, \delta_2 \in At^L_*$,

$$\delta_1 \prec_\sim \delta_2 \ \ \textit{iff} \ \ \delta_1 \vee \delta_2 \mid\!\sim \neg\delta_2.$$

Whilst it is true that if we defined $\prec_\sim$ in this way for $\mid\!\sim$ a *natural* consequence relation then $\prec_\sim$ *would* be irreflexive on the set $\mathcal{U}(\mid\!\sim)$ (since, for $\delta \in \mathcal{U}$, we would have $\delta \prec_{\sim} \delta$ iff $\delta \vee \delta \mathrel{|\!\sim} \neg\delta$, which implies $\delta \mathrel{|\!\sim} \neg\delta$ by LGE which in turn implies $\delta \mathrel{|\!\sim} \bot$ by P-RWE, contradicting $\delta \in \mathcal{U}(\mathrel{|\!\sim})$), it is not necessarily the case that $\prec_{\sim}$ would be transitive as the following example shows. Suppose $L = \{p, q, r\}$ and suppose we had a pairwise disjoint weakly admissible sequence t_1, t_2, t_3 with $p \wedge q \wedge r \in t_1$, $\neg p \wedge q \wedge r \in t_2$ and $p \wedge r \wedge q \in t_3$. Then it is easy to see that for the natural consequence relation $\mathrel{|\!\sim}$ defined by any such sequence we would have $p \wedge q \wedge r \prec_{\sim} \neg p \wedge q \wedge r$ and $\neg p \wedge q \wedge r \prec_{\sim} p \wedge r \wedge q$. However, by P-LLE and P-RWE, we have $(p \wedge q \wedge r) \vee (p \wedge r \wedge q) \mathrel{|\!\sim} \neg(p \wedge r \wedge q)$ iff $(p \wedge q \wedge r) \vee (p \wedge q \wedge r) \mathrel{|\!\sim} \neg(p \wedge q \wedge r)$ and so $p \wedge q \wedge r \prec_{\sim} p \wedge r \wedge q$ iff $p \wedge q \wedge r \prec_{\sim} p \wedge q \wedge r$. Since we have already shown that $\prec_{\sim}$ is irreflexive on $\mathcal{U}(\mathrel{|\!\sim})$ (and since clearly $p \wedge q \wedge r \mathrel{|\!\not\sim} \bot$ so $p \wedge q \wedge r \notin \mathcal{U}(\mathrel{|\!\sim})$) it must be that $p \wedge q \wedge r \not\prec_{\sim} p \wedge r \wedge q$ and so $\prec_{\sim}$ fails to be transitive. (More generally, for $\delta_1, \delta_2, \delta_3 \in At^L_*$ we may have $\delta_1 \prec_{\sim} \delta_2$, $\delta_1 \equiv \delta_3$ and $\delta_2 \prec_{\sim} \delta_3$ for a natural consequence relation. This could not happen in the rational case.) However we need not totally discard this definition of $\prec_{\sim}$ as a suitable preference order for natural consequence relations, for it turns out that $\prec_{\sim}$ defined above for $\mathrel{|\!\sim}$ a natural consequence relation *will* be transitive on certain subsets of $\mathcal{U}(\mathrel{|\!\sim})$ – those subsets in which all the permatoms are, in a sense to be explained below, "comparable" with each other. In this section we will concentrate on such sets, which we will call *permutation trees.*

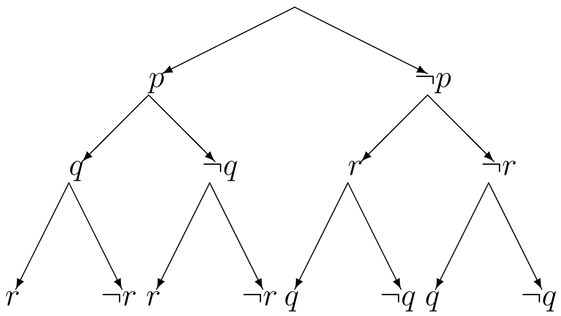
Definition 6.1 *A* permutation tree *(over L) is a non-empty set* $\mathcal{F} \subseteq At^L_*$ *of permatoms which satisfies the following conditions:*

- *(i). For any* $\delta \in \mathcal{F}$*, if* $\delta = q_1^{\epsilon_1} \wedge \ldots \wedge q_n^{\epsilon_n}$*, then for each* $1 \leq i \leq n$ *there exists some* $\delta' = r_1^{\upsilon_1} \wedge \ldots \wedge r_n^{\upsilon_n} \in \mathcal{F}$ *such that* $r_i = q_i$*,* $\upsilon_i = 1 - \epsilon_i$ *and* $r_j^{\upsilon_j} = q_j^{\epsilon_j}$ *for all* $j < i$*.*

- *(ii). For all distinct* $\delta_1, \delta_2 \in \mathcal{F}$*, if* $\delta_1 = q_1^{\epsilon_1} \wedge \ldots \wedge q_n^{\epsilon_n}$ *and* $\delta_2 = r_1^{\upsilon_1} \wedge \ldots \wedge r_n^{\upsilon_n}$ *then there exists some* i *such that* $1 \leq i \leq n$ *and we have* $q_i = r_i$*,* $\epsilon_i = 1 - \upsilon_i$ *and* $q_j^{\epsilon_j} = r_j^{\upsilon_j}$ *for all* $j < i$*.*

We denote the set of all permutation trees over L by P.

We use the term "permutation tree" to describe such sets since their elements may be thought of as corresponding to the paths through a binary tree in which each node has two edges out of it, labelled p and $\neg p$ for some $p \in L$, and in which every propositional variable appears precisely once in each path. So, for example, taking $L = \{p, q, r\}$, the set $\mathcal{F}_1 = \{p \wedge q \wedge r,\ p \wedge q \wedge \neg r,\ p \wedge \neg q \wedge r,\ p \wedge \neg q \wedge \neg r,\ \neg p \wedge r \wedge q,\ \neg p \wedge r \wedge \neg q,\ \neg p \wedge \neg r \wedge q,\ \neg p \wedge \neg r \wedge \neg q\}$ is a permutation tree for L which corresponds to the following tree diagram.



Another example of a permutation tree over L is the set At^L of atoms of L. Clearly for any $\mathcal{F} \in P$ we have $|\mathcal{F}| = 2^{|L|}$ and $\vdash \bigvee \mathcal{F}$. Also, by Proposition 3.15(i), we have $\mathcal{F}_1 = \mathcal{F}_2$ implies $\bigvee \mathcal{F}_1 \dot{\sim} \bigvee \mathcal{F}_2$. In the rest of this chapter, since for $\mathcal{F} \in P$ we will only ever be interested in $\bigvee \mathcal{F}$ up to $\dot{\sim}$– equivalence, we may leave this order unspecified.

We would now like to define the set of permutation trees for a given permatom δ to be the set of all $\mathcal{F} \in P$ to which δ belongs. The following more general definition will be more useful.

Definition 6.2 *For each $\theta \in SL$, the* set of permutation trees for θ, *F_θ, is defined by*

$$F_\theta = \{\mathcal{F} \in P \mid \theta \ddot{\sim} \bigvee \mathcal{F} \wedge \theta\}.$$

Clearly if $\theta \ddot{\sim} \phi$ then $F_\theta = F_\phi$. Also, if $\theta \in SL$ is non-contingent then $F_\theta = P$ since if $\vdash \theta$ then $\vdash \bigvee \mathcal{F} \wedge \theta$ for any $\mathcal{F} \in P$ and so $\theta \ddot{\sim} \bigvee \mathcal{F} \wedge \theta$ for any $\mathcal{F} \in P$, while if

$\vdash \neg\theta$ then $\vdash \neg(\bigvee \mathcal{F} \wedge \theta)$ for any $\mathcal{F} \in P$ and so again $\theta \ddot{\sim} \bigvee \mathcal{F} \wedge \theta$ for any $\mathcal{F} \in P$. What about if θ is contingent? A corollary of the forthcoming Lemma 6.5 will be that in this case a permutation tree for θ can always be found. That lemma will indicate the form that those trees can take. Before we give that result we shall give a useful equivalent way of defining F_θ. We need the following definition to express it:

Definition 6.3 *Given $\theta \in SL$ and $\mathcal{F} \in P$, we define the set $S_\theta^{\mathcal{F}} \subseteq At_*^L$ by*

$$S_\theta^{\mathcal{F}} = \{\delta \in \mathcal{F} \mid \delta \vdash \theta\}$$

Note that if $\mathcal{F} = At^L$ then we get $S_\theta^{\mathcal{F}} = S_\theta$ while for any $\mathcal{F} \in P$ we have $\theta \equiv \bigvee S_\theta^{\mathcal{F}}$. We may now state the following:

Proposition 6.4 *Let $\mathcal{F} \in P$ and $\theta \in SL$. Then $\mathcal{F} \in F_\theta$ iff $\bigvee S_\theta^{\mathcal{F}} \ddot{\sim} \theta$.*

Proof. We have $\mathcal{F} \in F_\theta$ iff $\bigvee \mathcal{F} \wedge \theta \ddot{\sim} \theta$. Hence it suffices to show that $\bigvee \mathcal{F} \wedge \theta \ddot{\sim} \bigvee S_\theta^{\mathcal{F}}$. But

$$\begin{aligned}
\bigvee \mathcal{F} \wedge \theta \quad & \dot{\sim} \quad (\bigvee S_\theta^{\mathcal{F}} \vee \bigvee S_{\neg\theta}^{\mathcal{F}}) \wedge \theta \\
& \dot{\sim} \quad (\bigvee S_\theta^{\mathcal{F}} \wedge \theta) \vee (\neg \bigvee S_\theta^{\mathcal{F}} \wedge \bigvee S_{\neg\theta}^{\mathcal{F}} \wedge \theta) \\
& \ddot{\sim} \quad \bigvee S_\theta^{\mathcal{F}} \wedge \theta \\
& \qquad \text{(since } \bigvee S_{\neg\theta}^{\mathcal{F}} \equiv \neg\theta \text{ and so } \vdash \neg(\neg \bigvee S_\theta^{\mathcal{F}} \wedge \bigvee S_{\neg\theta}^{\mathcal{F}} \wedge \theta)) \\
& \ddot{\sim} \quad \bigvee S_\theta^{\mathcal{F}} \qquad \text{by Theorem 3.7, since } \bigvee S_\theta^{\mathcal{F}} \vdash \theta.
\end{aligned}$$

Hence $\bigvee \mathcal{F} \wedge \theta \ddot{\sim} \bigvee S_\theta^{\mathcal{F}}$ as required. □

Given this proposition, it is now easy to see that, for $\delta \in At_*^L$ and $\mathcal{F} \in P$, $\mathcal{F} \in F_\delta$ iff $\delta \in \mathcal{F}$, and thus that the set of permutation trees for a given permatom is equal to the set of permutation trees to which that permatom belongs. For we have $\mathcal{F} \in F_\delta$ iff $\bigvee S_\delta^{\mathcal{F}} \ddot{\sim} \delta$ iff $\gamma \ddot{\sim} \delta$ where γ is that unique permatom in $\mathcal{F}$ such that $\gamma \vdash \delta$. If $\delta \in \mathcal{F}$ then $\delta = \gamma$ and the result is clear, while if $\delta \notin \mathcal{F}$ then $\delta \neq \gamma$ and so it is not the case that $\gamma \ddot{\sim} \delta$ as required (since, for e.g., no initial segment of γ is equal to δ). We now give the promised Lemma 6.5, the full power of which which will be needed in the proof of Theorem 6.21.

Lemma 6.5 *Let $\theta, \phi \in SL$ be jointly consistent and let $\delta \in T_{\theta\wedge\phi}$. Then $F_{\theta\wedge\phi} \cap F_\theta \cap F_\delta \neq \emptyset$.*

Proof. We define $\mathcal{F}' \in P$ by

$$\mathcal{F}' = cT((\delta \vee \neg\delta) \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0)^+$$

where $\mathcal{F}_0$ is some fixed arbitrary permutation tree over L. It should be clear (see Section 3.3 for a discussion of cT-trees) that $\mathcal{F}'$ so defined (indeed any set of the form $cT(\eta \wedge \bigvee \mathcal{F}_0)^+$ for any tautology η) is a legitimate permutation tree. We will show that $\mathcal{F}' \in F_{\theta\wedge\phi} \cap F_\theta \cap F_\delta$ which will suffice to prove the result. Note that $\delta \in T_{\theta\wedge\phi}$ implies that $\delta \vdash \theta \wedge \phi$. Let ψ now stand for any sentence such that $\delta \vdash \psi$. We have

$$\begin{aligned}
\bigvee \mathcal{F}' \wedge \psi \quad &\dot{\sim} \quad (\delta \vee \neg\delta) \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi \\
&\dot{\sim} \quad (\delta \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi) \ \vee \\
&\qquad \vee\ (\neg\delta \wedge \neg\delta \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi) \\
&\dot{\sim} \quad (\neg\delta \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi) \ \vee \\
&\qquad \vee\ (\delta \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi) \qquad (6.1)
\end{aligned}$$

using Proposition 3.2 (j) and (u). Now, since $\delta \vdash \psi$, we have

$$\delta \vdash ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi$$

and so, by Theorem 3.7,

$$\delta \ddot{\sim} \delta \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi.$$

Hence from (6.1), using Proposition 3.6, we get

$$\begin{aligned}
\bigvee \mathcal{F}' \wedge \psi \quad &\ddot{\sim} \quad (\neg\delta \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi) \vee \delta \\
&\dot{\sim} \quad (\neg\delta \wedge ((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi) \vee (\delta \wedge \delta) \\
&\dot{\sim} \quad \delta \vee (((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \psi), \qquad (6.2)
\end{aligned}$$

this last line following from Proposition 3.2(u), (j) and (h). Let us first show $\mathcal{F}' \in F_\theta$, i.e., $\bigvee \mathcal{F}' \wedge \theta \ddot{\sim} \theta$. Substituting θ for ψ in (6.2) gives

$$\bigvee \mathcal{F}' \wedge \theta \ddot{\sim} \delta \vee (((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \theta). \qquad (6.3)$$

Now, taking the right-hand disjunct from the right-hand side of the above we have

$$\begin{aligned}
&((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \theta \ \dot{\sim} \\
&\dot{\sim} \quad ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0 \wedge \theta) \vee (\neg(\theta \wedge \phi) \wedge \neg(\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0 \wedge \theta) \\
&\dot{\sim} \quad ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0) \vee (\neg(\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0 \wedge \theta) \\
&\dot{\sim} \quad ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0) \vee ((\neg\theta \vee \neg\phi) \wedge \bigvee \mathcal{F}_0 \wedge \theta) \\
&\dot{\sim} \quad ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0) \vee (\neg\theta \wedge \bigvee \mathcal{F}_0 \wedge \theta) \vee (\theta \wedge \neg\phi \wedge \bigvee \mathcal{F}_0 \wedge \theta) \\
&\dot{\sim} \quad ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0) \vee (\theta \wedge \neg\phi \wedge \bigvee \mathcal{F}_0) \vee (\neg\theta \wedge \bigvee \mathcal{F}_0 \wedge \theta) \\
&\ddot{\sim} \quad ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0) \vee (\theta \wedge \neg\phi \wedge \bigvee \mathcal{F}_0) \\
&\dot{\sim} \quad \theta \wedge ((\phi \wedge \bigvee \mathcal{F}_0) \vee (\neg\phi \wedge \bigvee \mathcal{F}_0)) \\
&\ddot{\sim} \quad \theta.
\end{aligned}$$

Hence, applying Proposition 3.6 to this and (6.3) gives us

$$\bigvee \mathcal{F}' \wedge \theta \ddot{\sim} \delta \vee \theta$$

and so we will have shown $\mathcal{F}' \in F_\theta$ if we can show $\delta \vee \theta \ddot{\sim} \theta$. If $\vdash \theta$ then this will clearly hold since in that case also $\vdash \delta \vee \theta$ while we assumed at the outset that $\nvdash \neg\theta$. Thus we may assume that θ is a contingent sentence and so $rT(\theta)$ is well defined. Since $\delta \in T_{\theta\wedge\phi}$ we have $\delta \in T_\theta$ and so $\delta = \tau \wedge \rho$ for some $\tau \in rT(\theta)^+$ and some (possibly empty) conjunction of literals ρ. Hence we have

$$\begin{aligned}
\delta \vee \theta &= (\tau \wedge \rho) \vee \theta \\
&\mathrel{\ddot{\sim}} (\tau \wedge \rho) \vee \bigvee rT(\theta)^+ \\
&\mathrel{\dot{\sim}} (\tau \vee \bigvee rT(\theta)^+) \wedge (\neg\tau \vee \rho \vee \bigvee rT(\theta)^+) \\
&\mathrel{\ddot{\sim}} \tau \vee \bigvee rT(\theta)^+ \\
&\mathrel{\dot{\sim}} \bigvee rT(\theta)^+ \mathrel{\ddot{\sim}} \theta
\end{aligned}$$

as required to show $\mathcal{F}' \in F_\theta$.

Next we will show that $\mathcal{F}' \in F_{\theta\wedge\phi}$, i.e., $\bigvee \mathcal{F}' \wedge \theta \wedge \phi \mathrel{\ddot{\sim}} \theta \wedge \phi$. Substituting $\theta \wedge \phi$ for ψ in (6.2) gives

$$\bigvee \mathcal{F}' \wedge \theta \wedge \phi \mathrel{\ddot{\sim}} \delta \vee (((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \theta \wedge \phi). \tag{6.4}$$

Taking the right-hand disjunct from the right-hand side of the above we have that

$$\begin{aligned}
&((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \theta \wedge \phi \mathrel{\dot{\sim}} \\
&\quad \mathrel{\dot{\sim}} ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0 \wedge \theta \wedge \phi) \vee (\neg(\theta \wedge \phi) \wedge \neg(\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0 \wedge \theta \wedge \phi) \\
&\quad \mathrel{\ddot{\sim}} ((\theta \wedge \phi) \wedge \bigvee \mathcal{F}_0 \wedge \theta \wedge \phi) \\
&\quad \mathrel{\dot{\sim}} \theta \wedge \phi \wedge \bigvee \mathcal{F}_0 \\
&\quad \mathrel{\ddot{\sim}} \theta \wedge \phi.
\end{aligned}$$

Hence, applying Proposition 3.6 to this and (6.4) gives us

$$\bigvee \mathcal{F}' \wedge \theta \wedge \phi \mathrel{\ddot{\sim}} \delta \vee (\theta \wedge \phi).$$

Hence we will have shown $\mathcal{F}' \in F_{\theta\wedge\phi}$ if we can show $\delta \vee (\theta \wedge \phi) \mathrel{\ddot{\sim}} \theta \wedge \phi$. This can be done in an exactly similar way to how we showed that $\delta \vee \theta \mathrel{\ddot{\sim}} \theta$ above (just replace "θ" by "$\theta \wedge \phi$" in the above short proof). Hence we do have $\mathcal{F}' \in F_{\theta\wedge\phi}$.

It remains to prove that $\mathcal{F}' \in F_\delta$, i.e., $\bigvee \mathcal{F}' \wedge \delta \ddot{\sim} \delta$. Substituting δ for ψ in (6.2) gives us

$$\begin{aligned} \bigvee \mathcal{F}' \wedge \delta &\ddot{\sim} \delta \vee (((\theta \wedge \phi) \vee \neg(\theta \wedge \phi)) \wedge \bigvee \mathcal{F}_0 \wedge \delta) \\ &\ddot{\sim} \delta \end{aligned}$$

as required, since for arbitrary $\chi_1, \chi_2 \in SL$ we have

$$\chi_1 \vee (\chi_2 \wedge \chi_1) \dot{\sim} \chi_1 \vee (\neg\chi_1 \wedge \chi_2 \wedge \chi_1) \ddot{\sim} \chi_1.$$

□

By putting $\phi = \theta$ in the above lemma we see that, if θ is contingent (indeed if θ is consistent), we can find a permutation tree $\mathcal{F}$ for θ which also belongs to F_δ, or equivalently contains δ, for any $\delta \in T_\theta$. We just set

$$\mathcal{F} = cT((\delta \vee \neg\delta) \wedge (\theta \vee \neg\theta) \wedge \bigvee \mathcal{F}_0)^+$$

where $\mathcal{F}_0$ is some fixed arbitrary permutation tree over L, i.e., we take $\mathcal{F}$ to contain the positive clauses (which, since we are dealing with a tautology, will be all of the clauses) of $cT((\delta \vee \neg\delta) \wedge (\theta \vee \neg\theta) \wedge \bigvee \mathcal{F}_0)$. Our next proposition is the following:

Proposition 6.6 *Let $\theta, \phi \in SL$. Then $F_{\theta\wedge\phi} \cap F_\theta \subseteq F_{\theta\wedge\neg\phi}$.*

Proof. Let $\mathcal{F} \in F_{\theta\wedge\phi} \cap F_\theta$. Then we have $\bigvee \mathcal{F} \wedge \theta \wedge \phi \ddot{\sim} \theta \wedge \phi$ and $\bigvee \mathcal{F} \wedge \theta \ddot{\sim} \theta$. We must show $\bigvee \mathcal{F} \wedge \theta \wedge \neg\phi \ddot{\sim} \theta \wedge \neg\phi$. Let G be any pre-ent over L. Then

$$\begin{aligned} Bel^G(\bigvee \mathcal{F} \wedge \theta \wedge \neg\phi) &= Bel^G(\bigvee \mathcal{F} \wedge \theta) - Bel^G(\bigvee \mathcal{F} \wedge \theta \wedge \phi) \\ &= Bel^G(\theta) - Bel^G(\theta \wedge \phi) \\ &= Bel^G(\theta \wedge \neg\phi). \end{aligned}$$

Hence $\bigvee \mathcal{F} \wedge \theta \wedge \neg\phi \ddot{\sim} \theta \wedge \neg\phi$ as required. □

In view of Proposition 6.6 we could equally well have written in Lemma 6.5 "$F_{\theta\wedge\neg\phi} \cap F_\theta \cap F_\delta \neq \emptyset$", since, by the proposition, $F_{\theta\wedge\phi} \cap F_\theta \subseteq F_{\theta\wedge\neg\phi}$ and $F_{\theta\wedge\neg\phi} \cap F_\theta \subseteq F_{\theta\wedge\neg\neg\phi} = F_{\theta\wedge\phi}$ (since $\psi \mathrel{\ddot{\sim}} \chi$ implies $F_\psi = F_\chi$). Hence $F_{\theta\wedge\phi} \cap F_\theta = F_{\theta\wedge\neg\phi} \cap F_\theta$. The next proposition shows us how we may express T_θ via the sets $S_\theta^{\mathcal{F}}$.

Proposition 6.7 *Let $\theta \in SL$. Then $T_\theta = \bigcup_{\mathcal{F}\in F_\theta} S_\theta^{\mathcal{F}}$.*

Proof. We first show $T_\theta \subseteq \bigcup_{\mathcal{F}\in F_\theta} S_\theta^{\mathcal{F}}$. If $T_\theta = \emptyset$ (equivalently $\vdash \neg\theta$) then this is clear, so suppose otherwise and let $\delta \in T_\theta$. Then $\delta \vdash \theta$ so it remains to prove that $\delta \in \mathcal{F}$ for some $\mathcal{F} \in F_\theta$. That this is true follows from the discussion right after Lemma 6.5. Hence $T_\theta \subseteq \bigcup_{\mathcal{F}\in F_\theta} S_\theta^{\mathcal{F}}$.
To show the converse let $\delta \in At_*^L$ now be any permatom such that $\delta \in S_\theta^{\mathcal{F}}$ for some $\mathcal{F}$ such that $\mathcal{F} \in F_\theta$. Then

$$\begin{aligned} \delta &\mathrel{|\!\!\ddot{\sim}} \delta \vee \bigvee (S_\theta^{\mathcal{F}} - \{\delta\}) \qquad \text{from Lemma 3.20(1)} \\ &\mathrel{\dot{\sim}} \bigvee S_\theta^{\mathcal{F}} \\ &\mathrel{\ddot{\sim}} \theta \qquad \text{from Proposition 6.4.} \end{aligned}$$

Hence $\delta \mathrel{|\!\!\ddot{\sim}} \theta$ which gives $\delta \in T_\theta$ as required. □

The next proposition gives us some more useful information about permutation trees.

Proposition 6.8 *Let τ be a conjunction of literals from distinct propositional variables in L. Then the following are true:*
(i). For all $\mathcal{F} \in P$, $\mathcal{F} \in F_\tau$ iff $\delta \mathrel{|\!\!\ddot{\sim}} \tau$ for all $\delta \in S_\tau^{\mathcal{F}}$, i.e., $S_\tau^{\mathcal{F}} \subseteq T_\tau$.
(ii). For each $p \in L$ such that $\pm p$ does not appear in τ and for $\epsilon \in \{0,1\}$, $F_{\tau\wedge p^\epsilon} \subseteq F_\tau$ and $F_{\tau\wedge p} = F_{\tau\wedge\neg p}$.

Proof. Let τ be a conjunction of literals from distinct propositional variables in L, say $\tau = q_1^{\epsilon_1} \wedge \ldots \wedge q_j^{\epsilon_j}$.

(i). Let $\mathcal{F} \in P$. The "only if" direction of part (i) follows from Proposition 6.7 which tells us $\bigcup_{\mathcal{F}\in F_\tau} S_\tau^{\mathcal{F}} \subseteq T_\tau$. To prove the "if" direction suppose $S_\tau^{\mathcal{F}} = \{\delta_1, \ldots, \delta_r\}$. Then, for each $i = 1, \ldots, r$, $\delta_i \mathrel{|\!\overset{..}{\sim}} \tau$ so $\delta_i = \tau \wedge \gamma_i$ for some conjunction γ_i of literals from the variables in $L - \{q_1, \ldots, q_j\}$. Hence

$$\bigvee S_\tau^{\mathcal{F}} \mathrel{\dot{\sim}} \bigvee_{i=1}^{r} (\tau \wedge \gamma_i) \mathrel{\dot{\sim}} \tau \wedge \bigvee_{i=1}^{r} \gamma_i.$$

Now, since all the δ_i belong to the same permutation tree $\mathcal{F}$, it should be clear that $\vdash \bigvee_{i=1}^{r} \gamma_i$ (indeed $\{\gamma_1, \ldots, \gamma_r\}$ now looks like a permutation tree over $L - \{q_1, \ldots, q_j\}$). Hence $\bigvee S_\tau^{\mathcal{F}} \mathrel{\overset{..}{\sim}} \tau$ and so $\mathcal{F} \in F_\tau$ by Proposition 6.4 as required.

(ii). Let $p \in L - \{q_1, \ldots, q_j\}$ and let $\mathcal{F} \in F_{\tau \wedge p^\epsilon}$. We must show $\mathcal{F} \in F_\tau$. Choose $\delta \in S_{\tau \wedge p^\epsilon}^{\mathcal{F}}$ (which is obviously non-empty). Then, by part (i) just proved, $\delta \mathrel{|\!\overset{..}{\sim}} \tau \wedge p^\epsilon$, say $\delta = q_1^{\epsilon_1} \wedge \ldots \wedge q_n^{\epsilon_n}$ with $q_{j+1}^{\epsilon_{j+1}} = p^\epsilon$. Let $\delta' \in S_\tau^{\mathcal{F}}$, say $\delta' = r_1^{\delta_1} \wedge \ldots \wedge r_n^{\epsilon_n}$. By part (i) proved above we will show $\mathcal{F} \in F_\tau$ if we can show $\delta' \mathrel{|\!\overset{..}{\sim}} \tau$. But $\delta' \in \mathcal{F}$ and so either $\delta' = \delta$, in which case clearly $\delta' \mathrel{|\!\overset{..}{\sim}} \tau$ as required, or $\delta' \neq \delta$, in which case we know (from part (ii) from the definition of permutation tree) that there exists $1 \leq i \leq n$ such that $r_i^{\delta_i} = q_i^{1-\epsilon_i}$ and $r_k^{\delta_k} = q_k^{\epsilon_k}$ for all $k < i$. If $i < j$ then $\delta' \not\models \tau$ and so $\delta' \notin S_\tau^{\mathcal{F}}$ – contradiction. Hence $i \geq j$ so $\delta' \mathrel{|\!\overset{..}{\sim}} \tau$ again as required.

The second part of (ii) now follows from the first part just proved and the remarks made after Proposition 6.6, since $F_{\tau\wedge p} = F_{\tau \wedge p} \cap F_\tau = F_{\tau \wedge \neg p} \cap F_\tau = F_{\tau \wedge \neg p}$ as required. □

We remark that it is *not* true, in general, that $F_{\theta\wedge\phi} \subseteq F_\theta$ for arbitrary $\theta, \phi \in SL$.

Now, given a permutation tree $\mathcal{F}$, consider how a pre-ent G over L would compute $Bel^G(\bigvee \mathcal{F} \wedge \theta)$ for any $\theta \in SL$. The effect of preceding θ with the sentence $\bigvee \mathcal{F}$ in this way is that, in evaluating $Bel^G(\bigvee \mathcal{F} \wedge \theta)$, G would, firstly, generate a complete picture of the world by "deciding" all the propositional variables of L in the order dictated by $\mathcal{F}$ before then deciding if θ is true in this world. The presence of $\bigvee \mathcal{F}$ here would, in fact, force Bel^G to behave like a λ-probability function. Precisely we have the following.

Proposition 6.9 *Let G be a pre-ent and $\mathcal{F} \in P$. If we define a function $Bel^G_{\mathcal{F}} : SL \to [0,1]^{(\lambda)}$ by setting, for $\theta \in SL$,*

$$Bel^G_{\mathcal{F}}(\theta) = Bel^G(\bigvee \mathcal{F} \wedge \theta),$$

then $Bel^G_{\mathcal{F}}$ is a λ-probability function.

Proof. We will show that, for $\theta \in SL$,

$$Bel^G_{\mathcal{F}}(\theta) = \sum_{\alpha \in S_\theta} Bel^G_{\mathcal{F}}(\alpha) \text{ and } \sum_{\alpha \in At^L} Bel^G_{\mathcal{F}}(\alpha) = 1. \tag{6.5}$$

This will suffice following the representation result for probability functions (which is easily extendable to λ-probability functions) discussed in Section 2.4.
To show the first of the above two conditions we have

$$\begin{aligned} Bel^G_{\mathcal{F}}(\theta) = Bel^G(\bigvee \mathcal{F} \wedge \theta) &= \sum_{t \Vdash \bigvee \mathcal{F}} G_{\bigvee \mathcal{F}}(\emptyset, t) \cdot Bel^G_t(\theta) \\ &= \sum_{t \in WL} G_{\bigvee \mathcal{F}}(\emptyset, t) \cdot Bel^G_t(\theta) \\ &\qquad \text{since for all } t,\ t \vdash \bigvee \mathcal{F}. \end{aligned}$$

We will now show that, for $t \in WL$, if $G_{\bigvee \mathcal{F}}(\emptyset, t) \neq 0$ then $t = t_\alpha$ for some $\alpha \in At^L$ (where, for each $\alpha \in At^L$, we define $t_\alpha = \{p^\epsilon \mid \alpha \vdash p^\epsilon\}$). Given t, we know that it is not the case that for all $\delta \in \mathcal{F}$, $t \vdash \neg\delta$ since if it were then we would have $t \vdash \bot$ contradicting the consistency (by definition of scenario) of t. Hence there exists $\delta_0 \in \mathcal{F}$ such that $t \nvdash \neg\delta_0$. Now if we also had $t \nvdash \delta_0$ then, since by definition we have $G_{\theta \vee \phi}(s, r) = 0$ whenever both $r \nvdash \theta$ and $r \nvdash \neg\theta$ for arbitrary θ, ϕ, s, r, we would have

$$G_{\bigvee \mathcal{F}}(\emptyset, t) = G_{\delta_0 \vee \bigvee(\mathcal{F} - \{\delta_0\})}(\emptyset, t) = 0.$$

Hence we have shown that if $G_{\bigvee \mathcal{F}}(\emptyset, t) \neq 0$ then $t \vdash \delta_0$ for some $\delta_0 \in \mathcal{F}$, and so we must have $t = t_\alpha$ where $\delta_0 \equiv \alpha \in At^L$ as required. So

$$Bel^G_{\mathcal{F}}(\theta) = \sum_{\alpha \in At^L} G_{\bigvee \mathcal{F}}(\emptyset, t_\alpha) \cdot Bel^G_{t_\alpha}(\theta).$$

For each $\alpha \in At^L$ we have either $\alpha \vdash \theta$, i.e., $t_\alpha \vdash \theta$, in which case $Bel^G_{t_\alpha}(\theta) = 1$, or $\alpha \vdash \neg\theta$, i.e., $t_\alpha \vdash \neg\theta$, in which case $Bel^G_{t_\alpha}(\neg\theta) = 1$ and so $Bel^G_{t_\alpha}(\theta) = 0$. Hence

$$Bel^G_{\mathcal{F}}(\theta) = \sum_{\alpha \in S_\theta} G_{\bigvee \mathcal{F}}(\emptyset, t_\alpha).$$

In particular, for any $\alpha \in At^L$

$$Bel^G_{\mathcal{F}}(\alpha) = \sum_{\beta \in S_\alpha} G_{\bigvee \mathcal{F}}(\emptyset, t_\beta) = G_{\bigvee \mathcal{F}}(\emptyset, t_\alpha)$$

hence

$$Bel^G_{\mathcal{F}}(\theta) = \sum_{\alpha \in S_\theta} Bel^G_{\mathcal{F}}(\alpha)$$

as required.

To show the second part of (6.5) we have

$$\sum_{\alpha \in At^L} Bel^G_{\mathcal{F}}(\alpha) = \sum_{\alpha \in S_\top} Bel^G_{\mathcal{F}}(\alpha) = Bel^G_{\mathcal{F}}(\top) = Bel^G(\bigvee \mathcal{F} \wedge \top) = 1$$

since $\bigvee \mathcal{F} \wedge \top$ is a tautology. Hence $Bel^G_{\mathcal{F}}$ is a probability function on SL. □

Corollary 6.10 *Let G be a pre-ent and $\mathcal{F} \in P$. Then for all $\theta, \phi \in SL$ we have if $\theta \vdash \phi$ then $Bel^G_{\mathcal{F}}(\theta) \leq Bel^G_{\mathcal{F}}(\phi)$ (and so if $\theta \equiv \phi$ then $Bel^G_{\mathcal{F}}(\theta) = Bel^G_{\mathcal{F}}(\phi)$).* □

Given $\mathcal{F} \in P$, Proposition 6.9 says, roughly, that by preceding everything by $\bigvee \mathcal{F}$ we turn a pre-ent's belief function into a probability function. Similarly the next proposition says that by preceding everything by $\bigvee \mathcal{F}$ we turn a natural consequence relation into a rational consequence relation. This result will prove very useful in our attempt to characterise natural consequence relations in terms of weakly admissible sequences, since it will frequently allow us to pass from a situation involving natural consequence relations to a situation where we are just talking about rational consequence relations.

Proposition 6.11 *Let $\mid\!\sim$ be a natural consequence relation on L and let $\mathcal{F} \in P$. If we define a binary relation $\mid\!\sim_{\mathcal{F}}$ on SL by setting, for $\theta, \phi \in SL$,*

$$\theta \mid\!\sim_{\mathcal{F}} \phi \textit{ iff } \bigvee \mathcal{F} \wedge \theta \mid\!\sim \phi,$$

then $\mid\!\sim_{\mathcal{F}}$ is a rational consequence relation. (Note that, by LGE, $\mid\!\sim_{\mathcal{F}}$ is independent of the order we take the permatoms in $\mathcal{F}$ to be in.)

Proof. We check each condition for rational consequence in turn.

<u>REF:</u> We have $\bigvee \mathcal{F} \wedge \theta \vdash \theta$ so, by SCL for $\mid\!\sim$ (see Lemma 5.11), $\bigvee \mathcal{F} \wedge \theta \mid\!\sim \theta$, i.e., $\theta \mid\!\sim_{\mathcal{F}} \theta$ as required.

<u>LLE:</u> Suppose $\theta \mid\!\sim_{\mathcal{F}} \phi$ and $\theta \equiv \psi$. Then we have $\bigvee \mathcal{F} \wedge \theta \mid\!\sim \phi$. For all G we have

$$\begin{aligned} Bel^G(\bigvee \mathcal{F} \wedge \theta) &= Bel^G_{\mathcal{F}}(\theta) \\ &= Bel^G_{\mathcal{F}}(\psi) \text{ (by Corollary 6.10)} \\ &= Bel^G(\bigvee \mathcal{F} \wedge \psi) \end{aligned}$$

hence $\bigvee \mathcal{F} \wedge \theta \,\ddot{\sim}\, \bigvee \mathcal{F} \wedge \psi$. Similarly, since $\theta \equiv \psi$ implies $\theta \wedge \phi \equiv \psi \wedge \phi$, we have $(\bigvee \mathcal{F} \wedge \theta) \wedge \phi \,\ddot{\sim}\, (\bigvee \mathcal{F} \wedge \psi) \wedge \phi$. Hence, by P-LLE we have $\bigvee \mathcal{F} \wedge \psi \mid\!\sim \phi$, i.e., $\psi \mid\!\sim_{\mathcal{F}} \phi$ as required.

<u>RWE:</u> Suppose $\theta \mid\!\sim_{\mathcal{F}} \phi$ and $\phi \vdash \psi$. Then $\bigvee \mathcal{F} \wedge \theta \mid\!\sim \phi$ and $\theta \wedge \phi \vdash \theta \wedge \psi$. We have, for all G,

$$\begin{aligned} Bel^G(\bigvee \mathcal{F} \wedge \theta \wedge \phi) &= Bel^G_{\mathcal{F}}(\theta \wedge \phi) \\ &\leq Bel^G_{\mathcal{F}}(\theta \wedge \psi) \text{ (by Corollary 6.10)} \\ &= Bel^G(\bigvee \mathcal{F} \wedge \theta \wedge \psi) \end{aligned}$$

hence, by P-RWE, $\bigvee \mathcal{F} \wedge \theta \mathrel{|\!\sim} \psi$, i.e., $\theta \mathrel{|\!\sim}_{\mathcal{F}} \psi$ as required.

AND: Suppose $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi$ and $\theta \mathrel{|\!\sim}_{\mathcal{F}} \psi$. Then $\bigvee \mathcal{F} \wedge \theta \mathrel{|\!\sim} \phi$ and, by RWE proved above, $\theta \mathrel{|\!\sim}_{\mathcal{F}} \neg\phi \vee \psi$, i.e., $\bigvee \mathcal{F} \wedge \theta \mathrel{|\!\sim} \neg\phi \vee \psi$. Then, by P-AND, $\bigvee \mathcal{F} \wedge \theta \mathrel{|\!\sim} \phi \wedge \psi$, i.e., $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi \wedge \psi$ as required.

CMO: Suppose $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi$ and $\theta \mathrel{|\!\sim}_{\mathcal{F}} \psi$. Then, by AND proved above, $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi \wedge \psi$, i.e., $\bigvee \mathcal{F} \wedge \theta \mathrel{|\!\sim} \phi \wedge \psi$. Then by P-CMO, $(\bigvee \mathcal{F} \wedge \theta) \wedge \phi \mathrel{|\!\sim} \psi$ and so, by LGE, $\bigvee \mathcal{F} \wedge (\theta \wedge \phi) \mathrel{|\!\sim} \psi$, i.e., $\theta \wedge \phi \mathrel{|\!\sim}_{\mathcal{F}} \psi$ as required.

RMO: Suppose $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi$ and $\theta \not\mathrel{|\!\sim}_{\mathcal{F}} \neg\psi$. Then $\bigvee \mathcal{F} \wedge \theta \not\mathrel{|\!\sim} \neg\psi$ and $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi$ gives us $\theta \mathrel{|\!\sim}_{\mathcal{F}} \neg\psi \vee \phi$ by RWE, i.e., $\bigvee \mathcal{F} \wedge \theta \mathrel{|\!\sim} \neg\psi \vee \phi$. Then by P-RMO we have $(\bigvee \mathcal{F} \wedge \theta) \wedge \psi \mathrel{|\!\sim} \phi$ and so, by LGE, $\bigvee \mathcal{F} \wedge (\theta \wedge \psi) \mathrel{|\!\sim} \phi$, i.e., $\theta \wedge \psi \mathrel{|\!\sim}_{\mathcal{F}} \phi$ as required.

OR: Suppose $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi$ and $\psi \mathrel{|\!\sim}_{\mathcal{F}} \phi$. We examine two separate cases:

Case (i): $\psi \not\mathrel{|\!\sim}_{\mathcal{F}} \theta$.

Then, by RWE and RMO, $\psi \wedge \neg\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi$. Hence, by LLE, $\neg\theta \wedge \psi \mathrel{|\!\sim}_{\mathcal{F}} \phi$. From Proposition 3.3 we have

$$\bigvee \mathcal{F} \wedge (\neg\theta \wedge \psi) \dot{\sim} \neg(\bigvee \mathcal{F} \wedge \theta) \wedge (\bigvee \mathcal{F} \wedge \psi)$$

and so, from this equivalence, LGE gives us

$$\neg(\bigvee \mathcal{F} \wedge \theta) \wedge (\bigvee \mathcal{F} \wedge \psi) \mathrel{|\!\sim} \phi.$$

By assumption, $\bigvee \mathcal{F} \wedge \theta \mathrel{|\!\sim} \phi$ so by P-OR, $(\bigvee \mathcal{F} \wedge \theta) \vee (\bigvee \mathcal{F} \wedge \psi) \mathrel{|\!\sim} \phi$ and so $\bigvee \mathcal{F} \wedge (\theta \vee \psi) \mathrel{|\!\sim} \phi$ by LGE, i.e., $\theta \vee \psi \mathrel{|\!\sim}_{\mathcal{F}} \phi$ as required.

Case (ii): $\psi \mathrel{|\!\sim}_{\mathcal{F}} \theta$.

Then we have $\bigvee \mathcal{F} \wedge \psi \mathrel{|\!\sim} \theta$. Also, by SCL for $\mathrel{|\!\sim}$, we have $\bigvee \mathcal{F} \wedge (\neg\psi \wedge \theta) \mathrel{|\!\sim} \theta$.

Therefore, as in case (i), by LGE we get $\neg(\bigvee \mathcal{F} \wedge \psi) \wedge (\bigvee \mathcal{F} \wedge \theta) \mid\!\sim \theta$ so, by P-OR, $(\bigvee \mathcal{F} \wedge \psi) \vee (\bigvee \mathcal{F} \wedge \theta) \mid\!\sim \theta$. Hence, by LGE, $\psi \vee \theta \mid\!\sim_{\mathcal{F}} \theta$ so $\theta \vee \psi \mid\!\sim_{\mathcal{F}} \theta$ by LLE. Since, by assumption, $\theta \mid\!\sim_{\mathcal{F}} \phi$ we get $(\theta \vee \psi) \wedge \theta \mid\!\sim_{\mathcal{F}} \phi$ by LLE. Hence, by RWE, $(\theta \vee \psi) \wedge \theta \mid\!\sim_{\mathcal{F}} \neg\theta \vee \phi$. Now, by SCL for $\mid\!\sim$,

$$\neg(\bigvee \mathcal{F} \wedge (\theta \vee \psi) \wedge \theta) \wedge (\bigvee \mathcal{F} \wedge (\theta \vee \psi) \wedge \neg\theta) \mid\!\sim \neg\theta \vee \phi.$$

So by P-OR,

$$(\bigvee \mathcal{F} \wedge (\theta \vee \psi) \wedge \theta) \vee (\bigvee \mathcal{F} \wedge (\theta \vee \psi) \wedge \neg\theta) \mid\!\sim \neg\theta \vee \phi$$

Hence, by LGE, $(\theta \vee \psi) \wedge (\theta \vee \neg\theta) \mid\!\sim_{\mathcal{F}} \neg\theta \vee \phi$. So,by LLE, $\theta \vee \psi \mid\!\sim_{\mathcal{F}} \neg\theta \vee \phi$ Using the earlier derived $\theta \vee \psi \mid\!\sim_{\mathcal{F}} \theta$ with AND and RWE now gives the result. □

Given a natural consequence relation $\mid\!\sim$ and $\mathcal{F} \in P$ then for each rule Ru for rational consequence we shall write $Ru^{\mathcal{F}}$ to mean "the rule Ru applied to the rational consequence relation $\mid\!\sim_{\mathcal{F}}$". For example OR$^{\mathcal{F}}$, AND$^{\mathcal{F}}$ etc. As an initial corollary to Proposition 6.11 we can now show that if we carry over the preference order which we defined for rational consequence relations (and which we discussed just before Definition 6.1) to natural consequence relations then this relation will be transitive on each permutation tree $\mathcal{F}$. In the proof of this and subsequent results it is useful to note that, given $\mathcal{F} \in P$ and $\delta_1, \delta_2 \in At_*^L$, if $\delta_i \in \mathcal{F}$ for $i = 1, 2$ then

$$\begin{aligned} \bigvee \mathcal{F} \wedge (\delta_1 \vee \delta_2) \;&\dot{\sim}\; ((\delta_1 \vee \delta_2) \vee \bigvee(\mathcal{F} - \{\delta_1, \delta_2\})) \wedge (\delta_1 \vee \delta_2) \\ &\qquad \text{by Proposition 3.15(i)} \\ &\ddot{\sim}\; \delta_1 \vee \delta_2. \end{aligned}$$

While

$$\begin{aligned} \bigvee \mathcal{F} \wedge (\delta_1 \vee \delta_2) \wedge \delta_2 \;&\dot{\sim}\; \bigvee \mathcal{F} \wedge (\delta_2 \vee \delta_1) \wedge \delta_2 \qquad \text{since } \delta_1 \vee \delta_2 \,\dot{\sim}\, \delta_2 \vee \delta_1 \\ &\ddot{\sim}\; \bigvee \mathcal{F} \wedge \delta_2 \end{aligned}$$

$$
\begin{aligned}
& \qquad \text{(using Proposition 3.6, since } (\delta_2 \vee \delta_1) \wedge \delta_2 \ddot{\sim} \delta_2) \\
\dot{\sim} \quad & (\delta_2 \vee \bigvee(\mathcal{F} - \{\delta_2\})) \wedge \delta_2 \\
\ddot{\sim} \quad & \delta_2 \\
\ddot{\sim} \quad & (\delta_2 \vee \delta_1) \wedge \delta_2 \\
\dot{\sim} \quad & (\delta_1 \vee \delta_2) \wedge \delta_2
\end{aligned}
$$

Hence $\bigvee \mathcal{F} \wedge (\delta_1 \vee \delta_2) \wedge \neg\delta_2 \ddot{\sim} (\delta_1 \vee \delta_2) \wedge \neg\delta_2$ and so, by P-LLE, for any natural consequence relation we have $\delta_1 \vee \delta_2 \mathrel{|\!\!\sim} \neg\delta_2$ iff $\bigvee \mathcal{F} \wedge (\delta_1 \vee \delta_2) \mathrel{|\!\!\sim} \neg\delta_2$, i.e., $\delta_1 \vee \delta_2 \mathrel{|\!\!\sim}_{\mathcal{F}} \neg\delta_2$.

Corollary 6.12 *Let $\mathrel{|\!\!\sim}$ be a natural consequence relation and let $\mathcal{F} \in P$. If we define a relation $\prec_\sim$ on At_*^L by, for $\delta_1, \delta_2 \in At_*^L$,*

$$\delta_1 \prec_\sim \delta_2 \quad \textit{iff} \quad \delta_1 \vee \delta_2 \mathrel{|\!\!\sim} \neg\delta_2,$$

then the restriction of $\prec_\sim$ to $\mathcal{F}$ is transitive.

Proof. Let $\delta_1, \delta_2, \delta_3 \in \mathcal{F}$ and suppose $\delta_1 \vee \delta_2 \mathrel{|\!\!\sim} \neg\delta_2$ and $\delta_2 \vee \delta_3 \mathrel{|\!\!\sim} \neg\delta_3$. Then, as above, we obtain $\delta_1 \vee \delta_2 \mathrel{|\!\!\sim}_{\mathcal{F}} \neg\delta_2$ and $\delta_2 \vee \delta_3 \mathrel{|\!\!\sim}_{\mathcal{F}} \neg\delta_3$. Since, by Proposition 6.11, $\mathrel{|\!\!\sim}_{\mathcal{F}}$ is a rational consequence relation, we may deduce (see the proof of Lemma 5.26 taking $\mathrel{|\!\!\sim}$ to be $\mathrel{|\!\!\sim}_{\mathcal{F}}$ there) that $\delta_1 \vee \delta_3 \mathrel{|\!\!\sim}_{\mathcal{F}} \neg\delta_3$ which gives us the required $\delta_1 \vee \delta_3 \mathrel{|\!\!\sim} \neg\delta_3$ by P-LLE again. □

6.3 Full Transitivity

In this section we introduce our new rule **(FT)**, make sure the rule is satisfied for $\mathrel{|\!\!\sim}_G$ for G a pre-ent (and hence sound also for $\mathrel{|\!\!\sim}_{\vec{\mathcal{U}}}$) and make sure the rule is satisfied by all rational consequence relations. It seems we are unable to present this rule in the same simple form as the rules we have seen so far for natural consequence. Rather, we give it using some auxiliary notation. As a first step along this route we now formally define what it means for one permatom to be comparable to another.

Definition 6.13 *Given $\delta_1, \delta_2 \in At^L_*$, we shall say that δ_1* is comparable to *δ_2 iff there exists some τ a (possibly empty) conjunction of literals from distinct propositional variables in L, some $p \in L$ which doesn't appear in τ, and some $\epsilon \in \{0,1\}$ such that $\delta_1 = \tau \wedge p^\epsilon \wedge \ldots$ and $\delta_2 = \tau \wedge p^{1-\epsilon} \wedge \ldots$*

Clearly the relation described above is symmetric and so we may speak of two permatoms "being comparable". However the relation is not transitive, for example, taking $L = \{p, q, r\}$, we have $p \wedge q \wedge r$ is comparable to $\neg p \wedge q \wedge r$ which in turn is comparable to $p \wedge r \wedge q$. However $p \wedge q \wedge r$ is not comparable to $p \wedge r \wedge q$. For another way of expressing comparability we have that δ_1 and δ_2 are comparable iff they are distinct and $F_{\delta_1} \cap F_{\delta_2} \neq \emptyset$, i.e., they have a permutation tree in common.

Given a consequence relation $\mid\!\sim$ on SL we define a binary relation $\prec^c_\sim$ on At^L_* by setting, for $\delta_1, \delta_2 \in At^L_*$,

$$\delta_1 \prec^c_\sim \delta_2 \textit{ iff } \delta_1 \vee \delta_2 \mid\!\sim \neg\delta_2 \textbf{ and } \delta_1 \text{ and } \delta_2 \text{ are comparable.}$$

We shall write $\delta_1 \lhd^c_\sim \delta_2$ to mean that δ_1 and δ_2 are comparable and $\delta_2 \not\prec^c_\sim \delta_1$, i.e., δ_1 and δ_2 are comparable and $\delta_2 \vee \delta_1 \mid\!\not\sim \neg\delta_1$. Note that the relation $\prec^c_\sim$ is not transitive since, as we have seen, the relation of being comparable is not transitive. Taking $\mid\!\sim$ to be a natural consequence relation, the relation $\prec^c_\sim$ provides a fairly intuitive and succinct way of expressing the fact that, according to $\mid\!\sim$, one permatom is more acceptable or preferred to another in the case when the two permatoms are comparable. What makes characterising natural consequence relations so difficult is that, for them, there appears to be no such appealing way of expressing this relation in the case when the two permatoms are *not* comparable.

We now introduce our rule **(FT)** (which stands for Full Transitivity) using the above notation. Let $\mid\!\sim$ be a consequence relation on SL and let $\prec^c_\sim$ be the binary relation on At^L_* defined from $\mid\!\sim$ as above.

(FT) Let $\delta_i \in At^L_*$ for $i = 1, 2, 3, 4$. If $\delta_1 \lhd^c_\sim \delta_2,\ \delta_2 \lhd^c_\sim \delta_3,\ \delta_3 \lhd^c_\sim \delta_4$, δ_1 is comparable to δ_4 and at least one of the above $\lhd^c_\sim$ is actually an occurrence of $\prec^c_\sim$, then $\delta_1 \prec^c_\sim \delta_4$.

That the relations $\mid\!\sim_{\vec{\mathcal{U}}}$ for $\vec{\mathcal{U}}$ a weakly admissible sequence of sets of permatoms, and, in turn, all rational consequence relations, satisfy the rule **(FT)** will, by results in Chapter 5, be implied by the following proposition.

Proposition 6.14 *Let G be a pre-ent over a language L. Then the relation $\mid\!\sim_G$ satisfies* **(FT)**.

Proof. First of all note that for comparable permatoms $\gamma_1, \gamma_2 \in At^L_*$ we have $Bel^G(\gamma_1 \vee \gamma_2) = Bel^G(\gamma_1) + Bel^G(\gamma_2)$ (applying Proposition 3.15(ii)) while

$$\begin{aligned} (\gamma_1 \vee \gamma_2) \wedge \neg\neg\gamma_2 &\ \dot{\sim}\ (\gamma_1 \vee \gamma_2) \wedge \gamma_2 \\ &\ \dot{\sim}\ (\gamma_2 \vee \gamma_1) \wedge \gamma_2 \qquad \text{using Proposition 3.15(i)} \\ &\ \dot{\sim}\ (\gamma_2 \wedge \gamma_2) \vee (\neg\gamma_2 \wedge \gamma_1 \wedge \gamma_2) \\ &\ \ddot{\sim}\ \gamma_2. \end{aligned}$$

Hence, for comparable permatoms $\gamma_1, \gamma_2 \in At^L_*$, we have

$$\begin{aligned} \gamma_1 \vee \gamma_2 \mid\!\sim_G \neg\gamma_2 \quad \text{iff} \quad & \textit{either } Bel^G(\gamma_1 \vee \gamma_2) = 0 \\ & \textit{or } Bel^G(\gamma_1 \vee \gamma_2) \neq 0 \text{ and } \frac{Bel^G((\gamma_1 \vee \gamma_2) \wedge \neg\neg\gamma_2)}{Bel^G(\gamma_1 \vee \gamma_2)} = O(\lambda) \\ \text{iff} \quad & \textit{either } Bel^G(\gamma_1 \vee \gamma_2) = 0 \\ & \textit{or } Bel^G(\gamma_1 \vee \gamma_2) \neq 0 \text{ and } \frac{Bel^G(\gamma_2)}{Bel^G(\gamma_1 \vee \gamma_2)} = O(\lambda) \\ \text{iff} \quad & \textit{either } Bel^G(\gamma_1 \vee \gamma_2) = 0 \\ & \textit{or } Bel^G(\gamma_1 \vee \gamma_2) \neq 0 \text{ and for the least } i \\ & \quad \text{such that } Bel^G(\gamma_1 \vee \gamma_2) \neq O(\lambda^i),\ Bel^G(\gamma_2) = O(\lambda^i). \end{aligned}$$

Now let $\mid\!\sim_G$ satisfy the hypotheses of the rule **(FT)**. We must show that $\delta_1 \vee \delta_4 \mid\!\sim_G \neg\delta_4$, i.e., that either $Bel^G(\delta_1 \vee \delta_4) = 0$ or, if $Bel^G(\delta_1 \vee \delta_4) \neq 0$, that $Bel^G(\delta_4) = O(\lambda^i)$ for the least i such that $Bel^G(\delta_1 \vee \delta_4) \neq O(\lambda^i)$. First we shall show that if $Bel^G(\delta_j) = 0$ for some $j \in \{1, 2, 3\}$ then $Bel^G(\delta_{j+1}) = 0$ and so $Bel^G(\delta_4) = 0$ which would clearly suffice. So suppose $Bel^G(\delta_j) = 0$. We know either $\delta_j \precsim^c \delta_{j+1}$ or $\delta_j \trianglelefteq^c_\sim \delta_{j+1}$. But if $\delta_j \trianglelefteq^c_\sim \delta_{j+1}$ then $\delta_{j+1} \not\precsim^c \delta_j$ and so it must be that $Bel^G(\delta_{j+1} \vee \delta_j) \neq 0$ and, for the least i such that $Bel^G(\delta_{j+1} \vee \delta_j) \neq O(\lambda^i)$, $Bel^G(\delta_j) \neq O(\lambda^i)$, contradicting $Bel^G(\delta_j) = 0$. Hence we must be in the situation where $\delta_j \precsim^c \delta_{j+1}$. But now if $Bel^G(\delta_{j+1}) \neq 0$ then $Bel^G(\delta_j \vee \delta_{j+1}) \neq 0$ and $\delta_j \precsim^c \delta_{j+1}$ implies that, for the least i such that $Bel^G(\delta_j \vee \delta_{j+1}) \neq O(\lambda^i)$, $Bel^G(\delta_{j+1}) = O(\lambda^i)$ which, since $Bel^G(\delta_j \vee \delta_{j+1}) = Bel^G(\delta_j) + Bel^G(\delta_{j+1})$, forces $Bel^G(\delta_j) \neq O(\lambda^i)$, again contradicting $Bel^G(\delta_j) = 0$. Hence it must be that $Bel^G(\delta_{j+1}) = 0$ as required. Hence we may assume now that $Bel^G(\delta_j) \neq 0$ for all $j = 1, \ldots, 4$. Now for arbitrary comparable permatoms γ_1, γ_2 such that $Bel^G(\gamma_j) \neq 0$ $(j = 1, 2)$ we have $\gamma_1 \precsim^c \gamma_2$ iff $Bel^G(\gamma_2) = O(\lambda^i)$ for the least i such that $Bel^G(\gamma_1 \vee \gamma_2) \neq O(\lambda^i)$. Since $Bel^G(\gamma_1 \vee \gamma_2) = Bel^G(\gamma_1) + Bel^G(\gamma_2)$ we have $Bel^G(\gamma_1 \vee \gamma_2) \neq O(\lambda^i)$ iff either $Bel^G(\gamma_1) \neq O(\lambda^i)$ or $Bel^G(\gamma_2) \neq O(\lambda^i)$. Hence we may see that $\gamma_1 \precsim^c \gamma_2$ iff the least i such that $Bel^G(\gamma_1) \neq O(\lambda^i)$ is strictly less than the least i such that $Bel^G(\gamma_2) \neq O(\lambda^i)$. For $j = 1, \ldots, 4$ let i_j be minimal such that $Bel^G(\delta_j) \neq O(\lambda^{i_j})$. Then the hypotheses of the rule **(FT)** tell us that we have $i_1 \leq i_2 \leq i_3 \leq i_4$ where at least one of the inequalities is strict. Hence we have $i_1 < i_4$ and so $\delta_1 \precsim^c \delta_4$ as required. □

Corollary 6.15 *(i). Let $\vec{\mathcal{U}} = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At^L_*$ be a weakly admissible sequence. Then the relation $\mid\!\sim_{\vec{\mathcal{U}}}$ on SL satisfies* **(FT)**.

(ii). Every rational consequence relation on SL satisfies **(FT)**.

Proof. Part (i) follows from Proposition 6.14 and Theorem 5.33 if $\mathcal{U}_i \neq \emptyset$ for some i while it is true trivially if $\mathcal{U}_i = \emptyset$ for all i (since in this case $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$ for *all*

$\theta, \phi \in SL$). Part (ii) follows from part (i) and the result (see Theorem 5.28) that every rational consequence relation is of the form $\mathrel{|\!\sim}_{\vec{\mathcal{U}}}$ for some admissible (and hence some weakly admissible) sequence $\vec{\mathcal{U}}$. □

Hence the rule **(FT)** is sound for $\mathrel{|\!\sim}_{\vec{\mathcal{U}}}$. We are as yet unable to show whether **(FT)** follows from the other rules for natural consequence. Thus we are bound to making the following definition:

Definition 6.16 *A* fully transitive (f. t.) natural consequence relation on *L is a natural consequence relation on L which satisfies the condition* **(FT)**.

Thus Proposition 6.14 says that $\mathrel{|\!\sim}_G$ is a f. t. natural consequence relation for every pre-ent G, while Corollary 6.15 tells us that $\mathrel{|\!\sim}_{\vec{\mathcal{U}}}$ is a f. t. natural consequence relation for every weakly admissible sequence $\vec{\mathcal{U}}$ and that the class of f. t. natural consequence relations includes as a sub-class the class of rational consequence relations. We now turn to our showing how *every* f. t. natural consequence relation is of the form $\mathrel{|\!\sim}_{\vec{\mathcal{U}}}$ for a weakly admissible sequence $\vec{\mathcal{U}}$.

6.4 The Representation Theorem

In this section we concentrate on showing how, from any f. t. natural consequence relation $\mathrel{|\!\sim}$, we can construct a weakly admissible sequence $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mathrel{|\!\sim})$ of sets of permatoms such that $\mathrel{|\!\sim} = \mathrel{|\!\sim}_{\vec{\mathcal{U}}}$. During this process we shall rely heavily on the framework set up in Section 6.2 which will often enable us to make use of our knowledge of rational consequence relations. So let $\mathrel{|\!\sim}$ be a given f. t. natural consequence relation. As we have already indicated, our first step is to set

$$\mathcal{U} = \mathcal{U}(\mathrel{|\!\sim}) = \{\delta \in At_*^L \mid \delta \mathrel{|\!\not\sim} \bot\},$$

i.e., let $\mathcal{U}$ be the set of permatoms which are consistent for $\mathrel{|\!\sim}$.

Given the binary relation $\prec^c_\sim$ defined from $\mathrel{|\!\sim}$ as in the previous section, we now define from $\mid\!\sim$ our full preference relation $\precsim^*$ on $\mathcal{U}$ in terms of $\precsim^c$ as follows. For $\delta_1, \delta_2 \in \mathcal{U}$,

$$\begin{aligned} \delta_1 \precsim^* \delta_2 \quad \textit{iff} \quad & \textit{either } \delta_1 \precsim^c \delta_2 \\ & \textit{or } \delta_1 \trianglelefteq^c \gamma \text{ and } \gamma \precsim^c \delta_2 \text{ for some } \gamma \in \mathcal{U} \\ & \textit{or } \delta_1 \precsim^c \gamma \text{ and } \gamma \trianglelefteq^c \delta_2 \text{ for some } \gamma \in \mathcal{U} \\ & \textit{or } \delta_1 \precsim^c \gamma \text{ and } \gamma \precsim^c \delta_2 \text{ for some } \gamma \in \mathcal{U}. \end{aligned}$$

We would now like to show that $\precsim^*$ is a strict partial order on the set $\mathcal{U}$, i.e., that $\precsim^*$ is transitive and irreflexive on $\mathcal{U}$. This is where the rule **(FT)** will come in. The next lemma (which, in fact, is the only place where we use **(FT)**) will make transitivity easier to see. Henceforth we shall write "$\delta_1 \trianglelefteq^c \delta_2 \trianglelefteq^c \delta_3$" instead of "$\delta_1 \trianglelefteq^c \delta_2,\ \delta_2 \trianglelefteq^c \delta_3$" etc.

Lemma 6.17 *Let $\mid\!\sim$ be a f. t. natural consequence relation and let $\mathcal{U} = \mathcal{U}(\mid\!\sim)$, $\precsim^c$ and $\precsim^*$ be defined from $\mid\!\sim$ as above. Let $\delta_1, \delta_2 \in \mathcal{U}$. Then $\delta_1 \precsim^* \delta_2$ iff there exist $\gamma_1, \ldots, \gamma_r \in \mathcal{U}$ (for some $r \geq 0$) such that $\delta_1 \trianglelefteq^c \gamma_1 \trianglelefteq^c \cdots \trianglelefteq^c \gamma_r \trianglelefteq^c \delta_2$ with at least one of the $\trianglelefteq^c$ a $\precsim^c$. (Thus $\precsim^*$ looks like the "transitive closure" of $\precsim^c$.)*

Proof. The "only if" direction is clear from the definition of $\precsim^*$ while for the "if" direction the result is clear in the case $r \leq 1$, again from the definition of $\precsim^*$. Consider the case $r = 2$, i.e., suppose there exist $\gamma_1, \gamma_2 \in \mathcal{U}$ such that

$$\delta_1 \trianglelefteq^c \gamma_1 \trianglelefteq^c \gamma_2 \trianglelefteq^c \delta_2$$

where at least one of the occurrences of $\trianglelefteq^c$ is actually an occurrence of $\precsim^c$. (So in fact we have either $\delta_1 \trianglelefteq^c \gamma_1 \precsim^* \delta_2$ or $\delta_1 \precsim^* \gamma_2 \trianglelefteq^c \delta_2$.) If δ_1 and δ_2 are comparable then, by the rule **(FT)**, we have $\delta_1 \precsim^c \delta_2$ and so $\delta_1 \precsim^* \delta_2$ as required. If δ_1 and δ_2 are not comparable then we must have either δ_1 and γ_2 are comparable or γ_1 and δ_2 are comparable. The reason for this is as follows. We know δ_1 and γ_1 are comparable, so there exists τ a (possibly empty) conjunction of literals, $p \in L$ and $\epsilon \in \{0,1\}$ such that $\delta_1 = \tau \wedge p^\epsilon \wedge \ldots$ and $\gamma_1 = \tau \wedge p^{1-\epsilon} \wedge \ldots$. Now we know γ_2 is comparable to γ_1. If γ_2 and δ_1 are comparable then we are done, so suppose γ_2 is not comparable to δ_1. Then all this forces γ_2 to be of the form $\tau \wedge p^\epsilon \wedge \ldots$. Also we have that δ_2 is comparable to γ_2, so, since we are assuming δ_2 is not comparable to δ_1, we are forced to conclude that $\delta_2 = \tau \wedge p^\epsilon \wedge \ldots$ and so δ_2 is comparable to γ_1 as required. Now, returning to the proof of the lemma, suppose we are in the situation where δ_1 and γ_2 are comparable. We have $\delta_1 \lhd^c_\sim \gamma_1 \lhd^c_\sim \gamma_2$ where none, one or both of the $\lhd^c_\sim$'s may be a $\prec^c_\sim$. Since $\delta_1, \gamma_1, \gamma_2$ are mutually comparable they must have at least one permutation tree in common. Let $\mathcal{F}$ be such a tree. Suppose we are in the situation where none of the $\lhd^c_\sim$'s is a $\prec^c_\sim$. Then we have $\gamma_1 \vee \delta_1 \not\mid\!\sim \neg\delta_1$ and $\gamma_2 \vee \gamma_1 \not\mid\!\sim \neg\gamma_1$. Hence, by P-LLE, $\gamma_1 \vee \delta_1 \not\mid\!\sim_{\mathcal{F}} \neg\delta_1$ and $\gamma_2 \vee \gamma_1 \not\mid\!\sim_{\mathcal{F}} \neg\gamma_1$. Then, since $\mid\!\sim_{\mathcal{F}}$ is a rational consequence relation, we have (from Lemma 5.25(4)) $\gamma_2 \vee \delta_1 \not\mid\!\sim_{\mathcal{F}} \neg\delta_1$ and so, again by P-LLE, $\gamma_2 \vee \delta_1 \not\mid\!\sim \neg\delta_1$. Hence we have $\delta_1 \lhd^c_\sim \gamma_2$. By again using the fact that $\mid\!\sim_{\mathcal{F}}$ forms a rational consequence relation it is easy to see that in the case where one or both of the $\lhd^c_\sim$'s is a $\prec^c_\sim$, we get $\delta_1 \prec^c_\sim \gamma_2$. Hence we have that $\delta_1 \lhd^c_\sim \gamma_2 \lhd^c_\sim \delta_2$ where at least one of these $\lhd^c_\sim$'s is a $\prec^c_\sim$, i.e., $\delta_1 \prec^*_\sim \delta_2$ as required. In the case where γ_1 and δ_2 are comparable we can follow the above line of reasoning to show $\delta_1 \lhd^c_\sim \gamma_1 \lhd^c_\sim \delta_2$ where at least one of these $\lhd^c_\sim$'s is a $\prec^c_\sim$, which again gives $\delta_1 \prec^*_\sim \delta_2$. We have shown that if there exist $\gamma_1, \gamma_2 \in \mathcal{U}$ such that $\delta_1 \lhd^c_\sim \gamma_1 \lhd^c_\sim \gamma_2 \lhd^c_\sim \delta_2$ with at least one $\lhd^c_\sim$ a $\prec^c_\sim$ then $\delta_1 \prec^*_\sim \delta_2$. This suffices to show the "if" direction since any chain $\delta_1 \lhd^c_\sim \gamma_1 \lhd^c_\sim \cdots \lhd^c_\sim \gamma_r \lhd^c_\sim \delta_2$ (with at least one $\lhd^c_\sim$ a $\prec^c_\sim$) may now be "whittled down" to $\delta_1 \prec^*_\sim \delta_2$. □

Lemma 6.18 *Let $\mid\!\sim$ be a f. t. natural consequence relation and let $\mathcal{U} = \mathcal{U}(\mid\!\sim)$ and $\prec^*_\sim$ be defined from $\mid\!\sim$ as above. Then the relation $\prec^*_\sim$ is transitive and irreflexive on $\mathcal{U}$.*

Proof. Given the representation of $\prec^*_\sim$ proved in Lemma 6.17, transitivity is easy to see. (In fact the relation is still transitive on the whole set At_*^L when we replace "…some $\gamma \in \mathcal{U}$ …" by "…some $\gamma \in At_*^L$ …" in each of the clauses in the definition of $\prec_\sim^*$.) To prove irreflexivity, we will show that, given $\delta \in \mathcal{U}$, each of the four clauses in the definition for $\delta \prec_\sim^* \delta$ is impossible. Firstly, $\delta \prec_\sim^c \delta$ is impossible since, by definition, no permatom is comparable to itself. Secondly suppose that there existed $\gamma \in At_*^L$ such that $\delta \lhd_\sim^c \gamma \prec_\sim^c \delta$. Then $\delta \lhd_\sim^c \gamma$ says that $\gamma \vee \delta \not\mid\!\sim \neg\delta$ while $\gamma \prec_\sim^c \delta$ says that $\gamma \vee \delta \mid\!\sim \neg\delta$, giving a contradiction. By a similar argument there can be no $\gamma \in At_*^L$ such that $\delta \prec_\sim^c \gamma \lhd_\sim^c \delta$. Finally suppose there was some $\gamma \in At_*^L$ such that $\delta \prec_\sim^c \gamma \prec_\sim^c \delta$. Let $\mathcal{F}$ be a permutation tree which δ and γ have in common. Then $\delta \prec_\sim^c \gamma$ says that $\delta \vee \gamma \mid\!\sim \neg\gamma$ and so $\delta \vee \gamma \mid\!\sim_{\mathcal{F}} \neg\gamma$, while $\gamma \prec_\sim^c \delta$ says that $\gamma \vee \delta \mid\!\sim \neg\delta$ and so $\gamma \vee \delta \mid\!\sim_{\mathcal{F}} \neg\delta$, equivalently (by LLE$^{\mathcal{F}}$) $\delta \vee \gamma \mid\!\sim_{\mathcal{F}} \neg\delta$. Hence, using AND$^{\mathcal{F}}$, we get $\delta \vee \gamma \mid\!\sim_{\mathcal{F}} \neg\delta \wedge \neg\gamma$ and so using this together with $\delta \vee \gamma \mid\!\sim_{\mathcal{F}} \delta \vee \gamma$ (REF$^{\mathcal{F}}$), AND$^{\mathcal{F}}$ and RWE$^{\mathcal{F}}$ yields $\delta \vee \gamma \mid\!\sim_{\mathcal{F}} \bot$ which itself implies (from Lemma 5.25(3)) that $\delta \mid\!\sim_{\mathcal{F}} \bot$ and so $\delta \mid\!\sim \bot$ by P-LLE. This, however, contradicts $\delta \in \mathcal{U}$. Hence $\delta \not\prec_\sim^* \delta$ as required. □

Hence we finally have in place a suitable preference order on the set $\mathcal{U}$. Now, analogously to the rational case, define a sequence of sets of permatoms $\mathcal{U}_1, \mathcal{U}_2, \ldots$ inductively by setting, for each $i = 1, 2, \ldots$,

$$\mathcal{U}_i = \{\delta \in At_*^L \mid \delta \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j \text{ and } \delta \text{ is minimal in } \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j \text{ under } \prec_\sim^*\}$$

By the finiteness of At_*^L there exists $k \geq 0$ such that $\mathcal{U}_i = \emptyset$ for all $i > k$. Hence we arrive at a finite sequence $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mid\!\sim) = \mathcal{U}_1, \mathcal{U}_2, \ldots, \mathcal{U}_k$, where $k \geq 0$, of pairwise disjoint, non-empty sets of permatoms with $\mathcal{U} = \bigcup_{i=1}^k \mathcal{U}_i$. It makes sense to talk about minimal elements of the sets $\mathcal{U} - \bigcup_{j<i} \mathcal{U}_i$ since, by Lemma 6.18, $\prec_\sim^*$ is a strict partial order on the set $\mathcal{U}$. We will shortly show that the sequence $\vec{\mathcal{U}}$ is weakly admissible but before that we need the following result.

Lemma 6.19 *Let* $\mathrel{|\!\sim}$ *be a f. t. natural consequence relation and let* $\mathcal{U} = \mathcal{U}(\mathrel{|\!\sim})$ *be defined from* $\mathrel{|\!\sim}$ *as above. Then, for all* $\theta \in SL$, $\mathcal{U} \cap T_\theta = \emptyset$ *iff* $\theta \mathrel{|\!\sim} \bot$.

Proof. Suppose $\mathcal{U} \cap T_\theta = \emptyset$. Choose $\mathcal{F} \in F_\theta$. Then, by Proposition 6.7, $\mathcal{U} \cap S_\theta^{\mathcal{F}} = \emptyset$, equivalently $\delta \mathrel{|\!\sim} \bot$ for all $\delta \in S_\theta^{\mathcal{F}}$. Now $\delta \in S_\theta^{\mathcal{F}}$ implies $\delta \in \mathcal{F}$, equivalently $\bigvee \mathcal{F} \wedge \delta \ddot{\sim} \delta$. Hence, by P-LLE we have, for all $\delta \in S_\theta^{\mathcal{F}}$, $\delta \mathrel{|\!\sim}_{\mathcal{F}} \bot$ which gives, using OR$^{\mathcal{F}}$ repeatedly, $\bigvee S_\theta^{\mathcal{F}} \mathrel{|\!\sim}_{\mathcal{F}} \bot$. Hence, by LLE$^{\mathcal{F}}$, we get $\theta \mathrel{|\!\sim}_{\mathcal{F}} \bot$ and so, since $\mathcal{F} \in F_\theta$, i.e., $\bigvee \mathcal{F} \wedge \theta \ddot{\sim} \theta$, we conclude that $\theta \mathrel{|\!\sim} \bot$ by P-LLE. For the converse direction we may just follow the above chain of reasoning backwards, noting that $\bigvee S_\theta^{\mathcal{F}} \mathrel{|\!\sim}_{\mathcal{F}} \bot$ implies $\delta \mathrel{|\!\sim}_{\mathcal{F}} \bot$ for all $\delta \in S_\theta^{\mathcal{F}}$, using Lemma 5.25(3) and LLE$^{\mathcal{F}}$. □

Proposition 6.20 *Let* $\mathrel{|\!\sim}$ *be a f. t. natural consequence relation on* L *and let the sequence* $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mathrel{|\!\sim})$ *be defined from* $\mathrel{|\!\sim}$ *as above. Then* $\vec{\mathcal{U}}$ *is weakly admissible.*

Proof. We check that the sequence $\mathcal{U}_1, \ldots, \mathcal{U}_k$ satisfies the condition **(WA)**. Let τ be a (possibly empty) conjunction of literals from distinct propositional variables in L and let p be a propositional variable which does not appear in τ. If $\mathcal{U}_i \cap T_\tau = \emptyset$ for all i then we are done, so suppose otherwise and let i be minimal such that $\mathcal{U}_i \cap T_\tau \neq \emptyset$. We are required to show $\mathcal{U}_i \cap (T_{\tau\wedge p} \cup T_{\tau\wedge\neg p}) \neq \emptyset$. If both $\mathcal{U} \cap T_{\tau\wedge p} = \emptyset$ and $\mathcal{U} \cap T_{\tau\wedge\neg p} = \emptyset$ then by Lemma 6.19 we must have both $\tau \wedge p \mathrel{|\!\sim} \bot$ and $\tau \wedge \neg p \mathrel{|\!\sim} \bot$. Now

$$\tau \wedge \neg p \;\ddot{\sim}\; \tau \wedge \neg p \wedge \neg p \;\dot{\sim}\; \neg(\tau \wedge p) \wedge (\tau \wedge \neg p)$$

from Proposition 3.3. Hence from $\tau \wedge \neg p \mathrel{|\!\sim} \bot$ and LGE we get $\neg(\tau \wedge p) \wedge (\tau \wedge \neg p) \mathrel{|\!\sim} \bot$. Hence, using P-OR with $\tau \wedge p \mathrel{|\!\sim} \bot$ we get $(\tau \wedge p) \vee (\tau \wedge \neg p) \mathrel{|\!\sim} \bot$ and so $\tau \mathrel{|\!\sim} \bot$ by P-LLE. Hence, by Lemma 6.19, $\mathcal{U} \cap T_\tau = \emptyset$, contradicting $\mathcal{U}_i \cap T_\tau \neq \emptyset$. Hence $\mathcal{U} \cap (T_{\tau\wedge p} \cup T_{\tau\wedge\neg p}) \neq \emptyset$. So let $\delta_0 \in At_*^L$ be a minimal element of $\mathcal{U} \cap (T_{\tau\wedge p} \cup T_{\tau\wedge\neg p})$ under $\prec_\sim^*$, say $\delta_0 \in T_{\tau\wedge p^{\epsilon_0}}$ where $\epsilon_0 \in \{0, 1\}$.

We will show that $\delta_0 \in \mathcal{U}_i$ which will suffice to prove the result. Suppose for contradiction that $\delta_0 \notin \mathcal{U}_i$. If $\delta_0 \in \mathcal{U}_j$ for some $j < i$ then $\mathcal{U}_j \cap T_\tau \neq \emptyset$ for some $j < i$ (since $T_{\tau \wedge p^{\epsilon_0}} \subseteq T_\tau$ by Proposition 5.18) contradicting the minimality of i. Hence it must be the case that there exists $\gamma \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ such that $\gamma \prec^*_\sim \delta_0$. We will now show that, under these assumptions, it cannot be the case that $\tau = \emptyset$. For suppose we *did* have $\tau = \emptyset$ and so $\delta_0 = p^{\epsilon_0} \wedge \ldots$ and δ_0 is minimal under $\prec^*_\sim$ in $\mathcal{U} \cap (T_p \cup T_{\neg p})$. Then (by Lemma 6.17) $\gamma \prec^*_\sim \delta_0$ implies that there exist $\nu_1, \ldots \nu_r \in \mathcal{U}$ (for some $r \geq 0$) such that $\gamma \lhd^c_\sim \nu_1 \lhd^c_\sim \ldots \lhd^c_\sim \nu_r \lhd^c_\sim \delta_0$ where at least one of the $\lhd^c_\sim$'s is actually a $\prec^c_\sim$. Each of the permatoms in this chain is comparable with its successor and so must begin with a $\pm p$. In particular we must have $\gamma \in \mathcal{U} \cap (T_p \cup T_{\neg p})$ and so δ_0 is not minimal in this set – contradiction. Hence for the rest of this proof we may assume $\tau \neq \emptyset$. Say $\tau = q_1^{\epsilon_1} \wedge \ldots \wedge q_l^{\epsilon_l}$ ($l > 0$). Given that δ_0 is not minimal under $\prec^*_\sim$ in $\mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ we can be in one of two situations: either there exists $\gamma \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ such that $\gamma \prec^c_\sim \delta_0$ or there is no such γ and the only elements which are keeping δ_0 out of $\mathcal{U}_i$ do so "indirectly", i.e., for any $\gamma_1 \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ such that $\gamma_1 \prec^*_\sim \delta_0$ we have γ_1 is not comparable to δ_0 and so there exists $\gamma_2 \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ such that $\gamma_1 \prec^c_\sim \gamma_2 \lhd^c_\sim \delta_0$. We examine these two cases separately.

Case (i): There exists $\gamma \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ such that $\gamma \prec^c_\sim \delta_0$

Firstly let us assume there exists $\gamma \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ such that $\gamma \prec^c_\sim \delta_0$ (so γ is comparable to δ_0 and $\gamma \vee \delta_0 \mathrel{|\!\sim} \neg\delta_0$). If $\gamma = \tau \wedge \pm p \wedge \ldots$ then we have $\gamma \prec^*_\sim \delta_0$ and $\gamma \in \mathcal{U} \cap (T_{\tau \wedge p} \cup T_{\tau \wedge \neg p})$ contradicting the minimality in this set of δ_0. Hence we must have $\gamma = q_1^{\epsilon_1} \wedge \ldots \wedge q_r^{1-\epsilon_r} \wedge \ldots$ for some $1 \leq r \leq l$. Now choose $\mathcal{F} \in F_{\delta_0} \cap F_\gamma$ (so $\delta_0 \in S_\tau^{\mathcal{F}}$). By P-LLE with $\gamma \vee \delta_0 \mathrel{|\!\sim} \neg\delta_0$ we have $\gamma \vee \delta_0 \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta_0$. We would like to show now that $\gamma \vee \nu \mathrel{|\!\sim}_{\mathcal{F}} \neg\nu$ for *all* $\nu \in S_\tau^{\mathcal{F}}$. So let $\nu \in S_\tau^{\mathcal{F}}$. If $\nu = \delta_0$ then we are done so suppose $\nu \neq \delta_0$. If we can show that $\nu \vee \delta_0 \mathrel{|\!\not\sim}_{\mathcal{F}} \neg\delta_0$ then from this and $\gamma \vee \delta_0 \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta_0$ we may apply a contrapositive form

of Lemma 5.25(4) to the rational consequence relation $\mid\!\sim_{\mathcal{F}}$ to obtain our desired conclusion. So suppose firstly that $\nu \notin \mathcal{U}$, i.e., $\nu \mid\!\sim \bot$. Then we have $\nu \mid\!\sim_{\mathcal{F}} \bot$ by P-LLE. Now if it were the case that $\nu \vee \delta_0 \mid\!\sim_{\mathcal{F}} \neg\delta_0$ then, applying Proposition 5.25(2) to the rational consequence relation $\mid\!\sim_{\mathcal{F}}$ would give $\nu \vee \delta_0 \mid\!\sim_{\mathcal{F}} \bot$. Then, applying Proposition 5.25(3) (and LLE$^{\mathcal{F}}$) to this would yield $\delta_0 \mid\!\sim_{\mathcal{F}} \bot$ and so $\delta_0 \mid\!\sim \bot$ by P-LLE, contradicting $\delta_0 \in \mathcal{U}$. Hence if $\nu \notin \mathcal{U}$ then $\nu \vee \delta_0 \not\mid\!\sim_{\mathcal{F}} \neg\delta_0$ as required. On the other hand suppose $\nu \in \mathcal{U}$. Then we have $\nu \vdash \tau \wedge p^{\epsilon}$ for some $\epsilon \in \{0,1\}$ and also $\nu \in \mathcal{F} \in F_{\delta_0}$. By repeated use of Proposition 6.8(ii) we know $F_{\delta_0} \subseteq F_{\tau\wedge p^{\epsilon_0}} = F_{\tau\wedge p^{\epsilon}}$. Hence $\nu \in \mathcal{F}$ for some $\mathcal{F} \in F_{\tau\wedge p^{\epsilon}}$ and so, by Proposition 6.7, $\nu \in T_{\tau\wedge p^{\epsilon}}$. Thus if $\nu \vee \delta_0 \mid\!\sim_{\mathcal{F}} \neg\delta_0$ then $\nu \vee \delta_0 \mid\!\sim \neg\delta_0$ by P-LLE, i.e., $\nu \precsim^c \delta_0$, which gives $\nu \precsim^* \delta_0$ which contradicts the minimality of δ_0 in $\mathcal{U} \cap (T_{\tau\wedge p} \cup T_{\tau\wedge\neg p})$. Hence in this case also we have $\nu \vee \delta_0 \not\mid\!\sim_{\mathcal{F}} \neg\delta_0$ as required. Thus we have shown $\gamma \vee \nu \mid\!\sim_{\mathcal{F}} \neg\nu$ for all $\nu \in S^{\mathcal{F}}_{\tau}$. Hence $\gamma \vee \nu \mid\!\sim_{\mathcal{F}} \gamma$ (using REF$^{\mathcal{F}}$, AND$^{\mathcal{F}}$, RWE$^{\mathcal{F}}$) for all $\nu \in S^{\mathcal{F}}_{\tau}$, and then repeatedly using OR$^{\mathcal{F}}$ followed by LLE$^{\mathcal{F}}$ gives us $\gamma \vee \bigvee S^{\mathcal{F}}_{\tau} \mid\!\sim_{\mathcal{F}} \gamma$. Now we have

$$\begin{aligned}
\bigvee \mathcal{F} \wedge (\gamma \vee \bigvee S^{\mathcal{F}}_{\tau}) &\mathrel{\ddot{\sim}} ((\gamma \vee \bigvee S^{\mathcal{F}}_{\tau}) \vee \bigvee(\mathcal{F} - (\{\gamma\} \cup S^{\mathcal{F}}_{\tau}))) \wedge (\gamma \vee \bigvee S^{\mathcal{F}}_{\tau}) \\
&\mathrel{\ddot{\sim}} \gamma \vee \bigvee S^{\mathcal{F}}_{\tau} \\
&\mathrel{\ddot{\sim}} \gamma \vee \tau \qquad (6.6)
\end{aligned}$$

since $\mathcal{F} \in F_{\delta_0} \subseteq F_{\tau}$ so $\tau \mathrel{\ddot{\sim}} \bigvee S^{\mathcal{F}}_{\tau}$ by Proposition 6.4. While also

$$\bigvee \mathcal{F} \wedge (\gamma \vee \bigvee S^{\mathcal{F}}_{\tau}) \wedge \gamma \mathrel{\ddot{\sim}} \bigvee \mathcal{F} \wedge \gamma \mathrel{\ddot{\sim}} \gamma \mathrel{\ddot{\sim}} (\gamma \vee \tau) \wedge \gamma. \qquad (6.7)$$

Hence from $\gamma \vee \bigvee S^{\mathcal{F}}_{\tau} \mid\!\sim_{\mathcal{F}} \gamma$ we may apply P-LLE to get $\gamma \vee \tau \mid\!\sim \gamma$. We will now show that $\gamma \precsim^* \eta$ for all $\eta \in T_{\tau}$, thereby showing that $\mathcal{U}_i \cap T_{\tau} = \emptyset$ which will give the required contradiction to show $\delta_0 \in \mathcal{U}_i$. So let $\eta \in T_{\tau}$, so $\eta = \tau \wedge \rho$ for some ρ and γ, η are comparable. Then, by the rule (A) from Lemma 5.11, from $\gamma \vee \tau \mid\!\sim \gamma$ we get $\gamma \vee (\tau \wedge \rho) \mid\!\sim \gamma$, i.e., $\gamma \vee \eta \mid\!\sim \gamma$. We have

$$(\gamma \vee \eta) \wedge \gamma \mathrel{\ddot{\sim}} \gamma \quad \mathrel{\ddot{\sim}} \quad \gamma \wedge \neg\eta$$

$$\begin{aligned} & \text{(by Theorem 3.7 since } \gamma, \eta \text{ comparable implies } \gamma \vdash \neg\eta) \\ \ddot{\sim} \quad & (\gamma \wedge \neg\eta) \vee (\neg\gamma \wedge \eta \wedge \neg\eta) \\ \dot{\sim} \quad & (\gamma \vee \eta) \wedge \neg\eta. \end{aligned}$$

Hence we may apply P-RWE to $\gamma \vee \eta \mid\!\sim \gamma$ to obtain $\gamma \vee \eta \mid\!\sim \neg\eta$ which means $\gamma \prec^c_\sim \eta$ and so $\gamma \prec^*_\sim \eta$ as required. Hence $\delta_0 \in \mathcal{U}_i \cap (T_{\tau\wedge p} \cup T_{\tau\wedge\neg p}) \neq \emptyset$.

<u>*Case (ii)*: For no $\gamma \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ do we have $\gamma \prec^c_\sim \delta_0$.</u>

Suppose for no $\gamma \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ do we have $\gamma \prec^c_\sim \delta_0$. Then $\delta_0 \notin \mathcal{U}_i$ implies that there must exist $\gamma_1, \gamma_2 \in \mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ such that $\gamma_1 \prec^c_\sim \gamma_2 \lhd^c_\sim \delta_0$ with γ_1 and δ_0 not comparable (otherwise $\gamma_1 \prec^c_\sim \delta_0$). We have that γ_2 and δ_0 are comparable and $\delta_0 \vee \gamma_2 \not\mid\!\sim \neg\gamma_2$. If $\gamma_2 = \tau \wedge \pm p \wedge \ldots$ then γ_2 must be a minimal element, under $\prec^*_\sim$, of $\mathcal{U} \cap (T_{\tau\wedge p} \cup T_{\tau\wedge\neg p})$ (since δ_0 is) and so, by case (i) proved above (just substituting γ_2 for δ_0), $\gamma_1 \prec^*_\sim \eta$ for all $\eta \in T_\tau$ and so $\mathcal{U}_i \cap T_\tau = \emptyset$ giving the required contradiction. So now suppose $\gamma_2 = q_1^{\epsilon_1} \wedge \ldots \wedge q_r^{1-\epsilon_r} \wedge \ldots$ for some $1 \leq r \leq l$. Choose $\mathcal{F} \in F_{\delta_0} \cap F_{\gamma_2}$. By P-LLE with $\delta_0 \vee \gamma_2 \not\mid\!\sim \neg\gamma_2$ we get $\delta_0 \vee \gamma_2 \not\mid\!\sim_{\mathcal{F}} \neg\gamma_2$. We would now like to show $\nu \vee \gamma_2 \not\mid\!\sim_{\mathcal{F}} \neg\gamma_2$ for *all* $\nu \in S^{\mathcal{F}}_\tau$. So let $\nu \in S^{\mathcal{F}}_\tau$. If $\nu = \delta_0$ then we are done so suppose $\nu \neq \delta_0$. As in case (i) we get $\nu \vee \delta_0 \not\mid\!\sim_{\mathcal{F}} \neg\delta_0$ and so, from this and $\delta_0 \vee \gamma_2 \not\mid\!\sim_{\mathcal{F}} \neg\gamma_2$ we may apply another contrapositive form of Lemma 5.25(4) to $\mid\!\sim_{\mathcal{F}}$ to obtain our desired conclusion. Given that $\nu \vee \gamma_2 \not\mid\!\sim_{\mathcal{F}} \neg\gamma_2$ for all $\nu \in S^{\mathcal{F}}_\tau$ we may then repeatedly apply the rule $\mathrm{DR}^{\mathcal{F}}$ followed by $\mathrm{LLE}^{\mathcal{F}}$ to obtain $\gamma_2 \vee \bigvee S^{\mathcal{F}}_\tau \not\mid\!\sim_{\mathcal{F}} \neg\gamma_2$. Now, similarly to equations (6.6) and (6.7) above, we can show

$$\bigvee \mathcal{F} \wedge (\gamma_2 \vee \bigvee S^{\mathcal{F}}_\tau) \ddot{\sim} \gamma_2 \vee \tau$$

and

$$\bigvee \mathcal{F} \wedge (\gamma_2 \vee \bigvee S^{\mathcal{F}}_\tau) \wedge \gamma_2 \ddot{\sim} (\gamma_2 \vee \tau) \wedge \gamma_2,$$

which together imply also

$$\bigvee \mathcal{F} \wedge (\gamma_2 \vee \bigvee S^{\mathcal{F}}_\tau) \wedge \neg\gamma_2 \ddot{\sim} (\gamma_2 \vee \tau) \wedge \neg\gamma_2.$$

Hence we may apply P-LLE to $\gamma_2 \vee \bigvee S_\tau^{\mathcal{F}} \not\mathrel{|\!\sim}_{\mathcal{F}} \neg\gamma_2$ to obtain $\gamma_2 \vee \tau \not\mathrel{|\!\sim} \neg\gamma_2$. Now let $\eta \in T_\tau$, so $\eta = \tau \wedge \rho$ for some ρ and γ_2, η are comparable. As in case (i) we will show that $\eta \notin \mathcal{U}_i$, proving the contradiction $\mathcal{U}_i \cap T_\tau = \emptyset$. By rule (B) from Lemma 5.11, from $\gamma_2 \vee \tau \not\mathrel{|\!\sim} \neg\gamma_2$ we get $\gamma_2 \vee (\tau \wedge \rho) \not\mathrel{|\!\sim} \neg\gamma_2$, i.e., $\gamma_2 \vee \eta \not\mathrel{|\!\sim} \neg\gamma_2$. Hence, by LGE, $\eta \vee \gamma_2 \not\mathrel{|\!\sim} \neg\gamma_2$, i.e., $\gamma_2 \lhd_\sim^c \eta$, and so we have $\gamma_1 \prec_\sim^c \gamma_2 \lhd_\sim^c \eta$ which gives $\gamma_1 \prec_\sim^* \eta$ and so $\eta \notin \mathcal{U}_i$ as required. □

We are now finally in a position to prove our representation result. Many of the steps involved in the following proof closely parallel those of the proof of the corresponding result (see Theorem 5.28) for rational consequence relations.

Theorem 6.21 *Let $\mathrel{|\!\sim}$ be a f. t. natural consequence relation. Then there exists a weakly admissible sequence $\vec{\mathcal{U}}$ such that $\mathrel{|\!\sim} = \mathrel{|\!\sim}_{\vec{\mathcal{U}}}$.*

Proof. Let $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mathrel{|\!\sim}) = \mathcal{U}_1, \ldots, \mathcal{U}_k \subseteq At_*^L$ be defined from $\mathrel{|\!\sim}$ as in the above process. We claim that this sequence, which is weakly admissible by Proposition 6.20, suffices, i.e., that we have, for all $\theta, \phi \in SL$,

$$\theta \mathrel{|\!\sim} \phi \quad \text{iff} \quad \textit{either} \quad \mathcal{U}_i \cap T_\theta = \emptyset \text{ for all } i$$
$$\textit{or} \quad \mathcal{U}_i \cap T_{\theta\wedge\neg\phi} = \emptyset \text{ for the least } i \text{ such that } \mathcal{U}_i \cap T_\theta \neq \emptyset$$

For the "only if" direction suppose $\theta \mathrel{|\!\sim} \phi$. If $\mathcal{U}_i \cap T_\theta = \emptyset$ for all i then we are done, so suppose otherwise and let i be minimal such that $\mathcal{U}_i \cap T_\theta \neq \emptyset$. If $T_{\theta\wedge\neg\phi} = \emptyset$ (equivalently $\vdash \neg(\theta \wedge \neg\phi)$) then the result is clear, so suppose otherwise and let $\delta \in T_{\theta\wedge\neg\phi}$. We will show that $\delta \notin \mathcal{U}_i$ and so $\mathcal{U}_i \cap T_{\theta\wedge\neg\phi} = \emptyset$. Choose $\mathcal{F} \in F_{\theta\wedge\phi} \cap F_\theta \cap F_\delta$. Such an $\mathcal{F}$ exists by Lemma 6.5 and the discussion following Proposition 6.6. Then, by P-LLE, $\theta \mathrel{|\!\sim} \phi$ iff $\theta \mathrel{|\!\sim}_{\mathcal{F}} \phi$. Now $\delta \in T_{\theta\wedge\neg\phi}$ implies $\delta \vdash \neg\phi$ and so by RWE$^{\mathcal{F}}$ we have $\theta \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$. Then, by LLE$^{\mathcal{F}}$, we get $\bigvee S_\theta^{\mathcal{F}} \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$. Now if $S_\theta^{\mathcal{F}} = \{\delta\}$ this means we have $\delta \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$ and so $\delta \mathrel{|\!\sim} \neg\delta$ by P-LLE which gives $\delta \mathrel{|\!\sim} \bot$ by P-RWE. Hence in this case we have $\delta \notin \mathcal{U}$ and so $\delta \notin \mathcal{U}_i$ as required. Now suppose $S_\theta^{\mathcal{F}} = \{\delta, \gamma_1, \ldots, \gamma_r\}$ where $r > 0$. Then we have

$\delta \vee \gamma_1 \vee \ldots \vee \gamma_r \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$ and so, by LLE$^{\mathcal{F}}$, $(\gamma_1 \vee \delta) \vee \ldots \vee (\gamma_r \vee \delta) \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$. Hence, using DR$^{\mathcal{F}}$ repeatedly, we must have $\gamma_l \vee \delta \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$ for some $1 \leq l \leq r$. Now if $\gamma_l \notin \mathcal{U}$ then $\gamma_l \mathrel{|\!\sim} \perp$ so $\gamma_l \mathrel{|\!\sim}_{\mathcal{F}} \perp$ by P-LLE. This together with $\gamma_l \vee \delta \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$ would give $\delta \mathrel{|\!\sim}_{\mathcal{F}} \perp$ (mainly by (2) and (3) of Lemma 5.25 applied to $\mathrel{|\!\sim}_{\mathcal{F}}$), and so $\delta \mathrel{|\!\sim} \perp$ by P-LLE. Hence again $\delta \notin \mathcal{U}$ which gives $\delta \notin \mathcal{U}_i$ as required. Now suppose $\gamma_l \in \mathcal{U}$. Then, by the minimality of i, we have $\gamma_l \notin \mathcal{U}_j$ for all $j < i$ (since $S^{\mathcal{F}}_\theta \subseteq T_\theta$ by Proposition 6.7 so $\gamma_l \in T_\theta$). Also, since $\gamma_l \vee \delta \mathrel{|\!\sim}_{\mathcal{F}} \neg\delta$ implies $\gamma_l \vee \delta \mathrel{|\!\sim} \neg\delta$ by P-LLE, we have $\gamma_l \prec^c_\sim \delta$ and so, if $\delta \in \mathcal{U}$ then δ is not minimal in $\mathcal{U} - \bigcup_{j<i} \mathcal{U}_j$ and so again $\delta \notin \mathcal{U}_i$ as required.

For the "if" direction, first suppose $\mathcal{U}_i \cap T_\theta = \emptyset$ for all i, i.e., $\mathcal{U} \cap T_\theta = \emptyset$. Then, by Lemma 6.19, we have $\theta \mathrel{|\!\sim} \perp$ and so, by P-RWE, $\theta \mathrel{|\!\sim} \phi$ as required. So now suppose $\mathcal{U} \cap T_\theta \neq \emptyset$ and let i be minimal such that $\mathcal{U}_i \cap T_\theta \neq \emptyset$. We will show that if $\theta \mathrel{|\!\not\sim} \phi$ then $\mathcal{U}_i \cap T_{\theta \wedge \neg\phi} \neq \emptyset$. By Corollary 5.35, since the sequence $\vec{\mathcal{U}}$ is weakly admissible, i is also minimal such that $\mathcal{U}_i \cap (T_{\theta\wedge\phi} \cup T_{\theta\wedge\neg\phi}) \neq \emptyset$. Let $\delta_0 \in \mathcal{U}_i \cap (T_{\theta\wedge\phi} \cup T_{\theta\wedge\neg\phi})$. Then clearly δ_0 must be minimal in $\mathcal{U} \cap (T_{\theta\wedge\phi} \cup T_{\theta\wedge\neg\phi})$. Choose $\mathcal{F} \in F_{\theta\wedge\phi} \cap F_\theta \cap F_{\delta_0}$ (and recall that any such $\mathcal{F}$ is also in $F_{\theta\wedge\neg\phi}$). Then, by P-LLE, $\theta \mathrel{|\!\not\sim} \phi$ iff $\theta \mathrel{|\!\not\sim}_{\mathcal{F}} \phi$ which is equivalent to $\bigvee S^{\mathcal{F}}_\theta \mathrel{|\!\not\sim}_{\mathcal{F}} \phi$ by LLE$^{\mathcal{F}}$. We know that δ_0 is a minimal element in $\mathcal{U} \cap S^{\mathcal{F}}_\theta$ under $\prec^*_\sim$, since $S^{\mathcal{F}}_\theta = S^{\mathcal{F}}_{\theta\wedge\phi} \cup S^{\mathcal{F}}_{\theta\wedge\neg\phi} \subseteq T_{\theta\wedge\phi} \cup T_{\theta\wedge\neg\phi}$ by Proposition 6.7. Let $\delta_1, \ldots, \delta_r \in At^L_*$ be the other (if any) permatoms which are minimal under $\prec^*_\sim$ in $\mathcal{U} \cap S^{\mathcal{F}}_\theta$. For all $\gamma \in S^{\mathcal{F}}_\theta - \{\delta_0, \delta_1, \ldots, \delta_r\}$ we have $\delta_j \vee \gamma \mathrel{|\!\sim}_{\mathcal{F}} \neg\gamma$ for some $j \in \{0, 1, \ldots, r\}$. This is clear if $\gamma \in \mathcal{U}$ (since otherwise γ would be one of the minimal elements) while if $\gamma \notin \mathcal{U}$ then $\gamma \mathrel{|\!\sim} \perp$, equivalently $\gamma \mathrel{|\!\sim}_{\mathcal{F}} \perp$, and so $\delta_0 \vee \gamma \mathrel{|\!\sim}_{\mathcal{F}} \neg\gamma$ by Lemma 5.25(6). Hence we may repeatedly apply Lemma 5.25(5) to $\bigvee S^{\mathcal{F}}_\theta \mathrel{|\!\not\sim}_{\mathcal{F}} \phi$ to obtain $\delta_0 \vee \delta_1 \vee \ldots \vee \delta_r \mathrel{|\!\not\sim}_{\mathcal{F}} \phi$, which means we must have $\delta_y \mathrel{|\!\not\sim}_{\mathcal{F}} \phi$, equivalently (by P-LLE) $\delta_y \mathrel{|\!\not\sim} \phi$, for some $0 \leq y \leq r$ (since otherwise we would be be able to derive $\delta_0 \vee \delta_1 \vee \ldots \vee \delta_r \mathrel{|\!\sim}_{\mathcal{F}} \phi$ by repeated use of OR$^{\mathcal{F}}$). If $\delta_y \vdash \phi$ then $\delta_y \mathrel{|\!\sim} \phi$ by SCL. Hence we must have $\delta_y \nvdash \phi$, i.e., $\delta_y \vdash \neg\phi$

and so $\delta_y \in T_{\theta\wedge\neg\phi}$. If $y = 0$ then $\delta_y = \delta_0 \in \mathcal{U}_i \cap T_{\theta\wedge\neg\phi}$. Hence $\mathcal{U}_i \cap T_{\theta\wedge\neg\phi} \neq \emptyset$ as required. If $y \neq 0$ then if $\delta_y \notin \mathcal{U}_i$ there must exist some $\lambda \in \mathcal{U} - \bigcup_{j<i}\mathcal{U}_j$ such that $\lambda \precsim^* \delta_y$. But then we have $\lambda \precsim^* \delta_y \trianglelefteq^c_\sim \delta_0$ so $\lambda \precsim^* \delta_0$ which implies $\delta_0 \notin \mathcal{U}_i$ – contradiction. Hence it must be the case that $\delta_y \in \mathcal{U}_i$ and so $\mathcal{U}_i \cap T_{\theta\wedge\neg\phi} \neq \emptyset$ as required. □

Thus we have characterised f. t. natural consequence relations in terms of weakly admissible sequences of sets of permatoms. We straight away give the following corollary.

Corollary 6.22 *Let $\mid\!\sim$ be a non-trivial f. t. natural consequence relation on L. Then there exists a pre-ent G (possibly over a larger language than L) such that, for all $\theta, \phi \in SL$, $\theta \mid\!\sim \phi$ iff $\theta \mid\!\sim_G \phi$.*

Proof. Let $\mid\!\sim$ be a non-trivial natural consequence relation on L and define $\vec{\mathcal{U}} = \vec{\mathcal{U}}(\mid\!\sim) = \mathcal{U}_1, \ldots, \mathcal{U}_k$ from $\mid\!\sim$ as in the above construction process. Then, by Theorem 6.21, for all $\theta, \phi \in SL$, $\theta \mid\!\sim \phi$ iff $\theta \mid\!\sim_{\vec{\mathcal{U}}} \phi$. If $\mathcal{U}_i = \emptyset$ for all $i = 1, \ldots, k$ then clearly $\mid\!\sim_{\vec{\mathcal{U}}}$ (and hence $\mid\!\sim$) is the trivial f. t. consequence relation on L – contradiction. Hence $\mathcal{U}_i \neq \emptyset$ for some i and so we may conclude from Theorem 5.33. □

Hence we have a representation result for f. t. natural consequence relations in terms of the family of relations $\mid\!\sim_G$ for G a pre-ent. The next section brings in one of our earlier results to characterise the sub-family of relations $\mid\!\sim_z$ for z an ent.

6.5 Ents and F.T. Natural Consequence

The results of the previous section have shown the correspondence between f. t. natural consequence relations and the family of consequence relations $\mid\!\sim_G$ for G

a pre-ent. A natural question to ask is: Does there exist a special sub-class of f. t. natural consequence relations which corresponds to the family of consequence relations $\mid\!\sim_z$ for z an *ent*? In this section we give a positive answer to this question. This answer will draw on the main result – Theorem 4.1 – of Chapter 4. It turns out that all we need to do to obtain this correspondence is add a single rule to the rules we already have for f. t. natural consequence.

Theorem 6.23 *Let z be an ent over L. Then the consequence relation $\mid\!\sim_z$ is a non-trivial f. t. natural consequence relation which, in addition, satisfies the following rule*

$$\frac{\theta \wedge \phi \mid\!\sim \bot}{\phi \wedge \theta \mid\!\sim \bot}\textit{(Consistency Commutativity(CCM))}$$

Conversely, given a non-trivial f. t. natural consequence relation $\mid\!\sim$ on L which also satisfies CCM, there exists an ent z (over a larger language than L) such that, for all $\theta, \phi \in SL$, $\theta \mid\!\sim \phi$ iff $\theta \mid\!\sim_z \phi$.

Proof. Let z be an ent over L. By Theorem 5.10 and Proposition 6.14 we know that, for any *pre-ent* G over L, $\mid\!\sim_G$ is a f. t. natural consequence relation while we also know that $\mid\!\sim_G$ is non-trivial (see the discussion following Definition 5.29). Hence certainly $\mid\!\sim_z$ is a non-trivial f. t. natural consequence relation. To show that $\mid\!\sim_z$ satisfies CCM we have, once again for any arbitrary pre-ent G and for any $\psi \in L$, $\psi \mid\!\sim_G \bot$ iff $Bel^G(\psi) = 0$, since if $\psi \mid\!\sim_G \bot$ and $Bel^G(\psi) \neq 0$ then, by definition of $\mid\!\sim_G$, we must have

$$\frac{Bel^G(\psi \wedge \neg\bot)}{Bel^G(\psi)} = \frac{Bel^G(\psi)}{Bel^G(\psi)} = 1 = O(\lambda)$$

– contradiction. Hence $\psi \mid\!\sim_G \bot$ implies $Bel^G(\psi) = 0$. The converse direction is immediate. So suppose $\theta \wedge \phi \mid\!\sim_z \bot$ and so $Bel^z(\theta \wedge \phi) = 0$. Then, by Theorem 2.11 together with Proposition 2.9, we have $Bel^z(\phi \wedge \theta) = 0$ which gives $\phi \wedge \theta \mid\!\sim_z \bot$ as required.

To show the converse direction let $\mid\!\sim$ be a non-trivial f. t. natural consequence relation on L which satisfies CCM. Then, by Corollary 6.22 there exists a pre-ent G (possibly over a larger language than L) such that, for all $\psi, \chi \in SL$, $\psi \mid\!\sim \chi$ iff $\psi \mid\!\sim_G \chi$. Now, for this G and any $\theta, \phi \in SL$ we have that $Bel^G(\theta \wedge \phi) = 0$ implies $Bel^G(\phi \wedge \theta) = 0$. To see this suppose $Bel^G(\theta \wedge \phi) = 0$. Then we must have $\theta \wedge \phi \mid\!\sim_G \perp$ and so (since we may clearly assume $\perp \in L$) $\theta \wedge \phi \mid\!\sim \perp$. Since $\mid\!\sim$ satisfies CCM this gives $\phi \wedge \theta \mid\!\sim \perp$ and so $\phi \wedge \theta \mid\!\sim_G \perp$. Hence, following the above discussion, we conclude $Bel^G(\phi \wedge \theta) = 0$ as required. Hence we see that Bel^G, on its restriction to SL, satisfies the hypotheses of Theorem 4.1 and so we may apply that theorem (even though G is defined over a larger language than L – see the discussion just before Lemma 4.2) to assert the existence of an ent z (defined over a larger language than L) such that, for all $\psi \in SL$, $Bel^z(\psi) = Bel^G(\psi)$. Thus we have, for all $\theta, \phi \in SL$, $\theta \mid\!\sim_z \phi$ iff $\theta \mid\!\sim_G \phi$ iff $\theta \mid\!\sim \phi$ as required. This concludes the proof. □

We close this section with the observation that every rational consequence relation satisfies CCM, which is a special case of LLE. So the class of f. t. natural consequence relations which satisfy CCM still contains all rational consequence relations. Example 5.7 shows that the converse is false.

6.6 Conclusion

In the first half of this thesis we have reviewed the pre-ent and ent models of belief and examined the logic of pre-ents and ents. In particular we have given a characterisation of the relation $\theta \mid\!\ddot{\sim} \phi$ iff $Bel(\theta) \leq Bel(\phi)$ for all pre-ents. We have also shown the essential difference, at the level of their belief functions, between the classes of pre-ents and ents. In the second half of this work we have defined a new class of consequence relations – that of fully transitive natural consequence relations – which is more general than the class of rational consequence relations, and shown how this class may be characterised in terms of pre-ents. We have also characterised a class which lies between the two in terms of ents. Much of the material in this thesis has been of a syntactic nature. It remains to find a truly adequate semantics both for the relation $\mid\!\ddot{\sim}$ and for f. t. natural consequence relations. Another outstanding problem is to show whether or not the rule **(FT)** follows from the rules for natural consequence. We would also like to be able to give this rule in a simpler form than it stands at the moment, but for now we content ourselves with the results presented here.

Bibliography

[1] E. W. Adams. *The Logic of Conditionals*. D. Reidel, Dordrecht, Netherlands, 1975.

[2] R. Booth and J. B. Paris. A note on the rational closure of knowledge bases with both positive and negative knowledge. *Journal of Logic, Language and Information*, 7(2):165–190, 1998.

[3] R. T. Cox. Probability, frequency and reasonable expectation. *American Journal of Physics*, 14(1):1–13, 1946.

[4] R. Gladstone. The ent model of belief. Master's thesis, Manchester University, 1993.

[5] J. Goguen. The logic of inexact concepts. *Synthese*, 19:325–373, 1969.

[6] J. Kolodner. *Case-Based Reasoning*. Morgan Kaufmann Publishers, San Mateo, CA, 1993.

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

[8] I. Maung and J. B. Paris. A note on the infeasibility of some inference processes. *International Journal of Intelligent Systems*, 5:595–604, 1990.

[9] T. Havránek P. Hájek and R. Jiroušek. *Uncertain Information Processing in Expert Systems*. CRC Press, 1992.

[10] J. B. Paris. *The Uncertain Reasoner's Companion.* Cambridge University Press, 1994.

[11] J. B. Paris and A. Vencovská. Belief formation by constructing models. In L. Dorst et al, editor, *Reasoning with uncertainty in robotics, International workshop, RUR '95*, volume 1093 of *Springer Lecture Notes in Artificial Intelligence*, pages 171–186.

[12] J. B. Paris and A. Vencovská. A model of belief. *Artificial Intelligence*, 64:197–241, 1993.

[13] J. Pearl. Fusion, propogation and structuring in belief networks. *Artificial Intelligence*, 29:241–288, 1986.

[14] J. Pearl. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufmann, San Mateo, CA, 1988.

[15] A. Robinson. *Non-standard Analysis.* North-Holland, Amsterdam, 1966.

[16] D. Lehmann S. Kraus and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44(1-2):167–207, 1990.

[17] G. Shafer. *A Mathematical Theory of Evidence.* Princeton University Press, Princeton, NJ, 1976.

[18] F. Voorbraak. Reasoning with uncertainty in ai. In L. Dorst et al, editor, *Reasoning with uncertainty in robotics, International workshop, RUR '95*, volume 1093 of *Springer Lecture Notes in Artificial Intelligence*, pages 52–90, 1995.