HELSINKI UNIVERSITY OF TECHNOLOGY
DIGITAL SYSTEMS LABORATORY
Series **A:** Research Reports ISSN 0783–5396
No. **21**; February 1993 ISBN 951-22-1407-5

WEAKENED NEGATIVE INTROSPECTION IN AUTOEPISTEMIC REASONING

Tomi Janhunen

Digital Systems Laboratory
Department of Computer Science
Helsinki University of Technology
Otaniemi, FINLAND

Helsinki University of Technology
Department of Computer Science
Digital Systems Laboratory
Otaniemi, Otakaari 1
FIN–02150 ESPOO, FINLAND

HELSINKI UNIVERSITY OF TECHNOLOGY
DIGITAL SYSTEMS LABORATORY
Series **A:** Research Reports
No. **21**; February 1993

ISSN 0783–5396
ISBN 951-22-1407-5

Weakened Negative Introspection in Autoepistemic Reasoning

Tomi Janhunen

Abstract: A new weakened scheme for negative introspection in autoepistemic reasoning is proposed. The scheme is formulated proof-theoretically in terms of enumerations of formulae. The resulting sets of conclusions based on this negative introspection principle, cautious autoepistemic expansions, generalize Moore style stable expansions such that every set of premises has a cautious expansion. A finitary characterization is developed for cautious expansions. The elimination of weakly grounded cautious expansions yields a subclass of cautious expansions: the $\mathcal{N}$-grounded cautious expansions. The least $\mathcal{N}$-grounded cautious expansion is shown to be unique and the intersection of all $\mathcal{N}$-grounded cautious expansions. This is in contrast with Moore style stable expansions. The least $\mathcal{N}$-grounded cautious expansion can be constructed iteratively as the least fixed point of a monotonic operator. Such a construction is exceptional, since autoepistemic expansions are nonconstructive in general. It is also shown that cautious autoepistemic logic is decidable. A comparison is made with some other schemes of weakened negative introspection. The given set of examples demonstrates that the proposed negative introspection principle is stronger than, e.g., that of the well-founded semantics.

Keywords: autoepistemic logic, nonmonotonic reasoning, weakened negative introspection

Printing: TKK Monistamo; Otaniemi 1993

Helsinki University of Technology
Department of Computer Science
Digital Systems Laboratory
Otaniemi, Otakaari 1
FIN–02150 ESPOO, FINLAND

Phone: $\dfrac{90}{+358\text{–}0}$ 4511
Telex: 125 161 htkk fi
Telefax: +358–0–465 077
E-mail: lab@saturn.hut.fi

Contents

List of Tables

List of Figures

1 Introduction

Nonmonotonic reasoning, NMR, has been a self-standing research field in artificial intelligence since the late seventies. NMR evolved from the development of knowledge bases and expert systems. An observation was made: nonmonotonic reasoning mechanisms are necessary to draw inferences in such systems. NMR is different from classical reasoning which is monotonic: as the set of premises grows, so does the set of conclusions that can be inferred. In nonmonotonic inference, some conclusions may be retracted if additional information is acquired. A wide variety of nonmonotonic systems has been introduced during the last two decades. Some of these are formalizations of the nonmonotonic reasoning principles – others may be merely seen as applications of NMR in a particular problem domain. Ginsberg [1987] describes these systems briefly and provides a good source of references. In the following, the various applications of nonmonotonic reasoning are shortly introduced. Their nonmonotonic nature is of course of special interest.

Inheritance networks describe how properties of objects are inherited between objects. An inheritance hierarchy consists of a taxonomy of objects, typically forming a tree-like structure. The relation between the objects is called `is-a` describing, e.g., that "Tweety is a bird". Each object inherits properties from the upper objects in the hierarchy. For example, we could infer that Tweety flies, as it is a bird. This scheme may be generalized by allowing multiple inheritance, which turns the network structure to an arbitrary directed graph rather than a tree. Nonmonotonicity comes into play when the absence of properties is presented by means of `isn't-a` relation. These generalizations lead to a problem: contradictory properties may be inherited form different ancestors. It is hard to even informally decide what properties should be inherited [Touretzky *et al.*, 1987].

A truth maintenance system consists of a network of dependencies between statements asserted by a reasoning system [Doyle, 1979]. The truth maintenance systems keeps track of the statements that should be believed or disbelieved. Therefore, each statement is considered either as being "in" or "out" in the system, respectively. Statements are dependent on their antecedents. If a fact is to be believed, some specific antecedents have to be "in" and "out". Updating a statement may result in further updates of the dependent statements in the network. A truth maintenance system behaves nonmonotonically, i.e. a statement may become "out" if some of its antecedents that should be "out" becomes "in". This demonstrates how nonmonotonicity is closely related to drawing negative inferences.

In *logic programming* negation is presented as failure to prove. This is a procedural solution for deriving negative information and it gives logic programming its nonmonotonic nature [Lloyd, 1984]. To prove `not p` an attempt to prove `p` is carried out by SLDNF-resolution. If this fails, `not p` has been proved. Suppose that we insert new *clauses* to the program such that `p` becomes provable. This causes the proof of `not p` to fail. Moreover, some earlier conclusions may be withdrawn. Later on logic programs have been provided with various formal semantics, e.g. program completion [Clark, 1978], stable model semantics [Gelfond and Lifschitz, 1988] and well founded semantics [Van Gelder *et al.*, 1991]. The negation as failure is a weaker notion of negation than the classical negation. Practical implementations of logic programming such as PROLOG are general programming tools and in wide use nowadays. Sterling and Shapiro [1986] discuss logic programming methodology at large in their book "The Art of PROLOG". Logic programming has a close connection to *deductive databases* [Przymusinska and Przymusinski, 1990]. Deductive databases can store and manipulate deductive rules of reasoning (intensional data) in addition to storing individual facts (extensional data). Logic programs can be seen as deductive databases: the *unit clauses* of a logic program form the extensional data and the rest of the clauses the intensional data.

Information contained in a conventional database is considered the available knowledge of the users making queries. However, if some specific fact cannot be inferred from the information present in the database, a user is tempted to conclude that the fact does not hold. This corresponds to making a *closed world assumption* [Reiter, 1978] about the information presented in the database. Reiter gives an example of a *closed world database* – a travel agent having a database of commercially available flights. If some specific flight is not present in the database, customers are told that "there is no such flight". This answer would definitely change, if the flight in question is inserted to the database. Thus making a closed world assumption is a form of nonmonotonic inference.

Abduction is the converse term of deduction. The general setting in abduction is that the premises supporting the given conclusion are to be found out. An example of abductive systems is an expert system for medical diagnosis. The target is to determine the possible diseases given the symptoms of a patient. The more symptoms are provided, the fewer diseases are possible as diagnosis. This reveals that diagnosis is also subject to nonmonotonicity.

We end our survey of applications of NMR with a couple of examples from every day life. *Common sense reasoning* has a nonmonotonic nature as well. Many times our line of reasoning has the following characteristics: we draw inferences on the basis there is no evidence to the contrary. This kind of inferences are by no means classical consequences of the premises. Gelfond [1988] further analyzes the nonmonotonic behaviour of common sense reasoning. As an example, imagine that we are standing on a lake shore, staring at the horizon. If we do not see a single boat, we easily adopt the assumption that "there are no boats in the lake". But suppose that – all of a sudden – a yacht comes into sight. This forces us to change our minds, retracting the former conclusion. This example is also related to the analysis of *vision* [Minsky, 1981] – though only a part of the scene is seen, inferences are drawn about the rest of the scene. A similar process is ongoing when one hears a sentence in *speech*. Default meanings are assigned to the parts of the sentence before it has been completely said. These meanings may change when the sentence is finished and its context gets fixed.

As the applications and examples covered so far suggest, nonmonotonicity is a frequently appearing phenomenon in knowledge representation systems. But what is the motivation of using NMR instead of standard monotonic logics? Is NMR absolutely necessary from the philosophical point of view? Why not representing knowledge using classical logic? The answer is effectiveness. Consider, for instance, the possibility of representing a closed world database by means of classical logic. We should explicitly store all negative information in our database. Such a database would be of tremendous size, as typically negative information outnumbers the positive. This would be impractical as real computer systems have a limited storage capacity. Therefore, from the point of view of the applications of NMR, nonmonotonic inference mechanisms are useful. They provide finitary ways of representing negative information as part of the logical system.

Formal approaches have been taken in nonmonotonic reasoning, too. These formalizations of NMR are described briefly in Section 1.1. They have abstracted away from particular problem domains – they rather provide us general principles how to draw nonmonotonic inferences.

1.1 Formal Approaches to NMR

One of the early attempts to formalize nonmonotonic reasoning was Reiter's *closed world assumption* CWA [Reiter, 1978]. It simply allows one to introduce the negations of atomic formulae that cannot be derived by the monotonic inference mechanism. This principle was already mentioned in conjunction with closed world databases. CWA may be semantically characterized by means of minimal models [Etherington, 1988].

Reiter [1980] proposed another formalism, called *default logic*. A default theory is a pair of a standard first order theory and a set of defaults of the form

$$\frac{\alpha : \beta_1, \ldots, \beta_n}{\gamma}.$$

Each default consists of a prerequisite α, justifications $\beta_1, ..., \beta_n$ and a consequent γ. An *extension* is the set of conclusions of a default theory defined by a fixed point equation. It is possible that defaults interact with each other. As a consequence, a default theory may have multiple extensions – the application order of defaults becomes significant.

McCarthy's *circumscription* is based on forcing minimal interpretations of theories [McCarthy, 1980]. Circumscription can be seen as a generalization of CWA. McCarthy has provided a syntactic representation of this minimization principle in a form of axiom schemata. Several forms of circumscription has been suggested during the past decade. The major difference is that the target of minimization varies. *Predicate circumscription* provides means of closing off the world with respect to particular predicates at a time. *Formula circumscription* [McCarthy, 1986] allows additional predicates to vary during the minimization procedure. Lifschitz [1987] has developed *pointwise circumscription* which allows finer control of the minimization procedure. *Domain-closure assumptions* are also frequently made in nonmonotonic reasoning. Closing the domain means that we acknowledge only the existence of individuals that are implied by the available information. This kind of assumption can be understood as circumscribing the domain of the individuals in question.

This work is centered around one of the nonmonotonic formalisms: *autoepistemic logic* suggested by Moore [1985]. Autoepistemic logic was designed as a reconstruction of McDermott and Doyle's nonmonotonic logics [McDermott and Doyle, 1980; McDermott, 1982] avoiding some of their difficulties. We proceed now to a more detailed description of autoepistemic logic. Moore's autoepistemic logic models beliefs of an ideally rational agent, which is capable of perfect positive and negative introspection. Informally said, the agent introspects the logical consequences of its beliefs and decides what it believes and does not believe. A monadic modal operator L is introduced in the *autoepistemic language* $\mathcal{L}_{ae}$ to represent the beliefs of the agent.

Given the set of premises as the initial assumptions of the agent, the agent's possible states of belief are of interest. Moore [1985] carries out semantical considerations on nonmonotonic logics: he requires these sets of beliefs of the agent to be both *semantically complete* and *sound* with respect to the set of premises. The semantical completeness corresponds to the ideality of the agent: it must believe all logical consequences of its beliefs. The soundness requirement forces the rationality of the agent, i.e. the beliefs of the agent are the logical consequences of the set of premises Σ. These semantical considerations lead Moore to define the agent's set of beliefs as follows:

$$\Delta = Cn(\Sigma \cup L\Delta \cup \neg L\overline{\Delta}). \tag{1}$$

The solutions Δ of this fixed-point equation are called stable (autoepistemic) expansions of the set of premises Σ. The agent believes the formulae in Δ, i.e. $L\Delta$ denotes the positive beliefs of the agent. Similarly, the agent does not believe the formulae in the complement of Δ with respect to $\mathcal{L}_{ae}$. The set $\neg L\overline{\Delta}$ therefore forms the negative beliefs of the agent. By the above equation, a stable expansion Δ is the set of logical consequences of the premises Σ and the beliefs of the agent $L\Delta \cup \neg L\overline{\Delta}$. As a consequence, stable expansions are self-referential.

The definition of stable expansions (1) is nonconstructive. It gives a specific condition on a whole set of formulae rather than a way of constructing such a set. Furthermore, these sets are infinite sets of formulae. This leads one to wonder whether there are any solutions to (1) in the first place. The answer to this question is positive in general – depending on the set of premises there may be one solution or even multiple solutions. However, there are sets of premises for which no expansion exists. This leads to some problems: if there are multiple solutions which one to choose? If there are none, what should the agent believe? We shall return to these questions later on.

There are concrete similarities between the various formalizations of NMR. A lot of translations have been discovered between the different forms of NMR. The connection between default logic and autoepistemic logic is addressed in [Konolige, 1988] and [Marek and Truszczyński, 1989]. Konolige presented the translation below. We provide it here as an example of the various translations. Defaults of the default theory are translated into autoepistemic logic as implications of the form

$$L\alpha \wedge \neg L\neg\beta_1 \wedge \ldots \wedge \neg L\neg\beta_n \rightarrow \gamma$$

where α is the prerequisites, $\beta_1, \ldots, \beta_n$ are the justifications and γ is the consequent of the default. In order to derive γ the agent has to believe the prerequisite and disbelieve the negations of the justifications. Gottlob [1992b] translates default logic into original autoepistemic logic. The existence of these translations suggest that default logic and autoepistemic logic are capable of carrying out very similar inferences. However, we have preferred autoepistemic logic for the sake of its uniform syntax for knowledge representation (autoepistemic formulae like the one above). In default logic, defaults form a separate syntactic category. Marek and Truszczyński [1990a; 1991] further introduce a variant of autoepistemic logic, the nonmonotonic modal logic $\mathcal{N}$, designed for default reasoning.

The stable model semantics of logic programs is in close relationship with autoepistemic reasoning [Gelfond and Lifschitz, 1988]. In fact, the various semantics of logic programs can be characterized by means of autoepistemic expansions. Marek and Subrahmanian have studied the interconnection of different semantics for general logic programs [Marek and Subrahmanian, 1992]. Gelfond and Lifschitz [1990] have introduced semantics for logic programs with classical negation. This semantics may be translated through default logic into autoepistemic logic. Przymusinski has extensively studied the relationship between the well-founded semantics of logic programs and 3-valued autoepistemic logic [Przymusinski, 1989; Przymusinski, 1991b; Przymusinski, 1991a].

Inheritance reasoning can be captured in autoepistemic logic [Gelfond and Przymusinska, 1990]. Kakas and Mancarella [1990] embed abduction into autoepistemic logic. Elkan [1990] gives a translation from truth maintenance systems. This interconnection has also been studied in [Reinfrank *et al.*, 1989] and [Fujiwara and S., 1989].

Since these concrete translations exist, the following conclusion is immediate: autoepistemic logic can model reasoning in other nonmonotonic formalisms. It seems that autoepistemic logic is a very general approach to NMR. This is why we have chosen it as our basic tool. The existence of translations allows us to concentrate on studying one nonmonotonic formalisms. The translations introduced above provide means of transferring new constructs and results to other nonmonotonic formalisms.

The development of nonmonotonic formalisms has been guided by benchmark problems presented in the literature. [Lifschitz, 1988] provides a summary of these problems with their intuitive solutions. There are dissimilarities in the expressiveness of the nonmonotonic formalisms – in the capability of representing specific problems of the domain. The formal properties of nonmonotonic consequence relations have also been of interest. Kraus, Lehman and Magidor [1990] present a list of general rules defining some structural properties of nonmonotonic consequence relations. Dix [1991] classifies the various semantics of logic programs by means of Kraus, Lehman and Magidor's rules.

As we pointed out earlier, stable expansions are not unique. In some cases there are several expansions of the given set of premises. There are sets of premises for which no expansion exists. Similar behaviour related to the uniqueness and existence of expansions arises in connection with other formalizations of NMR. In autoepistemic logic, the set of premises $\{\neg Lp \rightarrow q, \neg Lq \rightarrow p\}$ has two stable expansions. In one of them, p is believed and q is not believed and in the other p is not believed and q is believed. What should the correct conclusions be in situations like this? There are two basic approaches to this. In *cautious* reasoning, conclusions (or beliefs) that hold in every stable expansion are the correct ones. In our example, the agent believes none of p and q. To the contrary, the second approach selects one of the stable expansions as basis of the correct conclusions. This is called *brave* reasoning. But how should the expansion be selected? There are two choices in the above example. The choice between the alternative expansions seems to require some additional information, i.e. some domain specific knowledge about the propositions p and q.

Let us now proceed to the case, when there are no expansions at all. The set of premises $\{\neg Lp \rightarrow p\}$ does not have any stable expansions. If the agent believes p, it cannot infer p and should therefore disbelieve it. Similarly, if the agent does not believe p it can infer p. Thus it should believe p. This demonstrates how a stable state of beliefs cannot simply be reached. A solution to this anomalous situation is to consider the status of p as unknown. This principle is called *weakened negative introspection*. A new weakened negative introspection scheme is suggested in this work.

Nonmonotonic reasoning is subject to additional computational complexity, if compared to similar monotonic reasoning. This has caused objections to NMR for the sake of practical implementations. In fact, complexity results presented in the literature have been rather disappointing in the recent years. Reasoning in default logic is NP-hard even in very limited cases [Kautz and Selman, 1989; Stillman, 1990]. Later the decision problems of propositional default logics were shown to be complete problems on the second level of the polynomial time hierarchy [Gottlob, 1992a; Stillman, 1992].

Unfortunately, autoepistemic logic has a similar computational complexity. The decision problems of autoepistemic logic were shown to reside on the second level of polynomial time hierarchy by Niemelä [1990b]. This gave an upper bound for the complexity of the autoepistemic reasoning. Gottlob [1992a] provided the corresponding completeness results for both autoepistemic logic and Marek and Truszczyński's nonmonotonic modal logic $\mathcal{N}$.

The computational complexity of NMR may be understood to origin from two orthogonal sources. The one is determining the negative beliefs of the agent (or which defaults to apply) and the other is the decision problem of the underlying monotonic logic. These two problems are intermixed with each other in nonmonotonic formalisms. This informally justifies the position of the nonmonotonic formalisms on the second level of the polynomial time hierarchy [Gottlob, 1992a].

Despite the disappointing results concerning the computational complexity of nonmonotonic reasoning, efforts have been made to find tractable forms of nonmonotonic reasoning. Niemelä and Rintanen [1992] have demonstrated the impact of *stratification* on the complexity of nonmonotonic reasoning. They have shown that autoepistemic reasoning is tractable if the set of premises is stratified. A stratified set of premises may be partitioned to strata, i.e. to layers, such that – roughly speaking – propositions on a level depend on the previous levels only. This has a favorable effect on the computational complexity of determining the stable expansion. The introspection can be carried out in layers. The beliefs and disbeliefs determined by introspection at one layer are not affected by the introspection of the subsequent layers. As a result, a stratified set of premises has a unique stable expansion.

Stratification is an example of setting up syntactic restrictions on the logical language in order to reduce the computational complexity of the inference involved. Another example is logic programming where programs are composed of program clauses having a restricted syntax. Syntactic restrictions have typically an effect on the expressibility of the logical language. The more restrictions we impose on our logical language, the less expressive it is. At certain point these restrictions may become an obstacle, making the restricted formalism of no use in representing knowledge of interest. Generally speaking, syntactic restrictions can be seen as a tradeoff between expressibility of the formalism and the computational complexity of the required inference. Practical implementations of NMR call for a compromise that guarantees both sufficient capability of representing knowledge and acceptable computational complexity.

1.2 Research Problem

This work is concentrated on the study of weakened negative introspection in autoepistemic reasoning. The objective is that every set of premises has at least one autoepistemic expansion. The introspection of the set of premises $\{\neg Lp \rightarrow p\}$ demonstrated the strength of the negative introspection in the original autoepistemic logic – preventing the agent from a stable state of beliefs. Recall how weakened negative introspection handles this case: instead of disbelieving p, the belief status of p is left undetermined, though the agent cannot infer p. In light of this example, we expect that the existence of autoepistemic expansions can be ensured by means of weakened negative introspection.

The nonconstructive nature of autoepistemic logic calls also for a solution. It seems possible that weakened negative introspection provides one. We would like to find a general iterative construction for (possibly modified) stable autoepistemic expansions and therefore remove their nonconstructive nature. This seems a rather difficult task at the first glance, because the results of the introspection are often dependent on themselves. If such an iterative construction can be found, it offers a new basis for the development of algorithms to implement autoepistemic reasoning. Furthermore, it perhaps gives a better understanding of the computational aspects of autoepistemic reasoning. This is essential if new computationally feasible forms of autoepistemic reasoning are to be found.

The weakened negative introspection involves giving up the complete introspection capability of the agent to some extent. The weaker negative introspection scheme we adopt, the less accurate expansions we get. This has to be taken into account when designing new principles for negative introspection. For the sake of the accuracy of expansions, we should find as strong a principle as possible though we are to weaken the negative introspection with respect to Moore's original approach. Hence there is a tradeoff between the accuracy of autoepistemic expansions and the strength of the applied negative introspection principle. It seems reasonable that there is a similar connection between the computational complexity of autoepistemic reasoning and the strength of negative introspection.

The weakened negative introspection has been approached elsewhere in the works of Przymusinski [1991a] and Kalinski [1990]. These works have been centered around the formal semantics of logic programs. They both handle general logic programs that have been translated into autoepistemic logic. To the contrary, we intend to achieve a more general approach in this work, extending Moore's original autoepistemic logic. Both Przymusinski and Kalinski have provided iterative constructions for the stable expansions of restricted sets of premises. This suggest that iterativity really has a connection to weakened negative introspection.

1.3 Structure of the Work

We devote Section 2 to the formal development of CAEL. The autoepistemic language $\mathcal{L}_{ae}$ and the basic operations and notation for it are introduced. Then we return to the definition of stable autoepistemic expansions and further give definitions of Marek and Truszczyński's nonmonotonic modal logic $\mathcal{N}$ and Niemelä's enumeration-based expansions also explaining their close relationship.

In Section 2.2 we are ready to give the definition of cautious autoepistemic expansions. Here we also introduce some belief terminology and discuss the design criteria of *cautious autoepistemic logic* CAEL. The finitary characterization of cautious expansions is the subject of Section 2.3. Then the finitary characterization, i.e. full sets are defined. The full sets are shown to be in one-to-one correspondence with cautious expansions. In Section 2.4 some examples of cautious expansions are given.

We study groundedness properties of cautious expansions in Section 3. We use so called $\mathcal{N}$-groundedness condition, which is adopted from Marek and Truszczyński's nonmonotonic modal logic $\mathcal{N}$. A natural partial order $\sqsubseteq$ on cautious expansions is defined. In Section 3.1 we show that full sets are capable of capturing the notion of the $\mathcal{N}$-groundedness. We return to the partial order of the expansions in Section 3.2 and show that full sets can be used for testing whether two $\mathcal{N}$-grounded expansions are related by $\sqsubseteq$ or not. The same applies thereafter to determining the least $\mathcal{N}$-grounded cautious expansion.

In Section 3.3 it is shown that the least $\mathcal{N}$-grounded cautious expansion does always exist. In addition, it is the intersection of all $\mathcal{N}$-grounded cautious expansions. Therefore it fits well the framework of cautious autoepistemic reasoning. Furthermore, the proofs are based an iterative construction of the least $\mathcal{N}$-grounded cautious expansion. The decidability of cautious autoepistemic logic is addressed in Section 3.4. We provide decision procedures for the recursive consequence relations. The major sources of computational complexity are identified. Section 3.5 points out some of the differences to the other formalizations of weakened negative introspection in autoepistemic reasoning.

The conclusions on the results of this work are presented in Section 4. The further research work is also sketched.

2 Cautious Autoepistemic Logic

2.1 Basic Definitions

This section is devoted to the formal development of cautious autoepistemic logic. We begin by introducing the language of the logic, some notation and operations on sets of formulae. For the sake of simplicity, we shall restrict ourselves to the propositional case. Let $\mathcal{P}$ be the set of atomic propositions. The propositional language $\mathcal{L}$ is defined recursively as follows:

Definition 2.1 (Propositional language)
The propositional language $\mathcal{L}$ is the least set of formulae satisfying:
If $\phi \in \mathcal{P}$, then $\phi \in \mathcal{L}$.
If $\phi \in \mathcal{L}$, then $\neg\phi \in \mathcal{L}$.
If $\phi \in \mathcal{L}$ and $\psi \in \mathcal{L}$, then $\phi \wedge \psi \in \mathcal{L}$.

The leastness requirement is needed to ensure that $\mathcal{L}$ contains *exactly* the formulae that can be derived from $\mathcal{P}$ by the above rules. Apparently, we have chosen the classical negation $\neg$ and conjunction $\wedge$ as basic connectives of the logical language. The other connectives may be defined by means of these [Epstein, 1990]. So we let the disjunction $p \vee q$ and implication $p \rightarrow q$ denote $\neg(\neg p \wedge \neg q)$ and $\neg(p \wedge \neg q)$, respectively. Parenthesis are used whenever necessary for punctuation.

The propositional autoepistemic language $\mathcal{L}_{ae}$ is defined in the same manner. The difference is that belief formulae of the form $L\phi$ are included in the language. The formula $L\phi$ is read "ϕ is believed" or "the agent believes ϕ". The definition below allows nested belief operators to appear in the language. Furthermore, the belief operator may be applied to arbitrarily complex formulae. Some versions of autoepistemic logic presented in literature cover only a fragment of $\mathcal{L}_{ae}$, i.e. are restricted in syntax (c.f. [Przymusinski, 1991a]). However, we are to develop a general approach and therefore use the full autoepistemic language $\mathcal{L}_{ae}$.

Definition 2.2 (Propositional autoepistemic language)
The propositional autoepistemic language $\mathcal{L}_{ae}$ is the least set of formulae satisfying:
If $\phi \in \mathcal{P}$, then $\phi \in \mathcal{L}_{ae}$.
If $\phi \in \mathcal{L}_{ae}$, then $\neg\phi \in \mathcal{L}_{ae}$.
If $\phi \in \mathcal{L}_{ae}$ and $\psi \in \mathcal{L}_{ae}$, then $\phi \wedge \psi \in \mathcal{L}_{ae}$.
If $\phi \in \mathcal{L}_{ae}$, then $L\phi \in \mathcal{L}_{ae}$.

We adopt our basic operations and notation from the works of Niemelä (see the references). The formulae that begin with L are called *quasi-atomic* formulae. *Quasi-literals* are either quasi-atomic formulae or their negations. We shall also use terms *belief literal* and *belief formula* interchangeably. We define the function q-atom as follows: q-atom$(L\chi)$ = q-atom$(\neg L\chi) = L\chi$. Thus q-atom converts any quasi-literal to a quasi-atom by removing the possible negation sign. Subformulae of the form $L\phi$ are defined recursively as follows:

Definition 2.3 (L-subformulae)
1. $Sf^L(\phi) = \emptyset,$ *if* $\phi \in \mathcal{P}$
2. $Sf^L(\phi \wedge \psi) = Sf^L(\phi) \cup Sf^L(\psi)$
3. $Sf^L(\neg\phi) = Sf^L(\phi)$
4. $Sf^L(L\phi) = \{L\phi\} \cup Sf^L(\phi)$

The quasi-subformulae $Sf^{qL}(\phi)$ of the autoepistemic formula ϕ have a similar definition with the exception that $Sf^{qL}(L\phi) = \{L\phi\}$. Operations Sf^{qL}, Sf^L and q-atom are naturally extended to take sets of formulae as their arguments.

Definition 2.4 (Belief depth of autoepistemic formulae)
1. $BD(\phi) = 0,$ *if* $\phi \in \mathcal{P}$
2. $BD(\phi \wedge \psi) = max(BD(\phi), BD(\psi))$
3. $BD(\neg\phi) = BD(\phi)$
4. $BD(L\phi) = BD(\phi) + 1$

The belief depth of an autoepistemic formula $BD(\phi)$ is the maximal nesting level of belief operators appearing in the formula. Several proofs in this work use induction on the belief depth of formulae. We need a repertoire of additional operations on arbitrary sets of formulae:

Definition 2.5
For a set of formulae Δ *of* $\mathcal{L}_{ae}$*:*
1. $\neg\Delta = \{\neg\phi \mid \phi \in \Delta\}$
2. $\overline{\Delta} = \{\phi \mid \phi \in \mathcal{L}_{ae} - \Delta\}$
3. $L\Delta = \{L\phi \mid \phi \in \Delta\}$

Next we define Π, $[\_]^+$ and $[\_]^-$ each of which perform a selection from a set of quasi-literals.

Definition 2.6 (Selection)
Let Λ *be a set of quasi-atomic formulae and* Γ *any set of quasi-literals:*
1. $\Pi_\Lambda(\Gamma) = \{\chi \in \Gamma \mid q\text{-}atom(\chi) \in \Lambda\}$
2. $[\Gamma]^+ = \{L\chi \mid L\chi \in \Gamma\}$
3. $[\Gamma]^- = \{\neg L\chi \mid \neg L\chi \in \Gamma\}$.

We say that $\Pi_\Lambda(\Gamma)$ is the reduction of Γ with respect to Λ. Clearly, $\Pi_\Lambda(\Gamma) \subseteq \Gamma$, for all sets of quasi-atomic formulae Λ.

Having defined the two languages $\mathcal{L}$ and $\mathcal{L}_{ae}$, we turn our attention to the semantics. The standard propositional interpretations, models and the semantical consequence relation $\models$ define the semantics of propositional formulae [Epstein, 1990]. $\Sigma \models \phi$ means that ϕ is true in every model of Σ and therefore ϕ is a (semantical) consequence of Σ. The propositional calculus has a complete and sound proof theory. This means that $\models$ has a syntactic counterpart – the derivability relation $\vdash$ which is in one-to-one correspondence with the semantical consequence relation.

A slight extension is needed for the formulae in $\mathcal{L}_{ae}$. The question is how to define the truth of modal formulae. We introduce $\models_{ae}$ which is an extension of $\models$ in the sense that it treats all quasi-atomic formulae like atomic propositions. Propositional interpretations and models are extended by the same principle to autoepistemic interpretations and models. We use $Cn(\Sigma)$ to denote $\{\phi \mid \Sigma \models_{ae} \phi\}$. In the introduction, we mentioned that classical logic is monotonic. This holds, of course, for propositional case, i.e. the consequence relation $\models_{ae}$ is monotonic.

We now return to the definition of stable autoepistemic expansions [Moore, 1985]. A stable expansion Δ representing the set of beliefs of the agent contains exactly all the logical consequences of the premises Σ, the positive beliefs $L\Delta$ and negative beliefs $\neg L\overline{\Delta}$ of the agent:

Definition 2.7 (Stable expansion [Moore, 1985])
A set $\Delta \subseteq \mathcal{L}_{ae}$ is a stable autoepistemic expansion of the set of premises Σ if and only if

$$\Delta = AE_{\Sigma}(\Delta) = Cn(\Sigma \cup L\Delta \cup \neg L\overline{\Delta}). \tag{2}$$

Let us consider the set of premises $\Sigma = \{\neg Lp \rightarrow r\}$ as an example. As p is not a logical consequence of Σ, the agent does not believe it, i.e. $\neg Lp$ belongs to the stable expansion. Then r is believed by the agent as a logical consequence of Σ and $\neg Lp$. So far we have inferred that the agent should disbelieve p and believe r. But what is the connection of these two beliefs to a stable expansion – an infinite set of formulae? How can we verify that (2) is fulfilled in the first place? Unfortunately, answering to this kind of questions has to deferred until Section 2.3. In fact the two formulae $\neg Lp$ and Lr belong to the only stable expansion of Σ. To convince ourselves about the stability we extend the set of premises by these formulae – as it is done in (2) – and check that the agent has still the same beliefs.

Stable expansions are subject to nonmonotonic behaviour. We give here an informal explanation why this is the case. Suppose that Δ is a stable expansion of the set of premises Σ. If we extend the set of premises Σ to Σ', it is possible that (2) ceases to hold. We have to find another stable expansion Δ' satisfying (2) with Σ replaced with Σ'. The stable expansion Δ' is likely to contain some new formulae that do not belong to Δ, as $\models_{ae}$ is a monotonic consequence relation. It follows that the corresponding negative beliefs are removed from $\neg L\overline{\Delta'}$. As a consequence, some conclusions based on negative beliefs may have to be retracted and some formulae are dropped from Δ'. This affects again $\neg L\overline{\Delta'}$ and the process goes on. This demonstrates how stable expansions are difficult to reason about due to the self-reference.

As a concrete example of the nonmonotonic behaviour, reconsider the set of premises $\Sigma = \{\neg Lp \rightarrow r\}$ given above. If we extend Σ to $\Sigma' = \{\neg Lp \rightarrow r, p\}$, the proposition p becomes a logical consequence of Σ' and $\neg Lp$ no longer belongs to the stable expansion of Σ'. Thus the previous conclusion r is retracted and the agent disbelieves r.

Then consider the set of premises $\Sigma = \{Lp \rightarrow p\}$. Given Σ as the set of premises, the agent may end up in two different stable states of belief, i.e. there are two stable expansions of Σ. In the first one, the agent does not believe p. Therefore $\neg Lp$ belongs to the stable expansion. The stable state of beliefs seems possible as p is not a logical consequence of Σ and $\neg Lp$. The other stable expansion contains the formula Lp. This makes p a logical consequence of Σ resulting again in a stable state of beliefs.

The latter stable expansion has an unwanted property. The agent believes p, because it believes p. In other words, the belief in p is self-supporting. Konolige [1988] calls this kind of expansions weakly grounded in the set of premises. This type of expansions exist because of the direct positive introspection $L\Delta$ in Equation (2). McDermott and Doyle [1980; 1982] have proposed a scheme for nonmonotonic modal logics where a modal logic S is used for positive introspection:

$$\Delta = Cn_S(\Sigma \cup \neg L\overline{\Delta}).$$

Marek, Shvarts and Truszczyński [1991] have studied these logics in greater depth. Marek and Truszczyński [1990b] use the *modal logic of necessitation* $\mathcal{N}$ for realizing the positive introspection capability of the agent. The modal logic of necessitation is a nonstandard modal logic. It is simply classical propositional calculus equipped with the *necessitation rule* "$\phi \vdash L\phi$" of standard modal logics. We let $\Sigma \vdash_{\mathcal{N}} \phi$ denote that ϕ is derivable from Σ by the propositional inference and the necessitation rule. We denote the set of conclusions $\{\phi \mid \Sigma \vdash_{\mathcal{N}} \phi\}$ by $Cn_{\mathcal{N}}(\Sigma)$.

Definition 2.8 yields a subset of stable expansions called $\mathcal{N}$-expansions. In an $\mathcal{N}$-expansion positive beliefs are derivable from negative ones only by means of the necessitation rule. This is why *some* weakly grounded expansions are ruled out, but not all. As a counterexample, the set of premises $\Sigma = \{\neg L\neg Lp \rightarrow p\}$ has still an $\mathcal{N}$-expansion which contains a circularly supported belief p.

Definition 2.8 ($\mathcal{N}$-expansion [Marek and Truszczyński, 1990b])
A set $\Delta \subseteq \mathcal{L}_{ae}$ is an $\mathcal{N}$-expansion of the set of premises Σ if and only if

$$\Delta = IE_{\Sigma}(\Delta) = Cn_{\mathcal{N}}(\Sigma \cup \neg L\overline{\Delta}). \tag{3}$$

The derivability relation $\vdash_{\mathcal{N}}$ can be formalized as follows [Gottlob, 1992a]. The relation $\Sigma \vdash_{\mathcal{N}} \phi$ holds if and only if there is an *L-proof* $L\phi_1, ..., L\phi_n$ of ϕ from Σ such that for each $1 \leq i \leq n$, $\Sigma \cup \{L\phi_j \mid 1 \leq j < i\} \models_{ae} \phi_i$ and $\Sigma \cup \{L\phi_1, ..., L\phi_n\} \models_{ae} \phi$. Intuitively, the proof of ϕ involves application of the necessitation rule n times. The definition of L-proofs also implies that $\Sigma \vdash_{\mathcal{N}} \phi_i$ for each $1 \leq i \leq n$. A formula ϕ that is $\mathcal{N}$-derivable has several L-proofs (in fact infinitely many). Gottlob provides a result that minimal L-proofs of ϕ contain only members of $Sf^L(\Sigma \cup \{\phi\})$. The derivability relation $\vdash_{\mathcal{N}}$ is monotonic, as $\models_{ae}$ is.

Niemelä [1991] has introduced enumeration-based expansions. The objective is to incorporate the notion of a proof to autoepistemic reasoning. The intuitive idea is to carry out introspection in the order determined by an enumeration ($\mathcal{L}_{ae}$ is an enumerable language). Then the belief status of each formula is considered only once. Enumerations serve as the *autoepistemic proofs* of the formulae. The set Γ below is called the *enumeration base*. The set of quasi-literals $B^{\varepsilon}(\Sigma)$ is constructed by introspecting the formulae in Γ in the order determined by the enumeration ε.

Definition 2.9 (Enumeration-Based Expansion [Niemelä, 1991])
Let Σ be a set of formulae of $\mathcal{L}_{ae}$. Let $\varepsilon = \psi_1, \psi_2, ...$ be an enumeration of a subset Γ of $\mathcal{L}_{ae}$ and let $|\varepsilon|$ denote its length. Define $B^{\varepsilon}_0(\Sigma) = \emptyset$ and $B^{\varepsilon}_{i+1}(\Sigma)$ for $i = 0, 1, ..., |\varepsilon|$ as follows:

$$B^{\varepsilon}_{i+1}(\Sigma) = \begin{cases} B^{\varepsilon}_i(\Sigma) \cup \{L\psi_{i+1}\}, & \textit{if } \Sigma \cup B^{\varepsilon}_i(\Sigma) \models_{ae} \psi_{i+1} \\ B^{\varepsilon}_i(\Sigma) \cup \{\neg L\psi_{i+1}\}, & \textit{if } \Sigma \cup B^{\varepsilon}_i(\Sigma) \not\models_{ae} \psi_{i+1} \end{cases}$$

Finally let

$$B^{\varepsilon}(\Sigma) = \begin{cases} B^{\varepsilon}_{|\varepsilon|}(\Sigma), & \textit{if } |\varepsilon| \textit{ is finite} \\ \bigcup_{i=0}^{\infty} B^{\varepsilon}_i(\Sigma), & \textit{otherwise} \end{cases}$$

and the enumeration-based expansion

$$EE^{\varepsilon}(\Sigma) = \{\phi \mid \Sigma \cup B^{\varepsilon}(\Sigma) \models_{ae} \phi\}.$$

An enumeration ε is *L-hierarchic* if for each element ψ of ε it holds that each ϕ such that $L\phi \in Sf^L(\psi)$ appear before ψ in ε. The result of introspecting a formula ψ may depend on the results of introspecting its L-subformulae. If L-hierarchic enumerations are used in conjunction with enumeration-based expansions, the L-subformulae of each formula are introspected before introspecting the formula itself. This results in expansions that are tightly grounded in the set of premises. This is why Niemelä introduced the concept of L-hierarchicness.

Niemelä [1991] provides also a special acceptability condition on enumerations. An enumeration ε is Σ-*acceptable* if there is no i and no formula ϕ such that $\neg L\phi \in B_i^{\varepsilon}(\Sigma)$ but $\Sigma \cup B_i^{\varepsilon}(\Sigma) \models_{ae} \phi$. The Σ-acceptability condition is needed to rule out introspection sequences in which a decision to disbelieve a formula ϕ is made, but ϕ is derivable from $\Sigma \cup B^{\varepsilon}(\Sigma)$, i.e. ϕ should be believed. The sets of beliefs $EE^{\varepsilon}(\Sigma)$ generated by a Σ-acceptable enumerations of $\mathcal{L}_{ae}$ satisfy the fixed point equation (2), i.e. they are stable expansions. Furthermore, Niemelä has shown that in the propositional case the expansion $EE^{\varepsilon}(\Sigma)$ generated by a Σ-acceptable enumeration is an $\mathcal{N}$-expansion of Definition 2.8.

We shall also use the term Σ-acceptability for sets of quasi-literals. We say that a set of quasi-literals Γ is Σ-acceptable, if $L\chi \in \Gamma$ implies $\Sigma \cup \Gamma \models_{ae} \chi$ and $\neg L\chi \in \Gamma$ implies $\Sigma \cup \Gamma \not\models_{ae} \chi$. Thus we can say that an enumeration ε is Σ-acceptable if and only if $B^{\varepsilon}(\Sigma)$ is Σ-acceptable.

We denote a *concatenation* of two enumerations ε_1 and ε_2 by $\varepsilon_1 \cdot \varepsilon_2$ (assuming that ε_1 is finite) and an *empty enumeration* of formulae by λ.

2.2 Design of Cautious Autoepistemic Logic

We begin this section by examining an example. We shall study the introspection of the set of premises $\Sigma = \{\neg Lp \rightarrow p, \neg Lq \rightarrow r\}$. The three propositions p, q and r are not logical consequences of Σ, so at the first glance the agent should not believe in any of them. However, making such a decision to disbelieve them has some unwanted consequences. If the set of premises is extended by the formulae $\neg Lp$, $\neg Lq$ and $\neg Lr$ (as members of the expansion), p and r can be inferred and they should be believed instead of being disbelieved. This drives the agent to a state of beliefs that is not Σ-acceptable. Somehow the agent has been too brave when adopting such negative beliefs, i.e. the negative introspection has been too strong. Therefore, we should have a mechanism for examining the consequences of the negative introspection *beforehand* in order to prevent this kind of behaviour.

In the definition of cautious autoepistemic expansions below, enumeration-based introspection is utilized for checking that the logical consequences of the premises and the beliefs of the agent are not contradictory with the beliefs themselves, i.e. the beliefs of the agent are Σ-acceptable. We shall call this kind of expansions *cautious expansions* for convenience. The check-up by enumeration-based introspection is sufficient to prevent the above described behaviour of the agent. As we shall later see, stable expansions are also cautious expansions and therefore this definition generalizes Moore's original autoepistemic logic.

Definition 2.10 (Cautious autoepistemic expansion)
A pair of subsets $\langle\Delta^+,\Delta^-\rangle$ of $\mathcal{L}_{ae}$ is a cautious autoepistemic expansion of the set of premises Σ if and only if

$$\langle\Delta^+,\Delta^-\rangle = CAE_\Sigma(\Delta^+,\Delta^-) \tag{4}$$

where

1. $CAE_\Sigma(\Delta^+,\Delta^-) = \langle IP^+_\Sigma(\Delta^+,\Delta^-), IP^-_\Sigma(\Delta^+,\Delta^-)\rangle$,
2. $IP^+_\Sigma(\Delta^+,\Delta^-) = Cn(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$ *and*
3. $IP^-_\Sigma(\Delta^+,\Delta^-) = \{\phi \mid$ *there is no enumeration* ε *of* $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ *such that* $\phi \in EE^\varepsilon(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)\}$.

Let us now make some terminology involving cautious expansions precise. The set Δ^+ contains the formulae that are *believed* by the agent. The agent *disbelieves* every formula in Δ^-. Notice how a distinction is made between "not believing" and "disbelieving" a formula. Depending on the set of premises Σ, there may exist *undetermined* formulae that are neither believed or disbelieved by the agent. The term *belief status* of ϕ refers to its membership in one of the sets Δ^+, Δ^- or $\overline{\Delta^+ \cup \Delta^-}$.

It seems intuitive for the agent to believe every logical consequence of the premises and the results of the introspection. This is why the positive introspection is exactly the same as in the original autoepistemic logic. The negative introspection capability of the agent has been weakened, but in a very delicate way. We have tried to follow the principle that as strong negative introspection as possible should be made provided the set of beliefs of the agent remains Σ-acceptable.

The enumerations of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ are employed in negative introspection for checking that whatever the results of the introspection are, the formula under introspection ϕ will not be derivable. If this is the case, ϕ may be freely disbelieved. Otherwise a contradiction may arise with the results of the positive introspection. If there is an enumeration ε of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ such that $\phi \in EE^\varepsilon(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$, we say that ε *refutes* the formula ϕ and ε is a *refuting enumeration*. Thus the formulae in Δ^- are in this respect *irrefutable*. This scheme for negative introspection is proof-theoretical, i.e. an irrefutable formula ϕ does not have an autoepistemic proof from the set of premises and the beliefs of the agent.

Let us now informally examine, how the example at the beginning of this section is handled. The negative introspection uses enumerations that contain among others the formulae p, q and r. The introspection of p yields $\neg Lp$ in all enumerations. However, extending the set of premises Σ by $\neg Lp$ makes p derivable. Thus any set of beliefs containing $\neg Lp$ would not be Σ-acceptable. Therefore, the belief status of p is left undetermined by the agent. The introspection of q yields $\neg Lq$, too. Extending the set of premises by $\neg Lq$ does not make q derivable. Therefore, the agent disbelieves q. This affects in turn the status of r, which is now believed by the agent as a logical consequence of the premises and the beliefs of the agent.

The strength of the negative introspection affects the accuracy of the expansions. We argue that the relatively strong negative introspection has a favorable effect on the accuracy of the cautious expansions. In fact, we disbelieve every irrefutable formula. Proposition 2.11 ensures us that cautious autoepistemic expansions are well-defined.

Proposition 2.11 (Mutual exclusion)
Let $\langle\Delta^+, \Delta^-\rangle$ be a cautious autoepistemic expansion of the set of premises Σ. Then Δ^+ and Δ^- are disjoint.

Proof. We show that

$$\phi \in \Delta^+ = IP^+_\Sigma(\Delta^+, \Delta^-) \text{ implies } \phi \notin \Delta^- = IP^-_\Sigma(\Delta^+, \Delta^-). \tag{5}$$

So let $\phi \in IP^+_\Sigma(\Delta^+, \Delta^-)$ hold. By definition 2.10 ϕ belongs to the set of propositional consequences $Cn(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$, i.e. $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \models_{ae} \phi$. Then there is an enumeration ε of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ such that ϕ belongs to the expansion $EE^\varepsilon(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$ or equivalently,

$$\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \cup B^\varepsilon(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-) \models_{ae} \phi,$$

because $\models_{ae}$ is a monotonic consequence relation. By Definition 2.10 $\phi \notin IP^-_\Sigma(\Delta^+, \Delta^-)$. By (5) also

$$\phi \in \Delta^- = IP^-_\Sigma(\Delta^+, \Delta^-) \text{ implies } \phi \notin \Delta^+ = IP^+_\Sigma(\Delta^+, \Delta^-)$$

holds. Therefore positive and negative introspection are mutually exclusive. □

The cautious autoepistemic expansions of the set of premises Σ may be incomplete at the level of the beliefs of the agent. The undetermined formulae are allowed if necessary to preserve the Σ-acceptability. To the contrary, stable expansions are complete, i.e. each formula is either believed or disbelieved by the agent. Let us now prove that stable expansions are cautious as well.

Proposition 2.12 (Generalization)
All stable expansions are cautious expansions.

Proof. We show that stable expansions are in accordance with the requirements of Definition 2.10. Let Δ be a stable expansion. The agent believes the formulae in Δ and disbelieves the formulae in $\overline{\Delta}$. Therefore, let $\Delta^+ = \Delta$ and $\Delta^- = \overline{\Delta}$. The positive introspection in Definition 2.10 coincides now with Equation (1). As $\Delta^+ \cup \Delta^-$ equals to $\mathcal{L}_{ae}$, the negative introspection checks only the empty enumeration λ. By Definition 2.9 testing whether ϕ belongs to $EE^\lambda(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$, reduces to testing whether $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \not\models_{ae} \phi$. Thus $\Delta^- = \overline{\Delta^+}$ and this is clearly the case. □

Let us yet emphasize one important point. As it is apparent from the definition presented earlied, cautious autoepistemic logic uses two valued inference mechanism (classical propositional logic), even though the expansions have their three-valued character. Another possibility to represent incompleteness is to incorporate the third truth value to the underlying logic.

2.3 Finitary Characterization

Cautious expansions consist normally of two infinite sets of formulae. The exceptional case is the inconsistent cautious expansion, i.e. $\langle \mathcal{L}_{ae}, \emptyset \rangle$. Such an expansion is possible if the set of premises Σ is inconsistent. The infiniteness of cautious expansions leads to representational problems. The nonconstructive nature of the expansions makes the situation even worse – cautious expansions seem tedious to handle and reason about. A solution is to provide a *finitary characterization* for cautious expansions.

Stable and $\mathcal{N}$-expansions have finitary characterizations. Moore [1985] has shown that stable expansions are in one-to-one correspondence with their objective parts, i.e. their L-free parts. Niemelä [1990a] has developed a different approach – the *full sets*. He identifies a subset of the expansion that uniquely characterizes the expansions. Gottlob used Niemelä style full sets when he provided $\mathcal{N}$-expansions with a similar finitary characterization [Gottlob, 1992a]. Full sets provide a convenient way of reasoning about expansions and examining their properties.

Moore [1988] provided a finitary characterization similar to the full set approach, but he has used propositional valuations to define the construction. Moore gives an alternative characterization of stable expansions by means of complete S5-structures, too. This shows that autoepistemic logic has not totally drifted apart from the standard modal logics. But there is a difference, however. It is worthwhile to note that the semantics defined in Section 2.1 is "flat" compared to the traditional modal logics. Kripke-structures define the truth of the modal formulae in standard modal logics [Chellas, 1980]. Shvarts [1990] and Stärk [1990] have also provided finitary characterizations for stable expansions.

This subsection is devoted to extending Niemelä's full set approach to cautious expansions. However, we need some technical tools to achieve this. To show that a finitary characterization uniquely determines a cautious expansion, we need to prove that reasoning with respect to the finitary characterization coincides with the reasoning with respect to the whole expansion. In doing this, we need to remove belief literals from expansions and show that the remaining belief literals are sufficient to determine the expansion.

The first one of the supplementary propositions has been proved by Niemelä [1992]. The proof is based on direct model constructions. The import of this proposition is that we may reduce a set of quasi-literals Λ to $\Lambda' = \Pi_{Sf^{qL}(\Sigma\cup\{\phi\})}(\Lambda)$ without affecting the derivability of ϕ from $\Sigma \cup \Lambda$. The reduction is carried out with respect to $Sf^L(\Sigma \cup \{\phi\})$, i.e. exactly the L-subformulae appearing in the set of premises Σ and the formula under investigation ϕ. This is a very frequently used result throughout this work. The converse operation is easier – belief literals may always be added without affecting the derivability of ϕ, since $\models_{ae}$ is monotonic.

Proposition 2.13 ([Niemelä, 1992])
Let Σ be a set of sentences of $\mathcal{L}_{ae}$ and ϕ any formula. If there is a consistent set of quasi-literals Λ such that $\Sigma \cup \Lambda \models_{ae} \phi$, then $\Sigma \cup \Lambda' \models_{ae} \phi$ where $\Lambda' = \Pi_{Sf^{qL}(\Sigma\cup\{\phi\})}(\Lambda)$.

We need a similar result for the derivability relation $\vdash_{\mathcal{N}}$ when we show that the finitary characterization works for cautious expansions satisfying an additional groundedness requirement in Section 3. The proposition provides means of reducing L-proofs of formulae.

Proposition 2.14
Let Σ be a set of sentences of $\mathcal{L}_{ae}$ and ϕ any formula. If there is a set of negative belief literals Λ satisfying the consistency requirement

$$\neg L\chi \in \Lambda \textit{ implies } \Sigma \cup \Lambda \nvdash_{\mathcal{N}} \chi$$

and $\Sigma \cup \Lambda \vdash_{\mathcal{N}} \phi$, then $\Sigma \cup \Lambda' \vdash_{\mathcal{N}} \phi$ where $\Lambda' = \Pi_{Sf^L(\Sigma\cup\{\phi\})}(\Lambda)$. Furthermore, a minimal L-proof of ϕ contains only members of $Sf^L(\Sigma \cup \{\phi\})$.

Proof. Let $\Sigma \cup \Lambda \vdash_{\mathcal{N}} \phi$ hold. Then there is an L-proof $L\phi_1, ..., L\phi_n$ of ϕ from $\Sigma \cup \Lambda$ such that

$$\Sigma \cup \Lambda \cup \{L\phi_1, ..., L\phi_n\} \models_{ae} \phi. \tag{6}$$

Keeping the definition of the L-proofs in mind, each member ϕ_i $(1 \leq i \leq n)$ of the L-proof satisfies

$$\Sigma \cup \Lambda \cup \{L\phi_j \mid 1 \leq j < i\} \models_{ae} \phi_i \tag{7}$$

and $\Sigma \cup \Lambda \vdash_{\mathcal{N}} \phi_i$. Then the consistency requirement of the proposition yields that no $\neg L\phi_i$ is in Λ. Thus $\Lambda \cup \{L\phi_j \mid 1 \leq j < i\}$ $(1 \leq i \leq n)$ and $\Lambda \cup \{L\phi_1, ..., L\phi_n\}$ are consistent sets of belief literals. Let $L\phi'_1, ..., L\phi'_m$ $(m \leq n)$ be exactly those quasi-literals of $L\phi_1, ..., L\phi_n$ that belong to $Sf^L(\Sigma \cup \{\phi\})$ in the same order they appear in $L\phi_1, ..., L\phi_n$. Then we define a function n' for the values $k = 1, ..., m$ such that $L\phi'_k = L\phi_{n'(k)}$, i.e. n' gives the correspondence between the L-proof and its reduction with respect to $Sf^L(\Sigma \cup \{\phi\})$. An application of Proposition 2.13 to (6) yields that $\Sigma \cup \Lambda'' \models_{ae} \phi$ where

$$\begin{array}{rcll}
\Lambda'' & = & \Pi_{Sf^{qL}(\Sigma\cup\{\phi\})}(\Lambda \cup \{L\phi_1, ..., L\phi_n\}) & \text{[Proposition 2.13]} \\
& \subseteq & \Pi_{Sf^{L}(\Sigma\cup\{\phi\})}(\Lambda \cup \{L\phi_1, ..., L\phi_n\}) & \left[{\scriptstyle Sf^{qL}(\Sigma\cup\{\phi\})\subseteq Sf^L(\Sigma\cup\{\phi\})}\right] \\
& = & \Lambda' \cup \{L\phi'_1, ..., L\phi'_m\}. & \left[{\scriptstyle L\phi'_j \in Sf^L(\Sigma\cup\{\phi\})}\right]
\end{array}$$

By the monotonicity of $\models_{ae}$, $\Sigma \cup \Lambda' \cup \{L\phi'_1, ..., L\phi'_m\} \models_{ae} \phi$. Then consider any $L\phi'_k = L\phi_{n'(k)} \in Sf^L(\Sigma \cup \{\phi\})$. By Equation (7), it is the case that $\Sigma \cup \Lambda \cup \{L\phi_j \mid 1 \leq j < n'(k)\} \models_{ae} \phi'_k$. Proposition 2.13 yields now a set of belief literals

$$\begin{array}{rcll}
\Lambda^{(3)} & = & \Pi_{Sf^{qL}(\Sigma\cup\{\phi'_k\})}(\Lambda \cup \{L\phi_j \mid 1 \leq j < n'(k)\}) & \text{[Proposition 2.13]} \\
& \subseteq & \Pi_{Sf^{L}(\Sigma\cup\{\phi\})}(\Lambda \cup \{L\phi_j \mid 1 \leq j < n'(k)\}) & \\
& & \left[{\scriptstyle Sf^{qL}(\Sigma\cup\{\phi'_k\})\subseteq Sf^L(\Sigma\cup\{\phi\})}\right] & \\
& = & \Lambda' \cup \{L\phi'_j \mid 1 \leq j < k\}. & \\
& & \left[{\scriptstyle \{L\phi'_j \mid 1\leq j<k\}\subseteq Sf^L(\Sigma\cup\{\phi\})}\right] &
\end{array}$$

such that $\Sigma \cup \Lambda^{(3)} \models_{ae} \phi'_k$. It follows that $\Sigma \cup \Lambda' \cup \{L\phi'_j \mid 1 \leq j < k\} \models_{ae} \phi'_k$ by the monotonicity of $\models_{ae}$. But this is exactly the condition imposed on L-proofs. Therefore $L\phi'_1, ..., L\phi'_m$ forms an L-proof of ϕ from $\Sigma \cup \Lambda'$, i.e. $\Sigma \cup \Lambda' \vdash_{\mathcal{N}} \phi$ where $\Lambda' = \Pi_{Sf^L(\Sigma\cup\{\phi\})}(\Lambda)$.

Let us now show that a minimal L-proof of ϕ contains only members of $Sf^L(\Sigma \cup \{\phi\})$. Let $L\phi_1, ..., L\phi_n$ be a minimal L-proof of ϕ from $\Sigma \cup \Lambda$. Assume that some members of the L-proof do not belong to $Sf^L(\Sigma \cup \{\phi\})$. This implies there is a reduced L-proof $L\phi'_1, ..., L\phi'_m$ such that $m < n$ by the proof presented above. But then $L\phi_1, ..., L\phi_n$ is not minimal, a contradiction. □

The last one of the three reduction propositions is essential for the sake of the negative introspection. In fact, the infinite enumerations used for negative introspection can be replaced by their finite reductions when we are reasoning with respect to the finitary characterization.

Proposition 2.15 (Reduction of enumerations)
Let Σ be a set of sentences of $\mathcal{L}_{ae}$ and ϕ any formula such that $Sf^L(\Sigma \cup \{\phi\})$ is finite. Let Λ be a consistent set of quasi-literals and $\Lambda' = \Pi_{Sf^L(\Sigma\cup\{\phi\})}(\Lambda)$. Let ε be an enumeration of the formulae in $\mathcal{L}_{ae} - \{\psi \mid L\psi \in$ q-atom$(\Lambda)\}$ and ε' its reduction with respect to $Sf^L(\Sigma \cup \{\phi\})$. Then

$$\phi \in EE^{\varepsilon}(\Sigma \cup \Lambda) \text{ if and only if } \phi \in EE^{\varepsilon'}(\Sigma \cup \Lambda').$$

Proof. We begin by making clear the relationship between the two enumerations. Let $\varepsilon = \psi_1, \psi_2, ...$ be the possibly infinite enumeration of $\mathcal{L}_{ae} - \{\psi \mid L\psi \in \text{q-atom}(\Lambda)\}$. Its reduction $\varepsilon' = \psi'_1, \psi'_2, ...$ contains exactly those elements of ε that belong to the set $\{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\phi\})\}$ in the same order they appear in ε. Clearly, this enumeration is finite since $Sf^L(\Sigma \cup \{\phi\})$ is finite. Then we define function n' for values $i = 1, 2, ..., |\varepsilon'|$ by requiring $\psi'_i = \psi_{n'(i)}$. Intuitively, $n'(i)$ gives the location of the i^{th} element of ε' in the enumeration ε. We also define n' when $i = 0$ by letting $n'(0) = 0$ for proof-technical reasons.

At first we prove that Equation (8) holds for all i, $0 \le i \le |\varepsilon'|$, by induction on i.

$$B_i^{\varepsilon'}(\Sigma \cup \Lambda') = \Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^{\varepsilon}_{n'(i)}(\Sigma \cup \Lambda)) \tag{8}$$

Induction base: $i = 0$. Now

$$\begin{array}{rcll} B_0^{\varepsilon'}(\Sigma \cup \Lambda') & = & \emptyset & \text{[Definition 2.9]} \\ & = & B_0^{\varepsilon}(\Sigma \cup \Lambda) & \text{[Definition 2.9]} \\ & = & B^{\varepsilon}_{n'(0)}(\Sigma \cup \Lambda) & \text{[Definition of } n'\text{]} \\ & = & \Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^{\varepsilon}_{n'(0)}(\Sigma \cup \Lambda)). & \text{[Definition 2.6]} \end{array}$$

Thus Equation (8) holds when $i = 0$.

Inductive hypothesis: Equation (8) holds for $0 \le i < k$.

Induction step: $i = k$. The inductive hypothesis yields that

$$B_{k-1}^{\varepsilon'}(\Sigma \cup \Lambda') = \Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^{\varepsilon}_{n'(k-1)}(\Sigma \cup \Lambda)). \tag{9}$$

By the definition of the function n', none of the belief literals $L\psi_l$, $n'(k-1) < l < n'(k)$, belongs to $Sf^L(\Sigma \cup \{\phi\})$. Therefore,

$$(\neg)L\psi'_k \in \Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda)) \text{ iff } (\neg)L\psi'_k = (\neg)L\psi_{n'(k)}. \tag{10}$$

So it remains to be shown that

$$(\neg)L\psi'_k \in B_k^{\varepsilon'}(\Sigma \cup \Lambda') \text{ iff } (\neg)L\psi'_k \in \Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda)). \tag{11}$$

($\Rightarrow$) Let $L\psi'_k \in B_k^{\varepsilon'}(\Sigma \cup \Lambda')$. Then $\Sigma \cup \Lambda' \cup B_{k-1}^{\varepsilon'}(\Sigma \cup \Lambda') \models_{ae} \psi'_k$ follows by Definition 2.9. The substitution of Equation (9) yields that

$$\Sigma \cup \Lambda' \cup \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k-1)}(\Sigma \cup \Lambda)) \models_{ae} \psi'_k.$$

Since $\Lambda' \subseteq \Lambda$,

$$\Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k-1)}(\Sigma \cup \Lambda)) \subseteq B^{\varepsilon}_{n'(k-1)}(\Sigma \cup \Lambda) \subseteq B^{\varepsilon}_{n'(k)-1}(\Sigma \cup \Lambda)$$

and $\psi'_k = \psi_{n'(k)}$, $\Sigma \cup \Lambda \cup B^{\varepsilon}_{n'(k)-1}(\Sigma \cup \Lambda) \models_{ae} \psi_{n'(k)}$ by the monotonicity of $\models_{ae}$. By Definition 2.9, $L\psi_{n'(k)} \in B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda)$. Since $L\psi'_k \in Sf^L(\Sigma \cup \{\phi\})$, $L\psi'_k$ belongs to $\Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda))$.

Let $\neg L\psi'_k \in B_k^{\varepsilon'}(\Sigma \cup \Lambda')$. It holds that $\Sigma \cup \Lambda' \cup B_{k-1}^{\varepsilon'}(\Sigma \cup \Lambda') \not\models_{ae} \psi'_k$, by Definition 2.9. We may now substitute Equation 9. This yields

$$\Sigma \cup \Lambda' \cup \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k-1)}(\Sigma \cup \Lambda)) \not\models_{ae} \psi'_k. \tag{12}$$

Assume that $\Sigma \cup \Lambda \cup B^{\varepsilon}_{n'(k)-1}(\Sigma \cup \Lambda) \models_{ae} \psi_{n'(k)}$. The set $\Lambda \cup B^{\varepsilon}_{n'(k)-1}(\Sigma \cup \Lambda)$ is consistent by construction as ε does not contain elements ψ such that $L\psi \in$ q-atom(Λ). By Proposition 2.13 there is a set

$$\begin{array}{rcll}
\Lambda'' & = & \Pi_{Sf^{qL}(\Sigma \cup \{\psi_{n'(k)}\})}(\Lambda \cup B^{\varepsilon}_{n'(k)-1}(\Sigma \cup \Lambda)) & \text{[Proposition 2.13]} \\
& \subseteq & \Pi_{Sf^L(\Sigma \cup \{\phi\})}(\Lambda \cup B^{\varepsilon}_{n'(k)-1}(\Sigma \cup \Lambda)) & \\
& & \left[{\scriptstyle Sf^{qL}(\Sigma \cup \{\psi_{n'(k)}\}) \subseteq Sf^L(\Sigma \cup \{\phi\})}\right] & \\
& = & \Lambda' \cup \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k-1)}(\Sigma \cup \Lambda)) & \\
& & \left[{\scriptstyle L\psi_l \notin Sf^L(\Sigma \cup \{\phi\}), n'(k-1) < l < n'(k)}\right] &
\end{array}$$

such that $\Sigma \cup \Lambda'' \models_{ae} \psi_{n'(k)}$. It follows by the monotonicity of $\models_{ae}$ that

$$\Sigma \cup \Lambda' \cup \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k-1)}(\Sigma \cup \Lambda)) \models_{ae} \psi_{n'(k)}.$$

Since $\psi'_k = \psi_{n'(k)}$, this is a contradiction with (12). Therefore, it holds that $\Sigma \cup \Lambda \cup B^{\varepsilon}_{n'(k)-1}(\Sigma \cup \Lambda) \not\models_{ae} \psi_{n'(k)}$. Then $\neg L\psi_{n'(k)} \in B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda)$ by Definition 2.9. Since $L\psi'_k \in Sf^L(\Sigma \cup \{\phi\})$, $\neg L\psi'_k \in \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda))$.
($\Leftarrow$) Let $L\psi'_k \in \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda))$. It follows by (10) that

$$L\psi_{n'(k)} \in \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda)).$$

Let us now assume that $\neg L\psi'_k \in B_k^{\varepsilon'}(\Sigma \cup \Lambda')$. By the if-direction of (11) $\neg L\psi_{n'(k)} \in \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda))$ a contradiction. Hence, $L\psi'_k$ belongs to $B_k^{\varepsilon'}(\Sigma \cup \Lambda')$. Similarly, let $\neg L\psi'_k \in \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda))$. Equation (10) yields that $\neg L\psi_{n'(k)} \in \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda))$. Assume that $L\psi'_k$ is in $B_k^{\varepsilon'}(\Sigma \cup \Lambda')$. By Equation (11) $L\psi_{n'(k)} \in \Pi_{Sf^L(\Sigma \cup \{\phi\})}(B^{\varepsilon}_{n'(k)}(\Sigma \cup \Lambda))$, a contradiction. Hence, $\neg L\psi'_k \in B_k^{\varepsilon'}(\Sigma \cup \Lambda')$. This completes the proof of Equation (11). Therefore, the equality (8) holds.

We need the auxiliary results given in Equations (13) and (14). For all enumerations ε of $\mathcal{L}_{ae} - \{\chi \mid L\chi \in \text{q-atom}(\Lambda)\}$,

$$\Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^\varepsilon(\Sigma \cup \Lambda) - B^\varepsilon_{n'(|\varepsilon'|)}(\Sigma \cup \Lambda)) = \emptyset. \tag{13}$$

Equation (13) is justified since the enumeration ε does not contain elements ψ such that $L\psi \in Sf^L(\Sigma \cup \{\phi\})$ after the element $\psi_{n'(|\varepsilon'|)}$. The equation

$$B^{\varepsilon'}(\Sigma \cup \Lambda') = \Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^\varepsilon(\Sigma \cup \Lambda)) \tag{14}$$

is shown as follows:

$$\begin{array}{rcl}
\Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^\varepsilon(\Sigma \cup \Lambda)) & = & \Pi_{Sf^L(\Sigma\cup\{\phi\})}(B^\varepsilon_{n'(|\varepsilon'|)}(\Sigma \cup \Lambda)) \\
& & [\text{Equation (13)}] \\
& = & B^{\varepsilon'}_{|\varepsilon'|}(\Sigma \cup \Lambda') \\
& & [\text{Equation (8)}, i = |\varepsilon'|] \\
& = & B^{\varepsilon'}(\Sigma \cup \Lambda') \\
& & [\text{Definition 2.9}].
\end{array}$$

Now we are in position to prove Proposition 2.15. ($\Rightarrow$) Let $\phi \in EE^\varepsilon(\Sigma \cup \Lambda)$. Then $\Sigma \cup \Lambda \cup B^\varepsilon(\Sigma \cup \Lambda) \models_{ae} \phi$ holds. The set $\Lambda \cup B^\varepsilon(\Sigma \cup \Lambda)$ is consistent by the construction given in Definition 2.9, because ε does not contain elements ψ such that $L\psi \in \text{q-atom}(\Lambda)$. By Proposition 2.13 there exists the set Λ'' such that $\Sigma \cup \Lambda'' \models_{ae} \phi$ and

$$\begin{array}{rcll}
\Lambda'' & = & \Pi_{Sf^{qL}(\Sigma\cup\{\phi\})}(\Lambda \cup B^\varepsilon(\Sigma \cup \Lambda)) & [\text{Proposition 2.13}] \\
& \subseteq & \Pi_{Sf^L(\Sigma\cup\{\phi\})}(\Lambda \cup B^\varepsilon(\Sigma \cup \Lambda)) & [Sf^{qL}(\Sigma\cup\{\phi\}) \subseteq Sf^L(\Sigma\cup\{\phi\})] \\
& = & \Pi_{Sf^L(\Sigma\cup\{\phi\})}(\Lambda) \cup B^{\varepsilon'}(\Sigma \cup \Lambda') & [\text{Equation (14)}] \\
& = & \Lambda' \cup B^{\varepsilon'}(\Sigma \cup \Lambda'). & [\text{Definition of } \Lambda']
\end{array}$$

Thus $\Sigma \cup \Lambda' \cup B^{\varepsilon'}(\Sigma \cup \Lambda') \models_{ae} \phi$ by the monotonicity of $\models_{ae}$. Then ϕ belongs to $EE^{\varepsilon'}(\Sigma \cup \Lambda')$. ($\Leftarrow$) Let $\phi \in EE^{\varepsilon'}(\Sigma \cup \Lambda')$, i.e. $\Sigma \cup \Lambda' \cup B^{\varepsilon'}(\Sigma \cup \Lambda') \models_{ae} \phi$. Equation (14) implies that $B^{\varepsilon'}(\Sigma \cup \Lambda') \subseteq B^\varepsilon(\Sigma \cup \Lambda)$ and since Λ' is a subset of Λ, $\Sigma \cup \Lambda \cup B^\varepsilon(\Sigma \cup \Lambda) \models_{ae} \phi$ holds by the monotonicity of $\models_{ae}$. So ϕ belongs to $EE^\varepsilon(\Sigma \cup \Lambda)$ as desired. □

We may now give the exact definition of the finitary characterization. Niemelä [1992] has presented a similar construction for stable expansions. Two *recursive consequence relations* $\models_L$ and $=\!\mid_L$ are introduced. They provide the finitary mechanism for deciding whether a formula ϕ is believed or disbelieved by the agent, respectively. The *subbeliefs* $SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ of ϕ are determined before ϕ by both relations. The consequence relation $\models_L$ is in fact a neat formalization of the decision procedure described by Moore [1988]. Moore's description is rather informal and uses valuations to reduce theories. Niemelä uses $\not\models_L$ for negative introspection in conjunction with stable expansions. However, we have to use the *irrefutability relation* $\dashv_L$ to capture the weakened negative introspection capability of the agent. The reduced enumerations are used for the purpose of determining the refutability of formulae. In general, the two relations $\dashv_L$ and $\not\models_L$ are nonmonotonic. The set of quasi-literals Λ represents a cautious expansion.

Definition 2.16 (Recursive consequence relations)
For a set of premises Σ, a set of quasi-literals Λ and a formula ϕ:

1. $\langle \Sigma, \Lambda \rangle \models_L \phi$ *if and only if*
   $\Sigma \cup \Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi) \models_{ae} \phi$

2. $\langle \Sigma, \Lambda \rangle \dashv_L \phi$ *if and only if*
   there is no enumeration ε *of*
   $\{\chi \mid L\chi \in Sf^L(\Sigma \cup \{\phi\}) - \textit{q-atom}(\Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi))\}$
   such that $\phi \in EE^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi))$

where $SB_{\langle \Sigma, \Lambda \rangle}(\phi) = \{L\chi \in Sf^L(\phi) \mid \langle \Sigma, \Lambda \rangle \models_L \chi\} \cup \{\neg L\chi \in \neg Sf^L(\phi) \mid \langle \Sigma, \Lambda \rangle \dashv_L \chi\}$.

The *full sets* of Definition 2.17 form the finitary characterization for cautious expansions. As the auxiliary propositions at the beginning of this section suggest, we may drastically reduce the expansions and furthermore, without loosing any information. In the finitary characterization we have to consider only the belief status of formulae χ such that $L\chi \in Sf^L(\Sigma)$, i.e. the L-subformulae of the set of premises Σ. Most importantly, full sets are finite in size provided that $Sf^L(\Sigma)$ is finite. The two Σ-fullness conditions given in the definition are self-referential. It follows that Σ-full sets have a nonconstructive nature. Given a candidate Σ-full set we can verify that the two Σ-fullness conditions are satisfied. However, the two conditions do not provide a way of constructing such sets.

Definition 2.17 (Full sets)
For a set of premises Σ, a set $\Lambda \subseteq Sf^L(\Sigma) \cup \neg Sf^L(\Sigma)$ is Σ-full if the following conditions hold for every $L\chi \in Sf^L(\Sigma)$:

$$L\chi \in \Lambda \textit{ if and only if } \langle \Sigma, \Lambda \rangle \models_L \chi$$

$$\neg L\chi \in \Lambda \textit{ if and only if } \langle \Sigma, \Lambda \rangle \dashv_L \chi$$

We say that Σ-full sets are *closed with respect to subbeliefs*, i.e. $L\chi \in Sf^L(\Sigma)$ implies that $SB_{\langle \Sigma, \Lambda \rangle}(\chi) \subseteq \Lambda$. It is easy to see from Definition 2.17 that this is the case, since $Sf^L(\Sigma)$ is closed with respect to L-subformulae. The same property holds for the subbeliefs $SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ of an arbitrary formula ϕ by Definition 2.16.

Niemelä shows that the full sets characterizing stable expansions have the property that for each $L\chi \in Sf^L(\Sigma)$ either $L\chi$ or $\neg L\chi$ belongs to the Σ-full set. This is not surprising on the basis of the discussion in Section 2.2. The full sets characterizing cautious expansions share a somewhat different property, since undetermined formulae are allowed: for each $L\chi \in Sf^L(\Sigma)$, at most one of $L\chi$ and $\neg L\chi$ belongs to the Σ-full set characterizing a cautious expansion. Proposition 2.18 should make this clear.

Proposition 2.18 (Mutual exclusion)
For a set Σ of sentences of $\mathcal{L}_{ae}$, a set of quasi-literals Λ and a formula ϕ:

$$\langle\Sigma,\Lambda\rangle \models_L \phi \textit{ implies } \langle\Sigma,\Lambda\rangle \not\dashv_L \phi \tag{15}$$

or equivalently,

$$\langle\Sigma,\Lambda\rangle \dashv_L \phi \textit{ implies } \langle\Sigma,\Lambda\rangle \not\models_L \phi. \tag{16}$$

Proof. We prove (15). Let $\langle\Sigma,\Lambda\rangle \models_L \phi$ hold. Thus $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \phi$, by Definition 2.16. Then an arbitrary enumeration ε of

$$\{\chi \mid L\chi \in Sf^L(\Sigma \cup \{\phi\}) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))\}$$

can be selected such that $\phi \in EE^{\varepsilon}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$, i.e.

$$\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \cup B^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) \models_{ae} \phi$$

by the monotonicity of $\models_{ae}$. Thus $\langle\Sigma,\Lambda\rangle \not\dashv_L \phi$. □

We concentrate now on proving that full sets really form a finitary characterization for the cautious expansions. We shall formulate the exact the relationship between cautious expansions and the full sets characterizing them. In Proposition 2.19 we show that within the reduction $\Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-)$, the recursive consequence relations coincide with the results of positive and negative introspection with respect to the whole expansion.

Proposition 2.19
Let Σ be a set of sentences of $\mathcal{L}_{ae}$. If it holds that

$$\langle\Delta^+,\Delta^-\rangle = CAE_\Sigma(\Delta^+,\Delta^-) \textit{ and } \Lambda = \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-)$$

then for every $L\chi \in Sf^L(\Sigma)$,

$$\langle\Sigma,\Lambda\rangle \models_L \chi \textit{ if and only if } \chi \in IP^+_\Sigma(\Delta^+,\Delta^-) \tag{17}$$

$$\langle\Sigma,\Lambda\rangle \dashv_L \chi \textit{ if and only if } \chi \in IP^-_\Sigma(\Delta^+,\Delta^-) \tag{18}$$

Proof. Let $L\chi \in Sf^L(\Sigma)$. We prove that Equations (17) and (18) hold given that
$$SB_{\langle\Sigma,\Lambda\rangle}(\chi) = \Pi_{Sf^L(\chi)}(L\Delta^+ \cup \neg L\Delta^-). \tag{19}$$

The positive introspection is considered first.

($\Rightarrow$) Let $\langle\Sigma,\Lambda\rangle \models_L \chi$. Then $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \models_{ae} \chi$. Since both Λ and $SB_{\langle\Sigma,\Lambda\rangle}(\chi)$ are contained in $L\Delta^+ \cup \neg L\Delta^-$, $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \models_{ae} \chi$ by the monotonicity of $\models_{ae}$ and $\chi \in IP^+_\Sigma(\Delta^+,\Delta^-)$. ($\Leftarrow$) Let $\chi \in IP^+_\Sigma(\Delta^+,\Delta^-)$ hold, i.e. $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \models_{ae} \chi$. It follows from Proposition 2.11 that $L\Delta^+ \cup \neg L\Delta^-$ is consistent. By Proposition 2.13 there is a set Λ' such that $\Sigma \cup \Lambda' \models_{ae} \chi$ and

$$\begin{array}{rlll} \Lambda' & = & \Pi_{Sf^{qL}(\Sigma\cup\{\chi\})}(L\Delta^+ \cup \neg L\Delta^-) & [\text{Proposition 2.13}] \\ & \subseteq & \Pi_{Sf^{L}(\Sigma\cup\{\chi\})}(L\Delta^+ \cup \neg L\Delta^-) & [{\scriptstyle Sf^{qL}(\Sigma\cup\{\chi\})\subseteq Sf^L(\Sigma\cup\{\chi\})}] \\ & = & \Lambda \cup \Pi_{Sf^L(\chi)}(L\Delta^+ \cup \neg L\Delta^-) & [\text{Definition of } \Lambda] \\ & = & \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi). & [\text{Equation (19)}] \end{array}$$

Thus $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \models_{ae} \chi$ by the monotonicity of $\models_{ae}$. So $\langle\Sigma,\Lambda\rangle \models_L \chi$.

Negative introspection follows. Now

$$\begin{array}{rl} & (\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)) \cap \{\chi' \mid L\chi' \in Sf^L(\Sigma \cup \{\chi\})\} \\ = & \{\chi' \mid L\chi' \in Sf^L(\Sigma \cup \{\chi\})\} - (\Delta^+ \cup \Delta^-) \\ = & \{\chi' \mid L\chi' \in Sf^L(\Sigma \cup \{\chi\}) - \text{q-atom}(L\Delta^+ \cup \neg L\Delta^-)\} \\ = & \{\chi' \mid L\chi' \in Sf^L(\Sigma \cup \{\chi\}) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi))\}, \end{array} \tag{20}$$

since $L\chi \in Sf^L(\Sigma)$. ($\Rightarrow$) Let $\langle\Sigma,\Lambda\rangle \mathrel{=\!\!|}_L \chi$ hold. Then by the definition of $\mathrel{=\!\!|}_L$, there is no enumeration ε of

$$\{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\chi\}) - \text{q-atom}(\Lambda \cup \ SB_{\langle\Sigma,\Lambda\rangle}(\chi))\}, \tag{21}$$

such that $\chi \in EE^\varepsilon(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi))$. Assume that $\chi \notin IP^-_\Sigma(\Delta^+,\Delta^-)$. Then there is an enumeration ε' of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ such that χ belongs to $EE^{\varepsilon'}(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$. The set of belief literals $L\Delta^+ \cup \neg L\Delta^-$ is consistent by Proposition 2.11. By Equation (20) there is a reduction ε'' of ε' with respect to $Sf^L(\Sigma \cup \{\chi\}) = Sf^L(\Sigma)$ such that ε'' is an enumeration of (21). Now $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) = \Pi_{Sf^L(\Sigma\cup\{\chi\})}(L\Delta^+ \cup \neg L\Delta^-)$ by Equation (19). Proposition 2.15 yields that $\chi \in EE^{\varepsilon''}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi))$, a contradiction. Hence $\chi \in IP^-_\Sigma(\Delta^+,\Delta^-)$. ($\Leftarrow$) Let $\chi \in IP^+_\Sigma(\Delta^+,\Delta^-)$. Then there is no enumeration ε of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ such that $\chi \in EE^\varepsilon(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$. Assume that $\langle\Sigma,\Lambda\rangle \not\mathrel{=\!\!|}_L \chi$. Then there is an enumeration ε' of (21) such that $\chi \in EE^{\varepsilon'}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi))$. By Equation (20) the enumeration ε' can be extended to an enumeration ε'' of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ (by simply concatenating the rest of the elements to ε'') such that ε' is the reduction of ε'' with respect to $Sf^L(\Sigma \cup \{\chi\})$. By Proposition 2.15 $\chi \in EE^{\varepsilon''}(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$, a contradiction. Hence $\langle \Sigma, \Lambda \rangle \models\!\!\mid_L \chi$.

The rest of the proof is carried out by induction on $BD(\chi)$.

Induction base: $BD(\chi) = 0$. Then Equation (19) holds trivially, because $Sf^L(\chi) = \emptyset$. Thus Equations (17) and (18) are valid.

Inductive hypothesis: assume that for all $L\chi \in Sf^L(\Sigma)$ with $BD(\chi) < n$, Equations (17) and (18) hold.

Induction step: Let $BD(\chi) = n$. It is shown that Equation (19) holds. If $L\chi' \in SB_{\langle \Sigma, \Lambda \rangle}(\chi)$, then $\langle \Sigma, \Lambda \rangle \models_L \chi'$ and $L\chi' \in Sf^L(\chi)$. It follows that $L\chi' \in Sf^L(\Sigma)$, because $L\chi \in Sf^L(\Sigma)$. Since $BD(\chi') < n$, $\chi' \in IP^+_\Sigma(\Delta^+, \Delta^-)$ by the inductive hypothesis. Thus $\chi' \in \Delta^+$ and $L\chi' \in L\Delta^+$. If $\neg L\chi' \in SB_{\langle \Sigma, \Lambda \rangle}(\chi)$, then $\langle \Sigma, \Lambda \rangle \models\!\!\mid_L \chi'$ and $L\chi' \in Sf^L(\chi)$. It follows that $L\chi' \in Sf^L(\Sigma)$, because $L\chi \in Sf^L(\Sigma)$. Since $BD(\chi') < n$, $\chi' \in IP^-_\Sigma(\Delta^+, \Delta^-)$ by the inductive hypothesis. Thus $\chi' \in \Delta^-$ and $\neg L\chi' \in \neg L\Delta^-$. Since q-atom$(SB_{\langle \Sigma, \Lambda \rangle}(\chi)) \subseteq Sf^L(\chi)$, $SB_{\langle \Sigma, \Lambda \rangle}(\chi)$ is contained in $\Pi_{Sf^L(\chi)}(L\Delta^+ \cup \neg L\Delta^-)$. If it is the case that $L\chi' \in \Pi_{Sf^L(\chi)}(L\Delta^+ \cup \neg L\Delta^-)$, then $L\chi' \in Sf^L(\chi)$. In addition, it holds that $\chi' \in \Delta^+$ and $\chi' \in IP^+_\Sigma(\Delta^+, \Delta^-)$. Since $BD(\chi') < n$, $\langle \Sigma, \Lambda \rangle \models_L \chi'$ by the inductive hypothesis. Thus $L\chi' \in SB_{\langle \Sigma, \Lambda \rangle}(\chi)$. If $\neg L\chi'$ belongs to $\Pi_{Sf^L(\chi)}(L\Delta^+ \cup \neg L\Delta^-)$, then $L\chi' \in Sf^L(\chi)$. In addition, $\chi' \in \Delta^-$ and $\chi' \in IP^-_\Sigma(\Delta^+, \Delta^-)$. Since $BD(\chi') < n$, $\langle \Sigma, \Lambda \rangle \models\!\!\mid_L \chi'$ by the inductive hypothesis. Thus $\neg L\chi' \in SB_{\langle \Sigma, \Lambda \rangle}(\chi)$, the reduction $\Pi_{Sf^L(\chi)}(L\Delta^+ \cup \neg L\Delta^-)$ is contained in $SB_{\langle \Sigma, \Lambda \rangle}(\chi)$ and Equation (19) holds.

This implies that Equations Equations (17) and (18) hold when $BD(\chi) = n$ and we have completed the proof. □

Lemma 2.20 gives the mapping from a cautious expansion to a Σ-full set. It shows that the reduction $\Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-)$ is Σ-full.

Lemma 2.20
Let Σ be a set of sentences of $\mathcal{L}_{ae}$. If $\langle \Delta^+, \Delta^- \rangle = CAE_\Sigma(\Delta^+, \Delta^-)$, then

$$\Lambda = \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-) \tag{22}$$

is Σ-full.

Proof. Let $L\chi \in Sf^L(\Sigma)$.

$$\begin{array}{llll} L\chi \in \Lambda & \text{iff} & \chi \in \Delta^+ = IP^+_\Sigma(\Delta^+, \Delta^-) & \text{[Equation (22)]} \\ & \text{iff} & \langle \Sigma, \Lambda \rangle \models_L \chi & \text{[Proposition 2.19]} \end{array}$$

$$\begin{array}{llll} \neg L\chi \in \Lambda & \text{iff} & \chi \in \Delta^- = IP^-_\Sigma(\Delta^+, \Delta^-) & \text{[Equation (22)]} \\ & \text{iff} & \langle \Sigma, \Lambda \rangle \models\!\!\mid_L \chi & \text{[Proposition 2.19]} \end{array}$$

The Σ-fullness conditions are thus satisfied. $\square$

The remaining concern is how to reconstruct cautious expansions from Σ-full sets. Lemma 2.21 shows how this reconstruction takes place. The recursive consequence relations $\models_L$ and $=\!\mid_L$ provide the machinery to do this.

Lemma 2.21
Let Σ be a set of sentences of $\mathcal{L}_{ae}$ and Λ a Σ-full set. If

$$\Delta^+ = \{\phi \mid \langle \Sigma, \Lambda \rangle \models_L \phi\}$$

$$\Delta^- = \{\phi \mid \langle \Sigma, \Lambda \rangle =\!\mid_L \phi\},$$

then

$$\Lambda = \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-), \tag{23}$$

for every $\phi \in \mathcal{L}_{ae}$,

$$SB_{\langle \Sigma, \Lambda \rangle}(\phi) = \Pi_{Sf^L(\phi)}(L\Delta^+ \cup \neg L\Delta^-) \tag{24}$$

and $\langle \Delta^+, \Delta^- \rangle = CAE_\Sigma(\Delta^+, \Delta^-)$.

Proof. We first prove (23). If $L\chi \in \Lambda$, then $\langle \Sigma, \Lambda \rangle \models_L \chi$ holds by the Σ-fullness of Λ. Then $\chi \in \Delta^+$ and $L\chi \in L\Delta^+$. Since $L\chi \in Sf^L(\Sigma)$, $L\chi \in \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-)$. Conversely, if $L\chi \in \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-)$, then $L\chi \in \Delta^+$ and $L\chi \in Sf^L(\Sigma)$. So $\chi \in \Delta^+$ and hence $\langle \Sigma, \Lambda \rangle \models_L \chi$. Since $L\chi \in Sf^L(\Sigma)$ and Λ is a Σ-full set, $L\chi \in \Lambda$. Similarly, $\neg L\chi \in \Lambda$ if and only if $\neg L\chi \in \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-)$.

Then (24) is proved. If $L\chi \in SB_{\langle \Sigma, \Lambda \rangle}(\phi)$, then $L\chi \in Sf^L(\phi)$. It follows that $\langle \Sigma, \Lambda \rangle \models_L \chi$, $\chi \in \Delta^+$, $L\chi \in L\Delta^+$ and $L\chi \in \Pi_{Sf^L(\phi)}(L\Delta^+ \cup \neg L\Delta^-)$. Conversely, if $L\chi$ belongs to $\Pi_{Sf^L(\phi)}(L\Delta^+ \cup \neg L\Delta^-)$, then $L\chi \in L\Delta^+$ and $L\chi \in Sf^L(\phi)$. So $\chi \in \Delta^+$ and hence $\langle \Sigma, \Lambda \rangle \models_L \chi$. Since $L\chi \in Sf^L(\phi)$, $L\chi \in SB_{\langle \Sigma, \Lambda \rangle}(\phi)$. Similarly, $\neg L\chi \in SB_{\langle \Sigma, \Lambda \rangle}(\phi)$ if and only if $\neg L\chi$ belongs to $\Pi_{Sf^L(\phi)}(L\Delta^+ \cup \neg L\Delta^-)$.

In the following, it is shown that

$$\phi \in \Delta^+ \text{ if and only if } \phi \in IP^+_\Sigma(\Delta^+, \Delta^-). \tag{25}$$

($\Rightarrow$) Let $\phi \in \Delta^+$. Then $\langle \Sigma, \Lambda \rangle \models_L \phi$, i.e. $\Sigma \cup \Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi) \models_{ae} \phi$. By (23) and (24), $\Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi) \subseteq L\Delta^+ \cup \neg L\Delta^-$. By the monotonicity of $\models_{ae}$, $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \models_{ae} \phi$ and $\phi \in IP^+_\Sigma(\Delta^+, \Delta^-)$. ($\Leftarrow$) Let $\phi \in IP^+_\Sigma(\Delta^+, \Delta^-)$ hold, i.e. $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \models_{ae} \phi$. The set $L\Delta^+ \cup \neg L\Delta^-$ is consistent on the basis of Proposition 2.18. By Proposition 2.13 there is a set Λ' such that $\Sigma \cup \Lambda' \models_{ae} \phi$ and

$$\begin{array}{rcll}
\Lambda' & = & \Pi_{Sf^{qL}(\Sigma \cup \{\phi\})}(L\Delta^+ \cup \neg L\Delta^-) & [\text{Proposition 2.13}] \\
& \subseteq & \Pi_{Sf^L(\Sigma \cup \{\phi\})}(L\Delta^+ \cup \neg L\Delta^-) & [\scriptstyle Sf^{qL}(\Sigma \cup \{\phi\}) \subseteq Sf^L(\Sigma \cup \{\phi\})] \\
& = & \Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi). & [(23) \text{ and } (24)]
\end{array}$$

Thus by the monotonicity of $\models_{ae}$, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \phi$. It follows by Definition 2.16 that $\langle\Sigma,\Lambda\rangle \models_L \phi$ and $\phi \in \Delta^+$.

Finally, it is shown that

$$\phi \in \Delta^- \text{ if and only if } \phi \in IP^-_\Sigma(\Delta^+,\Delta^-). \tag{26}$$

Notice that

$$\begin{array}{rl} & (\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)) \cap \{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\phi\})\} \\ = & \{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\phi\})\} - (\Delta^+ \cup \Delta^-) \\ = & \{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\phi\}) - \text{q-atom}(L\Delta^+ \cup \neg L\Delta^-)\} \\ = & \{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\phi\}) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))\} \end{array} \tag{27}$$

by (23) and (24). ($\Rightarrow$) Let $\phi \in \Delta^-$. Then $\langle\Sigma,\Lambda\rangle \dashv_L \phi$, i.e. there is no enumeration ε of

$$\{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\phi\}) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))\}, \tag{28}$$

such that $\phi \in EE^\varepsilon(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$. Assume that $\phi \notin IP^-_\Sigma(\Delta^+,\Delta^-)$. Then there is an enumeration ε' of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ such that ϕ belongs to $EE^{\varepsilon'}(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$. The set $L\Delta^+ \cup \neg L\Delta^-$ is consistent by Proposition 2.18. By Equation (27) there is a reduction ε'' of ε' with respect to $Sf^L(\Sigma \cup \{\phi\})$ such that ε'' is an enumeration of (28). As

$$\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) = \Pi_{Sf^L(\Sigma\cup\{\phi\})}(L\Delta^+ \cup \neg L\Delta^-)$$

Proposition 2.15 yields that $\phi \in EE^{\varepsilon''}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$, a contradiction. Thus $\phi \in IP^-_\Sigma(\Delta^+,\Delta^-)$. ($\Leftarrow$) Let $\phi \in IP^-_\Sigma(\Delta^+,\Delta^-)$, i.e. there is no enumeration ε of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ such that $\chi \in EE^\varepsilon(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$. Let us assume that $\langle\Sigma,\Lambda\rangle \not\dashv_L \phi$. Then there is an enumeration ε' of (28) such that $\phi \in EE^{\varepsilon'}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$. By Equation (27) the enumeration ε' can be extended to an enumeration ε'' of $\mathcal{L}_{ae} - (\Delta^+ \cup \Delta^-)$ (by simply concatenating the rest of the elements to ε') such that ε' is the reduction of ε'' with respect to $Sf^L(\Sigma \cup \{\phi\})$. By Proposition 2.15 $\phi \in EE^{\varepsilon''}(\Sigma \cup L\Delta^+ \cup \neg L\Delta^-)$, a contradiction. So $\langle\Sigma,\Lambda\rangle \dashv_L \phi$ and $\phi \in \Delta^-$.

By combining the results in Equations (25) and (26)

$$\langle\Delta^+,\Delta^-\rangle = CAE_\Sigma(\Delta^+,\Delta^-).$$

□

The major result of this section is stated in Theorem 2.22. Cautious expansions are in one-to-one correspondence with the Σ-full sets characterizing them.

Theorem 2.22 (Finitary Characterizability)
Let Σ be the set of premises, $\mathcal{CE}$ the set of cautious expansions of Σ and $\mathcal{FS}$ the set of Σ-full sets. Define the function $FS_\Sigma : \mathcal{CE} \to \mathcal{FS}$ from cautious expansions to Σ-full sets by

$$FS_\Sigma(\langle \Delta^+, \Delta^- \rangle) = \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-).$$

and the function $CE_\Sigma : \mathcal{FS} \to \mathcal{CE}$ from Σ-full sets to cautious expansions by

$$CE_\Sigma(\Lambda) = \langle \{\phi \mid \langle \Sigma, \Lambda \rangle \models_L \phi\}, \{\phi \mid \langle \Sigma, \Lambda \rangle \dashv_L \phi\} \rangle.$$

Then both FS_Σ and CE_Σ are bijections and inverses of each other, i.e. for a cautious expansion $\langle \Delta^+, \Delta^- \rangle$,

$$\langle \Delta^+, \Delta^- \rangle = CE_\Sigma(FS_\Sigma(\langle \Delta^+, \Delta^- \rangle)).$$

Proof. The two functions CE_Σ and FS_Σ are based on Lemmas 2.20 and 2.21, respectively. If Σ has a cautious expansion, then there is a Σ-full set Λ by Lemma 2.20. Conversely, if there is a Σ-full set Λ, then there is a cautious autoepistemic expansion by Lemma 2.21.

At first, we show that CE_Σ is injective. Let Λ_1 and Λ_2 be two Σ-full sets such that $CE_\Sigma(\Lambda_1) = CE_\Sigma(\Lambda_2)$. By Lemma 2.21, the two Σ-full sets may be retrieved from cautious expansions by reduction with respect to $Sf^L(\Sigma)$, i.e. $\Lambda_1 = \Pi_{Sf^L(\Sigma)}(CE_\Sigma(\Lambda_1))$ and $\Lambda_2 = \Pi_{Sf^L(\Sigma)}(CE_\Sigma(\Lambda_2))$. It follows that $\Lambda_1 = \Lambda_2$ and CE_Σ is injective. Furthermore, $FS_\Sigma(CE_\Sigma(\Lambda_1)) = FS_\Sigma(CE_\Sigma(\Lambda_2))$. This implies that FS_Σ is the inverse function of CE_Σ when restricted to the image $CE_\Sigma(\mathcal{FS}) \subseteq \mathcal{CE}$.

It remains to be shown that FS_Σ is injective or equivalently, CE_Σ is surjective. Let $\langle \Delta_1^+, \Delta_1^- \rangle$ and $\langle \Delta_2^+, \Delta_2^- \rangle$ be two cautious expansions such that

$$\Lambda = FS_\Sigma(\langle \Delta_1^+, \Delta_1^- \rangle) = FS_\Sigma(\langle \Delta_2^+, \Delta_2^- \rangle). \tag{29}$$

In the sequel, we prove that the two expansions equal, i.e. $\langle \Delta_1^+, \Delta_1^- \rangle = \langle \Delta_2^+, \Delta_2^- \rangle$. Let ϕ be any formula of $\mathcal{L}_{ae}$. We use induction on $BD(\chi)$ to show that

$$\phi \in \Delta_1^+ \text{ if and only if } \phi \in \Delta_2^+ \tag{30}$$

and

$$\phi \in \Delta_1^- \text{ if and only if } \phi \in \Delta_2^- \tag{31}$$

or equivalently,

$$\phi \notin \Delta_1^- \text{ if and only if } \phi \notin \Delta_2^-. \tag{32}$$

Induction base: $BD(\phi) = 0$. We begin with Equation (30). ($\Rightarrow$) Let $\phi \in \Delta_1^+$. By Definition 2.10, $\Sigma \cup L\Delta_1^+ \cup \neg L\Delta_1^- \models_{ae} \phi$. The set $L\Delta_1^+ \cup \neg L\Delta_1^-$ is consistent by Proposition 2.11. An application of Proposition 2.13 yields that there is a set

$$\begin{array}{rcll} \Lambda' & = & \Pi_{Sf^{qL}(\Sigma\cup\{\phi\})}(L\Delta_1^+ \cup \neg L\Delta_1^-) & \text{[Proposition 2.13]} \\ & \subseteq & \Pi_{Sf^L(\Sigma)}(L\Delta_1^+ \cup \neg L\Delta_1^-) & [Sf^{qL}(\Sigma\cup\{\phi\})\subseteq Sf^L(\Sigma)] \\ & = & \Pi_{Sf^L(\Sigma)}(L\Delta_2^+ \cup \neg L\Delta_2^-) & \text{[Equation (29)]} \\ & \subseteq & L\Delta_2^+ \cup \neg L\Delta_2^- & \text{[Lemma 2.21]} \end{array}$$

such that $\Sigma \cup \Lambda' \models_{ae} \phi$. Then $\Sigma \cup L\Delta_2^+ \cup \neg L\Delta_2^- \models_{ae} \phi$ follows by the monotonicity of $\models_{ae}$. Hence, $\phi \in \Delta_2^+$. ($\Leftarrow$) The converse follows by symmetry.

Then consider Equation (32). ($\Rightarrow$) Let $\phi \notin \Delta_1^-$. By Definition 2.10 there is an enumeration ε_1 of $\mathcal{L}_{ae} - (\Delta_1^+ \cup \Delta_1^-)$ such that $\phi \in EE^{\varepsilon_1}(\Sigma \cup L\Delta_1^+ \cup \neg L\Delta_1^-)$. By Equation (27) ε_1 has a reduction ε with respect $Sf^L(\Sigma)$ such that ε is an enumeration of $\{\psi \mid L\psi \in Sf^L(\Sigma) - \text{q-atom}(\Lambda)\}$. Proposition 2.15 yields that $\phi \in EE^{\varepsilon}(\Sigma \cup \Lambda)$. The enumeration ε can be extended to an enumeration ε_2 of $\mathcal{L}_{ae} - (\Delta_2^+ \cup \Delta_2^-)$ such that ε is the reduction of ε_2 with respect to $Sf^L(\Sigma)$ on the basis of Equations (27) and (29). By Proposition 2.15 $\phi \in EE^{\varepsilon_2}(\Sigma \cup L\Delta_2^+ \cup \neg L\Delta_2^-)$ and $\phi \notin \Delta_2^-$. ($\Leftarrow$) The equivalence (32) follows by symmetry.

Inductive hypothesis:
Equations (30), (31) and (32) hold for ϕ having $BD(\phi) < k$.

Induction step: $BD(\phi) = k$. The first concern is to show that the subbeliefs of ϕ coincide in the both expansions, i.e.

$$\Pi_{Sf^L(\phi)}(L\Delta_1^+ \cup \neg L\Delta_1^-) = \Pi_{Sf^L(\phi)}(L\Delta_2^+ \cup \neg L\Delta_2^-) \tag{33}$$

holds. ($\subseteq$) Let $L\chi \in \Pi_{Sf^L(\phi)}(L\Delta_1^+ \cup \neg L\Delta_1^-)$. Then $L\chi \in Sf^L(\phi)$, $\chi \in \Delta_1^+$ and $BD(\chi) < k$. The inductive hypothesis yields that $\chi \in \Delta_2^+$. Because $L\chi$ belongs to $Sf^L(\phi)$, $L\chi \in \Pi_{Sf^L(\phi)}(L\Delta_2^+ \cup \neg L\Delta_2^-)$. Then let $\neg L\chi \in \Pi_{Sf^L(\phi)}(L\Delta_1^+ \cup \neg L\Delta_1^-)$ hold. It follows that $L\chi \in Sf^L(\phi)$, $\chi \in \Delta_1^-$ and $BD(\chi) < k$. Thus $\chi \in \Delta_2^-$ by the inductive hypothesis and $\neg L\chi$ belongs to $\Pi_{Sf^L(\phi)}(L\Delta_2^+ \cup \neg L\Delta_2^-)$. ($\supseteq$) This follows by symmetry. Notice also, that $SB_{\langle\Sigma,\Lambda\rangle}(\phi) = \Pi_{Sf^L(\phi)}(L\Delta_1^+ \cup \neg L\Delta_1^-)$ by Lemma 2.21.

Then consider Equation (30). ($\Rightarrow$) Let $\phi \in \Delta_1^+$. By Definition 2.10, it holds that $\Sigma \cup L\Delta_1^+ \cup \neg L\Delta_1^- \models_{ae} \phi$. The set $L\Delta_1^+ \cup \neg L\Delta_1^-$ is consistent by Proposition 2.11. By Proposition 2.13 there is a set

$$\begin{array}{rcll} \Lambda' & = & \Pi_{Sf^{qL}(\Sigma\cup\{\phi\})}(L\Delta_1^+ \cup \neg L\Delta_1^-) & \text{[Proposition 2.13]} \\ & \subseteq & \Pi_{Sf^L(\Sigma\cup\{\phi\})}(L\Delta_1^+ \cup \neg L\Delta_1^-) & [Sf^{qL}(\Sigma\cup\{\phi\})\subseteq Sf^L(\Sigma\cup\{\phi\})] \\ & = & \Pi_{Sf^L(\Sigma\cup\{\phi\})}(L\Delta_2^+ \cup \neg L\Delta_2^-) & \text{[Equations (29) and (33)]} \\ & \subseteq & L\Delta_2^+ \cup \neg L\Delta_2^- & \text{[Lemma 2.21]} \end{array}$$

such that $\Sigma \cup \Lambda' \models_{ae} \phi$. It follows that $\Sigma \cup L\Delta_2^+ \cup \neg L\Delta_2^- \models_{ae} \phi$ by the monotonicity of $\models_{ae}$. Therefore, $\phi \in \Delta_2^+$.($\Leftarrow$) The equivalence in (30) follows by symmetry.

Equation (32) is the remaining one. ($\Rightarrow$) So let $\phi \notin \Delta_1^-$. By Definition 2.10 there is an enumeration ε_1 of $\mathcal{L}_{ae} - (\Delta_1^+ \cup \Delta_1^-)$ such that ϕ belongs to $EE^{\varepsilon_1}(\Sigma \cup L\Delta_1^+ \cup \neg L\Delta_1^-)$. The enumeration ε_1 has a reduction ε with respect $Sf^L(\Sigma \cup \{\phi\})$ such that ε is an enumeration of

$$\{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\phi\}) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))\}$$

by Equation (27). Proposition 2.15 yields that $\phi \in EE^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$. The enumeration ε can be extended to an enumeration ε_2 of $\mathcal{L}_{ae} - (\Delta_2^+ \cup \Delta_2^-)$ such that ε is the reduction of ε_2 with respect to $Sf^L(\Sigma\{\phi\})$ by Equations (27), (29) and (33). Then $\phi \in EE^{\varepsilon_2}(\Sigma \cup L\Delta_2^+ \cup \neg L\Delta_2^-)$ by Proposition 2.15. Therefore, $\phi \notin \Delta_2^-$. ($\Leftarrow$) Again, the equivalence (32) follows by symmetry.

Equations (30) and (31) imply the equality of the two cautious expansions. Then the function FS_Σ is injective and both CE_Σ and FS_Σ are bijections and inverses of each other. Moreover, the sets $\mathcal{FS}$ and $\mathcal{CE}$ are in one-to-one correspondence and their cardinalities are equal, i.e. $|\mathcal{FS}| = |\mathcal{CE}|$. $\Box$

We may now use Σ-full sets for reasoning about cautious expansions and testing what kind of expansions there exist given the set of premises. In Section 2.4 we give examples of cautious expansions.

2.4 Examples of Cautious Expansions

The purpose of this subsection is to demonstrate reasoning in CAEL. The cautious expansions are compared with stable and $\mathcal{N}$-expansions on the basis of the existence of expansions. By Theorem 2.22 cautious expansions are in one-to-one correspondence with full sets. Therefore we need to consider the consistent subsets of $Sf^L(\Sigma) \cup \neg Sf^L(\Sigma)$ only and check whether Σ-fullness conditions are satisfied. We provide tables in which existing full sets have been marked with "•" for each of the variants under examination. The Σ-fullness conditions are checked for cautious expansions only. By the results of Niemelä and Gottlob the full sets are applicable to stable and $\mathcal{N}$-expansions as well (the recursive consequence relations differ). The examples we present are rather simple, but in the essence they point out the dissimilarities of the expansions under study.

Example 1 (Positive Loop)
The set of premises $\Sigma = \{Lp \rightarrow p\}$.

	$\emptyset$	$\{Lp\}$	$\{\neg Lp\}$
Stable		•	•
$\mathcal{N}$			•
Cautious		•	•

Table 1: Expansions of $\{Lp \rightarrow p\}$

In Example 1 there is a stable expansion corresponding to the full set $\{Lp\}$. This expansion is weakly grounded in the set of premises. The reason why p is believed is that p is believed. This kind of a situation has been considered counterintuitive in the literature. Thus Marek and Truszczyński proposed $\mathcal{N}$-expansions, requiring that positive beliefs depend on the negative ones only. There is only a single $\mathcal{N}$-expansion – the one corresponding to the full set $\{\neg Lp\}$. By Proposition 2.12 stable expansions are also cautious expansions. So there are weakly grounded cautious expansions as well. In this example, the weakened introspection mechanism has no effect – cautious expansions are exactly the same.

The set $\Lambda = \emptyset$ is not Σ-full. Assume Λ is Σ-full. Then as $\neg Lp \notin \Lambda$, it must be the case that $\langle \Sigma, \Lambda \rangle \not\dashv_L p$. By the definition of $\dashv_L$, there is an enumeration ε of $\{p\}$ such that $p \in EE^{\varepsilon}(\Sigma)$. The only possible enumeration of $\{p\}$ is $\varepsilon_1 = p$. Then $B^{\varepsilon_1}(\Sigma) = \{\neg Lp\}$. However, $\Sigma \cup B^{\varepsilon_1}(\Sigma) \not\models_{ae} p$. Thus, $p \notin EE^{\varepsilon_1}(\Sigma \cup \emptyset)$, which is a contradiction.

The set $\Lambda = \{Lp\}$ is Σ-full. Now $Lp \in \Lambda$. Because $SB_{\langle \Sigma, \Lambda \rangle}(p) = \emptyset$ and $\Sigma \cup \{Lp\} \models_{ae} p$, $\langle \Sigma, \Lambda \rangle \models_L p$ holds.

The set $\Lambda = \{\neg Lp\}$ is Σ-full. Now $\neg Lp \in \Lambda$. We have to consider the empty enumeration λ only. Then trivially $B^{\lambda}(\Sigma \cup \{\neg Lp\}) = \emptyset$. As $\Sigma \cup \{\neg Lp\} \not\models_{ae} p$, it follows that $p \notin EE^{\lambda}(\Sigma \cup \{\neg Lp\})$. Thus $\langle \Sigma, \Lambda \rangle \dashv_L p$.

Example 2 (Negative Loop)
The set of premises $\Sigma = \{\neg Lp \rightarrow p\}$.

	$\emptyset$	$\{Lp\}$	$\{\neg Lp\}$
Stable			
$\mathcal{N}$			
Cautious	•		

Table 2: Expansions of $\{\neg Lp \rightarrow p\}$

Example 2 is tricky. There are no stable expansions and therefore no $\mathcal{N}$-expansions. If the agent believes p, then it cannot infer p. But then it should disbelieve p, which leads to contradiction. On the other hand, if the agent disbelieves p, it can infer p. Therefore it should believe p and again the agent becomes inconsistent. Adding a formula of the form above to a set of premises – provided that p is a fresh proposition – prevents the agent from reaching a stable state of beliefs, i.e. there are no stable expansions for such sets of premises. This is a rather unpleasant property of stable expansions. To the contrary, there is a cautious expansion. The belief status of p remains undetermined. This example demonstrates how the weakened negative introspection comes into play by keeping the beliefs of the agent Σ-acceptable and realizing a stable state of beliefs of the agent.

The set $\Lambda = \emptyset$ is Σ-full. Because $\Sigma \cup \emptyset \not\models_{ae} p$ and $BD(p) = 0$, $\langle \Sigma, \Lambda \rangle \not\models_L p$. It holds also that $\neg Lp \notin \Lambda$. We need to consider the enumeration $\varepsilon_1 = p$ only. The enumeration-based introspection yields $B^{\varepsilon_1}(\Sigma) = \{\neg Lp\}$. However, $\Sigma \cup B^{\varepsilon_1}(\Sigma) \models_{ae} p$ and so $p \in EE^{\varepsilon_1}(\Sigma)$. Hence $\langle \Sigma, \Lambda \rangle \not\dashv_L p$.

The set $\Lambda = \{Lp\}$ is not Σ-full. Assume that Λ is Σ-full. Then $Lp \in \Lambda$ implies $\langle \Sigma, \Lambda \rangle \models_L p$. Since p does not have any subbeliefs, $\Sigma \cup \{Lp\} \models_{ae} p$, a contradiction.

The set $\Lambda = \{\neg Lp\}$ is not Σ-full. Assume that Λ is Σ-full. Then $\neg Lp \in \Lambda$ implies $\langle \Sigma, \Lambda \rangle \dashv_L p$. Since recursive consequence relations are mutually exclusive, $\langle \Sigma, \Lambda \rangle \not\models_L p$. As $SB_{\langle \Sigma, \Lambda \rangle}(p) = \emptyset$, $\Sigma \cup \{\neg Lp\} \not\models_{ae} p$, a contradiction.

Example 3
The set of premises $\Sigma = \{\neg L\neg p \rightarrow p\}$.

	$\emptyset$	$\{L\neg p\}$	$\{\neg L\neg p\}$
Stable			$\bullet$
$\mathcal{N}$			$\bullet$
Cautious			$\bullet$

Table 3: Expansions of $\{\neg L\neg p \rightarrow p\}$

The set $\Lambda = \emptyset$ is not Σ-full. Assume that Λ is Σ-full. Then $\langle \Sigma, \Lambda \rangle \not\dashv_L \neg p$ implies there is an enumeration ε of $\{\neg p\}$ such that $\neg p \in EE^{\varepsilon}(\Sigma \cup \emptyset)$. The only possible enumeration of $\{\neg p\}$ is $\varepsilon_1 = \neg p$. Then $B^{\varepsilon_1}(\Sigma) = \{\neg L\neg p\}$. Since $\Sigma \cup B^{\varepsilon_1}(\Sigma) \not\models_{ae} \neg p$, $\neg p \notin EE^{\varepsilon_1}(\Sigma)$, a contradiction

The set $\Lambda = \{L\neg p\}$ is not Σ-full. Assume that Λ is Σ-full. Then $L\neg p \in \Lambda$ implies $\langle \Sigma, \Lambda \rangle \models_L \neg p$. Because $SB_{\langle \Sigma, \Lambda \rangle}(\neg p) = \emptyset$ and $\Sigma \cup \{L\neg p\} \models_{ae} \neg p$, a contradiction results.

The set $\Lambda = \{\neg L\neg p\}$ is Σ-full. Now $\neg L\neg p \in \Lambda$. We need to consider the empty enumeration λ only. Trivially $B^{\lambda}(\Sigma \cup \{\neg L\neg p\}) = \emptyset$. Since $\neg p$ is not a logical consequence of $\Sigma \cup \{\neg L\neg p\}$, $\neg p \notin EE^{\lambda}(\Sigma \cup \{\neg L\neg p\})$. Therefore, $\langle \Sigma, \Lambda \rangle \dashv_L \neg p$ holds.

Example 4
The set of premises $\Sigma = \{\neg Lp \rightarrow q, \neg Lq \rightarrow p, \neg Lp \wedge \neg Lq \rightarrow r\}$.

	$\emptyset$	$\{\neg Lp, Lq\}$	$\{Lp, \neg Lq\}$
Stable		•	•
$\mathcal{N}$		•	•
Cautious	•	•	•

Table 4: Expansions of $\{\neg Lp \rightarrow q, \neg Lq \rightarrow p, \neg Lp \wedge \neg Lq \rightarrow r\}$

The table is not complete, only the interesting candidates for Σ-full sets have been presented. This example has two stable and $\mathcal{N}$-expansions representing two complementary states of beliefs. There is an additional cautious expansion corresponding to the full set $\emptyset$. The belief status of both p and q is left undetermined, since there are two competing choices. An interesting feature of the cautious reasoning is the fact that $\neg Lr$ belongs to the cautious expansion corresponding to the full set $\Lambda = \emptyset$. Though there are the two competing choices, p and q are never disbelieved simultaneously. Therefore, it is reasonable to disbelieve r.

The set $\Lambda = \emptyset$ is Σ-full. Now $Lp \notin \Lambda$. Since $BD(p) = 0$ and $\Sigma \cup \emptyset \models_{ae} p$, $\langle \Sigma, \Lambda \rangle \not\models_L p$. It holds also that $\neg Lp \notin \Lambda$. Consider the enumeration $\varepsilon_1 = q, p$. Then $B^{\varepsilon_1}(\Sigma) = \{\neg Lq, Lp\}$ and $\Sigma \cup B^{\varepsilon_1}(\Sigma) \models_{ae} p$. Thus $p \in EE^{\varepsilon_1}(\Sigma)$ and ε_1 is a refuting enumeration. Therefore, $\langle \Sigma, \Lambda \rangle \not\dashv_L p$. The cases that $Lq \notin \Lambda$ implies $\langle \Sigma, \Lambda \rangle \not\models_L q$ and $\neg Lq \notin \Lambda$ implies $\langle \Sigma, \Lambda \rangle \not\dashv_L q$ follow by symmetry.

The set $\Lambda = \{Lp, \neg Lq\}$ is Σ-full. Now $Lp \in \Lambda$. Since $BD(p) = 0$ and $\Sigma \cup \{Lp, \neg Lq\} \models_{ae} p$, it holds that $\langle \Sigma, \Lambda \rangle \models_L p$. Consider next $\neg Lq \in \Lambda$. We need to check the empty enumeration λ only. Trivially $B^{\lambda}(\Sigma \cup \{Lp, \neg Lq\}) = \emptyset$. As $\Sigma \cup \{Lp, \neg Lq\} \not\models_{ae} q$, $q \notin EE^{\lambda}(\Sigma \cup \{Lp, \neg Lq\})$. Hence $\langle \Sigma, \Lambda \rangle \dashv_L q$.

The set $\Lambda = \{\neg Lp, Lq\}$ is Σ-full. This follows by the symmetry from the former one.

3 Grounded Cautious Expansions

Reconsider the case when there are several autoepistemic expansions of the given set of premises. According to the cautious reasoning strategy, the intersection of the expansions should be determined. In this section we study the properties of the intersection of cautious expansions. We begin by defining a partial order on cautious expansions.

Definition 3.1 (Partial order on cautious expansions)
Let $\mathcal{E}_1 = \langle \Delta_1^+, \Delta_1^- \rangle$ and $\mathcal{E}_2 = \langle \Delta_2^+, \Delta_2^- \rangle$ be cautious expansions. Then $\mathcal{E}_1$ is smaller than or equal to $\mathcal{E}_2$, denoted by $\mathcal{E}_1 \sqsubseteq \mathcal{E}_2$, if and only if $\Delta_1^+ \subseteq \Delta_2^+$ and $\Delta_1^- \subseteq \Delta_2^-$.

The relation $\sqsubseteq$ is straightforwardly reflexive, transitive and antisymmetric as $\subseteq$ is. Therefore, it is a partial order. In Definition 3.2 we make precise the notion of the least cautious expansion. If the least cautious expansion exists in a set of cautious expansions $\mathcal{CE}$, it is contained in every other cautious expansion. Thus it is the intersection of all cautious expansions in $\mathcal{CE}$ by the properties of subset inclusion. Our aim is to find out the conditions guaranteeing that the least cautious expansion exists and how to decide whether a given formula belongs to it. The formulae that belong to the least cautious expansion of the given set of premises are the conclusions that correspond to the cautious reasoning strategy.

Definition 3.2 (Least expansion)
Let $\mathcal{CE}$ be a set of cautious expansions. A cautious expansion $\mathcal{E} \in \mathcal{CE}$ is the least cautious expansion of $\mathcal{CE}$ if and only if $\mathcal{E} \sqsubseteq \mathcal{E}'$ for every $\mathcal{E}'$ in $\mathcal{CE}$.

The uniqueness of the least cautious expansion follows essentially from the antisymmetry of $\sqsubseteq$. Suppose there are two least cautious expansions $\mathcal{E}_1$ and $\mathcal{E}_2$ in $\mathcal{CE}$. It follows by Definition 3.2 that $\mathcal{E}_1 \sqsubseteq \mathcal{E}_2$ and $\mathcal{E}_2 \sqsubseteq \mathcal{E}_1$. Since $\sqsubseteq$ is an antisymmetric relation, $\mathcal{E}_1 = \mathcal{E}_2$, i.e. the two expansions are equal.

Let us now reconsider our examples. In Example 1 the cautious expansions corresponding to the Σ-full sets $\{Lp\}$ and $\{\neg Lp\}$ are exceptional. The two expansions are incomparable by means of the relation $\sqsubseteq$ – both expansions contain formulae that are not contained in the other. It follows that the cautious expansions of the set of premises $\{Lp \rightarrow p\}$ do not have the least element. The cautious expansion corresponding to the Σ-full set $\{Lp\}$ is self-supporting. We show that if this kind of cautious expansions are ruled out, the remaining cautious expansions do have the least element which is also their intersection.

The elimination is carried out by imposing a special groundedness condition on cautious expansions. The groundedness requirement of cautious expansions is similar to that of $\mathcal{N}$-expansions. Therefore, we call it $\mathcal{N}$-groundedness. The modal logic $\mathcal{N}$ is used for positive introspection. As a consequence of Definition 3.3, we have defined a subclass of cautious expansions: the $\mathcal{N}$-grounded cautious expansions.

Definition 3.3 ($\mathcal{N}$-Groundedness)
Let $\langle \Delta^+, \Delta^- \rangle$ be a cautious autoepistemic expansion. If it also satisfies the equation

$$\Delta^+ = Cn_{\mathcal{N}}(\Sigma \cup \neg L\Delta^-),$$

then the cautious expansion is $\mathcal{N}$-grounded.

3.1 Finitary Groundedness Condition

The full sets introduced in Subsection 2.3 form our basic tool for reasoning about cautious expansions. Therefore, our interest is to find out how the $\mathcal{N}$-groundedness property is reflected in the finitary characterization. Gottlob [1992a] has provided a groundedness requirement for full sets characterizing $\mathcal{N}$-expansions. It turns out, that this condition is also applicable in the framework of the cautious expansions.

Definition 3.4 (Finitary $\mathcal{N}$-groundedness)
Let Λ be a Σ-full set. It is $\mathcal{N}$-grounded if and only if it satisfies for all $L\chi \in Sf^L(\Sigma)$

$$L\chi \in \Lambda \text{ if and only if } \Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi.$$

The latter groundedness requirement resembles the former: the positive part $[\Lambda]^+$ of the Σ-full set Λ is derivable from the negative part $[\Lambda]^-$ and the premises only, i.e. $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} [\Lambda]^+$. A set of proof technical tools is needed in the proofs of this section. They are presented in Lemmata 3.5-3.8. The essential content of the lemmata is the relation between the enumerations used for introspection and the derivability relation $\vdash_{\mathcal{N}}$. We say that a set of quasi-literals Γ is *generated by an enumeration* from a set of premises Σ if the set can be built by carrying introspection on Σ in an order determined by an enumeration ε of $\{\psi \mid L\psi \in \text{q-atom}(\Gamma)\}$, i.e. $\Gamma = B^\varepsilon(\Sigma)$. The fundamental property of $\mathcal{N}$-grounded full sets is that they can be generated by enumerations from the set of premises Σ.

Lemma 3.5
Let Σ be a set of formulae of $\mathcal{L}_{ae}$ and Λ an $\mathcal{N}$-grounded Σ-full set. Then there is an enumeration ε of $\{\psi \mid L\psi \in$ q-atom$(\Lambda)\}$ such that $B^\varepsilon(\Sigma) = \Lambda$.

Proof. At first, consider any $\neg L\chi \in \Lambda$. It follows that $\langle\Sigma,\Lambda\rangle \dashv_L \chi$ by the definition of Σ-full sets. In Proposition 2.18 it was proved that the recursive consequence relations are mutually exclusive. Therefore, $\langle\Sigma,\Lambda\rangle \not\models_L \chi$, i.e. $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \not\models_{ae} \chi$. Then $\Sigma \cup \Lambda \not\models_{ae} \chi$ follows, because $\models_{ae}$ is monotonic. So we can conclude that

$$\neg L\chi \in \Lambda \text{ implies } \Sigma \cup \Lambda \not\models_{ae} \chi. \tag{34}$$

Let $\varepsilon^- = \psi_1, ..., \psi_m$ be any enumeration of $\{\psi \mid \neg L\psi \in \Lambda\}$. We then use induction on i to show that for all $1 \leq i \leq m$

$$B_i^{\varepsilon^-}(\Sigma) = \{\neg L\psi_1, ..., \neg L\psi_i\}. \tag{35}$$

Induction base: $i = 0$. As $\neg L\psi_1 \in \Lambda$, $\Sigma \cup \Lambda \not\models_{ae} \psi_1$ by Equation (34). This implies $\Sigma \not\models_{ae} \psi_1$, because $\models_{ae}$ is monotonic. So $\neg L\psi_1$ is in $B_1^{\varepsilon^-}(\Sigma)$. Inductive hypothesis: Equation (35) holds for $i < m$. Induction step: $i = m$. Because $\neg L\psi_m \in \Lambda$, $\Sigma \cup \Lambda \not\models_{ae} \psi_m$. By inductive hypothesis $B_{m-1}^{\varepsilon^-}(\Sigma) = \{\neg L\psi_1, ..., \neg L\psi_{m-1}\}$ which clearly is a subset of Λ. Since $\models_{ae}$ is monotonic, $\Sigma \cup B_{m-1}^{\varepsilon^-}(\Sigma) \not\models_{ae} \psi_m$ and $\neg L\psi_m \in B_m^{\varepsilon^-}(\Sigma)$. It follows that $[\Lambda]^- = B^{\varepsilon^-}(\Sigma)$.

We then consider the positive belief literals in Λ and show that they can be generated by an enumeration, too. As Λ is an $\mathcal{N}$-grounded Σ-full set, it holds for each $L\chi \in \Lambda$ that $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi$ by Definition 3.4. Proposition 2.14 yields that a minimal L-proof of χ from $\Sigma \cup [\Lambda]^-$ contains only members of $Sf^L(\Sigma \cup \{\chi\})$. Let ε_j^+ be an enumeration of

$$\{\psi \mid L\psi \in \Lambda \text{ and } \psi \text{ has a minimal L-proof of length } j\}$$

and $0 \leq j \leq k$ where k is the length of the longest minimal L-proof of any formula ψ such that $L\psi \in \Lambda$. The concatenation of these enumerations yields a composite enumeration $\varepsilon^+ = \psi'_1, ..., \psi'_n = \varepsilon_0^+ \cdot ... \cdot \varepsilon_k^+$ which is an enumeration of $\{\psi \mid L\psi \in \Lambda\}$. We claim that

$$B_i^{\varepsilon^+}(\Sigma \cup [\Lambda]^-) = \{L\psi'_1, ..., L\psi'_i\} \tag{36}$$

for values of i, $1 \leq i \leq n$. To show this, we use induction on i once more. Induction base: $i = 1$. At first, we show that ψ'_1 belongs always to ε_0^+. Assume that ψ'_1 has a minimal L-proof of length greater that zero. Then there is an element ψ' such that $L\psi'$ belongs to the L-proof of ψ'_1. By proposition 2.14, $L\psi' \in Sf^L(\Sigma \cup \{\psi'_1\}) = Sf^L(\Sigma)$ as a member of the minimal L-proof of ψ'_1. It follows that $L\psi' \in \Lambda$, as ψ' has an L-proof from $\Sigma \cup [\Lambda]^-$ and Λ is an $\mathcal{N}$-grounded Σ-full set. But then ψ' has a shorter minimal L-proof than that of ψ'_1 and certainly appears before ψ'_1 in ε^+, a contradiction. Hence ψ'_1 belongs to the enumeration ε_0^+. The element ψ'_1 is the first element of ε_0^+, because it is the first element of ε^+. As ψ'_1 has an L-proof of length 0, $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \psi'_1$ implies $\Sigma \cup [\Lambda]^- \models_{ae} \psi'_1$. Therefore

$L\psi_1' \in B_1^{\varepsilon^+}(\Sigma \cup [\Lambda]^-)$. Inductive hypothesis: Equation (36) holds for $i < l$. Induction step: $i = l$. As $L\psi_l' \in \Lambda$, it satisfies $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \psi_l'$. Let ε_p^+ be the enumeration to which ψ_l' belongs. Then $L\psi_l'$ has a minimal L-proof of length p. Let $L\phi_1, ..., L\phi_p$ be the minimal L-proof of ψ_l' from $\Sigma \cup [\Lambda]^-$, i.e. $\Sigma \cup [\Lambda]^- \cup \{L\phi_1, ..., L\phi_p\} \models_{ae} \psi_l'$. As a member of the minimal L-proof of ψ_l', each $L\phi_j$ belongs to $Sf^L(\Sigma \cup \{\psi_l'\}) = Sf^L(\Sigma)$. Furthermore, each ϕ_j satisfies $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \phi_j$. It follows that $\{L\phi_1, ..., L\phi_p\} \subseteq \Lambda$. Each ϕ_j has an L-proof shorter than l so they appear before both ε_p^+ and ψ_l' in ε^+. Therefore, by the inductive hypothesis, $\{L\phi_1, ..., L\phi_p\} \subseteq B_{l-1}^{\varepsilon^+}(\Sigma \cup [\Lambda]^-)$. It follows that $\Sigma \cup [\Lambda]^- \cup B_{l-1}^{\varepsilon^+}(\Sigma \cup [\Lambda]^-) \models_{ae} \psi_l'$, by the monotonicity of $\models_{ae}$. Therefore, $L\psi_l' \in B_l^{\varepsilon^+}(\Sigma \cup [\Lambda]^-)$. This completes the inductive step, yielding $[\Lambda]^+ = B^{\varepsilon^+}(\Sigma \cup [\Lambda]^-)$.

Now $\Lambda = [\Lambda]^- \cup [\Lambda]^+ = [\Lambda]^- \cup B^{\varepsilon^+}(\Sigma \cup [\Lambda]^-) = B^{\varepsilon^-}(\Sigma) \cup B^{\varepsilon^+}(\Sigma \cup B^{\varepsilon^-}(\Sigma)) = B^{\varepsilon^- \cdot \varepsilon^+}(\Sigma)$. Therefore, the catenation of the two enumerations ε^+ and ε^-, $\varepsilon^- \cdot \varepsilon^+$ generates the $\mathcal{N}$-grounded Σ-full set Λ. $\Box$

The enumerations that generate $\mathcal{N}$-grounded Σ-full sets can be used as building blocks of L-proofs. To handle arbitrary formulae (rather than members of $Sf^L(\Sigma)$) by means of generating enumerations we need Lemma 3.6. It provides us a result that subbeliefs of formulae also have generating enumerations.

Lemma 3.6
Let Σ be a set of formulae of $\mathcal{L}_{ae}$ and ϕ any formula. If Λ is an $\mathcal{N}$-grounded Σ-full set, then there is an enumeration ε of $\{\psi \mid L\psi \in q\text{-}atom(SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda)\}$ such that $B^\varepsilon(\Sigma \cup \Lambda) = SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$.

Proof. Like before, we build an enumeration meeting the condition of the lemma explicitly. Let $\varepsilon = \psi_1, ..., \psi_n$ be an L-hierarchic enumeration of

$$\{\psi \mid L\psi \in \text{q-atom}(SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda)\}.$$

Clearly, each $L\psi_j$ is a member of $Sf^L(\phi)$. It is shown by induction on i that for each i, $1 \leq i \leq n$,

$$(\neg)L\psi_i \in B^\varepsilon(\Sigma \cup \Lambda) \text{ if and only if } (\neg)L\psi_i \in (SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda). \qquad (37)$$

Induction base: $i = 1$. We start by considering the subbeliefs of ψ_1. Let $L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\psi_1)$. It follows that $\langle\Sigma, \Lambda\rangle \models_L \chi$. Assume now that $L\chi \notin \Lambda$. Then $L\chi \notin Sf^L(\Sigma)$ and $L\chi \in Sf^L(\psi_1) \subseteq Sf^L(\phi)$. Therefore, $L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ and χ belongs to ε, i.e. $\chi = \psi_j$ for some j. Since ε is L-hierarchic and $L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\psi_1)$, $j < 1$ which is a contradiction. Therefore $L\chi \in \Lambda$. Similarly, $\neg L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\psi_1)$ implies that $\neg L\chi \in \Lambda$. Thus we have established that $SB_{\langle\Sigma,\Lambda\rangle}(\psi_1) \subseteq \Lambda$.

The equivalence in (37) is shown next. ($\Rightarrow$) Let $L\psi_1 \in B^\varepsilon(\Sigma \cup \Lambda)$. Then $\Sigma \cup \Lambda \models_{ae} \psi_1$. The monotonicity of $\models_{ae}$ yields $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\psi_1) \models_{ae} \psi_1$, i.e. $\langle\Sigma,\Lambda\rangle \models_L \psi_1$. As $L\psi_1 \in Sf^L(\phi)$, $L\psi_1 \in SB_{\langle\Sigma,\Lambda\rangle}(\phi)$. Furthermore, $L\psi_1 \notin \Lambda$, because ψ_1 is a member of ε. ($\Leftarrow$) Let $L\psi_1 \in SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$ hold. Then $\langle\Sigma,\Lambda\rangle \models_L \psi_1$ and $L\psi_1 \in Sf^L(\phi)$. By the definition of $\models_L$, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\psi_1) \models_{ae} \psi_1$. As $SB_{\langle\Sigma,\Lambda\rangle}(\psi_1)$ is a subset of Λ, it follows that $\Sigma \cup \Lambda \models_{ae} \psi_1$ and $L\psi_1 \in B^\varepsilon(\Sigma \cup \Lambda)$.

Then we consider the negated version of the equivalence. ($\Rightarrow$) Let $\neg L\psi_1 \in B^\varepsilon(\Sigma \cup \Lambda)$ hold. Assume that $\neg L\psi_1 \notin SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$. Then $L\psi_1$ belongs to $SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$ and $L\psi_1 \in B^\varepsilon(\Sigma \cup \Lambda)$, a contradiction and hence $\neg L\psi_1 \in SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$. ($\Leftarrow$) Similarly, $\neg L\psi_1 \in SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$ implies $\neg L\psi_1 \in B^\varepsilon(\Sigma \cup \Lambda)$.

Inductive hypothesis: Equation (37) holds for $i < k$.

Induction step: $i = k$. Let us then consider the subbeliefs of ψ_k. Let $L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\psi_k)$. It follows that $\langle\Sigma,\Lambda\rangle \models_L \chi$. Assume now that $L\chi \notin \Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$. Then $L\chi \notin Sf^L(\Sigma \cup \{L\psi_1, ..., L\psi_{k-1}\})$ and $L\chi \in Sf^L(\psi_k) \subseteq Sf^L(\phi)$. Therefore, $L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ and χ belongs to ε, i.e. $\chi = \psi_j$ for some j. Since ε is L-hierarchic and $L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\psi_k)$, $j < k$ which is a contradiction. Therefore, $L\chi \in \Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$. Similarly, $\neg L\chi \in SB_{\langle\Sigma,\Lambda\rangle}(\psi_k)$ implies that $\neg L\chi \in \Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$. So it has been established that $SB_{\langle\Sigma,\Lambda\rangle}(\psi_k) \subseteq \Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$.

($\Rightarrow$) Let $L\psi_k \in B^\varepsilon(\Sigma \cup \Lambda)$. Then $\Sigma \cup \Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda) \models_{ae} \psi_k$. The set $B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$ is a subset of $SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ by the inductive hypothesis. The set $\Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$ is a consistent set of beliefs literals as q-atom(Λ) and q-atom($B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$) do not have common elements by construction. Proposition 2.13 yields that there is a set

$$
\begin{array}{rcll}
\Lambda' & = & \Pi_{Sf^{qL}(\Sigma \cup \{\psi_k\})}(\Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)) & [\text{Proposition 2.13}] \\
& \subseteq & \Pi_{Sf^{L}(\Sigma \cup \{\psi_k\})}(\Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)) & [\scriptstyle Sf^{qL}(\Sigma \cup \{\psi_k\}) \subseteq Sf^L(\Sigma \cup \{\psi_k\})] \\
& \subseteq & \Pi_{Sf^{L}(\Sigma \cup \{\psi_k\})}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) & [\text{Inductive hypothesis}] \\
& = & \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\psi_k) & [\scriptstyle SB_{\langle\Sigma,\Lambda\rangle}(\psi_k) \subseteq SB_{\langle\Sigma,\Lambda\rangle}(\phi)]
\end{array}
$$

such that $\Sigma \cup \Lambda' \models_{ae} \psi_k$. It follows that $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\psi_k) \models_{ae} \psi_k$, i.e. $\langle\Sigma,\Lambda\rangle \models_L \psi_k$, by the monotonicity of $\models_{ae}$. As $L\psi_k \in Sf^L(\phi)$, $L\psi_k$ belongs to $SB_{\langle\Sigma,\Lambda\rangle}(\phi)$. Again, $L\psi_k \notin \Lambda$, because ψ_k is a member of ε. ($\Leftarrow$) Let $L\psi_k \in SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$. Then $L\psi_k \in Sf^L(\phi)$ and $\langle\Sigma,\Lambda\rangle \models_L \psi_k$, i.e. $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\psi_k) \models_{ae} \psi_k$. As $SB_{\langle\Sigma,\Lambda\rangle}(\psi_k)$ is contained in $\Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda)$, $\Sigma \cup \Lambda \cup B^\varepsilon_{k-1}(\Sigma \cup \Lambda) \models_{ae} \psi_k$. Hence, $L\psi_k \in B^\varepsilon(\Sigma \cup \Lambda)$.

Then consider the negated version of the equivalence. ($\Rightarrow$) Let $\neg L\psi_k \in B^\varepsilon(\Sigma \cup \Lambda)$ hold. Assume that $\neg L\psi_k \notin SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$. Then $L\psi_k$ belongs to $SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$ and $L\psi_k \in B^\varepsilon(\Sigma \cup \Lambda)$, a contradiction and hence

$\neg L\psi_k \in SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$. ($\Leftarrow$) Similarly, $\neg L\psi_k \in SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$ implies $\neg L\psi_k \in B^\varepsilon(\Sigma \cup \Lambda)$. The inductive case has therefore been handled, and we have shown that $B^\varepsilon(\Sigma \cup \Lambda) = SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$. □

Lemma 3.7 provides us the condition on which the difference of a set of quasi-literals and its arbitrary subset can be generated by means of an enumeration.

Lemma 3.7
Let Σ be a set of formulae of $\mathcal{L}_{ae}$, Λ a set of quasi-literals satisfying

$$\neg L\chi \in \Lambda \textit{ implies } \Sigma \cup \Lambda \not\models_{ae} \chi \tag{38}$$

and ε an enumeration of $\{\psi \mid L\psi \in q\text{-}atom(\Lambda)\}$ such that $\Lambda = B^\varepsilon(\Sigma)$. Then for every subset Λ' of Λ, there is an enumeration ε_d of

$$\{\psi \mid L\psi \in q\text{-}atom(\Lambda - \Lambda')\}$$

such that $\Lambda - \Lambda' = B^{\varepsilon_d}(\Sigma \cup \Lambda')$.

Proof. Let $\varepsilon = \psi_1, \dots, \psi_n$ be the enumeration generating the set of quasi-literals Λ from Σ. As Λ' is a subset of Λ, there is a reduction $\varepsilon_d = \psi'_1, \dots, \psi'_m$ of ε with respect to $\{\psi \mid L\psi \in \text{q-atom}(\Lambda - \Lambda')\}$. The order of the elements is preserved in the reduction.

In the sequel, we show by induction on i that for all $1 \leq i \leq m$,

$$(\neg)L\psi'_i \in B^{\varepsilon_d}(\Sigma \cup \Lambda') \text{ if and only if } (\neg)L\psi'_i \in \Lambda - \Lambda', \tag{39}$$

i.e. the difference of the sets Λ and Λ' is generated by the enumeration ε_d.

Induction base: $i = 0$. ($\Rightarrow$) Let $L\psi'_1 \in B^{\varepsilon_d}(\Sigma \cup \Lambda')$ hold. Then $\Sigma \cup \Lambda' \models_{ae} \psi'_1$. Since $\Lambda' \subseteq \Lambda$, $\Sigma \cup \Lambda \models_{ae} \psi'_1$ follows by the monotonicity of $\models_{ae}$. Assume that $L\psi'_1 \notin \Lambda$. As $L\psi'_1 \in \text{q-atom}(\Lambda)$, $\neg L\psi'_1 \in \Lambda$. This implies $\Sigma \cup \Lambda \not\models_{ae} \psi'_1$ by Equation (38). But this is a contradiction and thus $L\psi'_1 \in \Lambda$. As $L\psi'_1 \in \text{q-atom}(\Lambda - \Lambda')$, $L\psi'_1 \in \Lambda - \Lambda'$. ($\Leftarrow$) Let $L\psi'_1 \in \Lambda - \Lambda'$ and $\psi_j = \psi'_1$ be the corresponding element in the enumeration ε. As Λ is generated by the enumeration ε from Σ, $\Sigma \cup B^\varepsilon_{j-1}(\Sigma) \models_{ae} \psi_j$. In fact, j is the first index such that $L\psi_j \in \text{q-atom}(\Lambda - \Lambda')$. Therefore, the preceding elements generated by ε, namely $B^\varepsilon_{j-1}(\Sigma)$, are contained in Λ'. Thus $\Sigma \cup \Lambda' \models_{ae} \psi_j$ follows by the monotonicity of $\models_{ae}$. Because $\psi_j = \psi'_1$, $L\psi_1 \in B^{\varepsilon_d}(\Sigma \cup \Lambda')$. The negated version of the equivalence follows easily from the fact that $\text{q-atom}(B^{\varepsilon_d}(\Sigma \cup \Lambda'))$ equals to $\text{q-atom}(\Lambda - \Lambda')$.

Inductive hypothesis: Equation (39) holds for $i < k$.

Induction step: $i = k$. ($\Rightarrow$) Let $L\psi'_k \in B^{\varepsilon_d}(\Sigma \cup \Lambda')$ hold. It follows that $\Sigma \cup \Lambda' \cup B^{\varepsilon_d}_{k-1}(\Sigma \cup \Lambda') \models_{ae} \psi'_k$. The inductive hypothesis yields that $\Lambda' \cup$

$B^{\varepsilon_d}_{k-1}(\Sigma \cup \Lambda') \subseteq \Lambda$. Therefore, $\Sigma \cup \Lambda \models_{ae} \psi'_k$, since $\models_{ae}$ is monotonic. Assume that $L\psi'_k \notin \Lambda$. As $L\psi'_k \in$ q-atom(Λ), $\neg L\psi'_k \in \Lambda$. Equation (38) implies now that $\Sigma \cup \Lambda \not\models_{ae} \psi'_1$. But then we have derived a contradiction and $L\psi'_k \in \Lambda$. As $L\psi'_k \in$ q-atom$(\Lambda - \Lambda')$, $L\psi'_k \in \Lambda - \Lambda'$. ($\Leftarrow$) Let $L\psi'_k \in \Lambda - \Lambda'$ and $\psi_j = \psi'_k$ be the matching element in the enumeration ε. As Λ is generated by the enumeration ε from Σ, $\Sigma \cup B^{\varepsilon}_{j-1}(\Sigma) \models_{ae} \psi_j$. By the inductive hypothesis, $B^{\varepsilon}_{j-1}(\Sigma) \subseteq \Lambda' \cup B^{\varepsilon_d}_{k-1}(\Sigma \cup \Lambda')$. Then $\Sigma \cup \Lambda' \cup B^{\varepsilon_d}_{k-1}(\Sigma \cup \Lambda') \models_{ae} \psi'_k$ by the monotonicity of $\models_{ae}$. This implies $L\psi'_k \in B^{\varepsilon_d}(\Sigma \cup \Lambda')$ as desired. The negated version of the equivalence follows straightforwardly by the principle given in the induction base. We have therefore established that $B^{\varepsilon_d}(\Sigma \cup \Lambda') = \Lambda - \Lambda'$.
□

The last one of the four technical lemmas dictates the relationship between enumerations and L-proofs. If a set of quasi-literals Γ has a generating enumeration from Σ, then the positive belief literals $[\Gamma]^+$ serve as an L-proof from $\Sigma \cup [\Gamma]^-$. This is the general proof strategy that we are to take advantage of in several following proofs.

Lemma 3.8
Let Λ be a finite set of quasi-literals. If there is an enumeration ε of

$$\{\psi \mid L\psi \in q\text{-}atom(\Lambda)\}$$

such that $B^{\varepsilon}(\Sigma) = \Lambda$, then $L\chi \in \Lambda$ implies $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi$ and for every $\phi \in \mathcal{L}_{ae}$, $\Sigma \cup \Lambda \models_{ae} \phi$ implies $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \phi$.

Proof. Let $\varepsilon = \psi_1, ..., \psi_n$ be the enumeration generating Λ from Σ and ϕ any formula. In the following, we prove by induction on i that for all i, $1 \leq i \leq n$,

$$\Sigma \cup B^{\varepsilon}_i(\Sigma) \models_{ae} \phi \text{ implies } \Sigma \cup [B^{\varepsilon}_i(\Sigma)]^- \vdash_{\mathcal{N}} \phi \tag{40}$$

such that $[B^{\varepsilon}_i(\Sigma)]^+$ is an L-proof of ϕ from $\Sigma \cup [B^{\varepsilon}_i(\Sigma)]^-$. The order in which formulae of $[B^{\varepsilon}_i(\Sigma)]^+$ appear in the L-proof is determined by the enumeration ε. (This is also assumed in the rest of the proof).

Induction base: Let $\Sigma \models_{ae} \phi$ hold. Then ϕ may be derived from Σ without the necessitation rule, i.e. $[B^{\varepsilon}_0(\Sigma)]^+ = \emptyset$ is an empty L-proof of ϕ from $\Sigma \cup [B^{\varepsilon}_0(\Sigma)]^- = \Sigma$.

Inductive hypothesis: Equation (40) holds for $i < k$.

Induction step: $i = k$. Let $\Sigma \cup B^{\varepsilon}_k(\Sigma) \models_{ae} \phi$ hold. By separating the positive and negative belief literals we get

$$\Sigma \cup [B^{\varepsilon}_k(\Sigma)]^- \cup [B^{\varepsilon}_k(\Sigma)]^+ \models_{ae} \phi. \tag{41}$$

Then we separately consider the two possible cases. Firstly, if $B^{\varepsilon}_k(\Sigma) = B^{\varepsilon}_{k-1}(\Sigma) \cup \{L\psi_k\}$ then $\Sigma \cup B^{\varepsilon}_{k-1}(\Sigma) \models_{ae} \psi_k$. The inductive hypothesis may now be applied to ψ_k yielding that $[B^{\varepsilon}_{k-1}(\Sigma)]^+$ is an L-proof of ψ_k from $\Sigma \cup [B^{\varepsilon}_{k-1}(\Sigma)]^-$. Now $[B^{\varepsilon}_{k-1}(\Sigma)]^- = [B^{\varepsilon}_{k}(\Sigma)]^-$. As (41) holds, $[B^{\varepsilon}_{k}(\Sigma)]^+$ forms an L-proof of ϕ from $\Sigma \cup [B^{\varepsilon}_{k}(\Sigma)]^-$.

Secondly, if $B^{\varepsilon}_{k}(\Sigma) = B^{\varepsilon}_{k-1}(\Sigma) \cup \{\neg L\psi_k\}$, then $\Sigma \cup B^{\varepsilon}_{k-1}(\Sigma) \not\models_{ae} \psi_k$. Let j, $1 \leq j < k$ be the largest number such that $L\psi_j \in B^{\varepsilon}_{k-1}(\Sigma)$. We suppose now that such j exists, and consider later the opposite case. It follows that

$$\{\neg L\psi_{j+1}, ..., \neg L\psi_k\} \subseteq B^{\varepsilon}_{k}(\Sigma). \tag{42}$$

The inductive hypothesis is then applied to ψ_j. Thus $[B^{\varepsilon}_{j-1}(\Sigma)]^+$ is an L-proof of ψ_j from $\Sigma \cup [B^{\varepsilon}_{j-1}(\Sigma)]^-$. Because $\vdash_{\mathcal{N}}$ is monotonic and (42) holds $[B^{\varepsilon}_{j-1}(\Sigma)]^+$ is an L-proof of ψ_j from $\Sigma \cup [B^{\varepsilon}_{k}(\Sigma)]^-$, too. As $[B^{\varepsilon}_{k}(\Sigma)]^+ = [B^{\varepsilon}_{j}(\Sigma)]^+$ and (41), $[B^{\varepsilon}_{k}(\Sigma)]^+$ is an L-proof of ϕ from $\Sigma \cup [B^{\varepsilon}_{k}(\Sigma)]^-$. Suppose then, that there is no j described above. Then $B^{\varepsilon}_{k}(\Sigma) = [B^{\varepsilon}_{k}(\Sigma)]^-$ and $[B^{\varepsilon}_{k}(\Sigma)]^+ = \emptyset$ is an L-proof of ϕ from $\Sigma \cup [B^{\varepsilon}_{k}(\Sigma)]^-$ on the basis of Equation (41).

It is now easy to establish that for every $\phi \in \mathcal{L}_{ae}$, $\Sigma \cup \Lambda \models_{ae} \phi$ implies $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \phi$, by simply pointing out that $\Lambda = B^{\varepsilon}_{n}(\Sigma)$. If $L\chi \in \Lambda$, there is a number j such that $L\chi = L\psi_j \in B^{\varepsilon}_{j}(\Sigma)$, as Λ is generated by ε from Σ. Therefore, $\Sigma \cup B^{\varepsilon}_{j-1}(\Sigma) \models_{ae} \psi_j$. Equation (40) yields that $[B^{\varepsilon}_{j-1}(\Sigma)]^+$ is an L-proof of ψ_j from $\Sigma \cup [B^{\varepsilon}_{j-1}(\Sigma)]^-$, i.e. $\Sigma \cup [B^{\varepsilon}_{j-1}(\Sigma)]^- \vdash_{\mathcal{N}} \psi_j$. As $[B^{\varepsilon}_{j-1}(\Sigma)]^- \subseteq [\Lambda]^-$ and $\psi_j = \chi$, $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi$ follows by the monotonicity of $\vdash_{\mathcal{N}}$. □

In the sequel, we prove that $\mathcal{N}$-groundedness of the cautious expansions corresponds exactly to the $\mathcal{N}$-groundedness of full sets. Proposition 3.9 is crucial in establishing this. It shows how the positive consequence relation $\models_L$ conforms to the $\mathcal{N}$-groundedness condition when we reason with respect to an $\mathcal{N}$-grounded full set. In fact, we could have defined a new recursive consequence relation corresponding to the right hand side of the equivalence in Proposition 3.9 and shown that it coincides with $\models_L$.

Proposition 3.9
Let Σ be a set of formulae of $\mathcal{L}_{ae}$ and ϕ any formula. If Λ is an $\mathcal{N}$-grounded Σ-full set, then

$$\langle \Sigma, \Lambda \rangle \models_L \phi \textit{ if and only if } \Sigma \cup [\Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi)]^- \vdash_{\mathcal{N}} \phi.$$

Proof. ($\Rightarrow$) Let $\langle \Sigma, \Lambda \rangle \models_L \phi$, i.e. $\Sigma \cup \Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi) \models_{ae} \phi$ hold. Because Λ is an $\mathcal{N}$-grounded Σ-full set, it can be generated by an enumeration, say ε, by Lemma 3.5. The belief literals in $SB_{\langle \Sigma, \Lambda \rangle}(\phi) - \Lambda$ can also be generated by an enumeration by Lemma 3.6. We denote this enumeration by ε_s. The catenation of the two enumerations yields a set of quasi-literals

$$\begin{aligned} B^{\varepsilon\cdot\varepsilon_s}(\Sigma) &= B^{\varepsilon}(\Sigma) \cup B^{\varepsilon_s}(\Sigma \cup B^{\varepsilon}(\Sigma)) \\ &= \Lambda \cup B^{\varepsilon_s}(\Sigma \cup \Lambda) \\ &= \Lambda \cup (SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda) \\ &= \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi). \end{aligned}$$

Thus the set of belief literals $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ is generated by the enumeration $\varepsilon \cdot \varepsilon_s$. Hence, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \phi$ implies $\Sigma \cup [\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)]^- \vdash_{\mathcal{N}} \phi$ by Lemma 3.8.

($\Leftarrow$) Let $\Sigma \cup [\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)]^- \vdash_{\mathcal{N}} \phi$ hold. By Proposition 2.14 a minimal L-proof of ϕ contains only members of $Sf^L(\Sigma \cup \{\phi\})$. Thus,

$$\Sigma \cup [\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)]^- \cup \{L\phi_1, ..., L\phi_n\} \models_{ae} \phi$$

where each $L\phi_i$ $(1 \leq i \leq n)$ belongs to $Sf^L(\Sigma \cup \{\phi\})$. We show by induction on i that

$$\{L\phi_1, ..., L\phi_i\} \subseteq \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi). \tag{43}$$

Induction base: $i = 1$. Since $L\phi_1$ is the first member of the L-proof, it holds that $\Sigma \cup [\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)]^- \models_{ae} \phi_1$. Then $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \phi_1$ follows, as $\models_{ae}$ is monotonic. On the basis of Proposition 2.18, $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ is a consistent set of beliefs literals. Thus by Proposition 2.13, there is a set

$$\begin{aligned} \Lambda' &= \Pi_{Sf^{qL}(\Sigma\cup\{\phi_1\})}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) && [\text{Proposition 2.13}] \\ &\subseteq \Pi_{Sf^{L}(\Sigma\cup\{\phi_1\})}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) && [\scriptstyle Sf^{qL}(\Sigma\cup\{\phi_1\})\subseteq Sf^{L}(\Sigma\cup\{\phi_1\})] \\ &= \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi_1) && [\scriptstyle L\phi_1 \in Sf^{L}(\Sigma\cup\{\phi\})] \end{aligned}$$

such that $\Sigma \cup \Lambda' \models_{ae} \phi_1$. Thus $\Sigma \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi_1) \models_{ae} \phi$, i.e. $\langle\Sigma, \Lambda\rangle \models_L \phi_1$ by the monotonicity of $\models_{ae}$. Since $L\phi_1 \in Sf^L(\Sigma \cup \{\phi\})$, $L\phi_1$ is contained in $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$.

Inductive hypothesis: The elements $\{L\phi_1, ..., L\phi_{k-1}\}$ of the L-proof are contained in $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$.

Induction step: $i = k$. Since $L\phi_k$ is the k^{th} member of the L-proof, it holds that $\Sigma \cup [\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)]^- \cup \{L\phi_1, ..., L\phi_{k-1}\} \models_{ae} \phi_k$. By the inductive hypothesis, the set $\{L\phi_1, ..., L\phi_{k-1}\}$ is a subset of $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$. Therefore, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \phi_k$ follows by the monotonicity of $\models_{ae}$. Recalling that $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ is a consistent set of beliefs literals, Proposition 2.13 yields that there is a set

$$\begin{aligned} \Lambda' &= \Pi_{Sf^{qL}(\Sigma\cup\{\phi_k\})}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) && [\text{Proposition 2.13}] \\ &\subseteq \Pi_{Sf^{L}(\Sigma\cup\{\phi_k\})}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) && [\scriptstyle Sf^{qL}(\Sigma\cup\{\phi_k\})\subseteq Sf^{L}(\Sigma\cup\{\phi_k\})] \\ &= \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi_k) && [\scriptstyle L\phi_k \in Sf^{L}(\Sigma\cup\{\phi\})] \end{aligned}$$

such that $\Sigma \cup \Lambda' \models_{ae} \phi_k$. Thus $\Sigma \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi_k) \models_{ae} \phi_k$ by the monotonicity of $\models_{ae}$. Then $\langle\Sigma,\Lambda\rangle \models_L \phi_k$ and as $L\phi_k \in Sf^L(\Sigma \cup \{\phi\})$, $L\phi_k$ belongs to $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi_k)$. Now we know that Equation (43) holds for $i = n$ as well and $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \phi$ follows by the monotonicity of $\models_{ae}$. Thus $\langle\Sigma,\Lambda\rangle \models_L \phi$. □

We present now two lemmata similar to those in Section 2.3. Recall that cautious expansions are in one-to-one correspondence with full sets by Theorem 2.22. Firstly, Lemma 3.10 provides us the way of reducing an $\mathcal{N}$-grounded cautious expansion to the corresponding $\mathcal{N}$-grounded Σ-full set. Secondly, we can extend an $\mathcal{N}$-grounded Σ-full set to the corresponding $\mathcal{N}$-grounded cautious expansion by means of Lemma 3.11.

Lemma 3.10
Let Σ be a set of formulae of $\mathcal{L}_{ae}$. If $\langle\Delta^+,\Delta^-\rangle = CAE_\Sigma(\Delta^+,\Delta^-)$ is $\mathcal{N}$-grounded, then $\Lambda = \Pi_{Sf^L(\Sigma)}(L\Delta^+ \cup \neg L\Delta^-)$ is $\mathcal{N}$-grounded.

Proof. The Σ-fullness of Λ is already proven in Lemma 2.20. In the following we show that Λ (as defined above) satisfies the finitary $\mathcal{N}$-groundedness condition given in Definition 3.4. But first we check that $\neg L\Delta^-$ is consistent in the sense of Proposition 2.14. Let $\neg L\chi \in \neg L\Delta^-$. Assume that $\Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \chi$. Since $\langle\Delta^+,\Delta^-\rangle$ is $\mathcal{N}$-grounded, $\Delta^+ = \{\phi \mid \Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \phi\}$. Thus $\chi \in \Delta^+$ and $L\chi \in L\Delta^+$. Since Δ^+ and Δ^- are disjoint, $\neg L\chi \notin \neg L\Delta^-$, a contradiction. Hence $\Sigma \cup \neg L\Delta^- \not\vdash_{\mathcal{N}} \chi$.

Let $L\chi \in Sf^L(\Sigma)$. ($\Rightarrow$) Let $L\chi \in \Lambda$. Then $\chi \in \Delta^+$ and $\Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \chi$ since $\langle\Delta^+,\Delta^-\rangle$ is $\mathcal{N}$-grounded. By Proposition 2.14, $\Sigma \cup \Pi_{Sf^L(\Sigma)}(\neg L\Delta^-) \vdash_{\mathcal{N}} \chi$, i.e. $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi$. ($\Leftarrow$) Let $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi$. Since $[\Lambda]^- \subseteq \neg L\Delta^-$, it follows that $\Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \chi$ by the monotonicity of $\vdash_{\mathcal{N}}$. As $\langle\Delta^+,\Delta^-\rangle$ is $\mathcal{N}$-grounded, $\chi \in \Delta^+$ and $L\chi \in \Lambda$. Hence Λ satisfies the finitary $\mathcal{N}$-groundedness condition of Definition 3.4. □

Lemma 3.11
Let Σ be a set of formulae of $\mathcal{L}_{ae}$ and Λ an $\mathcal{N}$-grounded Σ-full set. If $\Delta^+ = \{\phi \mid \langle\Sigma,\Lambda\rangle \models_L \phi\}$ and $\Delta^- = \{\phi \mid \langle\Sigma,\Lambda\rangle =\mid_L \phi\}$, then $\langle\Delta^+,\Delta^-\rangle = CAE_\Sigma(\Delta^+,\Delta^-)$ is $\mathcal{N}$-grounded.

Proof. In Lemma 2.21 $\langle\Delta^+,\Delta^-\rangle$ was shown to be a cautious autoepistemic expansion. In the sequel, it is proved that

$$\phi \in \Delta^+ \text{ if and only if } \Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \phi.$$

($\Rightarrow$) Let $\phi \in \Delta^+$. Then $\langle\Sigma,\Lambda\rangle \models_L \phi$ and $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \phi$. Since Λ is $\mathcal{N}$-grounded, by Proposition 3.9, $\Sigma \cup [\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)]^- \vdash_{\mathcal{N}} \phi$. By Lemma 2.21, it holds that $[\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)]^- \subseteq \neg L\Delta^-$. $\Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \phi$ follows by the monotonicity of $\vdash_{\mathcal{N}}$.

($\Leftarrow$) The only-if direction is shown by induction on length of the L-proofs.

Induction base: Let $\Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \phi$ and ϕ have an L-proof of length 0. Because the necessitation rule is not needed in the proof of ϕ, $\Sigma \cup \neg L\Delta^- \models_{ae} \phi$. It follows by the monotonicity of $\models_{ae}$ that $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \models_{ae} \phi$. Thus $\phi \in IP_\Sigma^+(\Delta^+, \Delta^-)$ and $\phi \in \Delta^+$.

Inductive hypothesis:
Only-if-direction holds for ϕ having an L-proof shorter than k.

Induction step: Let $\Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \phi$ and ϕ have an L-proof of length k. It follows that $\Sigma \cup \neg L\Delta^- \cup \{L\phi_1, ..., L\phi_k\} \models_{ae} \phi$ and for each element $L\phi_j$ of the L-proof $\Sigma \cup \neg L\Delta^- \vdash_{\mathcal{N}} \phi_j$ holds. Since the L-proof of $L\phi_j$ has a length less than k, $\phi_j \in \Delta^+$ by inductive hypothesis. Thus $L\phi_j \in L\Delta^+$. It follows that $\{L\phi_1, ..., L\phi_k\} \subseteq L\Delta^+$. This allows us to conclude that $\Sigma \cup L\Delta^+ \cup \neg L\Delta^- \models_{ae} \phi$ by the monotonicity of $\models_{ae}$. Thus ϕ belongs to $IP_\Sigma^+(\Delta^+, \Delta^-)$ and $\phi \in \Delta^+$. □

Theorem 3.12 ending this section states the one-to-one correspondence between $\mathcal{N}$-grounded Σ-full sets and $\mathcal{N}$-grounded cautious expansions.

Theorem 3.12 (Correspondence)
Let Σ be a set of premises, $\mathcal{CE}$ the set of $\mathcal{N}$-grounded cautious expansions of Σ and $\mathcal{FS}$ the set of $\mathcal{N}$-grounded Σ-full sets. Let FS_Σ and CE_Σ be defined similarly as in Theorem 2.22. Then both FS_Σ and CE_Σ are bijections and inverses of each other.

Proof. If $\mathcal{E}$ is an $\mathcal{N}$-grounded cautious expansion of Σ, then the corresponding Σ-full set $FS_\Sigma(\mathcal{E})$ is $\mathcal{N}$-grounded by Lemma 3.10. The other way round, if Λ is an $\mathcal{N}$-grounded Σ-full set, then the corresponding cautious expansion $CE_\Sigma(\Lambda)$ is $\mathcal{N}$-grounded by Lemma 3.11. The functions FS_Σ and CE_Σ coincide with the ones presented in Theorem 2.22 but their domains are restricted to $\mathcal{N}$-grounded cautious expansions and $\mathcal{N}$-grounded Σ-full sets, respectively. Therefore FS_Σ and CE_Σ are bijections and inverses of each other by Theorem 2.22. □

3.2 On the Partial Order of the Expansions

In this section we return to the partial order of cautious expansions. We show that the subset inclusion of $\mathcal{N}$-grounded full sets coincides with the partial order between $\mathcal{N}$-grounded cautious expansions given in Definition 3.1. As a consequence, the least $\mathcal{N}$-grounded cautious expansion corresponds to the least $\mathcal{N}$-grounded full set (providing of course that such a full set exists). In Propositions 3.13 and 3.14 we prove that the recursive consequence relations act monotonically on $\mathcal{N}$-grounded full sets.

Proposition 3.13
Let Σ be a set of formulae of $\mathcal{L}_{ae}$, ϕ any formula and Λ' and Λ be two $\mathcal{N}$-grounded Σ-full sets such that $\Lambda' \subseteq \Lambda$. If also $SB_{\langle\Sigma,\Lambda'\rangle}(\phi) \subseteq SB_{\langle\Sigma,\Lambda\rangle}(\phi)$, then

$$\langle\Sigma,\Lambda\rangle \not\dashv_L \phi \text{ implies } \langle\Sigma,\Lambda'\rangle \not\dashv_L \phi.$$

Proof. Let $\langle\Sigma,\Lambda\rangle \not\dashv_L \phi$ hold and let ε_f be an enumeration generating the $\mathcal{N}$-grounded Σ-full set Λ. Such an enumeration exists on the basis of Lemma 3.5. In addition, Lemma 3.6 yields that the difference $SB_{\langle\Sigma,\Lambda\rangle}(\phi) - \Lambda$ can be generated by means of an enumeration. We denote this enumeration by ε_s. The concatenation of ε_f and ε_s generates the set $B^{\varepsilon_f\cdot\varepsilon_s}(\Sigma)$ that equals to $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$. Since $\langle\Sigma,\Lambda\rangle \not\dashv_L \phi$ holds, there is an enumeration ε_r of $\{\psi \mid L\psi \in Sf^L(\Sigma\cup\{\phi\}) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))\}$ such that $\phi \in EE^{\varepsilon_r}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$, i.e. $\Sigma \cup B^{\varepsilon_f\cdot\varepsilon_s\cdot\varepsilon_r}(\Sigma) \models_{ae} \phi$.

We then verify that the conditions of Lemma 3.7 are met. Firstly, the set of quasi-literals $\Lambda' \cup SB_{\langle\Sigma,\Lambda'\rangle}(\phi)$ is clearly contained in $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$. On the other hand, let $\neg L\chi \in \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$. Then $L\chi \in Sf^L(\Sigma\cup\{\phi\})$ and $\langle\Sigma,\Lambda\rangle \dashv_L \chi$. The mutual exclusion of the recursive consequence relations yields $\langle\Sigma,\Lambda\rangle \not\models_L \chi$. Assume now, that $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \models_{ae} \chi$. As $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)$ is a consistent set of beliefs literals, Proposition 2.13 yields that there is a set

$$\begin{array}{rcll}
\Lambda'' & = & \Pi_{Sf^{qL}(\Sigma\cup\{\phi\})}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) & [\text{Proposition 2.13}] \\
 & \subseteq & \Pi_{Sf^{L}(\Sigma\cup\{\phi\})}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) & [\scriptstyle Sf^{qL}(\Sigma\cup\{\phi\})\subseteq Sf^L(\Sigma\cup\{\phi\})] \\
 & = & \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) & [\scriptstyle L\chi\in Sf^L(\Sigma\cup\{\phi\})]
\end{array}$$

such that $\Sigma \cup \Lambda'' \models_{ae} \chi$. Since $\models_{ae}$ is monotonic, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \models_{ae} \chi$, or equivalently $\langle\Sigma,\Lambda\rangle \models_L \chi$, which is a contradiction. Therefore it holds that $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi) \not\models_{ae} \chi$. By Lemma 3.7 there is an enumeration ε_d that generates the difference $(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi)) - (\Lambda' \cup SB_{\langle\Sigma,\Lambda'\rangle}(\phi))$ Secondly, $B^{\varepsilon_f\cdot\varepsilon_s\cdot\varepsilon_r}(\Sigma) = B^{\varepsilon_d\cdot\varepsilon_r}(\Sigma \cup \Lambda' \cup SB_{\langle\Sigma,\Lambda'\rangle}(\phi))$ holds. Then there is an enumeration $\varepsilon'_r = \varepsilon_d \cdot \varepsilon_r$ such that

$$\Sigma \cup \Lambda' \cup SB_{\langle\Sigma,\Lambda'\rangle}(\phi) \cup B^{\varepsilon'_r}(\Sigma \cup \Lambda' \cup SB_{\langle\Sigma,\Lambda'\rangle}(\phi)) \models_{ae} \phi,$$

i.e. $\phi \in EE^{\varepsilon'_r}(\Sigma \cup \Lambda' \cup SB_{\langle\Sigma,\Lambda'\rangle}(\phi))$. As ε'_r is an enumeration of

$$\{\psi \mid L\psi \in Sf^L(\Sigma\cup\{\phi\}) - \text{q-atom}(\Lambda' \cup SB_{\langle\Sigma,\Lambda'\rangle}(\phi))\}$$

it follows that $\langle\Sigma,\Lambda'\rangle \not\dashv_L \phi$. □

Proposition 3.14
Let Σ be a set of formulae of $\mathcal{L}_{ae}$. Let Λ' and Λ be two $\mathcal{N}$-grounded Σ-full sets such that $\Lambda' \subseteq \Lambda$. Then for all $\phi \in \mathcal{L}_{ae}$:

$$\langle\Sigma,\Lambda'\rangle \models_L \phi \text{ implies } \langle\Sigma,\Lambda\rangle \models_L \phi \tag{44}$$

$$\langle \Sigma, \Lambda' \rangle \dashv_L \phi \textit{ implies } \langle \Sigma, \Lambda \rangle \dashv_L \phi. \tag{45}$$

Proof. We prove this by induction on the $BD(\phi)$.

Induction base: $BD(\phi)$=0. Consider Equation (44) first. Let $\langle \Sigma, \Lambda' \rangle \models_L \phi$ hold. Since Λ' is an $\mathcal{N}$-grounded Σ-full set, $\Sigma \cup [\Lambda' \cup SB_{\langle \Sigma, \Lambda' \rangle}(\phi)]^- \vdash_{\mathcal{N}} \phi$ follows by Proposition 3.9. As $BD(\phi)$=0, $SB_{\langle \Sigma, \Lambda' \rangle}(\phi) = \emptyset$ and therefore $\Sigma \cup [\Lambda']^- \vdash_{\mathcal{N}} \phi$. Since $\Lambda' \subseteq \Lambda$, $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \phi$ holds by the monotonicity of $\vdash_{\mathcal{N}}$. Again, $SB_{\langle \Sigma, \Lambda \rangle}(\phi) = \emptyset$ and $\langle \Sigma, \Lambda \rangle \models_L \phi$ follows by the $\mathcal{N}$-groundedness of Λ and Proposition 3.9. Then consider Equation (45). We show the converse, i.e. $\langle \Sigma, \Lambda \rangle \not\dashv_L \phi$ implies $\langle \Sigma, \Lambda' \rangle \not\dashv_L \phi$. So let $\langle \Sigma, \Lambda \rangle \not\dashv_L \phi$ hold. Because $BD(\phi) = 0$, both $SB_{\langle \Sigma, \Lambda \rangle}(\phi)$ and $SB_{\langle \Sigma, \Lambda' \rangle}(\phi)$ are empty. Since Λ' and Λ are $\mathcal{N}$-grounded Σ-full sets such that $\Lambda' \subseteq \Lambda$, $\langle \Sigma, \Lambda' \rangle \not\dashv_L \phi$ follows by Proposition 3.13.

Inductive hypothesis:
Equations (44) and (45) hold for ϕ having $BD(\phi)$ less than k.

Induction step: $BD(\phi) = k$. We need the result that

$$SB_{\langle \Sigma, \Lambda' \rangle}(\phi) \subseteq SB_{\langle \Sigma, \Lambda \rangle}(\phi) \tag{46}$$

holds. Let $L\chi \in SB_{\langle \Sigma, \Lambda' \rangle}(\phi)$. Then $\langle \Sigma, \Lambda' \rangle \models_L \chi$. Since $BD(\phi) < k$, $\langle \Sigma, \Lambda \rangle \models_L \chi$ follows by the inductive hypothesis. Therefore $L\chi \in SB_{\langle \Sigma, \Lambda \rangle}(\phi)$. Let $\neg L\chi \in SB_{\langle \Sigma, \Lambda' \rangle}(\phi)$. Then $\langle \Sigma, \Lambda' \rangle \dashv_L \chi$. The fact that $BD(\chi) < k$ and the inductive hypothesis yield $\langle \Sigma, \Lambda \rangle \dashv_L \chi$. Therefore $\neg L\chi \in SB_{\langle \Sigma, \Lambda \rangle}(\phi)$.

Then consider Equation (44). Let $\langle \Sigma, \Lambda' \rangle \models_L \phi$. It follows by Proposition 3.9 that $\Sigma \cup [\Lambda' \cup SB_{\langle \Sigma, \Lambda' \rangle}(\phi)]^- \vdash_{\mathcal{N}} \phi$. Moreover $\Sigma \cup [\Lambda \cup SB_{\langle \Sigma, \Lambda \rangle}(\phi)]^- \vdash_{\mathcal{N}} \phi$ holds by (46) and the monotonicity of $\vdash_{\mathcal{N}}$. Then consider Equation (45). We show the converse, i.e. $\langle \Sigma, \Lambda \rangle \not\dashv_L \phi$ implies $\langle \Sigma, \Lambda' \rangle \not\dashv_L \phi$. So let $\langle \Sigma, \Lambda \rangle \not\dashv_L \phi$ hold. Since $\Lambda' \subseteq \Lambda$, both Λ' and Λ are $\mathcal{N}$-grounded and on the other hand (46) holds $\langle \Sigma, \Lambda' \rangle \not\dashv_L \phi$ follows by Proposition 3.13. □

Finally, we prove the correspondence between the subset inclusion of $\mathcal{N}$-grounded full sets and the partial order $\sqsubseteq$ on cautious expansions in Proposition 3.15.

Proposition 3.15
Let Σ be a set of premises. Let Λ_1 and Λ_2 be two $\mathcal{N}$-grounded Σ-full sets and $\mathcal{E}_1$ and $\mathcal{E}_2$ the corresponding $\mathcal{N}$-grounded cautious expansions $CE_\Sigma(\Lambda_1)$ and $CE_\Sigma(\Lambda_2)$, respectively. Then

$$\Lambda_1 \subseteq \Lambda_2 \textit{ if and only if } \mathcal{E}_1 \sqsubseteq \mathcal{E}_2.$$

Proof. Let $\mathcal{E}_1 = \langle \Delta_1^+, \Delta_1^- \rangle$ and $\mathcal{E}_2 = \langle \Delta_2^+, \Delta_2^- \rangle$.

($\Rightarrow$) Let $\Lambda_1 \subseteq \Lambda_2$. Let $\phi \in \Delta_1^+$. By Lemma 3.11, $\langle \Sigma, \Lambda_1 \rangle \models_L \phi$. By Proposition 3.14, $\langle \Sigma, \Lambda_2 \rangle \models_L \phi$. By Lemma 3.11, $\phi \in \Delta_2^+$. Then, we consider the disbeliefs. Let $\phi \in \Delta_1^-$. By Lemma 3.11, $\langle \Sigma, \Lambda_1 \rangle \dashv_L \phi$. By Proposition 3.14, $\langle \Sigma, \Lambda_2 \rangle \dashv_L \phi$. By Lemma 3.11, $\phi \in \Delta_2^-$.

($\Leftarrow$) Let $\mathcal{E}_1 \sqsubseteq \mathcal{E}_2$ hold. Now, if $L\chi \in \Lambda_1$, $\langle \Sigma, \Lambda_1 \rangle \models_L \chi$. This implies $\chi \in \Delta_1^+$ by Lemma 3.11. As $\Delta_1^+ \subseteq \Delta_2^+$, $\chi \in \Delta_2^+$. Since $L\chi \in Sf^L(\Sigma)$, $L\chi \in \Lambda_2$ by Lemma 3.10. Similarly, $\neg L\chi \in \Lambda_1$ implies $\neg L\chi \in \Lambda_2$. $\square$

3.3 Least Grounded Expansion

Our target is now to prove that the least $\mathcal{N}$-grounded cautious expansion always exists for the given set of premises Σ. In Section 3.1 the full sets were shown to capture the notion of $\mathcal{N}$-groundedness. In addition, the least $\mathcal{N}$-grounded full set corresponds to the least $\mathcal{N}$-grounded cautious expansion by Proposition 3.15. Therefore, it remains to be shown that such a full set always exists. In this section we provide an iterative method for constructing the least $\mathcal{N}$-grounded Σ-full set.

A slightly different pair of consequence relations (with respect to the recursive ones) is utilized in the iterative procedure. The difference is that the derivation of subbeliefs of formulae has been left implicit. The consequence relations $\models_{ae}$ and $\dashv_{ae}$ introduced in Definition 3.16 are the non-recursive counterparts of $\models_L$ and $\dashv_L$. The iterative method presented later constructs essentially a Σ-full set.

Definition 3.16
For a set Σ of sentences of $\mathcal{L}_{ae}$, a set of quasi-literals Γ and a formula ϕ:

1. *$\langle \Sigma, \Gamma \rangle \models_{ae} \phi$ if and only if $\Sigma \cup \Gamma \models_{ae} \phi$*

2. *$\langle \Sigma, \Gamma \rangle \dashv_{ae} \phi$ if and only if there is no enumeration ε of*
$$\{\chi \mid L\chi \in Q - q\text{-}atom(\Gamma)\}$$
such that $\phi \in EE^{\varepsilon}(\Sigma \cup \Gamma)$,

where Q is a fixed and finite set of quasi-atoms satisfying $Sf^L(\Sigma) \subseteq Q$ and $Sf^L(Q) \subseteq Q$.

The above consequence relations are defined relative to the set of quasi-literals Q defined above. The set Q is required to contain $Sf^L(\Sigma)$ and to be closed with respect to the L-subformulae. The reason for introducing Q is that we are restricting the construction to a certain part of the cautious expansion. This idea will be clarified in Proposition 3.25. The presence of some fixed Q satisfying the requirements above is implicit in the propositions of this section. In Proposition 3.17, we make sure that the two relations $\models_{ae}$ and $\dashv_{ae}$ exclude each other and therefore allow consistent introspection in the iterative construction.

Proposition 3.17 (Mutual exclusion)
For a set Σ of sentences of $\mathcal{L}_{ae}$, a set of quasi-literals Γ and a formula ϕ:

$$\langle \Sigma, \Gamma \rangle \models_{ae} \phi \textit{ implies } \langle \Sigma, \Gamma \rangle \not\dashv_{ae} \phi \tag{47}$$

or equivalently,

$$\langle \Sigma, \Gamma \rangle \dashv_{ae} \phi \textit{ implies } \langle \Sigma, \Gamma \rangle \not\models_{ae} \phi. \tag{48}$$

Proof. Let $\langle \Sigma, \Gamma \rangle \models_{ae} \phi$. By Definition 3.16, $\Sigma \cup \Gamma \models_{ae} \phi$. Let ε be any enumeration of $\{\psi \mid L\psi \in Q - \text{q-atom}(\Gamma)\}$. Then $\Sigma \cup \Gamma \cup B^{\varepsilon}(\Sigma \cup \Gamma) \models_{ae} \phi$ follows by the monotonicity of $\models_{ae}$. It follows that $\phi \in EE^{\varepsilon}(\Sigma \cup \Gamma)$, for some ε. Thus $\langle \Sigma, \Gamma \rangle \not\dashv_{ae} \phi$. $\square$

We formalize the iterative construction by defining a parametrized operator $FCAE_{\Sigma,Q}$ taking a consistent set of belief literals as its argument.

Definition 3.18 (Operator for iteration)
Given a set Σ of sentences of $\mathcal{L}_{ae}$, define an operator $FCAE_{\Sigma,Q}$ taking a consistent set of quasi-literals $\Gamma \subseteq Q \cup \neg Q$ as its argument:

$$FCAE_{\Sigma,Q}(\Gamma) = \{L\psi \in Q \mid \langle \Sigma, \Gamma \rangle \models_{ae} \psi\} \cup \{\neg L\psi \in \neg Q \mid \langle \Sigma, \Gamma \rangle \dashv_{ae} \psi\}.$$

The argument Γ represents a part of the cautious expansions under construction. At each step of the iteration, we carry out introspection on the basis of the previous step. The set Q gives clearly the scope of the introspection. We expect the construction to be complete when the operator reaches a fixed point, i.e. the application of $FCAE_{\Sigma,Q}$ yields its argument unchanged. We would like the iteration to reach its fixed point in finite number of steps. To show this, we first prove that $FCAE_{\Sigma,Q}$ is monotonic. More precisely, it should be non-decreasing with respect to the subset inclusion. It turns out that the relation $\dashv_{ae}$ acts monotonically in the construction, though it is nonmonotonic in general!

The fixed points of $FCAE_{\Sigma,Q}$ share a very important property given in Proposition 3.19, namely they can be generated by enumerations. We understand that the operator $FCAE_{\Sigma,Q}$ is not the only possibility of iteratively constructing full sets. Essential properties of the operators are the monotonicity and generation by enumerations. Therefore, the later proofs will be dependent on these. This gives more flexibility, e.g. when writing algorithms for cautious autoepistemic reasoning. The operator $FCAE_{\Sigma,Q}$ carries out the construction layer by layer. This is not necessarily the best choice from the computational point of view.

Proposition 3.19 (Monotonicity of iteration)
Let Γ_0 be a consistent set of quasi-literals satisfying q-atom$(\Gamma_0) \subseteq Q$ and two special conditions, namely

$$\Gamma_0 = B^{\varepsilon_0}(\Sigma) \text{ and } \Gamma_0 \subseteq FCAE_{\Sigma,Q}(\Gamma_0) \tag{49}$$

where ε_0 is an enumeration of $\{\psi \mid L\psi \in$ q-atom$(\Gamma_0)\}$ generating the set Γ_0. Then define a sequence of repeated applications of $FCAE_{\Sigma,Q}$ to Γ_0 by letting

$$\Gamma_{i+1} = FCAE_{\Sigma,Q}(\Gamma_i)$$

when $i > 0$. As a consequence, the members of the sequence $\Gamma_0, \Gamma_1, \ldots$ are generated by finite enumerations ε_i $(i \geq 0)$ such that each ε_i is an enumeration of $\{\psi \mid L\psi \in$ q-atom$(\Gamma_i)\}$ and it is prefixed by ε_{i-1}, i.e. $\varepsilon_i = \varepsilon_{i-1} \cdot \varepsilon_i^d$ for some enumeration ε_i^d. Moreover, the sequence Γ_i $(i \geq 0)$ is non-decreasing, i.e. $\Gamma_i \subseteq \Gamma_{i+1}$ for all $i \geq 0$.

Proof. The sets Γ_i $(i \geq 0)$ are consistent sets of belief literals. This is clear because Γ_0 is consistent and by Proposition 3.17, the consequence relations $\models_{ae}$ and $=\!\mid_{ae}$ are mutually exclusive. Therefore the sequence is well-defined. The proposition is proved by induction on i.

Induction base: $i = 0$. The base case is straightforward by the conditions in (49). The set Γ_0 is generated by an enumeration of

$$\{\psi \mid L\psi \in \text{q-atom}(\Gamma_0)\}$$

and furthermore, $\Gamma_0 \subseteq FCAE_{\Sigma,Q}(\Gamma_0) = \Gamma_1$. These conditions have to verified for each Γ_0 separately, if the results of this proposition are to hold.

Inductive hypothesis: The sets Γ_i $(0 \leq i < k)$ are generated by enumerations ε_i $(0 \leq i < k)$, respectively, and $\Gamma_i \subseteq \Gamma_{i+1}$ holds for all $i = 0, 1, \ldots, k-1$.

Induction step: $i = k$. The set Γ_{k-1} is generated by the enumeration ε_{k-1} of

$$\{\psi \mid L\psi \in \text{q-atom}(\Gamma_{k-1})\}.$$

We extend this enumeration to an enumeration that generates Γ_k. By the inductive hypothesis Γ_{k-1} is a subset of Γ_k. Let $\varepsilon_d = \psi_1, \ldots, \psi_n$ be an arbitrary enumeration of $\{\psi \mid L\psi \in \text{q-atom}(\Gamma_k - \Gamma_{k-1})\}$. We then show that ε_d exactly generates the difference $\Gamma_k - \Gamma_{k-1}$. Consider first the positive beliefs in the difference. Let $L\psi_j \in \Gamma_k - \Gamma_{k-1}$ hold. Then $L\psi_j \in FCAE_{\Sigma,Q}(\Gamma_{k-1})$. By the definition of $FCAE_{\Sigma,Q}$, $\Sigma \cup \Gamma_{k-1} \models_{ae} \psi_j$. The monotonicity of $\models_{ae}$ yields now that $\Sigma \cup \Gamma_{k-1} \cup B_{j-1}^{\varepsilon_d}(\Sigma \cup \Gamma_{k-1}) \models_{ae} \psi_j$, i.e. $L\psi_j \in B_j^{\varepsilon_d}(\Sigma \cup \Gamma_{k-1})$. A similar inspection is made for the negative beliefs in $\Gamma_k - \Gamma_{k-1}$. So let $\neg L\psi_j \in \Gamma_k - \Gamma_{k-1}$. Then $\neg L\psi_j \in FCAE_{\Sigma,Q}(\Gamma_{k-1})$ and $\langle \Sigma, \Gamma_{k-1} \rangle =\!\mid_{ae} \psi_j$ holds. Let us now assume that $\Sigma \cup \Gamma_{k-1} \cup B^{\varepsilon_d}(\Sigma \cup \Gamma_{k-1}) \models_{ae} \psi_j$. The enumeration ε_d can be extended to an enumeration $\varepsilon_d \cdot \varepsilon_e$ of

$$\{\psi \mid L\psi \in Q - \text{q-atom}(\Gamma_{k-1})\}$$

such that $\Sigma \cup \Gamma_{k-1} \cup B^{\varepsilon_d \cdot \varepsilon_e}(\Sigma \cup \Gamma_{k-1}) \models_{ae} \psi_j$ by the monotonicity of $\models_{ae}$. But this is clearly a refuting enumeration and thus $\langle \Sigma, \Gamma_{k-1} \rangle \not\dashv_{ae} \psi_j$, which is a contradiction. Therefore, $\Sigma \cup \Gamma_{k-1} \cup B^{\varepsilon_d}(\Sigma \cup \Gamma_{k-1}) \not\models_{ae} \psi_j$. The monotonicity of $\models_{ae}$ yields

$$\Sigma \cup \Gamma_{k-1} \cup B^{\varepsilon_d}_{j-1}(\Sigma \cup \Gamma_{k-1}) \not\models_{ae} \psi_j$$

and $\neg L\psi_j \in B^{\varepsilon_d}_j(\Sigma \cup \Gamma_{k-1})$. We have now established that $\Gamma_k - \Gamma_{k-1} \subseteq B^{\varepsilon_d}(\Sigma \cup \Gamma_{k-1})$. Since $\text{q-atom}(\Gamma_k - \Gamma_{k-1}) = \text{q-atom}(B^{\varepsilon_d}(\Sigma \cup \Gamma_{k-1}))$, $\Gamma_k - \Gamma_{k-1} = B^{\varepsilon_d}(\Sigma \cup \Gamma_{k-1})$. Therefore, $\varepsilon_k = \varepsilon_{k-1} \cdot \varepsilon_d$ generates the set Γ_k, i.e. $\Gamma_k = B^{\varepsilon_k}(\Sigma)$, and ε_{k-1} is a prefix of ε_k.

We then turn our attention to the monotonicity of the sequence. Recall that by the inductive hypothesis $\Gamma_{k-1} \subseteq \Gamma_k$. Let $L\chi \in \Gamma_k$ hold, i.e. $L\chi \in FCAE_{\Sigma,Q}(\Gamma_{k-1})$. It follows that $\langle \Sigma, \Gamma_{k-1} \rangle \models_{ae} \chi$. Since $\Gamma_{k-1} \subseteq \Gamma_k$, we get $\langle \Sigma, \Gamma_k \rangle \models_{ae} \chi$ by the monotonicity of $\models_{ae}$. Now, $L\chi \in FCAE_{\Sigma,Q}(\Gamma_k)$ and $L\chi \in \Gamma_{k+1}$. Then let $\neg L\chi \in \Gamma_k$, i.e. $\neg L\chi \in FCAE_{\Sigma,Q}(\Gamma_{k-1})$. It follows that $\langle \Sigma, \Gamma_{k-1} \rangle \dashv_{ae} \chi$. Assume now that $\langle \Sigma, \Gamma_k \rangle \not\dashv_{ae} \chi$. Then there is an enumeration ε_r of $\{\psi \mid L\psi \in Q - \text{q-atom}(\Gamma_k)\}$ such that $\chi \in EE^{\varepsilon_r}(\Sigma \cup \Gamma_k)$, i.e. $\Sigma \cup \Gamma_k \cup B^{\varepsilon_r}(\Sigma \cup \Gamma_k) \models_{ae} \chi$. Since $\Gamma_k - \Gamma_{k-1}$ is generated by the enumeration ε_d, so we have that $\Sigma \cup \Gamma_{k-1} \cup B^{\varepsilon_d \cdot \varepsilon_r}(\Sigma \cup \Gamma_{k-1}) \models_{ae} \chi$ where $\varepsilon_d \cdot \varepsilon_r$ is an enumeration of $\{\psi \mid L\psi \in Q - \text{q-atom}(\Gamma_{k-1})\}$. Therefore, $\langle \Sigma, \Gamma_{k-1} \rangle \not\dashv_{ae} \chi$, a contradiction. Hence $\langle \Sigma, \Gamma_k \rangle \dashv_{ae} \chi$ and $\neg L\chi \in \Gamma_{k+1}$. It has been established that $\Gamma_k \subseteq \Gamma_{k+1}$, i.e. the sequence is monotonic. □

The iterations may be started from any Γ_0 satisfying the conditions in (49). The varying starting points of iteration have been introduced for the sake of generality of the results. By the first condition, there has to be an enumeration ε_0 that generates the set Γ_0. Furthermore, the first application of $FCAE_{\Sigma,Q}$ must yield a superset of Γ_0. The inductive proof extends these properties over the whole iteration sequence. The monotonicity of the sequence may now be conveniently used when proving that an iteration reaches its fixed point in a finite number of steps.

Proposition 3.20 (Finite iterations)
The sequence given in Proposition 3.19 reaches a fixed point of the operator $FCAE_{\Sigma,Q}$ in finite number of steps.

Proof. If the sequence reaches a fixed point of $FCAE_{\Sigma,Q}$, Γ_{n+1} equals to Γ_n for some n. Assume there is no such n. Then at each step of iteration at least one quasi-literal χ such that q-atom(χ) $\in Q$ is inserted to the set Γ_i. But this implies that infinitely many quasi-literals are added during the construction. Thus Q must be infinite, a contradiction with Definition 3.16. Therefore, the fixed point is reached in finite number of steps. □

We let fp($FCAE_{\Sigma,Q}, \Gamma_0$) denote the fixed point of $FCAE_{\Sigma,Q}$ when iteration is started from Γ_0 satisfying the initial conditions.

The next objective is to verify that the iterative method yields correct results. As our intention is to construct full sets, we should check that the fixed points of the iteration satisfy the fullness conditions of Definition 2.17. However, our slight generalization has to be taken into account. Recalling the definition of the iteration, we relativized the procedure with respect to the set Q. Therefore, we show in Proposition 3.24 that fixed points of $FCAE_{\Sigma,Q}$ are closed with respect to the recursive consequence relations $\models_L$ and $\dashv_L$. This property implies straightforwardly that the reduct of the fixed point with respect to $Sf^L(\Sigma)$ is Σ-full.

In Proposition 3.23 we show that closures with respect to the non-recursive and recursive consequence relations coincide. Lemmata 3.21 and 3.22 are needed in the proofs.

Lemma 3.21
Let Γ be a consistent set of quasi-literals such that q-atom(Γ) $\subseteq Q$, $L\chi \in Q$ and $\Gamma' = \Pi_{Sf^L(\Sigma\cup\{\chi\})}(\Gamma)$. Let $\varepsilon = \psi_1, ..., \psi_n$ be an enumeration of

$$\{\psi \mid L\chi \in Sf^L(\Sigma \cup \{\chi\}) - \textit{q-atom}(\Gamma')\}.$$

Then

$$B^\varepsilon(\Sigma \cup \Gamma') = B^\varepsilon(\Sigma \cup \Gamma).$$

Proof. We show by induction on i that

$$(\neg)L\psi_i \in B_i^\varepsilon(\Sigma \cup \Gamma') \text{ iff } (\neg)L\psi_i \in B_i^\varepsilon(\Sigma \cup \Gamma).$$

Induction base: ($\Rightarrow$) Let $L\psi_1 \in B_1^\varepsilon(\Sigma \cup \Gamma')$. Then $\Sigma \cup \Gamma' \models_{ae} \psi_1$. By the monotonicity of $\models_{ae}$, $\Sigma \cup \Gamma \models_{ae} \psi_1$, i.e. $L\psi_1 \in B_1^\varepsilon(\Sigma \cup \Gamma)$. ($\Leftarrow$) To show the converse, let $L\psi_1 \in B_1^\varepsilon(\Sigma \cup \Gamma)$. It follows that $\Sigma \cup \Gamma \models_{ae} \psi_1$. Since Γ is consistent, by Proposition 2.13 there is a set

$$\begin{array}{rcll} \Gamma'' & = & \Pi_{Sf^{qL}(\Sigma\cup\{\psi_1\})}(\Gamma) & \text{[Proposition 2.13]} \\ & \subseteq & \Pi_{Sf^L(\Sigma\cup\{\chi\})}(\Gamma) & [\scriptstyle L\psi_1 \in Sf^L(\Sigma\cup\{\chi\}) \text{ and } Sf^{qL}(\Sigma\cup\{\chi\}) \subseteq Sf^L(\Sigma\cup\{\chi\})] \\ & = & \Gamma' & \end{array}$$

such that $\Sigma \cup \Gamma'' \models_{ae} \psi_1$. Thus $\Sigma \cup \Gamma' \models_{ae} \psi_1$ follows by the monotonicity of $\models_{ae}$ and $L\psi_1 \in B_1^\varepsilon(\Sigma \cup \Gamma')$. The negated version of the equivalence follows easily from the fact that q-atom($B_1^\varepsilon(\Sigma \cup \Gamma')$) = q-atom($B_1^\varepsilon(\Sigma \cup \Gamma)$).

Inductive hypothesis: The equivalences hold for $i < k$.

Induction step: ($\Rightarrow$) Let $L\psi_k \in B^{\varepsilon}_k(\Sigma \cup \Gamma')$. By the definition of enumeration based introspection, $\Sigma \cup \Gamma' \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma') \models_{ae} \psi_k$. Then it follows that $\Sigma \cup \Gamma' \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma) \models_{ae} \psi_k$ by the inductive hypothesis. The monotonicity of $\models_{ae}$ yields that $\Sigma \cup \Gamma \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma) \models_{ae} \psi_k$, i.e. $L\psi_k \in B^{\varepsilon}_k(\Sigma \cup \Gamma)$. ($\Leftarrow$) Let $L\psi_k \in B^{\varepsilon}_k(\Sigma \cup \Gamma)$. It follows that $\Sigma \cup \Gamma \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma) \models_{ae} \psi_k$. $\Gamma \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma)$ is consistent because Γ is consistent, q-atom(Γ) is disjoint from q-atom($B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma)$) and $B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma)$ is consistent by construction. Proposition 2.13 yields that there is a set

$$\begin{aligned}
\Gamma'' &= \Pi_{Sf^{qL}(\Sigma \cup \{\psi_k\})}(\Gamma \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma)) && \text{[Proposition 2.13]}\\
&\subseteq \Pi_{Sf^{L}(\Sigma \cup \{\chi\})}(\Gamma \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma))\\
&\quad \left[L\psi_k \in Sf^L(\Sigma \cup \{\chi\}) \text{ and } Sf^{qL}(\Sigma \cup \{\chi\}) \subseteq Sf^L(\Sigma \cup \{\chi\})\right]\\
&= \Gamma' \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma)
\end{aligned}$$

such that $\Sigma \cup \Gamma'' \models_{ae} \psi_k$. Thus $\Sigma \cup \Gamma' \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma) \models_{ae} \psi_k$ by the monotonicity of $\models_{ae}$. By the inductive hypothesis, $\Sigma \cup \Gamma' \cup B^{\varepsilon}_{k-1}(\Sigma \cup \Gamma') \models_{ae} \psi_k$. Therefore $L\psi_k \in B^{\varepsilon}_k(\Sigma \cup \Gamma')$. The negated version of the equivalence follows straightforwardly, as q-atom($B^{\varepsilon}_k(\Sigma \cup \Gamma')$) = q-atom($B^{\varepsilon}_k(\Sigma \cup \Gamma)$). □

Lemma 3.22
Let Γ be a consistent set of quasi-literals and $\Lambda = \Pi_{Sf^L(\Sigma)}(\Gamma)$ its reduction with respect to $Sf^L(\Sigma)$. Let $L\chi \in Q$ such that $SB_{\langle\Sigma,\Lambda\rangle}(\chi) = \Pi_{Sf^L(\chi)}(\Gamma)$. Then

$$\langle\Sigma, \Gamma\rangle \models_{ae} \chi \textit{ if and only if } \langle\Sigma, \Lambda\rangle \models_L \chi \tag{50}$$

and

$$\langle\Sigma, \Gamma\rangle \mathrel{=\!\!|}_{ae} \chi \textit{ if and only if } \langle\Sigma, \Lambda\rangle \mathrel{=\!\!|}_L \chi. \tag{51}$$

Proof. The two reductions can be combined into one such that

$$\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) = \Pi_{Sf^L(\Sigma \cup \{\chi\})}(\Gamma). \tag{52}$$

We begin by proving the equivalence in (50). ($\Rightarrow$) Let $\Sigma \cup \Gamma \models_{ae} \chi$. The set Γ is a consistent set of quasi-literals. Proposition 2.13 yields that there is a set Γ' such that $\Gamma' = \Pi_{Sf^{qL}(\Sigma \cup \{\chi\})}(\Gamma)$ and $\Sigma \cup \Gamma' \models_{ae} \chi$. Since $Sf^{qL}(\Sigma \cup \{\chi\}) \subseteq Sf^L(\Sigma \cup \{\chi\})$ and on the other hand (52) holds, $\Gamma' \subseteq \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi)$. The monotonicity of $\models_{ae}$ yields that $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \models_{ae} \chi$, i.e. $\langle\Sigma, \Lambda\rangle \models_L \chi$. ($\Leftarrow$) Let $\langle\Sigma, \Lambda\rangle \models_L \chi$. Then $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \models_{ae} \chi$ holds. Because $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi)$ is a subset of Γ, $\langle\Sigma, \Gamma\rangle \models_{ae} \chi$ follows by the monotonicity of $\models_{ae}$.

Then we consider the equivalence in (51). ($\Rightarrow$). Let $\langle\Sigma, \Lambda\rangle \not\mathrel{=\!\!|}_L \chi$ hold. By the definition of $\mathrel{=\!\!|}_L$ there is an enumeration ε of

$$\{\psi \mid L\psi \in Sf^L(\Sigma \cup \{\chi\}) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi))\} \tag{53}$$

such that $\chi \in EE^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi))$, i.e.

$$\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \cup B^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi)) \models_{ae} \chi.$$

Because Equation (52) holds and ε is an enumeration of (53), then by Lemma 3.21

$$B^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi)) = B^{\varepsilon}(\Sigma \cup \Gamma). \tag{54}$$

A substitution of Equation (54) yields $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \cup B^{\varepsilon}(\Sigma \cup \Gamma) \models_{ae} \chi$. As $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \subseteq \Gamma$, it follows that $\Sigma \cup \Gamma \cup B^{\varepsilon}(\Sigma \cup \Gamma) \models_{ae} \chi$ by the monotonicity of $\models_{ae}$. The enumeration ε can now be extended to an enumeration ε' of

$$\{\psi \mid L\psi \in Q - \text{q-atom}(\Gamma)\} \tag{55}$$

such that by the monotonicity of $\models_{ae}$, $\Sigma \cup \Gamma \cup B^{\varepsilon'}(\Sigma \cup \Gamma) \models_{ae} \chi$. This implies that ε' is a refuting enumeration, i.e. $\chi \in EE^{\varepsilon'}(\Sigma \cup \Gamma)$. Therefore, $\langle\Sigma, \Gamma\rangle \not\dashv_{ae} \chi$ which is a contradiction. Hence $\langle\Sigma, \Lambda\rangle \dashv_L \chi$.

($\Leftarrow$) Let $\langle\Sigma, \Lambda\rangle \dashv_L \chi$ hold. Assuming that $\langle\Sigma, \Gamma\rangle \not\dashv_{ae} \chi$, there is an enumeration ε of (55) such that $\chi \in EE^{\varepsilon}(\Sigma \cup \Lambda)$, i.e. $\Sigma \cup \Gamma \cup B^{\varepsilon}(\Sigma \cup \Gamma) \models_{ae} \chi$. This enumeration can be extended to an enumeration ε' of

$$\mathcal{L}_{ae} - \{\psi \mid L\psi \in \text{q-atom}(\Gamma)\}$$

by concatenating the remaining elements. Thereby $\Sigma \cup \Gamma \cup B^{\varepsilon'}(\Sigma \cup \Gamma) \models_{ae} \chi$ holds, since $\models_{ae}$ is monotonic. Let ε'' be the reduction of ε' with respect to $Sf^{L}(\Sigma \cup \{\chi\})$. It follows that ε'' is an enumeration of (53). Because $\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) = \Pi_{Sf^{L}(\Sigma\cup\{\chi\})}(\Gamma)$, Proposition 2.15 yields exactly that

$$\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \cup B^{\varepsilon''}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi)) \models_{ae} \chi.$$

Thereby it follows that $\chi \in EE^{\varepsilon''}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi))$, i.e. $\langle\Sigma, \Lambda\rangle \not\dashv_L \chi$. □

Proposition 3.23 (Closure under recursive consequence relations) *Let Γ be a consistent set of quasi-literals satisfying the closure conditions*

$$L\chi \in \Gamma \textit{ if and only if } \langle\Sigma, \Gamma\rangle \models_{ae} \chi \tag{56}$$

and

$$\neg L\chi \in \Gamma \textit{ if and only if } \langle\Sigma, \Gamma\rangle \dashv_{ae} \chi \tag{57}$$

for all $L\chi \in Q$. Then the set Γ is closed with respect to the recursive consequence relations, i.e. it holds for all $L\chi \in Q$

$$L\chi \in \Gamma \textit{ if and only if } \langle\Sigma, \Lambda\rangle \models_{L} \chi \tag{58}$$

and

$$\neg L\chi \in \Gamma \textit{ if and only if } \langle\Sigma, \Lambda\rangle \dashv_{L} \chi \tag{59}$$

where $\Lambda = \Pi_{Sf^{L}(\Sigma)}(\Gamma)$.

Proof. The proof is carried out by induction on $BD(\chi)$. Let $L\chi \in Q$.

Induction base: $BD(\chi) = 0$. Equation (58) is proved first. As $Sf^L(\chi) = \emptyset$,

$$SB_{\langle\Sigma,\Lambda\rangle}(\chi) = \emptyset = \Pi_{\emptyset}(\Gamma). \tag{60}$$

Then $L\chi \in \Gamma$ holds if and only if $\langle\Sigma,\Gamma\rangle \models_{ae} \chi$ holds, by the closure condition in (56). As Γ is a consistent set of belief literals and (60) holds, Lemma 3.22 can be applied. Thus $\langle\Sigma,\Gamma\rangle \models_{ae} \chi$ if and only if $\langle\Sigma,\Lambda\rangle \models_L \chi$. Let us then consider the equivalence (59). The second closure condition (57) yields that $\neg L\chi \in \Gamma$ if and only if $\langle\Sigma,\Gamma\rangle =\!\mid_{ae} \chi$. By Lemma 3.22, $\langle\Sigma,\Gamma\rangle =\!\mid_{ae} \chi$ if and only if $\langle\Sigma,\Lambda\rangle =\!\mid_L \chi$.

Inductive hypothesis:
The equivalences (58) and (59) hold for formulae χ having $BD(\chi) < k$.

Induction step: $BD(\chi) = k$. Let us begin by showing that

$$SB_{\langle\Sigma,\Lambda\rangle}(\chi) = \Pi_{Sf^L(\chi)}(\Gamma). \tag{61}$$

($\subseteq$) Let $L\chi' \in SB_{\langle\Sigma,\Lambda\rangle}(\chi)$. Then $L\chi' \in Sf^L(\chi)$ and $\langle\Sigma,\Lambda\rangle \models_L \chi'$. Since $Sf^L(Q) \subseteq Q$ and $BD(\chi') < k$ allowing the application of the inductive hypothesis, $L\chi' \in \Gamma$ and $L\chi' \in \Pi_{Sf^L(\chi)}(\Gamma)$. In the same fashion, $\neg L\chi' \in SB_{\langle\Sigma,\Lambda\rangle}(\chi)$ implies, $L\chi' \in \Pi_{Sf^L(\chi)}(\Gamma)$. ($\supseteq$) Let $L\chi' \in \Pi_{Sf^L(\chi)}(\Gamma)$. Then $L\chi' \in Sf^L(\chi)$ and $BD(\chi') < k$. By the inductive hypothesis, $\langle\Sigma,\Lambda\rangle \models_L \chi'$ and $L\chi' \in SB_{\langle\Sigma,\Lambda\rangle}(\chi)$. Similarly, $\neg L\chi' \in \Pi_{Sf^L(\chi)}(\Gamma)$ implies that $\neg L\chi'$ belongs to $SB_{\langle\Sigma,\Lambda\rangle}(\chi)$.

By the closure condition (56), $L\chi \in \Gamma$ if and only if $\langle\Sigma,\Gamma\rangle \models_{ae} \chi$. As Γ is a consistent set of belief literals and (61) holds, Lemma 3.22 yields that $\langle\Sigma,\Gamma\rangle \models_{ae} \chi$ if and only if $\langle\Sigma,\Lambda\rangle \models_L \chi$. What remains to be shown is the equivalence (59). $\neg L\chi \in \Gamma$ holds if and only if $\langle\Sigma,\Gamma\rangle =\!\mid_{ae} \chi$ holds by the closure condition (57). By Lemma 3.22, $\langle\Sigma,\Gamma\rangle =\!\mid_{ae} \chi$ if and only if $\langle\Sigma,\Lambda\rangle =\!\mid_L \chi$.
□

Proposition 3.24 (Properties of the iterative construction)
Let Σ be a set of sentences of $\mathcal{L}_{ae}$. Let $\Gamma = fp(FCAE_{\Sigma,Q}, \Gamma_0)$ and $\Lambda = \Pi_{Sf^L(\Sigma)}(\Gamma)$. Then Γ is closed with respect to the recursive consequence relations. Furthermore, Λ is Σ-full.

Proof. The fixed points of $FCAE_{\Sigma,Q}$ satisfy the closure conditions (56) and (57) by Definition 3.18. Therefore Γ satisfies Equations (58) and (59), i.e. Γ is closed with respect to the recursive consequence relations. The Σ-fullness of Λ becomes clear if we restrict the two equivalences in (58) and (59) to $Sf^L(\Sigma)$ instead of the whole Q. What remains is exactly the Σ-fullness conditions for Λ given in Definition 2.17. □

Proposition 3.25 clarifies the choice of Q. The iterative method constructs the reduction of the cautious expansion with respect to Q. Recall that we demanded the set Q to contain all elements of $Sf^L(\Sigma)$. This is crucial for the iterative construction of the Σ-full set.

Proposition 3.25 (Reduction property)
Let Σ be a set of sentences of $\mathcal{L}_{ae}$. Let $\Gamma = fp(FCAE_{\Sigma,Q}, \Gamma_0)$ and $\Lambda = \Pi_{Sf^L(\Sigma)}(\Gamma)$ the Σ-full set characterizing the corresponding cautious expansion $CE_\Sigma(\Lambda) = \langle \Delta^+, \Delta^- \rangle$. Then $\Gamma = \Pi_Q(L\Delta^+ \cup \neg L\Delta^-)$.

Proof. Let $L\chi \in Q$. Now,

$$
\begin{array}{rlll}
L\chi \in \Gamma & \text{iff} & \langle \Sigma, \Lambda \rangle \models_L \chi & \text{[Proposition 3.23]} \\
& \text{iff} & \chi \in \Delta^+ & \text{[Lemma 2.21]} \\
& \text{iff} & L\chi \in L\Delta^+ & \\
& \text{iff} & L\chi \in \Pi_Q(L\Delta^+ \cup \neg L\Delta^-) & \text{[Definition 2.6]}
\end{array}
$$

and

$$
\begin{array}{rlll}
\neg L\chi \in \Gamma & \text{iff} & \langle \Sigma, \Gamma \rangle \mathrel{=\!\!|}_L \chi & \text{[Proposition 3.23]} \\
& \text{iff} & \chi \in \Delta^- & \text{[Lemma 2.21]} \\
& \text{iff} & \neg L\chi \in \neg L\Delta^- & \\
& \text{iff} & \neg L\chi \in \Pi_Q(L\Delta^+ \cup \neg L\Delta^-). & \text{[Definition 2.6]}
\end{array}
$$

□

Suppose that we enlarge Q (but still keep it finite and L-subformulae closed) gradually towards the set $L\mathcal{L}_{ae}$. As a consequence of this, the fixed point fp($FCAE_{\Sigma,Q}, \Gamma_0$) approaches the set $L\Delta^+ \cup \neg L\Delta^-$ on the basis of Proposition 3.25. See what happens to Proposition 3.23 at the limit: $L\chi \in L\Delta^+ \cup \neg L\Delta^-$, i.e. $\chi \in \Delta^+$ if and only if $\langle \Sigma, \Lambda \rangle \models_L \chi$. Similarly, $\chi \in \Delta^-$ if and only if $\langle \Sigma, \Lambda \rangle \mathrel{=\!\!|}_L \chi$. Hence, we have derived exactly the relationship between the finitary characterization and the expansions presented in Lemma 2.21. This speaks for the combability of the different consequence relations presented in this work. To this point we have examined the general iteration. Now we fix the initial point of iteration to the empty set $\emptyset$, which fulfills the initial conditions given in (49). The empty set is generated by the empty enumeration λ from Σ. A consequence of this has been stated in Corollary 3.26. Recall that this claim does not hold for stable expansions in light of Example 2.

Corollary 3.26 (Existence of expansions)
Let Σ be a set of formulae of $\mathcal{L}_{ae}$. Then there is at least one cautious expansion of Σ.

Proof. As mentioned above, $\emptyset$ meets the initial conditions of (49). By Proposition 3.20 the fixed point $\Lambda = \mathrm{fp}(FCAE_{\Sigma,Sf^L(\Sigma)}, \emptyset)$ is reached in a finite number of steps. Furthermore, the fixed point Λ is Σ-full by Proposition 3.23. By Theorem 2.22 there is the corresponding cautious expansion. □

So far we have shown that the iterative method constructs the reduction of a cautious expansion determined by the set Q. One of the remaining questions is whether the corresponding cautious expansion meets the $\mathcal{N}$-groundedness condition. We show in Proposition 3.27 that it indeed does. This result was expected: at each step of iteration, the newly added beliefs are certainly grounded in the beliefs of the previous step.

Proposition 3.27 (Groundedness)
Let Σ be a set of formulae of $\mathcal{L}_{ae}$. Then $\Lambda = fp(FCAE_{\Sigma,Sf^L(\Sigma)}, \emptyset)$ is an $\mathcal{N}$-grounded Σ-full set.

Proof. In this proof Q is restricted to $Sf^L(\Sigma)$. Recall that $\emptyset$ fulfills the initial conditions given in (49). Let $\Lambda_0 = \emptyset$ and $\Lambda_{i+1} = FCAE_{\Sigma,Sf^L(\Sigma)}(\Lambda_i)$ when $i > 0$. We show that the sequence Λ_i ($i \geq 0$) is $\mathcal{N}$-grounded. The set Λ is Σ-full by Proposition 3.24. By Definition 3.4, Λ is $\mathcal{N}$-grounded if it satisfies for all $L\chi \in Sf^L(\Sigma)$

$$L\chi \in \Lambda \text{ if and only if } \Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi.$$

($\Rightarrow$) Let $L\chi \in \Lambda$. Suppose that the iteration reaches its fixed point in n steps. Then $\Lambda = \Lambda_n$. Recall that Λ_n is generated by the enumeration ε_n constructed in Proposition 3.19, i.e. $\Lambda = B^{\varepsilon_n}(\Sigma)$ and ε_n is an enumeration of $\{\psi \mid L\psi \in \text{q-atom}(\Lambda)\}$. Therefore, by Lemma 3.8, $L\chi \in \Lambda$ implies $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi$.

($\Rightarrow$) Let $\Sigma \cup [\Lambda]^- \vdash_{\mathcal{N}} \chi$ hold. Then $\Sigma \cup [\Lambda]^- \cup \{L\chi_1, ..., L\chi_m\} \models_{ae} \chi$. We may assume that the L-proof $L\chi, ..., L\chi_m$ is a minimal one and therefore it contains only members of $Sf^L(\Sigma \cup \{\chi\})$ by Proposition 2.14. We show by induction on j that $\{L\chi_1, ..., L\chi_j\} \subseteq \Lambda$ where $1 \leq j \leq m$.

Induction base: $j = 1$. By the definition of the L-proofs $\Sigma \cup [\Lambda]^- \models_{ae} \chi_1$. The monotonicity of $\models_{ae}$ yields $\Sigma \cup \Lambda \models_{ae} \chi_1$ and $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi_1) \models_{ae} \chi_1$. Thus $\langle\Sigma, \Lambda\rangle \models_L \chi_1$ and as Λ is Σ-full, $L\chi_1 \in \Lambda$.

Inductive hypothesis: $\{L\chi_1, ..., L\chi_{k-1}\} \subseteq \Lambda$.

Induction step: $j = k$. Then $\Sigma \cup [\Lambda]^- \cup \{L\chi_1, ..., L\chi_{k-1}\} \models_{ae} \chi_k$ holds, as χ_k is a member of the L-proof of χ. By the inductive hypothesis and the monotonicity of $\models_{ae}$, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi_k) \models_{ae} \chi_k$. Therefore, $\langle\Sigma, \Lambda\rangle \models_L \chi_k$ and $L\chi_k \in \Lambda$.

So the L-proof of χ is contained in Λ implying that $\Sigma \cup \Lambda \models_{ae} \chi$, as $\models_{ae}$ is monotonic. For the same reason, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \models_{ae} \chi$, i.e. $\langle\Sigma,\Lambda\rangle \models_L \chi$ and $L\chi \in \Lambda$. $\square$

Next we show that $\mathrm{fp}(FCAE_{\Sigma,Sf^L(\Sigma)},\emptyset)$ is contained in any other $\mathcal{N}$-grounded cautious expansion. Having derived this, we are able to prove that the fixed point corresponds to the least cautious expansion. This is in fact a direct consequence from the choice of the empty set as the starting point of the iteration.

Proposition 3.28 (Containment)
Let Σ be a set of sentences of $\mathcal{L}_{ae}$. Then $fp(FCAE_{\Sigma,Sf^L(\Sigma)},\emptyset)$ is contained in any other $\mathcal{N}$-grounded Σ-full set.

Proof. Let Λ be an arbitrary $\mathcal{N}$-grounded Σ-full set. By Lemma 3.5, Λ is generated by an enumeration ε of $\{\psi \mid L\psi \in \text{q-atom}(\Lambda)\}$. Let Λ_i's be defined similarly as in Proposition 3.27. We show by induction on i that $\Lambda_i \subseteq \Lambda$.

Induction base: $i = 0$. Trivially, $\Lambda_0 = \emptyset \subseteq \Lambda$.

Inductive hypothesis: $\Lambda_i \subseteq \Lambda$ for $i < k$.

Induction step: $i = k$. Let $L\chi \in \Lambda_k = FCAE_{\Sigma,Sf^L(\Sigma)}(\Lambda_{k-1})$ hold. Then $\Sigma \cup \Lambda_{k-1} \models_{ae} \chi$. By the inductive hypothesis and the monotonicity of $\models_{ae}$, $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi) \models_{ae} \chi$. Thus $\langle\Sigma,\Lambda\rangle \models_L \chi$. As Λ is Σ-full, $L\chi \in \Lambda$.

Let $\neg L\chi \in \Lambda_k$. Then $\neg L\chi \in FCAE_{\Sigma,Sf^L(\Sigma)}(\Lambda_{k-1})$ i.e. $\langle\Sigma,\Lambda_{k-1}\rangle \dashv_{ae} \chi$. Assume now that $\langle\Sigma,\Lambda\rangle \not\dashv_L \chi$. As Λ is subbelief closed and $L\chi \in Sf^L(\Sigma)$, $SB_{\langle\Sigma,\Lambda\rangle}(\chi) \subseteq \Lambda$. Then there is an enumeration ε_r of

$$\{\psi \mid L\psi \in Sf^L(\Sigma) - \text{q-atom}(\Lambda)\} \tag{62}$$

such that $\chi \in EE^{\varepsilon_r}(\Sigma \cup \Lambda)$, i.e. $\Sigma \cup \Lambda \cup B^{\varepsilon_r}(\Sigma \cup \Lambda) \models_{ae} \chi$. As Λ is generated by the enumeration ε, $\Lambda = B^{\varepsilon}(\Sigma)$. By the inductive hypothesis $\Lambda_{k-1} \subseteq \Lambda$. Furthermore, by Proposition 3.19 the set Λ_{k-1} is generated by an enumeration, say ε_{k-1}. Consider then a quasi-literal $\neg L\chi' \in \Lambda$. It follows that $\langle\Sigma,\Lambda\rangle \dashv_L \chi'$. Then by Proposition 2.18, it holds that $\langle\Sigma,\Lambda\rangle \not\models_L \chi'$, i.e. $\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\chi') \not\models_{ae} \chi'$. As Λ is subbelief closed and $L\chi' \in Sf^L(\Sigma)$, $SB_{\langle\Sigma,\Lambda\rangle}(\chi') \subseteq \Lambda$. Therefore, $\Sigma \cup \Lambda \not\models_{ae} \chi'$. We have thus shown that Λ_{k-1} and Λ satisfy the conditions of Lemma 3.7. Then there is an enumeration ε_d of $\{\psi \mid L\psi \in \text{q-atom}(\Lambda - \Lambda_{k-1})\}$ generating the difference $\Lambda - \Lambda_{k-1}$ such that $\Lambda - \Lambda_{k-1} = B^{\varepsilon_d}(\Sigma \cup \Lambda_{k-1})$. This yields that $\Sigma \cup B^{\varepsilon_{k-1}\cdot\varepsilon_d\cdot\varepsilon_r}(\Sigma) \models_{ae} \chi$, or equivalently

$$\Sigma \cup \Lambda_{k-1} \cup B^{\varepsilon_d\cdot\varepsilon_r}(\Sigma \cup \Lambda_{k-1}) \models_{ae} \chi.$$

Now $\varepsilon_d \cdot \varepsilon_r$ is an enumeration of $\{\psi \mid L\psi \in Sf^L(\Sigma) - \text{q-atom}(\Lambda_{k-1})\}$. Moreover, it refutes χ. This implies that $\langle\Sigma,\Lambda_{k-1}\rangle \not\dashv_{ae} \chi$, which is a contradiction. Therefore, $\langle\Sigma,\Lambda\rangle \dashv_L \chi$ and $\neg L\chi \in \Lambda$, as Λ is Σ-full. Thus $\Lambda_k \subseteq \Lambda$ holds.

We have now established that $\Lambda_i \subseteq \Lambda$ for all $i \geq 0$. Therefore, the fixed point

$$\Lambda_n = \mathrm{fp}(FCAE_{\Sigma,Sf^L(\Sigma)}, \emptyset)$$

is also contained in Λ. □

Theorem 3.29 states the fundamental result of this section: the least $\mathcal{N}$-grounded cautious expansion always exists. This result is in contrast with the original autoepistemic logic. By Example 2 we know that stable expansions do not exist for every set of premises.

Theorem 3.29 (Characterization of the iteration)
Let Σ be the set of premises. Then $\Lambda = fp(FCAE_{\Sigma,Sf^L(\Sigma)}, \emptyset)$ is the least $\mathcal{N}$-grounded Σ-full set and $CE_\Sigma(\Lambda)$ is the least $\mathcal{N}$-grounded cautious expansion. Furthermore, both Λ and $CE_\Sigma(\Lambda)$ are unique.

Proof. Consider the leastness of Λ first. Suppose Λ' is an $\mathcal{N}$-grounded Σ-full set such that $\Lambda' \subseteq \Lambda$. By Proposition 3.28, $\Lambda \subseteq \Lambda'$. Hence $\Lambda' = \Lambda$ and Λ is the least $\mathcal{N}$-grounded Σ-full set. Let $\mathcal{E}'$ be any $\mathcal{N}$-grounded cautious expansion of Σ. Then $\mathcal{E}' = CE_\Sigma(FS_\Sigma(\mathcal{E}'))$ by Theorem 3.12. As Λ is the least $\mathcal{N}$-grounded Σ-full set, $\Lambda \subseteq FS_\Sigma(\mathcal{E}')$. By Proposition 3.15 it holds that $CE_\Sigma(\Lambda) \sqsubseteq \mathcal{E}'$ and thus $CE_\Sigma(\Lambda)$ is the least $\mathcal{N}$-grounded cautious expansion.

Both Λ and $CE_\Sigma(\Lambda)$ are the least elements in the respective partial orders. The antisymmetry of partial orders implies the uniqueness of these elements.
□

3.4 Computational Aspects

We shall now turn our attention to computability and computational complexity. Davis [1983] and Balcázar [1988] address these concepts in further detail. We say that a function is *computable* if it can be computed by some *model of computation* in a finite time. *Turing machines* are an example of such models of computation. From the practical point of view, to show that a given function is computable we should find an algorithm that computes the function in a finite number of steps. The term *decidability* is closely related to computability. A *decision problem* involves making a membership test - deciding whether an element belongs to a set satisfying a certain criteria or not. Cautious autoepistemic logic is decidable if we can determine the membership of a given formula in expansions by means of an algorithm.

The subsequent propositions are based on the fact that classical propositional logic is decidable, i.e. the relation $\models$ is decidable. *Truth tables* and *analytic tableaux* [Smullyan, 1968] provide means of solving the propositional decision problem in the general case. These methods are also applicable to $\models_{ae}$, because this extended consequence relation treats quasi-literals as atomic propositions.

```
Function PosL(Σ, Λ, φ): Boolean;
Var
   Σ: a finite subset of L_ae;
   Λ,SB: finite sets of quasi-literals;
   φ: a formula of L_ae;
Begin
   SB:=∅;
   For Lχ ∈ Sf^L(φ) Do
      If PosL(Σ, Λ, χ) Then
         SB:=SB∪{Lχ}
      Elif NegL(Σ, Λ, χ) Then
         SB:=SB∪{¬Lχ};
   If Σ ∪ Λ ∪ SB ⊨_ae φ Then
      PosL:=True
   Else
      PosL:=False;
End
```

Figure 1: Decision procedure for $\models_L$

In Proposition 3.30 we show that the recursive consequence relations are decidable. The proof is based on algorithms PosL and NegL presented in Figures 1 and 2. The algorithms have been written in pascal-like syntax. A few notes on the statements of the language is needed. The **Exit**-statement causes a function to return its value immediately. The execution of the innermost loop is continued when a **Continue**-statement is reached. **For**-loops choose elements from sets in an arbitrary order, but the elements of an enumeration are processed in the order determined by the enumeration. We use $En(\Gamma)$ to denote the set of all enumerations of $\{\psi \mid L\psi \in \Gamma\}$.

Proposition 3.30 (Decidability of the consequence relations)
Let Σ be a finite set of formulae of $\mathcal{L}_{ae}$ and Λ be a finite set of quasi-literals. Then the recursive consequence relations $\models_L$ and $\dashv_L$ are decidable.

Proof. We prove the assertion by induction on $BD(\phi)$.

Induction base: $BD(\phi) = 0$. As $BD(\phi) = 0$, $Sf^L(\phi) = \emptyset$ and $SB_{\langle\Sigma,\Lambda\rangle}(\phi) = \emptyset$. In both Algorithms PosL and NegL, **For**-loop is not executed and SB remains empty. Let us first consider Algorithm PosL. To test whether $\langle\Sigma, \Lambda\rangle \models_L \phi$

Function NegL(Σ, Λ, ϕ): Boolean;
  Σ: a finite subset of $\mathcal{L}_{ae}$;
  Λ,SB,BE: finite sets of quasi-literals;
  ε: a finite enumeration of formulae;
  ϕ, χ: formulae of $\mathcal{L}_{ae}$.
Begin
  SB:=$\emptyset$;
  For $L\chi \in Sf^L(\phi)$ **Do**
    If PosL(Σ, Λ, χ) **Then**
      SB:=SB$\cup\{L\chi\}$
    Elif NegL(Σ, Λ, χ) **Then**
      SB:=SB$\cup\{\neg L\chi\}$;
  NegL:=False;
  For $\varepsilon \in En(Sf^L(\Sigma \cup \{\phi\}) -$ q-atom$(\Lambda \cup \mathrm{SB}))$ **Do**
  Begin
    BE:=Introspect($\Sigma \cup \Lambda \cup \mathrm{SB}, \varepsilon$);
    If $\Sigma \cup \Lambda \cup \mathrm{SB} \cup \mathrm{BE} \models_{ae} \phi$ **Then**
      Exit
  End
  NegL:=True;
End

Figure 2: Decision procedure for $=\!\!|_L$

holds one has to check whether $\Sigma \cup \Lambda \models_{ae} \phi$ holds on the basis of the definition of $\models_L$. Because SB equals to $\emptyset$, Algorithm PosL checks exactly this. Because $\models_{ae}$ is decidable, $\models_L$ is decidable in this restricted case. Then consider Algorithm NegL which is the decision procedure for $\dashv_L$. To check whether $\langle \Sigma, \Lambda \rangle \dashv_L \phi$ holds, every enumeration ε of $\{\psi \mid L\psi \in Sf^L(\Sigma) - \text{q-atom}(\Lambda)\}$ must be examined. This is done by means of the second **For**-loop in Algorithm NegL. Because the enumeration base is finite, there are only finitely many enumerations to check. For each enumeration ε, the set $B^\varepsilon(\Sigma \cup \Lambda)$ is constructed. The function Introspect for constructing this set is given in Figure 3. This set of belief literals is computable as decisions are based on $\models_{ae}$ and the enumeration ε is finite. The variable BE is assigned the return value of Introspect. For each enumeration ε, it is verified that $\phi \notin EE^\varepsilon(\Sigma \cup \Lambda)$ holds. A test is made whether ϕ is a logical consequence of $\Sigma \cup \Lambda \cup \text{BE}$. This is decidable as $\models_{ae}$ is. Therefore, $\dashv_L$ is decidable when $BD(\phi) = 0$.

Function Introspect(Σ, ε): a finite set of quasi-literals;
  Σ: a finite subset of $\mathcal{L}_{ae}$;
  EB: a finite set of quasi-literals;
  ε: a finite enumeration of formulae;
  χ: a formula of $\mathcal{L}_{ae}$.
Begin
  EB:=$\emptyset$;
  For $\chi \in \varepsilon$ **Do**
    If $\Sigma \cup \text{EB} \models_{ae} \chi$ **Then**
      EB:=EB$\cup\{L\chi\}$
    Else
      EB:=EB$\cup\{\neg L\chi\}$;
  Introspect:=EB;
End

Figure 3: Enumeration-based introspection

Inductive hypothesis:
The relations $\models_L$ and $\dashv_L$ are decidable when $BD(\phi) < k$.

Induction step: $BD(\phi) = k$. The formulae $L\chi \in Sf^L(\phi)$ satisfy $BD(\chi) < k$. By the inductive hypothesis, the both relations $\models_L$ and $\dashv_L$ are decidable for such formulae. Therefore, the set of subbeliefs of ϕ, $SB_{\langle\Sigma,\Lambda\rangle}(\phi)$, is computable. This set is computed by the first **For**-loop in the both of functions PosL and NegL. Since $Sf^L(\phi)$ is finite, the loop is executed only finite number of times. After the execution of the loop, SB is assigned to $SB_{\langle\Sigma,\Lambda\rangle}(\phi)$. Checking whether $\langle \Sigma, \Lambda \rangle \models_L \phi$ holds, requires a single propositional test in the function PosL, namely $\Sigma \cup \Lambda \cup \text{SB} \models_{ae} \phi$. This is decidable as $\models_{ae}$ is. Then consider the function NegL. To check whether $\langle\Sigma,\Lambda\rangle \mathrel{=\!\!|}_L \phi$ holds we must examine every enumeration ε of $\{\psi \mid L\psi \in Sf^L(\Sigma) - \text{q-atom}(\Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))\}$ and verify that $\phi \notin EE^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$. This is achieved by the second **For**-loop in the function NegL. Firstly, there are only finitely many enumerations to check, as the enumeration base is finite. Testing whether $\phi \notin EE^{\varepsilon}(\Sigma \cup \Lambda \cup SB_{\langle\Sigma,\Lambda\rangle}(\phi))$ is decidable by arguments similar to those in the induction base. Therefore, $\models_L$ and $\mathrel{=\!\!|}_L$ are decidable for the formulae of arbitrary belief depth. □

The remaining concern is whether full sets are computable. In Proposition 3.31 we show that the least $\mathcal{N}$-grounded full set is computable. A more general result is given Proposition 3.32: all Σ-full sets are computable.

Function Neg(Σ, Γ, Q, ϕ): Boolean;
  Σ: a finite subset of $\mathcal{L}_{ae}$;
  Γ,BE: finite sets of quasi-literals;
  Q: a finite set of quasi-atoms satisfying $Sf^L(\Sigma) \subseteq Q$ and $Sf^L(Q) \subseteq Q$;
  ε: a finite enumeration of formulae;
  ϕ: a formula of $\mathcal{L}_{ae}$;
Begin
  Neg:=False;
  For $\varepsilon \in En(Q - \text{q-atom}(\Gamma))$ **Do**
  Begin
    BE:=Introspect($\Sigma \cup \Gamma, \varepsilon$);
    If $\Sigma \cup \Gamma \cup \text{BE} \models_{ae} \phi$ **Then**
      Exit
  End
  Neg:=True;
End

Figure 4: Decision procedure for $\mathrel{=\!\!|}_{ae}$

Proposition 3.31
Let Σ be a finite set of formulae of $\mathcal{L}_{ae}$ Then the least $\mathcal{N}$-grounded Σ-full set is computable.

Proof. The iterative method presented in Section 3.3 is available for the construction of the least $\mathcal{N}$-grounded Σ-full sets by Theorem 3.29. The consequence relation $\mathrel{=\!\!|}_{ae}$ is decidable by similar arguments as its recursive counterpart. The function Neg in Figure 4 implements $\mathrel{=\!\!|}_{ae}$. To show whether $\langle\Sigma,\Gamma\rangle \mathrel{=\!\!|}_{ae} \chi$ holds, all the enumerations ε of $\{\psi \mid L\psi \in Q - \text{q-atom}(\Gamma)\}$ have to be checked. Because Q is finite, there are only finitely many enumerations to examine. The set $B^{\varepsilon}(\Sigma \cup \Gamma)$ is computed by the function Introspect which is computable. The test $\Sigma \cup \Gamma \cup \mathrm{EB} \models_{ae} \chi$ is decidable. This implies that $\dashv|_{ae}$ is a decidable consequence relation. As a consequence the operator $FCAE_{\Sigma,Q}$ is computable. Algorithm FCAE in Figure 5 takes the monotonicity of $FCAE_{\Sigma,Q}$ into account and tests only the formulae in the difference $Sf^L(\Sigma) - \text{q-atom}(\Lambda)$. By Proposition 3.20, the fixed points of $FCAE_{\Sigma,Q}$ are reached in finite number of steps. Therefore, the fixed point, i.e. the least $\mathcal{N}$-grounded Σ-full set, is computable. □

Function FCAE(Σ): a set of quasi-literals;
  Σ: a finite subset of $\mathcal{L}_{ae}$;
  $\Lambda, \Lambda^+, \Lambda^-$: sets of quasi-literals;
  χ: a formula of $\mathcal{L}_{ae}$;
Begin
  $\Lambda := \emptyset$;
  Repeat
    $\Lambda^+ := \emptyset, \Lambda^- := \emptyset$;
    For $L\chi \in Sf^L(\Sigma) - \text{q-atom}(\Lambda)$ **Do**
      If $\Sigma \cup \Lambda \models_{ae} \chi$ **Then**
        $\Lambda^+ := \Lambda^+ \cup \{L\chi\}$
      Elif $\Sigma \cup \Lambda \dashv|_{ae} \chi$ **Then**
        $\Lambda^- := \Lambda^- \cup \{\neg L\chi\}$;
    $\Lambda := \Lambda \cup \Lambda^+ \cup \Lambda^-$;
  Until $\Lambda^+ \cup \Lambda^- = \emptyset$;
End

Figure 5: Construction of the least $\mathcal{N}$-grounded full set

Proposition 3.32
Let Σ be a finite set of formulae of $\mathcal{L}_{ae}$. Then the set of all Σ-full sets is computable.

Proof. Σ-full sets can be computed by exhaustively checking all the candidates, i.e. all consistent subsets of $Sf^L(\Sigma) \cup \neg Sf^L(\Sigma)$. This is the strategy in the function FullSets in Figure 7. The function Full in Figure 6 implements the test whether a set of quasi-literals is Σ-full. By Proposition 3.30 the recursive consequence relations are decidable. As $Sf^L(\Sigma)$ is finite, the candidate Σ-full sets are finite. If each quasi-literal belonging to the set $Sf^L(\Sigma)$ satisfies the conditions of Definition 2.17, the candidate is accepted as a Σ-full set. It is therefore decidable, whether a set of quasi-literals is Σ-full. Furthermore, the set of all Σ-full sets is computable. □

```
Function Full(Σ, Λ): Boolean;
  Σ: a finite subset of ℒ_ae;
  Λ: a set of quasi-literals;
  χ: a formula of ℒ_ae;
Begin
  Full:=False;
  For Lχ ∈ Sf^L(Σ) Do
    If Lχ ∈ Λ Then
      If ⟨Σ, Λ⟩ ⊨_L χ Then
        Continue Else Exit
    Elif ¬Lχ ∈ Λ Then
      If ⟨Σ, Λ⟩ ⫤_L χ Then
        Continue Else Exit
    Elif ⟨Σ, Λ⟩ ⊨_L χ or ⟨Σ, Λ⟩ ⫤_L χ Then
      Exit;
  Full:=True;
End
```

Figure 6: Test for Σ-fullness

```
Function FullSets(Σ): a set of sets of quasi-literals;
  Σ: a finite subset of ℒ_ae;
  Λ: a set of quasi-literals;
Begin
  FullSets:=∅;
  For Λ ∈Candidates(Sf^L(Σ)) Do
    If Full(Σ, Λ) Then
      FullSets:=FullSets∪{Λ};
End
```

Figure 7: Computation of the Σ-full sets

Theorem 3.33 (Decidability of cautious autoepistemic logic)
Let Σ be a finite set of formulae. Then the problems whether

1. *a formula ϕ belongs to one of the cautious expansions of Σ*
2. *a formula ϕ belongs to all cautious expansions of Σ*

are decidable.

Proof. The Σ-full sets are computable by Proposition 3.32. Theorem 2.22 yields that cautious expansions are in one-to-one correspondence with Σ-full sets. The recursive consequence relations provide means to test whether a given formula belongs to the cautious expansion determined by a Σ-full set. By Proposition 3.30, these relations are decidable. Therefore, the membership problem of cautious expansions is decidable. This holds also for the least $\mathcal{N}$-grounded cautious expansion, because it is computable by Proposition 3.31. $\square$

We have left the exact complexity analysis of the iterative construction outside this work. The two sources of computational complexity are clear, however. The one is the propositional decision problem and the other the enumerations employed in the negative introspection. Syntactic restrictions provide means of reducing the former source of computational complexity. The latter requires still further research to be done. There are several ways to reduce the number of enumerations that have to be examined when a refuting enumeration is searched for.

Notice that recursive consequence relations are rather impractical from the point of view of implementations as they always require the computation of the subbeliefs of the formula under investigation. The same applies to the sketches of algorithms presented in the preceding proofs: they are enough to point out that cautious autoepistemic logic is decidable, but are computationally inefficient. The iterative construction suggest a better implementation strategy. Full sets (or arbitrary parts of expansions) are constructed by gradual addition of quasi-literals. The proofs presented in section 3.3 were intentionally based on generating enumerations and the closure under the non-recursive consequence relations. These requirements have to be met by the optimized versions of the operator $FCAE_{\Sigma,Q}$.

The iterative construction provides another way of optimizing the decision procedure. The intermediate results of computation, i.e. the sets Γ_i ($i \geq 0$) defined in conjunction with the operator $FCAE_{\Sigma,Q}$ can be used for testing whether a formula is believed or disbelieved. Thus it is not always necessary to compute the complete finitary characterization of the least autoepistemic expansion to determine the belief status of an autoepistemic formula. The formulae that remain undetermined are the hardest: their belief status becomes certain as late as the fixed point has been reached.

Given the set of premises, the iteration may be understood as a compilation process of the knowledge base. The result of the compilation, i.e. the corresponding full set can then be used as basis for answering queries. This requires implementation of the recursive consequence relations in a form or another. A possibility is to use $FCAE_{\Sigma,Q}$ for constructing the subbeliefs of any formula ϕ. This is easily achieved by setting Q to $Sf^L(\Sigma \cup \{L\phi\})$ and testing the belief status of ϕ within in the fixed point. However, this requires an additional proof that $\mathcal{N}$-grounded Σ-full sets are valid as initial points of iteration. This should be straightforward as $\mathcal{N}$-grounded Σ-full sets are generated by enumerations. By varying the initial point, other $\mathcal{N}$-grounded full sets than the least one can be constructed iteratively.

3.5 Comparison with Other Approaches

It is time to compare cautious autoepistemic logic with other approaches taking advantage of the weakened negative introspection. We shall refer to our repertoire of examples to point out, why the approach presented in this work has interesting properties.

First of all, it should be emphasized that the least $\mathcal{N}$-grounded cautious expansion is different from the intersection of stable expansions. In Example 4 this difference becomes clear. The set of premises $\Sigma = \{\neg Lp \rightarrow q, \neg Lq \rightarrow q, \neg Lp \wedge \neg Lq \rightarrow r\}$ has two stable expansions corresponding to Σ-full sets $\{\neg Lp, Lq\}$ and $\{Lp, \neg Lq\}$. The intersection of these stable expansions contains for instance the formula $\neg Lp \vee \neg Lq$. However, this formula is not contained in the least $\mathcal{N}$-grounded cautious expansion. Thus Example 4 demonstrates that conclusions based on the least $\mathcal{N}$-grounded cautious expansion are weaker than the ones based on the intersection of stable expansions.

Przymusinski [1991b] has proposed an autoepistemic logic of closed beliefs AEL_{circ}. In his work circumscription is used for weakening the negative introspection capability of the agent. Przymusinski introduced a nonmonotonic consequence relation $\models_{circ}$ requiring falsity in all minimal models:

$$\Sigma \models_{circ} \neg L\phi \text{ if and only if } \phi \text{ is false in all minimal models of } \Sigma.$$

For example, the set of premises $\Sigma = \{Lp \rightarrow p\}$ has a unique minimal model $\emptyset$. Therefore $\Sigma \models_{circ} \neg Lp$, i.e. p is not believed in Przymusinski's approach. The circumscriptive approach has been designed to provide a characterization of the semantics of general logic programs by means of autoepistemic expansions.

In fact the least expansion in this circumscriptive approach coincides with the leading semantics of general logic programs called the *well-founded semantics* [Van Gelder *et al.*, 1991].

At the early stages of this work we considered the possibility of extending Przymusinski's approach to a larger class of autoepistemic theories than general logic programs having been translated into autoepistemic logic. In Example 3 the set of premises $\Sigma = \{\neg L\neg p \rightarrow p\}$ has among others propositional models $\{L\neg p, p\}$, $\{L\neg p\}$ and $\{p\}$. The last two of these are minimal with respect to the subset inclusion. Hence $\neg p$ is not disbelieved, as $\neg p$ is true in the minimal model $\{L\neg p\}$. The circumscriptive agent would therefore end up in a state where the beliefs status of p is left undetermined. In light of this example, it is not clear how to extend Przymusinski's approach to a larger class of autoepistemic theories in a natural way. The minimal models are not well suited for reasoning whether, e.g., a negated atomic proposition should be disbelieved. Stable and cautious expansions of the given set of premises are more accurate: Σ has both stable and cautious expansion corresponding to the Σ-full set $\{\neg L\neg p\}$.

Kalinski's weak autoepistemic reasoning [1992] utilizes standard modal logic KD45 for both positive and negative introspection. His work is also centered around semantics of logic programs. Kalinski defines his introspection scheme in the following way:

$$\begin{aligned}
\mathcal{E}^{+}_{\Sigma}(T,F) &= \{\phi \mid \phi \in \mathcal{L}, \vdash_{KD45} \Sigma \cup L(T) \cup \neg L(F) \rightarrow \phi\}\\
\mathcal{E}^{-}_{\Sigma}(T,F) &= \{\phi \mid \phi \in \mathcal{L},\ \textit{there is no } \Sigma\textit{-consistent } \langle T',F'\rangle \textit{ such that}\\
&\qquad \langle T,F\rangle \subseteq \langle T',F'\rangle \textit{ and } \vdash_{KD45} \Sigma \cup L(T') \cup \neg L(F') \rightarrow \phi\}.
\end{aligned}$$

The pair $\langle T',F'\rangle$ is Σ-consistent if and only if there is a KD45-model M with

$$M, w \models \Sigma \cup L(T') \cup \neg L(F')$$

where w is a possible world. The sets T and F are subsets of $\mathcal{L}$, and they represent the the objective (L-free) part of the weak autoepistemic expansion. The two operators $\mathcal{E}^{+}_{\Sigma}$ and $\mathcal{E}^{-}_{\Sigma}$ form the body of an iterative method for computing the least weak autoepistemic expansion. The iterations are started from the pair $\langle\emptyset,\emptyset\rangle$. However, this scheme of weakened negative introspection yields different expansions than the least $\mathcal{N}$-grounded cautious expansion in CAEL. To show this, let us re-examine Example 1 where the set of premises was $\Sigma = \{Lp \rightarrow p\}$. Firstly, p is not derivable from Σ in KD45. Secondly, there is a Σ-consistent pair $\langle\{p\},\emptyset\rangle$ such that $\langle\emptyset,\emptyset\rangle \subseteq \langle\{p\},\emptyset\rangle$ and $\vdash_{KD45} \Sigma \cup \{Lp\} \rightarrow p$. To verify the Σ- consistency of $\langle\{p\},\emptyset\rangle$ consider for example the KD45-model $M = \langle W, R, V\rangle$ where the set of possible worlds $M = \{w_1\}$, the accessibility relation $R = \{\langle w_1, w_1\rangle\}$, the valuation function V assigns p true in $\{w_1\}$ and $M, w_1 \models \Sigma \cup \{Lp\}$ holds (in fact any reflexive KD45-model would do). As a consequence, the belief status of p is left undetermined. The least $\mathcal{N}$-grounded cautious expansion of Σ contains the formula $\neg Lp$, i.e. the agent disbelieves p. In light of the last two examples, the enumeration-based negative introspection is stronger than that based on circumscription or modal logic KD45.

Example 2 is related to a problem in the field of logic programming. General logic programs consist of *clauses* of the form $\mathtt{p}_0$ `:-` $\mathtt{p}_1$`,...,`$\mathtt{p}_m$`,not` $\mathtt{p}_{m+1}$`,...,not` $\mathtt{p}_{m+n}$ where $\mathtt{p}_0$ is called the *head* of the clause and the rest of the formulae form the *body* of the clause. General logic programs can be translated into autoepistemic logic by replacing negation as failure by disbelief. The general logic program corresponding to the set of premises $\Sigma = \{\neg Lp \rightarrow p\}$ is `p :- not p`. This program is anomalous, as it causes for example standard PROLOG interpreters to loop forever (actually this program loops till the interpreter runs out of memory). In a way, the "infinite computation" of the interpreter means that the truth value of p is not known. In the least $\mathcal{N}$-grounded cautious expansion corresponding to the Σ-full set $\emptyset$, the belief status of p is left undetermined. The well-founded semantics of general logics programs agrees about this conclusion.

However, the least $\mathcal{N}$-grounded cautious expansion is different from the well founded semantics in general. Example 4 demonstrates this interesting dissimilarity. The well-founded semantics of a general logic program P is defined by means of an operator W_P. Again, an iterative construction is given. W_P operates on *partial models* I of instantiated general logic programs. Partial models can be presented as sets of *literals*, i.e. atomic propositions or their negations. We say that an instantiated program clause is *satisfied* by a partial interpretation I, if the head is true in I or some of the subgoals of the body is false in I; it is *falsified* if the head is false and all subgoals of the body are true. The operator W_P is defined by the equation

$$W_P(I) = T_P(I) \cup \neg U_P(I).$$

The operator $T_P(I)$ gives the heads of the program clauses whose body evaluates to true under the partial interpretation I. The greatest *unfounded set* of P with respect to I is given by $U_P(I)$. The unfounded sets which contain positive atoms are defined as follows. Let A be an unfounded set and p its element. Then each program clause C whose head is p meets at least one of the following two conditions: *either* some subgoal of the body of C is false in I *or* some positive subgoal of the body appears in A. The greatest unfounded set $U_P(I)$ is the union of all unfounded sets of P with respect to I. Intuitively, the atoms in an unfounded set cannot be derived by means of any clause if the atoms in the unfounded set are set false in the partial model under construction. This principle is somewhat similar to the negative introspection in cautious autoepistemic logic, but there is a difference, as we shall see.

The well-founded semantics of a general logic program P is defined as the least fixed point of W_P when the iteration is started from the the empty partial model. A reader interested in further details of the well-founded semantics is referred to [Van Gelder *et al.*, 1991]. Let us now return to the example. Recall that $\Sigma = \{\neg Lp \rightarrow q, \neg Lq \rightarrow p, \neg Lp \wedge \neg Lq \rightarrow r\}$ was the set of premises considered in Example 4. The corresponding general logic program P contains the following clauses:

```
p :- not q
q :- not p
r :- not p, not q.
```

Let $I_0 = \emptyset$ be the initial partial model as required by the well-founded semantics. Firstly, $T_P(I_0) = \emptyset$, because the partial model I_0 does not satisfy any of the bodies of the clauses in P. Then consider any *nonempty* subset S of $\{p, q, r\}$. The partial interpretation I_0 does not falsify any subgoal of the program clauses in P, as required by the first unfoundedness condition. Moreover, there are no positive occurrences of the atoms of S in P, so that the second condition for unfounded sets cannot hold. Therefore S cannot be an unfounded set. It follows that $\emptyset$, fulfilling the unfoundedness conditions vacuously, is the greatest unfounded set of P with respect to I_0 and $W_P(I_0) = I_0$, i.e. I_0 is the least fixed point of W_P. Thus, the well-founded semantics leaves the truth values of all three atoms p, q and r undetermined. In contrast to this, $\neg Lr$ belongs to the least $\mathcal{N}$-grounded cautious expansion corresponding to the Σ-full set $\emptyset$. This is intuitive, as both p and q are never disbelieved *simultaneously* by the agent.

The analyzed examples have demonstrated the strength of enumeration-based negative introspection if compared to other schemes of weakened negative introspection. It seems possible that cautious expansions offer an alternative semantics for general logic programs. We expect that a generalization to logic programs with classical negation can be accomplished. This will be one of the interesting topics in the further research work.

4 Conclusions

We have presented a new class of autoepistemic expansions in the propositional case, extending Moore's original approach. These expansions are called cautious expansions. A new scheme has been introduced for weakening the negative introspection capability of the agent. The modified negative introspection takes advantage of enumeration-based introspection. This mechanism is employed to examine the consequences of making negative introspection decisions. The leading principle is to strive for as strong negative introspection of the set of premises Σ as possible, but maintain the beliefs of the agent Σ-acceptable.

The new class of autoepistemic expansions can be characterized in a finitary way. In this work we have adapted Niemelä style full sets for characterizing cautious expansions. The full set approach is a convenient technique for reasoning about cautious expansions. Furthermore, its generality makes it even more practical – the major forms of autoepistemic logics do have finitary characterizations based on the notion of full sets. Therefore, full sets provide us compact means of comparing expansions. The proofs of correctness of the finitary characterization have been included in this work. Moreover, cautious expansions are shown to be in one-to-one correspondence with full sets.

The notion of $\mathcal{N}$-groundedness has been introduced to exclude cautious expansions having self-supporting positive beliefs. $\mathcal{N}$-groundedness is reflected to the finitary characterization: full sets are given a similar groundedness requirement that is in harmony with the $\mathcal{N}$-groundedness of the cautious expansions. The subset inclusion of $\mathcal{N}$-grounded full sets was shown to correspond to the subset inclusion of cautious expansions. Therefore, the least $\mathcal{N}$-grounded Σ-full set corresponds to the least $\mathcal{N}$-grounded cautious expansion.

In the rest of the work, we have concentrated on examining the properties of $\mathcal{N}$-grounded cautious expansions. One of the important results is that the least $\mathcal{N}$-grounded cautious expansion does always exist. Furthermore, we have provided an iterative construction for the least $\mathcal{N}$-grounded cautious expansion. This is rather exceptional, as stable and cautious expansions are nonconstructive in general. In the iterative construction we have identified the strongly negative beliefs, i.e. the ones that preserve their belief status after insertion to expansion under construction.

A comparison has been made with the other approaches utilizing weakened negative introspection. They are all centered around the well-founded semantics of general logic programs. A perception is immediate: the other formalizations have a weaker notion of negative introspection. This has an effect on the accuracy of the expansions in the general case. The form of negative introspection in cautious autoepistemic logic seems very promising: a tradeoff with accuracy is made if it is necessary for maintaining the acceptability of the cautious expansion. This behaviour has been demonstrated in the examples. It seems possible to give an interesting semantics for general logic programs (or even a larger class of logic programs) by means of the cautious expansions. We should stress the simplicity of the principle used for weakening the negative introspection, leading to a class of well-behaving autoepistemic expansions.

We have shown that cautious autoepistemic logic is decidable, i.e. there exists an algorithm that can carry out the membership test of cautious expansions in a finite time. The decidability analysis lays the basis for the full complexity analysis of cautious autoepistemic logic. The most significant source of computational complexity is evidently the enumeration-based introspection mechanism. The number of enumerations increases steeply when larger and larger sets of premises are considered. However, we have optimization strategies in mind to reduce the number of enumerations that have to be checked.

Cautious autoepistemic expansions provide a reasonable solution to the problems related to the existence of expansions. First of all, there is always the least $\mathcal{N}$-grounded cautious expansion. Secondly, it is the intersection of all $\mathcal{N}$-grounded cautious expansions. This implies that knowledge representation systems based on the least $\mathcal{N}$-grounded cautious expansion can at least provide the answer "I don't know" rather than failing to give an answer. This would be the case if stable expansions were the basis of analyzing sets of premises like $\{\neg Lp \rightarrow p\}$. We pointed out earlier that this kind of a robustness is lacked from, e.g. standard PROLOG interpreters. It is up to the programmer's expertise to avoid such logic programs.

4.1 Further Research Topics

We give here some interesting topics of further research work related to cautious autoepistemic logic. The exact complexity analysis of cautious autoepistemic reasoning is required. It is possible that different types of enumerations used for negative introspection give rise to cautious autoepistemic logics of different computational complexity. But whatever class of enumerations is selected, the iterative construction should be preserved.

Another direction is the study of the various translations of formalizations of NMR into cautious autoepistemic logic. We expect that the general principle we have introduced handles anomalous situations appearing in the other formalizations, given a suitable translation. The enumerations employed by the negative introspection provide a kind of a loop detection mechanism. For example inheritance networks with `is-a` and `isn't-a` links are prone to conflict situations that call for a loop analysis. Perhaps we can solve these conflict situations by means of cautious expansions. Another target is to provide an alternative formal semantics for logic programs with negation as failure to prove or even classical negation. The connection to the well-founded semantics should be clarified. The abductive systems used for diagnosis and truth maintenance systems seem interesting, too.

We have avoided syntactic restrictions in our approach. This kind of generality is advantageous: nested modalities can be used in the sets of premises. We expect that other monotonic logics can replace propositional logic as the underlying logic in cautious autoepistemic reasoning so that the resulting nonmonotonic logic is still decidable. Such a generalization should be rather straightforward as we have not made any restrictions specific to propositional logic. The essential requirement for the underlying consequence relation is decidability. Moreover, quantification in the modal context cannot be allowed and the belief operator L must be a fresh modal operator. These restrictions were imposed by Niemelä when he provided a similar construction for the original autoepistemic logic.

Acknowledgements

This report contains the results of my work in Digital Systems Laboratory at Helsinki University of Technology from March 1992 to January 1993. I am very grateful to professor Leo Ojala for arranging this opportunity to work in his laboratory. I am greatly indebted to senior assistant Ilkka Niemelä who has substantially contributed to the results presented in this work. He has also assisted in proofreading the manuscript. I also thank Laboratory Engineer Johan Lilius for arranging an excellent working environment and all my colleagues for the creative atmosphere.

References

[Balcázar *et al.*, 1988] J.L. Balcázar, I. Díaz, and J. Gabarró. *Structural Complexity I.* Springer-Verlag, Berlin, 1988.

[Chellas, 1980] B.F. Chellas. *Modal Logic: an Introduction.* Cambridge University Press, Cambridge, 1980.

[Clark, 1978] K.L. Clark. Negation as failure. In H. Gallaire and J. Minker, editors, *Logic and Data Bases*, pages 293–322. Plenum Press, New York, 1978.

[Davis and Weyuker, 1983] M. Davis and E. Weyuker. *Computability, Complexity, and Languages.* Fundamentals of Theoretical Computer Science. Academic Press, 1983.

[Dix, 1991] Jürgen Dix. Classifying semantics of logic programs. *Proceedings of the 1st International Workshop on Logic Programming and Nonmonotonic Reasoning*, pages 160–180, 1991.

[Doyle, 1979] J. Doyle. A truth maintenance system. *Artificial Intelligence*, 12:231–272, 1979.

[Elkan, 1990] C. Elkan. A rational reconstruction of nonmonotonic truth maintenance systems. *Artificial Intelligence*, 43:219–234, 1990.

[Epstein, 1990] R. Epstein. *The Semantic Foundations of Logic*, volume 35 of *Nijhoff International Philosophy Series.* Kluwer Academic Publishers, The Netherlands, 1990.

[Etherington, 1988] D.W. Etherington. *Reasoning with Incomplete Information.* Pitman, London, 1988.

[Fujiwara and S., 1989] Y. Fujiwara and Honiden S. Relating TMS to autoepistemic logic. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pages 1199–1205, Detroit, USA, August 1989. Morgan Kaufmann Publishers.

[Gelfond and Lifschitz, 1988] M. Gelfond and V. Lifschitz. The stable model semantics for logic programming. In *Proceedings of the 5th International Conference on Logic Programming*, pages 1070–1080, Seattle, USA, August 1988. The MIT Press.

[Gelfond and Lifschitz, 1990] M. Gelfond and V. Lifschitz. Logic programs with classical negation. In *Proceedings of the 7th International Conference on Logic Programming*, pages 579–597, Jerusalem, Israel, June 1990. The MIT Press.

[Gelfond and Przymusinska, 1990] M. Gelfond and H. Przymusinska. Formalization of inheritance reasoning in autoepistemic logic. *Fundamenta Informaticae*, 13:430–443, 1990.

[Gelfond, 1988] M. Gelfond. Autoepistemic logic and formalization of commonsense reasoning, preliminary report. In *Proceedings of the 2nd Workshop on Non-Monotonic Reasoning*, pages 176–186, Grassau, FRG, June 1988. Springer-Verlag.

[Ginsberg, 1987] M.L. Ginsberg, editor. *Readings in Nonmonotonic Reasoning*. Morgan Kaufmann Publishers, Los Altos, 1987.

[Gottlob, 1992a] G. Gottlob. Complexity results for nonmonotonic logics. *Journal of Logic and Computation*, 2(3):397–425, June 1992.

[Gottlob, 1992b] G. Gottlob. The power of beliefs or translating default logic into standard autoepistemic logic. Presented at the European Workshop on Theoretical Foundations of Knowledge Representation and Reasoning held in conjunction with ECAI'92, August 3–4, Vienna, Austria, 1992.

[Kakas and Mancarella, 1990] A.C. Kakas and P. Mancarella. Generalized stable models: a semantics for abduction. In *Proceedings of the 9th European Conference on Artificial Intelligence*, pages 385–391, Stockholm, Sweden, August 1990. Pitman Publishing.

[Kalinski, 1990] J. Kalinski. Autoepistemic expansions with incomplete belief introspection. In *Proceedings of the 1st International Workshop on Nonmonotonic and Inductive Logic*, pages 223–232, Karlsruhe, Germany, December 1990. Springer-Verlag. Lecture Notes in Artificial Intelligence 543.

[Kalinski, 1992] J. Kalinski. Weak autoepistemic reasoning. Presented at the European Workshop on Theoretical Foundations of Knowledge Representation and Reasoning held in conjunction with ECAI'92, August 3–4, Vienna, Austria, 1992.

[Kautz and Selman, 1989] H. Kautz and B. Selman. Hard problems for simple default logics. In *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, pages 189–197, Toronto, Canada, May 1989. Morgan Kaufmann Publishers.

[Konolige, 1988] K. Konolige. On the relation between default and autoepistemic logic. *Artificial Intelligence*, 35:343–382, 1988.

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

[Lifschitz, 1987] V. Lifschitz. Pointwise circumscription. In M.L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 179–193. Morgan Kaufmann Publishers, Los Altos, 1987.

[Lifschitz, 1988] V. Lifschitz. Benchmark problems for formal nonmontonic reasoning. *Proceedings of the 2nd Workshop on Non-Monotonic Reasoning*, pages 202–219, 1988.

[Lloyd, 1984] J.W. Lloyd. *Foundations of Logic Programming*. Springer-Verlag, Berlin, 1984.

[Marek and Subrahmanian, 1992] W. Marek and V.S. Subrahmanian. The relationship between stable, supported, default and autoepistemic semantics for general logic programs. *Theoretical Computer Science*, 103:365–386, 1992.

[Marek and Truszczyński, 1989] W. Marek and M. Truszczyński. Relating autoepistemic and default logics. In *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, pages 276–288, Toronto, Canada, May 1989.

[Marek and Truszczyński, 1990a] W. Marek and M. Truszczyński. Modal logic for default reasoning. In *Preprints of the 3rd International Workshop on Nonmonotonic Reasoning*, pages 247–258, South Lake Tahoe, California, May 1990.

[Marek and Truszczyński, 1990b] W. Marek and M. Truszczyński. Modal logic for default reasoning. *Annals of Mathematics and Artificial Intelligence*, 1:275–302, 1990.

[Marek *et al.*, 1991] W. Marek, G.F. Shvarts, and M. Truszczyński. Modal nonmonotonic logics: Ranges, characterization, computation. In *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 395–404, Cambridge, MA, USA, April 1991. Morgan Kaufmann Publishers.

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

[McCarthy, 1986] J. McCarthy. Applications of circumscription to formalizing commonsense knowledge. *Artificial Intelligence*, 28:89–116, 1986.

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

[McDermott, 1982] D. McDermott. Non-monotonic logic II. *Journal of the ACM*, 29:33–57, 1982.

[Minsky, 1981] M. Minsky. A framework for representing knowledge. In J. Haugeland, editor, *Mind Design*, pages 95–128. The MIT Press, Cambridge, 1981.

[Moore, 1985] R.C. Moore. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25:75–94, 1985.

[Moore, 1988] R.C. Moore. Autoepistemic logic. In P. Smets, E.H. Mamdani, and D. Dubois, editors, *Non-Standard Logics for Automated Reasoning*, pages 105–136. Academic Press, London, 1988.

[Niemelä and Rintanen, 1992] I. Niemelä and J. Rintanen. On the impact of stratification on the complexity of nonmonotonic reasoning. In *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning*, pages 627–638, Cambridge, MA, USA, October 1992. Morgan Kaufmann Publishers.

[Niemelä, 1990a] I. Niemelä. A new characterization of autoepistemic reasoning. In *Proceedings of the Finnish Artificial Intelligence Symposium*, pages 64–73, Oulu, Finland, June 1990. Finnish Artificial Intelligence Society.

[Niemelä, 1990b] I. Niemelä. Towards automatic autoepistemic reasoning. In *Proceedings of the European Workshop on Logics in Artificial Intelligence—JELIA'90*, pages 428–443, Amsterdam, The Netherlands, September 1990. Springer-Verlag.

[Niemelä, 1991] I. Niemelä. Constructive tightly grounded autoepistemic reasoning. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 399–404, Sydney, Australia, August 1991. Morgan Kaufmann Publishers.

[Niemelä, 1992] I. Niemelä. On the decidability and complexity of autoepistemic reasoning. *Fundamenta Informaticae*, 17(1,2):117–155, 1992.

[Przymusinska and Przymusinski, 1990] H. Przymusinska and T.C. Przymusinski. Semantic issues in deductive databases and logic programs. In R. Banerji, editor, *Formal Techniques in Artificial Intelligence*, pages 321–367. North-Holland, Amsterdam, 1990.

[Przymusinski, 1989] T.C. Przymusinski. Three-valued formalizations of nonmonotonic reasoning and logic programming. In *Proceedings of the 1st International Conference on Principles of Knowledge Representation and Reasoning*, pages 341–348, Toronto, Canada, May 1989. Morgan Kaufmann Publishers.

[Przymusinski, 1991a] T.C. Przymusinski. Autoepistemic logics of closed beliefs and logic programming. In *Proceedings of the 1st International Workshop on Logic Programming and Non-monotonic Reasoning*, pages 3–20, Washington, D.C., USA, July 1991. The MIT Press.

[Przymusinski, 1991b] T.C. Przymusinski. Three-valued nonmonotonic formalisms and semantics of logic programming. *Artificial Intelligence*, 49:309–343, 1991.

[Reinfrank *et al.*, 1989] M. Reinfrank, O. Dressler, and G. Brewka. On the relation between truth maintenance and autoepistemic logic. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence*, pages 1206–1212, Detroit, Michigan, August 1989. Morgan Kaufmann Publishers.

[Reiter, 1978] R. Reiter. On closed world data bases. In H. Gallaire and J. Minker, editors, *Logic and Data Bases*, pages 55–76. Plenum Press, New York, 1978.

[Reiter, 1980] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[Shvarts, 1990] G. Shvarts. Autoepistemic modal logics. In *Proceedings of the 3rd Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 97–109, Pacific Grove, USA, March 1990. Morgan Kaufmann Publishers.

[Smullyan, 1968] R.M. Smullyan. *First-Order Logic*. Springer-Verlag, Berlin, 1968.

[Stärk, 1990] R.F. Stärk. On the existence of fixpoints in Moore's autoepistemic logic and the non-monotonic logic of McDermott and Doyle. In *Proceedings of the 4th Workshop on Computer Science Logic*, pages 354–365, Heidelberg, Germany, October 1990. Springer-Verlag.

[Sterling and Shapiro, 1986] L. Sterling and E. Shapiro. *The Art of Prolog*. MIT Series in logic programming. The MIT Press, London, 1986.

[Stillman, 1990] J. Stillman. It's not my default: the complexity of membership problems in restricted propositional default logics. In *Proceedings of the 8th National Conference on Artificial Intelligence*, pages 571–578, Boston, Massachusetts, USA, July 1990. The MIT Press.

[Stillman, 1992] J. Stillman. The complexity of propositional default logics. In *Proceedings of the 10th National Conference on Artificial Intelligence*, pages 794–800, San Jose, California, USA, July 1992. The MIT Press.

[Touretzky *et al.*, 1987] D.S. Touretzky, J.F. Horty, and R.H. Thomason. A clash of intuitions: the current state of nonmonotonic multiple inheritance systems. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pages 476–482, Milan, Italy, August 1987. Morgan Kaufmann Publishers.

[Truszczyński, 1991] M. Truszczyński. Modal interpretations of default logic. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 393–398, Sydney, Australia, August 1991. Morgan Kaufmann Publishers.

[Van Gelder *et al.*, 1991] A. Van Gelder, K.A. Ross, and J.S. Schlipf. The well-founded semantics for general logic programs. *Journal of the ACM*, 38(3):620–650, July 1991.