A Knowledge Representation Framework Based on Autoepistemic Logic of Minimal Beliefs

Teodor C. Przymusinski*
Department of Computer Science
University of California
Riverside, CA 92521
(teodor@cs.ucr.edu)

Abstract

In recent years, various formalizations of non-monotonic reasoning and different semantics for normal and disjunctive logic programs have been proposed, including autoepistemic logic, circumscription, CWA, $GCWA$, $ECWA$, epistemic specifications, stable, well-founded, stationary and static semantics of normal and disjunctive logic programs.

In this paper we introduce a simple non-monotonic knowledge representation framework which isomorphically contains all of the above mentioned non-monotonic formalisms and semantics as special cases and yet is significantly more expressive than each one of these formalisms considered individually. The new formalism, called the *AutoEpistemic Logic of minimal Beliefs*, $AELB$, is obtained by augmenting Moore's autoepistemic logic, AEL, with an additional *minimal belief* operator, $\mathcal{B}$, which allows us to explicitly talk about minimally entailed formulae.

The existence of such a uniform framework not only results in a new powerful non-monotonic formalism but also allows us to compare and better understand mutual relationships existing between different non-monotonic formalisms and semantics and enables us to provide simpler and more natural definitions of some of them. It also naturally leads to new, even more expressive and flexible formalizations and semantics.

1 Introduction

Moore's autoepistemic logic AEL (Moore 1985) is obtained by augmenting classical propositional logic with a modal operator $\mathcal{L}$. The intended meaning of the modal atom $\mathcal{L}F$ is "F is provable" or "F is logically derivable" (in the stable autoepistemic expansion). Thus Moore's modal operator $\mathcal{L}$ can be viewed as a "knowledge operator" which allows us to reason about formulae *known* to be true in the expansion. However, usually, in addition to reasoning about facts which are known to be true, we also need to reason about those that are only *believed* to be true, where what is believed or not believed is determined by a specific non-monotonic formalism. In particular, we may want to express beliefs based on *minimal entailment* or *circumscription* and thus may need a modal "belief operator" $\mathcal{B}$ with the intended meaning of the modal atom $\mathcal{B}F$ given by "F true in all minimal models" or "F is minimally entailed" (in the expansion).

For example, consider a scenario in which: (1) you plan to rent a movie if you believe that you will not go to a baseball game (bg) and will not go to a football game (fg), but, (2) you do not plan to buy buy tickets to either of the games if you don't know for sure that you will go to see it. We could describe the initial scenario as follows:

$$\mathcal{B}(\neg goto(bg) \wedge \neg goto(fg)) \supset rent\_movie$$
$$\neg\mathcal{L}goto(bg) \wedge \neg\mathcal{L}goto(fg) \supset \neg buy\_tickets.$$

Assuming that this is all you know and that your beliefs are based on minimal entailment (circumscription), you should rent a movie because you believe that you will not go to see any games (i.e., $\neg goto(bg) \wedge \neg goto(fg)$ holds in all minimal models) and you should not buy tickets because you don't know that you will see any of the games (i.e., neither $goto(bg)$ nor $goto(fg)$ is provable).

Suppose now that you learn that you will either go a baseball game or to a football game (i.e., $goto(bg) \vee goto(fg)$). In the new scenario you should no longer plan to rent a movie (because $\neg goto(bg) \wedge \neg goto(fg)$ no longer holds in all minimal models) but you still do not intend to buy any tickets, because you don't know yet which game you are going to see (i.e., neither $goto(bg)$ nor $goto(fg)$ is provable).

However, when you eventually learn that you actually go to a baseball game (i.e., $goto(bg)$) you no longer believe in not buying tickets because you now know that you are going to see a specific game (i.e., $goto(bg)$ is provable).

Observe, that in the above example the roles played by the knowledge and belief operators are quite different and one cannot be substituted by the other. In particular, we cannot replace the premise $\mathcal{B}(\neg goto(bg) \wedge \neg goto(fg))$ in the first implication by $\mathcal{L}(\neg goto(bg) \wedge \neg goto(fg))$ because that would result in $rent\_movie$ not being true in first scenario. Similarly, we cannot replace it by $\neg\mathcal{L}goto(bg) \wedge \neg\mathcal{L}goto(fg)$ because that would result in $rent\_movie$ being true in the second scenario.

In order to be able to explicitly reason about minimal beliefs, we introduce a new non-monotonic formalism, called the *AutoEpistemic Logic of minimal Beliefs*,

*Partially supported by the National Science Foundation grant #IRI-9313061. Revised on March 24, 1994.

$AELB$, obtained by augmenting Moore's autoepistemic logic, AEL, with an additional *minimal belief* operator, $\mathcal{B}$. The resulting non-monotonic knowledge representation framework turns out to be rather simple and yet quite powerful. We prove that many of the recently introduced non-monotonic formalisms and semantics for normal and disjunctive logic programs are *isomorphically embeddable* into $AELB$. In particular this applies to autoepistemic logic (Moore 1985); circumscription (McCarthy 1980; Lifschitz 1985); CWA (Reiter 1978); $GCWA$ (Minker 1982); $ECWA$ (Gelfond, Przymusinska, & Przymusinski 1989); epistemic specifications (Gelfond 1992); stable, well-founded, stationary and static semantics of normal and disjunctive logic programs (Gelfond & Lifschitz 1988; Van Gelder, Ross, & Schlipf 1990; Przymusinski 1991c; Gelfond & Lifschitz 1990; Przymusinski 1994). At the same time the AutoEpistemic Logic of Minimal Beliefs, $AELB$, is significantly more expressive than each one of these formalisms considered individually.

The existence of such a unifying framework allows us to provide simpler and more natural definitions of several non-monotonic formalisms and semantics and it enables us to compare and better understand mutual relationships existing between them. It also naturally leads to new more expressive and flexible formalizations and semantics.

2 Language

The language of the *AutoEpistemic Logic of minimal Beliefs*, $AELB$, is a propositional modal language, $\mathcal{K}_{\mathcal{L},\mathcal{B}}$, with standard connectives ($\vee$, $\wedge$, $\supset$, $\neg$) and two modal operators $\mathcal{L}$ and $\mathcal{B}$, called *knowledge* and *belief* operators, respectively. The atomic formulae of the form $\mathcal{L}F$ (respectively, $\mathcal{B}F$), where F is an arbitrary formula of $\mathcal{K}_{\mathcal{L},\mathcal{B}}$, are called *knowledge atoms* (respectively, *belief atoms*). Knowledge and belief atoms are jointly referred to as *introspective atoms*.

The formulae of $\mathcal{K}_{\mathcal{L},\mathcal{B}}$ in which neither $\mathcal{L}$ nor $\mathcal{B}$ occurs are called *objective* and the set of all such formulae is denoted by $\mathcal{K}$. Similarly, the set of all formulae of $\mathcal{K}_{\mathcal{L},\mathcal{B}}$ in which only $\mathcal{L}$ (respectively, only $\mathcal{B}$) occurs is denoted by $\mathcal{K}_{\mathcal{L}}$ (respectively, $\mathcal{K}_{\mathcal{B}}$). Any theory T in the language $\mathcal{K}_{\mathcal{L},\mathcal{B}}$ will be called an *autoepistemic theory*.

The intended meaning of $\mathcal{L}F$ is *"F is known"*, or, more precisely, *"F can be logically inferred"*, i.e., $T \models F$. The intended meaning of $\mathcal{B}F$ is *"F is believed"*, or, more precisely, *"F can be non-monotonically inferred"*, i.e., $T \models_{nm} F$, where $\models_{nm}$ denotes a fixed non-monotonic inference relation. In general, different non-monotonic inference relations, $\models_{nm}$, can be used. In this paper we use the *minimal model entailment*, $T \models_{\min} F$, or, more precisely, *circumscription* (McCarthy 1980; Lifschitz 1985) which *minimizes* all objective atoms and *fixes* all the introspective (knowledge and belief) atoms[1], i.e.:

$$T \models_{\min} F \equiv CIRC(T; \mathcal{K}) \models F.$$

In other words, the precise intended meaning of belief atoms $\mathcal{B}F$ is *"F is minimally entailed"* by the theory, i.e., $T \models_{\min} F$. We assume the following two simple axiom schemata describing the arguably obvious properties of belief atoms:

Consistency Axiom: For any formula F:

$$\mathcal{B}F \supset \neg\mathcal{B}\neg F. \tag{1}$$

Conjunctive Belief Axiom: For any formulae F and G:

$$\mathcal{B}(F \wedge G) \quad \equiv \quad \mathcal{B}F \wedge \mathcal{B}G\ . \tag{2}$$

The first axiom states that if a formula F is believed then the formula $\neg F$ is *not* believed. The second axiom states that the conjunction $F \wedge G$ of formulae F and G is believed if and only if both F and G are believed. We assume that all theories implicitly *include* the axioms (1) and (2) and therefore when we talk about the set of logical consequences of a given theory T we actually have in mind the set $Con^*(T)$ of all logical consequences of the theory T *augmented* with the axioms (1) and (2):

$$Con^*(T) = Con(T \cup \{(1)\} \cup \{(2)\}).$$

Analogous axioms could be as well assumed about the knowledge atoms $\mathcal{L}F$ but they are in fact *automatically* satisfied in all static autoepistemic expansions which are defined in the next section. Additional axioms that can enhance the expressiveness of our logic are discussed later in Section 4.3.

3 Static Autoepistemic Expansions

Like Moore's autoepistemic logic, AEL, the autoepistemic logic of minimal beliefs, $AELB$, models the set of beliefs that an ideally rational and introspective agent should hold given a set of premises T. It does so by defining *static autoepistemic expansions* T^* of T, which constitute plausible sets of such rational beliefs.

Definition 3.1 (Static Autoepistemic Expansion) *An autoepistemic theory T^* is called a static autoepistemic expansion of an autoepistemic theory T if it satisfies the following fixed-point equation:*

$$T^* = Con^*(T \cup \{\mathcal{L}F : T^* \models F\} \cup \{\neg\mathcal{L}F : T^* \not\models F\} \cup \cup \{\mathcal{B}F : T^* \models_{\min} F\}). \ \square$$

The definition of static autoepistemic expansions is based on the idea of building an expansion T^* of a theory T by augmenting T with: *(i)* knowledge atoms $\mathcal{L}F$ that satisfy the condition that the formula F is logically implied by T^*, *(ii)* negations $\neg\mathcal{L}F$ of the remaining knowledge atoms, and, *(iii)* belief atoms $\mathcal{B}F$ which satisfy the condition that the formula F is minimally entailed by T^*. Consequently,

[1] The reason that we treat objective and introspective atoms differently is that the objective atoms A represent objective, *ground-level* information which, according to the principle of closed world assumption, is minimized in order to arrive at minimal beliefs $\mathcal{B}A$. On the other hand, the introspective atoms $\mathcal{L}F$ and $\mathcal{B}F$ intuitively describe *meta-level* information, namely, a plausible rational *scenario*, which is not subject to minimization.

the definition of static expansions *enforces* the intended meaning of introspective atoms described in the previous section. Note that negations $\neg \mathcal{B}F$ of (the remaining) belief atoms are not *explicitly* added to the expansion but some of them will be forced in by the Consistency Axiom (1).

Observe that the first part of the definition of static expansions is identical to the definition of stable autoepistemic expansions in Moore's autoepistemic logic, AEL. However, as we now show, the addition of belief atoms $\mathcal{B}F$ results in a *much more powerful non-monotonic logic* which contains, as special cases, several other well-known non-monotonic formalisms.

3.1 Circumscription

To begin with, one easily sees that propositional circumscription (and thus also CWA, $GCWA$ and $ECWA$ (Reiter 1978; Minker 1982; Gelfond, Przymusinska, & Przymusinski 1989)) can be properly embedded into $AELB$.

Proposition 3.1 (Embeddability of Circumscription) *Propositional circumscription, CWA, $GCWA$ and $ECWA$ are all properly embeddable into the autoepistemic logic of minimal beliefs, $AELB$. More precisely, if T is any objective theory, i.e., a theory which does not contain any introspective atoms $\mathcal{L}F$ and $\mathcal{B}F$, then T has a unique static expansion T^* and any objective formula F is logically implied by the circumscription $CIRC(T)$ of T if and only if T^* logically implies the belief atom $\mathcal{B}F$:*

$$CIRC(T) \models F \;\equiv\; T^* \models \mathcal{B}F. \;\square$$

3.2 Moore's Autoepistemic Logic

Since the first part of the definition of static autoepistemic expansions is identical to the definition of *stable autoepistemic expansions* in Moore's autoepistemic logic, AEL, it is easy to see that AEL is also properly *embeddable* into the autoepistemic logic of minimal beliefs, $AELB$.

Proposition 3.2 (Embeddability of Autoepistemic Logic) *Moore's autoepistemic logic, AEL, is properly embeddable into the autoepistemic logic of minimal beliefs, $AELB$. More precisely, for any autoepistemic theory T in the language $\mathcal{K}_\mathcal{L}$, i.e., for any theory that does not use belief atoms $\mathcal{B}F$, there is a one-to-one correspondence between stable autoepistemic expansions and static autoepistemic expansions of T.* □

In other words, the restriction, $AELB_\mathcal{L}$, of the autoepistemic logic of minimal beliefs, $AELB$, to the language $\mathcal{K}_\mathcal{L}$, i.e., its restriction to theories using only the knowledge operator $\mathcal{L}$, is *isomorphic* to Moore's autoepistemic logic, AEL. Thus, as its acronym suggests, $AELB$ constitutes an extension of Moore's AEL obtained by adding the belief operator $\mathcal{B}$.

3.3 Autoepistemic Logic of Purely Minimal Beliefs

While the restriction $AELB_\mathcal{L}$ of $AELB$ to the language $\mathcal{K}_\mathcal{L}$ is isomorphic to Moore's autoepistemic logic, the restriction $AELB_\mathcal{B}$ of $AELB$ to the language $\mathcal{K}_\mathcal{B}$, i.e., its restriction to theories using only the belief operator $\mathcal{B}$, constitutes an entirely new logic, which can be called the *autoepistemic logic of purely minimal beliefs*. It turns out that $AELB_\mathcal{B}$ has some quite natural and interesting properties. We first introduce the belief closure operator Ψ_T.

Definition 3.2 (Belief Closure Operator) *For any autoepistemic theory T define the belief closure operator Ψ_T by the formula:*

$$\Psi_T(S) = Con^*(T \cup \{\mathcal{B}F : S \models_{\min} F\}),$$

where S is an arbitrary autoepistemic theory. □

Thus $\Psi_T(S)$ augments the theory T with all those belief atoms $\mathcal{B}F$ for which F is minimally entailed by S. We first prove the restricted monotonicity of the belief closure operator Ψ_T.

Theorem 3.1 (Monotonicity of the Belief Operator) *Suppose that the theories T' and T'' are extensions of an autoepistemic theory T obtained by adding some belief atoms $\mathcal{B}F$ to T. If $T' \subseteq T''$ then $\Psi_T(T') \subseteq \Psi_T(T'')$.* □

From Theorem 3.1 and the well-known result of Tarski, ensuring the existence of least fixed points of monotonic operators, we easily conclude that for any autoepistemic theory T there is a unique theory S which is the least fixed point of the operator Ψ_T, i.e., satisfies the property $S = \Psi_T(S) = Con^*(T \cup \{\mathcal{B}F : S \models_{\min} F\})$. We now need the next result ensuring the existence of unique static autoepistemic expansions of theories which are fixed points of the operator Ψ_T.

Theorem 3.2 (Uniqueness of Expansions) *Suppose that T is an autoepistemic theory in the language $\mathcal{K}_\mathcal{B}$ and S is a fixed point of the operator Ψ_T, i.e., $\Psi_T(S) = S$. Then S has a unique static autoepistemic expansion $\widetilde{S}$ which is also a static autoepistemic expansion of T itself.* □

From Theorems 3.1 and 3.2 we deduce the following important result.

Theorem 3.3 (Least Static Autoepistemic Expansions) *Every autoepistemic theory T in the language $\mathcal{K}_\mathcal{B}$ has the least (in the sense of inclusion) static autoepistemic expansion $\overline{T}$.*

The expansion $\overline{T}$ can be constructed as follows. Let $T^0 = T$ and suppose that T^α has already been defined for any ordinal number $\alpha < \beta$. If $\beta = \alpha + 1$ is a successor ordinal then define:

$$T^{\alpha+1} = \Psi_T(T^\alpha) = Con^*(T \cup \{\mathcal{B}F : T^\alpha \models_{\min} F\}).$$

Else, if β is a limit ordinal, define $T^\beta = \bigcup_{\alpha<\beta} T^\alpha$. The sequence $\{T^\alpha\}$ is monotonically increasing and thus has a unique fixed point $T^\lambda = \Psi_T(T^\lambda)$, for some ordinal λ. Now define $\overline{T} = \widetilde{T^\lambda}$. □

The existence of least static autoepistemic expansions of theories in $AELB_\mathcal{B}$ sharply contrasts with the properties of stable autoepistemic expansions in AEL which typically do not have least elements. Observe that the *least* static autoepistemic expansion of T contains those and only those formulae which are true in *all* static autoepistemic expansions of T. The following theorem significantly extends Theorem 3.3 and provides a complete characterization of *all* static autoepistemic expansions of a theory T in the language $\mathcal{K}_\mathcal{B}$.

Theorem 3.4 (Characterization Theorem) *A theory T^* is a static autoepistemic expansion of a theory T in $\mathcal{K}_\mathcal{B}$ if and only if T^* is the least static autoepistemic expansion $\overline{T'}$ of a theory $T' = T \cup \{\mathcal{B}F_s : s \in S\}$ satisfying the condition that $T^* \models_{\min} F_s$, for every $s \in S$. In particular, the least static autoepistemic expansion $\overline{T}$ of T is obtained when the set $\{\mathcal{B}F_s : s \in S\}$ is empty.* □

4 Semantics of Logic Programs

We already know that Circumscription, Moore's Autoepistemic Logic and the Autoepistemic Logic of Purely Minimal Beliefs are all properly embeddable into the Autoepistemic Logic of Minimal Beliefs, $AELB$. We will now show that major semantics defined for normal and disjunctive *logic programs* are also embeddable into $AELB$. In the next section we will discuss two other non-monotonic formalisms embeddable into $AELB$.

4.1 Stable Semantics

Since Moore's autoepistemic logic, AEL, is isomorphic to the subset $AELB_\mathcal{L}$ of $AELB$, it follows from the results of Gelfond and Lifschitz (Gelfond & Lifschitz 1988) that stable semantics of logic programs can be obtained by means of a suitable translation of a logic program into an autoepistemic theory. Namely, for a logic program P consisting of clauses:

$$A \leftarrow B_1, ..., B_m, not\, C_1, ..., not\, C_n$$

define $T_{\neg\mathcal{L}}(P)$ to be its translation into the autoepistemic theory consisting of formulae:

$$B_1 \wedge ... \wedge B_m \wedge \neg\mathcal{L}C_1 \wedge ... \wedge \neg\mathcal{L}C_n \supset A.$$

The translation $T_{\neg\mathcal{L}}(P)$ is obtained therefore by replacing the *negation by default* $not\,C$ by $\neg\mathcal{L}C$ which has the intended meaning *"C is not known to be true"*.

Theorem 4.1 (Embeddability of Stable Semantics) *There is a one-to-one correspondence between stable models $\mathcal{M}$ of the program P and static autoepistemic expansions T^* of $T_{\neg\mathcal{L}}(P)$. Namely, for any objective atom A we have:*

$$A \in \mathcal{M} \quad \textit{iff} \quad \mathcal{L}A \in T^*$$

$$\neg A \in \mathcal{M} \quad \textit{iff} \quad \neg\mathcal{L}A \in T^*. \ \square$$

4.2 Stationary and Well-Founded Semantics

Similarly, it follows from the results obtained in (Przymusinski 1994) that the stationary (or partial stable) and the well-founded semantics of logic programs can be obtained by means of a suitable translation of a logic program into an autoepistemic theory. Namely, for a logic program P consisting of clauses:

$$A \leftarrow B_1, ..., B_m, not\, C_1, ..., not\, C_n$$

define $T_{\mathcal{B}\neg}(P)$ to be its translation into the autoepistemic theory consisting of formulae:

$$B_1 \wedge ... \wedge B_m \wedge \mathcal{B}\neg C_1 \wedge ... \wedge \mathcal{B}\neg C_n \supset A.$$

The translation $T_{\mathcal{B}\neg}(P)$ is obtained therefore by replacing the *negation by default* $not\,C$ by $\mathcal{B}\neg C$ which has the intended meaning *"C is believed to be false"* or *"$\neg C$ is minimally entailed"*.

Theorem 4.2 (Embeddability of Stationary Semantics) *There is a one-to-one correspondence between stationary (or partial stable) models $\mathcal{M}$ of the program P and static autoepistemic expansions T^* of $T_{\mathcal{B}\neg}(P)$. Namely, for any objective atom A we have:*

$$A \in \mathcal{M} \quad \textit{iff} \quad \mathcal{B}A \in T^*$$

$$\neg A \in \mathcal{M} \quad \textit{iff} \quad \mathcal{B}\neg A \in T^*.$$

Since the well-founded model $\mathcal{M}_0$ of the program P coincides with the least stationary model of P (Przymusinski 1991c), it corresponds to the least static autoepistemic expansion $\overline{T}$ of $T_{\mathcal{B}\neg}(P)$, whose existence is guaranteed by Theorem 3.3.

Moreover, (total) stable models $\mathcal{M}$ of P correspond to those static autoepistemic expansions T^* of $T_{\mathcal{B}\neg}(P)$ that satisfy the condition that for all objective atoms A, either $\mathcal{B}A \in T^*$ or $\mathcal{B}\neg A \in T^*$. □

Analogous result applies to the translation $T_{\neg\mathcal{B}}(P)$ defined by:

$$B_1 \wedge ... \wedge B_m \wedge \neg\mathcal{B}C_1 \wedge ... \wedge \neg\mathcal{B}C_n \supset A.$$

However, for disjunctive programs (discussed below) the two translations $T_{\mathcal{B}\neg}(P)$ and $T_{\neg\mathcal{B}}(P)$ lead to different results.

4.3 Semantics of Disjunctive Programs

As it was the case with normal logic programs, static expansions can be used to define the semantics of *disjunctive logic programs* (see (Lobo, Minker, & Rajasekar 1992) for an overview of disjunctive logic programming). In particular, we can extend the transformation $T_{\mathcal{B}\neg}(P)$ to any disjunctive logic program P consisting of clauses:

$$A_1 \vee ... \vee A_l \leftarrow B_1, ..., B_m, not\, C_1, ..., not\, C_n$$

by translating it into the autoepistemic theory consisting of formulae:

$$B_1 \wedge ... \wedge B_m \wedge \mathcal{B}\neg C_1 \wedge ... \wedge \mathcal{B}\neg C_n \supset A_1 \vee ... \vee A_l.$$

It turns out that this transformation immediately leads to the *static semantics* of disjunctive logic programs defined in (Przymusinski 1994):

Theorem 4.3 (Embeddability of Static Semantics) *There is a one-to-one correspondence between static expansions of the disjunctive logic program P, as defined in (Przymusinski 1994), and static autoepistemic expansions of its translation $T_{\mathcal{B}\neg}(P)$.* □

Although static semantics for disjunctive programs has a number of important advantages it is by far not the only semantics for disjunctive programs that can be derived by means of a suitable translation of a logic program into the autoepistemic logic of minimal beliefs, $AELB$. The expressive power of $AELB$ allows us to obtain other well-known semantics for disjunctive programs by simply using a different transformation and/or assuming additional axioms. To illustrate this claim let us consider the following three natural axioms:

Disjunctive Belief Axiom: For any formulae F and G:

$$\textbf{(DBA)} \quad \mathcal{B}(F \vee G) \equiv \mathcal{B}F \vee \mathcal{B}G.$$

Disjunctive Knowledge Axiom: For any formulae F and G:

$$\textbf{(DKA)} \quad \mathcal{L}(F \vee G) \equiv \mathcal{L}F \vee \mathcal{L}G.$$

Generalized Closed World Assumption: For any positive formula F:

$$\textbf{(GCWA)} \quad \mathcal{L}\mathcal{B}\neg F \supset \neg F.$$

The last axiom intuitively says that if we know that we believe in the falsity of a (positive) formula F then F is indeed false. It turns out that both the *disjunctive stationary semantics* introduced in (Przymusinski 1991b) and the *(partial or total) disjunctive stable semantics* introduced in (Przymusinski 1991c; Gelfond & Lifschitz 1990) can be expressed by means of these axioms.

Theorem 4.4 (Embeddability of Disjunctive Stationary Semantics) *There is a one-to-one correspondence between stationary expansions of a disjunctive program P and static autoepistemic expansions of its translation $T_{\mathcal{B}\neg}(P)$ augmented with the axioms (DBA) and $(GCWA)$.*

Theorem 4.5 (Embeddability of Disjunctive Stable Semantics) *There is a one-to-one correspondence between disjunctive partial stable models of a disjunctive program P and static autoepistemic expansions of its translation $T_{\mathcal{B}\neg}(P)$ augmented with the axioms (DBA), (DKA) and $(GCWA)$. Moreover, (total) disjunctive stable models of P correspond to those static autoepistemic expansions T^* of $T_{\mathcal{B}\neg}(P)$ that satisfy the condition that for all objective atoms A, either $\mathcal{B}A \in T^*$ or $\mathcal{B}\neg A \in T^*$.* □

4.4 Programs with Strong Negation

The negation operator *not* A used in logic programs does not represent the *classical negation*, but rather a non-monotonic negation by default. Gelfond and Lifschitz pointed out (Gelfond & Lifschitz 1990) that in logic programming, as well as in other areas of non-monotonic reasoning, it is often useful to use *both* the non-monotonic negation and a different negation, $-A$, which they called "classical negation" but which can perhaps more appropriately be called "strong negation" (Alferes & Pereira 1992). They also extended the stable model semantics to the class of *extended logic programs* with strong negation.

It is easy to add strong negation to the autoepistemic logic of minimal beliefs, $AELB$. All one needs to do is to augment the original objective language $\mathcal{K}$ with new *objective* propositional symbols "$-A$" with the intended meaning that "*$-A$ is the strong negation of A*" and assume the following *strong negation axiom* schema:

$$\textbf{(SNA)} \quad A \wedge -A \supset false, \text{ or, equivalently, } -A \supset \neg A.$$

Observe that, as opposed to classical negation $\neg$, the law of excluded middle $A \vee -A$ is not assumed. As pointed out by Bob Kowalski, the proposition A may describe the property of being *"good"* while proposition $-A$ describes the property of being *"bad"*. The strong negation axiom states that things cannot be both good and bad. We do not assume, however, that things must always be either good or bad.

Since this method of defining strong negation applies to *all* autoepistemic theories, it applies, in particular, to normal and disjunctive logic programs (see also (Alferes & Pereira 1992)). Moreover, the following theorem shows that the resulting general framework provides a strict *generalization* of the original approach proposed by Gelfond-Lifschitz.

Theorem 4.6 (Embeddability of Extended Stable Semantics) *There is a one-to-one correspondence between stable models $\mathcal{M}$ of an extended logic program P with strong negation, as defined in (Gelfond & Lifschitz 1990), and static autoepistemic expansions T^* of its translation $T_{\neg\mathcal{L}}(P)$ into autoepistemic theory under which a strong negation of an atom A is translated into $-A$.* □

5 Combining Knowledge and Belief

In most of the results presented so far the theories under consideration used only one of the introspective operators, either the belief operator $\mathcal{B}$ or the knowledge operator $\mathcal{L}$. However, the greatest expressive power of the autoepistemic logic of minimal beliefs, $AELB$, is achieved when both of these operators are used in combination. We have already seen examples of such combined use of the two operators in Theorems 4.4 and 4.5, both of which involved the axiom $(GCWA)$. In this section we first give an example and then we discuss two specific application areas, namely, logic programming and epistemic specifications,

Example 5.1 We first revisit the example informally discussed in the Introduction.

Scenario 1: You rent a movie if you believe that you do not go to a baseball game (bg) and do not go to a football game (fg). You do not buy tickets to a game if you don't know that you will go to see it.

$$\begin{array}{rcl} \mathcal{B}\neg goto(bg) \wedge \mathcal{B}\neg goto(fg) & \supset & rent\_movie \\ \neg\mathcal{L}goto(bg) \wedge \neg\mathcal{L}goto(fg) & \supset & dont\_buy\_tickets \end{array}$$

This theory has a unique static autoepistemic expansion in which you rent a movie, because you believe that you will not go to see any games (i.e., $\neg goto(bg) \wedge \neg goto(fg)$ holds in all minimal models) and you do not buy tickets because you don't know you that will go to see any of the games (i.e., neither $goto(bg)$ nor $goto(fg)$ are provable).

Scenario 2: Now, suppose that you learn that you either go to see a baseball game or go to see a football game, i.e., $goto(bg) \vee goto(fg)$. The new theory has a unique static autoepistemic expansion in which you believe you should *not* rent a movie[2] and you still do not buy any tickets, because you don't know yet which game you are going to see (i.e., neither $goto(bg)$ nor $goto(fg)$ are provable).

Scenario 3: Finally, suppose that you learn that you actually go to see a baseball game, i.e., $goto(bg)$. The new theory has a unique static autoepistemic expansion in which you still believe you should *not* rent any movies but you no longer believe in not buying game tickets because you know now that you are going to see a specific game.

Observe, that we cannot replace the premise $\mathcal{B}\neg goto(bg) \wedge \mathcal{B}\neg goto(fg)$ in the first implication by $\mathcal{L}\neg goto(bg) \wedge \mathcal{L}\neg goto(fg)$ because that would result in $rent\_movie$ not being true in Scenario 1. Similarly, we cannot replace it by $\neg\mathcal{L}goto(bg) \wedge \neg\mathcal{L}goto(fg)$ because that would result in $rent\_movie$ becoming true in Scenario 2. We also cannot replace the premise $\neg\mathcal{L}goto(bg) \wedge \neg\mathcal{L}goto(fg)$ in the second implication by $\neg\mathcal{B}goto(bg) \wedge \neg\mathcal{B}goto(fg)$ or by $\mathcal{B}\neg goto(bg) \wedge \mathcal{B}\neg goto(fg)$, because it would no longer imply that we should not buy tickets in Scenario 2. Thus the roles of the two operators are quite different and one cannot be substituted by the other. □

5.1 Combining Stable and Well-Founded Negation in Logic Programs

As we have seen in the previous section, both stable and well-founded negation in logic programs can be obtained by translating the non-monotonic negation *not* C into introspective literals $\neg\mathcal{L}C$ and $\mathcal{B}\neg C$, respectively. However, the existence of both types of introspective literals in $AELB$ allows us to *combine both types of negation* in one epistemic theory consisting of formulae of the form:

$$B_1 \wedge ... \wedge B_m \wedge \neg\mathcal{L}C_1 \wedge ... \wedge \neg\mathcal{L}C_k \wedge \mathcal{B}\neg C_{k+1} \wedge ... \wedge \mathcal{B}\neg C_n \supset$$

$$\supset A_1 \vee \ldots \vee A_l.$$

Such an epistemic theory may be viewed as representing a more *general disjunctive logic program* which permits the simultaneous use of both types of negation. In such logic programs, the first k negative premises represent *stable negation* and the remaining ones represent the *well-founded negation*. The ability to use both types of negation significantly increases the expressibility of logic programs. For instance, the previous Example 5.1 is a special case of such generalized programs.

[2]This follows from the fact that the expansion obviously implies $\mathcal{B}(goto(bg) \vee goto(fg))$ and thus, by the Consistency Axiom (1), it also contains $\neg\mathcal{B}(\neg goto(bg) \wedge \neg goto(fg)) \equiv \neg\mathcal{B}(\neg goto(bg)) \vee \neg\mathcal{B}(\neg goto(fg))$, which implies that $\neg rent\_movie$ holds in all minimal models.

5.2 Epistemic Specifications

Epistemic specifications were recently introduced in (Gelfond 1992) using a rather complex language of belief sets and world views which includes two operators, $\mathbf{K}F$ and $\mathbf{M}F$, called belief and possibility operators, respectively. As an illustration of the expressive power of the Autoepistemic Logic of Minimal Beliefs, $AELB$, we now demonstrate that epistemic specifications can be also *isomorphically embedded* as a proper subset of $AELB$, and thus, in particular, we show that epistemic specifications can be defined entirely in the language of classical propositional logic.

We show that Gelfond's belief operator $\mathbf{K}F$ can be defined as $\mathcal{LB}F$ and thus have the intended meening "F is known to be believed". On the other hand, the possibility operator $\mathbf{M}F$ is proved to be equivalent to $\neg\mathbf{K}\neg F$, or, equivalently, to $\neg\mathcal{LB}\neg F$. The translation provides therefore an example of a *nested use* of the belief and knowledge operators, $\mathcal{B}$ and $\mathcal{L}$ (see also the axiom $(GCWA)$ in Section 4.3).

Due to the space limitation, we assume familiarity with epistemic specifications. Let G be a database describing Gelfond's epistemic specification. Define $T(G)$ to be its translation into autoepistemic logic of minimal beliefs, $AELB$, obtained by:

(i) Replacing, for all *objective* atoms A, the classical negation symbol $\neg A$ by the strong negation symbol $-A$. We assume that the objective language $\mathcal{K}$ was first augmented with strong negation atoms $-A$ as described in Section 4.4.

(ii) Eliminating Gelfond's "possibility" operator $\mathbf{M}$ by replacing every expression of the form $\mathbf{M}F$ by the expression $\neg\mathbf{K}\neg F$, where $\mathbf{K}$ is Gelfond's "belief" operator.

(iii) Finally, eliminating Gelfond's "belief" operator $\mathbf{K}$ by replacing every expression of the form $\mathbf{K}F$ by the autoepistemic formula $\mathcal{LB}F$.

The substitution (i) is motivated by the fact that in his paper Gelfond uses the classical negation symbol $\neg A$ when in fact he refers to *strong negation* $-A$. The substitution allows us to reserve the standard negation symbol $\neg A$ for true classical negation. The substitution (ii) is motivated by the fact that Gelfond's "possibility" operator $\mathbf{M}F$ can now be shown to be *equivalent* to $\neg\mathbf{K}\neg F$, and, vice versa, $\mathbf{K}F$ can be shown to be equivalent to $\neg\mathbf{M}\neg F$. The last substitution (iii) leads to a complete translation into an autoepistemic theory. It replaces $\mathbf{K}F$ by the formula $\mathcal{LB}F$ with the intended meaning *"F is known to be believed"*. Equivalently, its intended meaning can be described by *"F is known to be true in all minimal models"*.

Now we can show that epistemic specifications are isomorphically embeddable into the autoepistemic logic of minimal beliefs, $AELB$. The limited size of this abstract does not allow us to provide complete details.

Theorem 5.1 (Embeddability of Epistemic Specifications) *Epistemic specifications are isomorphically embeddable into the autoepistemic logic of minimal beliefs, $AELB$.*

More precisely, there is a one-to-one correspondence between world views V of an epistemic specification G and static autoepistemic expansions T^* of its translation $T(G)$ into $AELB$. Moreover, there is a one-to-one correspondence between belief sets B of a world view V and minimal models $\mathcal{M}$ of the corresponding static expansion T^* of $T(G)$. □

Gelfond's paper contains several intersting examples of epistemic specifications which now can be easily translated into the simpler language of $AELB$.

6 Conclusion

We introduced an extension, $AELB$, of Moore's autoepistemic logic, AEL, and showed that it provides a powerful knowledge representation framework unifying several well-known non-monotonic formalisms and semantics for normal and disjunctive logic programs. It allows us to compare different formalisms, better understand mutual relationships existing between them and introduce simpler and more natural definitions of some of them.

The proposed formalism significantly differs from other formalisms based on the notion of minimal beliefs. In particular, it it is different from the circumscriptive autoepistemic logic introduced in (Przymusinski 1991a) and the logic of minimal beliefs and negation as failure proposed in (Lifschitz 1992). The formalism is also quite flexible by allowing various extensions and modifications, including the use of a different formalism defining the *meaning of beliefs* and introduction of *additional axioms*. For example, by using the *weak* minimal model entailment, instead of the standard minimal model entailment, in the definition of belief atoms $\mathcal{B}F$, one can ensure that disjunctions are treated *inclusively* rather than *exclusively*. Other forms of circumscriptions as well as other non-monotonic formalisms can be used to define the meaning of belief atoms. By using such modifications one may be able to tailor the formalism to fulfill the needs of different application domains.

References

Alferes, J. J., and Pereira, L. M. 1992. On logic program semantics with two kinds of negation. In Apt, K., ed., *International Joint Conference and Symposium on Logic Programming*, 574–588. MIT Press.

Gelfond, M., and Lifschitz, V. 1988. The stable model semantics for logic programming. In Kowalski, R., and Bowen, K., eds., *Proceedings of the Fifth Logic Programming Symposium*, 1070–1080. Cambridge, Mass.: Association for Logic Programming.

Gelfond, M., and Lifschitz, V. 1990. Logic programs with classical negation. In *Proceedings of the Seventh International Logic Programming Conference, Jerusalem, Israel*, 579–597. Cambridge, Mass.: Association for Logic Programming.

Gelfond, M.; Przymusinska, H.; and Przymusinski, T. 1989. On the relationship between circumscription and negation as failure. *Journal of Artificial Intelligence* 38:75–94.

Gelfond, M. 1992. Logic programming and reasoning with incomplete information. Technical report, University of Texas at El Paso.

Lifschitz, V. 1985. Computing circumscription. In *Proceedings IJCAI-85*, 121–127. Los Altos, CA: American Association for Artificial Intelligence.

Lifschitz, V. 1992. Minimal belief and negation as failure. Research report, University of Texas at Austin.

Lobo, J.; Minker, J.; and Rajasekar, A. 1992. *Foundations of Disjunctive Logic Programming*. Cambridge, Massachusetts: MIT Press.

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

Minker, J. 1982. On indefinite data bases and the closed world assumption. In *Proc. 6-th Conference on Automated Deduction*, 292–308. New York: Springer Verlag.

Moore, R. 1985. Semantic considerations on non-monotonic logic. *Journal of Artificial Intelligence* 25:75–94.

Przymusinski, T. C. 1991a. Autoepistemic logics of closed beliefs and logic programming. In Nerode, A.; Marek, W.; and Subrahmanian, V., eds., *Proceedings of the First International Workshop on Logic Programming and Non-monotonic Reasoning, Washington, D.C., July 1991*, 3–20. Cambridge, Mass.: MIT Press.

Przymusinski, T. C. 1991b. Semantics of disjunctive logic programs and deductive databases. In Delobel, C.; Kifer, M.; and Masunaga, Y., eds., *Proceedings of the Second International Conference on Deductive and Object-Oriented Databases DOOD'91*, 85–107. Munich, Germany: Springer Verlag.

Przymusinski, T. C. 1991c. Stable semantics for disjunctive programs. *New Generation Computing Journal* 9:401–424. (Extended abstract appeared in: Extended stable semantics for normal and disjunctive logic programs. *Proceedings of the 7-th International Logic Programming Conference, Jerusalem*, pages 459–477, 1990. MIT Press.).

Przymusinski, T. C. 1994. Static semantics for normal and disjunctive logic programs. *Annals of Mathematics and Artificial Intelligence*. (in print).

Reiter, R. 1978. On closed-world data bases. In Gallaire, H., and Minker, J., eds., *Logic and Data Bases*. New York: Plenum Press. 55–76.

Van Gelder, A.; Ross, K. A.; and Schlipf, J. S. 1990. The well-founded semantics for general logic programs. *Journal of the ACM*. (to appear). Preliminary abstract appeared in Seventh ACM Symposium on Principles of Database Systems, March 1988, pp. 221–230.