INFSYS
RESEARCH
REPORT

INSTITUT FÜR INFORMATIONSSYSTEME

ABTEILUNG WISSENSBASIERTE SYSTEME

EMBEDDING NON-GROUND LOGIC PROGRAMS INTO AUTOEPISTEMIC LOGIC FOR KNOWLEDGE BASE COMBINATION

JOS DE BRUIJN THOMAS EITER
AXEL POLLERES HANS TOMPITS

INFSYS RESEARCH REPORT 1843-08-12
NOVEMBER 2008

Institut für Informationssysteme
Abtg. Wissensbasierte Systeme
Technische Universität Wien
Favoritenstraße 9-11
A-1040 Wien, Austria
Tel: +43-1-58801-18405
Fax: +43-1-58801-18493
sek@kr.tuwien.ac.at
www.kr.tuwien.ac.at

TECHNISCHE UNIVERSITÄT WIEN

INFSYS Research Report

INFSYS Research Report 1843-08-12, November 2008

Embedding Non-Ground Logic Programs into Autoepistemic Logic for Knowledge Base Combination

Jos de Bruijn[1] Thomas Eiter[2] Axel Polleres[3] Hans Tompits[2]

Abstract. In the context of the Semantic Web, several approaches to the combination of ontologies, given in terms of theories of classical first-order logic, and rule bases have been proposed. They either cast rules into classical logic or limit the interaction between rules and ontologies. Autoepistemic logic (AEL) is an attractive formalism which allows to overcome these limitations, by serving as a uniform host language to embed ontologies and nonmonotonic logic programs into it. For the latter, so far only the propositional setting has been considered. In this paper, we present three embeddings of normal and three embeddings of disjunctive non-ground logic programs under the stable model semantics into first-order AEL. While the embeddings all correspond with respect to objective ground atoms, differences arise when considering non-atomic formulas and combinations with first-order theories. We compare the embeddings with respect to stable expansions and autoepistemic consequences, considering the embeddings by themselves, as well as combinations with classical theories. Our results reveal differences and correspondences of the embeddings and provide useful guidance in the choice of a particular embedding for knowledge combination.

Keywords: First-order autoepistemic logic, knowledge combination, ontologies, logic programming, rules, stable model semantics

---

[1]Faculty of Computer Science, Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100 Bolzano, Italy; email: `debruijn@inf.unibz.it`.

[2]Institut für Informationssysteme, Technische Universität Wien, Favoritenstraße 9-11, A-1040 Vienna, Austria; email: `(eiter,tompits)@kr.tuwien.ac.at`.

[3]Digital Enterprise Research Institute, National University of Ireland, Galway, IDA Business Park, Lower Dangan Galway, Ireland; email: `axel.polleres@deri.org`.

Acknowledgements: This work was partially funded by the European Commission projects Knowledge Web (IST 507482), DIP (IST 507483), and REWERSE (IST 506779), by the Austrian Science Fund (FWF) projects P17212-N04 and P21840, by the CICyT of Spain project TIC-2003-9001, and by the Science Foundation Ireland under Grant No. SFI/02/CE1/I131.

An extended abstract of this article appeared in the *Proceedings of the 20th International Joint Conference on Artificial Intelligence* (*IJCAI 2007*), AAAI Press, pp. 304–309.

Contents

1 Introduction

In the context of the ongoing discussion around combinations of rules and ontologies for the Semantic Web, there have been several proposals for integrating classical knowledge bases (ontologies) and rule bases (logic programs). Generally speaking, all these approaches try to define a reasonable semantics for a combined knowledge base consisting of a classical component and a rules component.

Two trends are currently observable. On the one hand, approaches such as SWRL [Horrocks *et al.*, 2005]—and indeed the earlier CARIN approach [Levy and Rousset, 1998], which is one of the pioneering approaches in combining rules and description logic (DL) ontologies—extend the (DL) ontology with Horn formulas in a classical framework. This approach is straightforward, but does not allow for nonmonotonic rules. On the other hand, existing approaches which do allow nonmonotonic rules either (a) distinguish between "classical" and "rules" predicates and limit the domain of interpretation (e.g., as done by Rosati [2006]) or (b) restrict the interaction to ground entailment (e.g., Eiter *et al.* [2008] follow this method). The main distinction between these approaches is the type of interaction between the classical knowledge base on the one hand and the rule base on the other (cf. de Bruijn *et al.* [2006] for an examination of this issue).

As for combination, a classical theory[1] and a logic program should be viewed as complementary descriptions of the same domain. Therefore, a syntactic separation between predicates defined in these two components should not be enforced. Furthermore, it is desirable to neither restrict the interaction between the classical and the rules components nor impose any syntactic or semantic restrictions on the individual components. That is, the classical component may be an arbitrary theory Φ of some first-order language with equality, and the rules component may be an arbitrary non-ground normal or disjunctive logic program P, interpreted using, e.g., the common stable model semantics [Gelfond and Lifschitz, 1988; Gelfond and Lifschitz, 1991]. The goal is a combined theory, $\iota(\Phi, P)$, in a uniform logical formalism. Naturally, this theory should amount to Φ if P is empty, and to P if Φ is empty. Therefore, such a combination must provide faithful embeddings $\sigma(\Phi)$ and $\tau(P)$ of Φ and P, respectively, such that $\sigma(\Phi) = \iota(\Phi, \emptyset)$ and $\tau(P) = \iota(\emptyset, P)$. In turn, knowledge combination may be carried out on top of $\sigma(\cdot)$ and $\tau(\cdot)$, where in the simplest case one may choose $\iota(\Phi, P) = \sigma(\Phi) \cup \tau(P)$.

This raises the following questions: (a) which uniform formalism is suitable and (b) which embeddings are suitable and, furthermore, how do potentially suitable embeddings relate to each other and how do they behave under knowledge combination?

Concerning the first question, Motik and Rosati [2007] use a variant of Lifschitz's bimodal nonmonotonic *logic of minimal knowledge and negation-as-failure* (MKNF) [Lifschitz, 1991]. While the proposed embeddings of the first-order (FO) theory and the logic program are both faithful in the sense described above, the particular combination proposed by Motik and Rosati is only one among many possible methods and MKNF is only one possible underlying formalism for such combinations (we discuss these issues in more detail in Section 7). Indeed, de Bruijn *et al.* [2007] use quantified equilibrium logic (QEL) [Pearce and Valverde, 2005] as a host formalism. Unlike Motik and Rosati, de Bruijn et al. do not propose a new semantics for combinations, but rather show that QEL can capture the semantics of combinations by Rosati [2006] and can be used, for example, to define notions of equivalence of combinations.

Autoepistemic logic (AEL) [Moore, 1985], which extends classical logic with a single nonmono-

[1] Most description logic ontologies can be viewed as theories of classical first-order logic; see Borgida [1996] and Sattler *et al.* [2003].

tonic modal belief operator, being essentially the nonmonotonic variant of the modal logic **kd45** [Shvarts, 1990; Marek and Truszczyński, 1993a], is an attractive candidate for serving as a uniform host formalism for combinations. Compared to other well-known nonmonotonic formalisms, like Reiter's default logic [Reiter, 1980], FO-AEL offers a uniform language in which (nonmonotonic) rules themselves can be expressed at the object level. This conforms with the idea of treating an ontology and a logic program together as a unified theory. Furthermore, in FO-AEL we can decide, depending on the context, whether (the negation of) a particular atomic formula should be interpreted nonmonotonically simply by including a modal operator. This enables us to use the same predicate in both a monotonic and a nonmonotonic context. This is in contrast to circumscription [McCarthy, 1986], in which one has to decide, for the entire theory, which predicates are to be minimized.

Embedding a classical theory in AEL is trivial, and several embeddings of logic programs in AEL have been described [Gelfond and Lifschitz, 1988; Marek and Truszczynski, 1993b; Lifschitz and Schwarz, 1993; Chen, 1993; Przymusinski, 1991a]. However, all these embeddings have been developed for the propositional case only, whereas we need to deal with non-ground theories and programs. This requires us to consider *first-order autoepistemic logic* (FO-AEL) [Konolige, 1991; Kaminski and Rey, 2002; Levesque and Lakemeyer, 2000], and non-ground versions of these embeddings. We consider the semantics for FO-AEL as defined by Konolige [1991], because it faithfully extends first-order logic with equality (other variants are discussed in Section 7).

Our contribution in this paper is twofold:

(1) We define several embeddings of non-ground logic programs into FO-AEL, taking into account subtle issues of quantification in FO-AEL. In more detail, we present three embeddings τ_{HP}, τ_{EB}, and τ_{EH} for normal logic programs which extend respective embeddings for the propositional case [Gelfond, 1987; Gelfond and Lifschitz, 1988; Marek and Truszczynski, 1993b; Chen, 1993; Lifschitz, 1994], and three embeddings $\tau_{HP}^{\vee}$, $\tau_{EB}^{\vee}$, and $\tau_{EH}^{\vee}$ for disjunctive logic programs, where $\tau_{HP}^{\vee}$ and $\tau_{EH}^{\vee}$ extend embeddings considered in the ground case [Przymusinski, 1991a; Marek and Truszczynski, 1993b]. We show that all these embeddings are faithful in the sense that the stable models of the logic program P and the sets of objective ground atoms in the stable expansions of the embeddings $\tau_{\chi}(P)$ are in one-to-one correspondence (Theorem 5.3). However, the embeddings behave differently on formulas beyond ground atoms, in some cases already for simple ground formulas. This, in turn, may impact the behavior of the embeddings when used in combinations of logic programs and classical theories. This raises the question under which conditions the embeddings differ and under which conditions they correspond. Of particular interest for knowledge combination is how these embeddings behave relatively to each other in combinations with classical theories.

(2) To answer these questions, we conduct two comparative studies of the behavior of the various embeddings. We consider three classes of programs: ground, safe, and arbitrary logic programs under the stable model semantics.

(a) We first determine correspondences between the stable expansions of different embeddings τ_{χ} beyond ground atomic formulas (Propositions 5.4-5.6), and present inclusion relations between the sets of consequences of the embeddings (Theorem 5.11). These results already allow to draw a few conclusions on the behavior of embeddings in combinations.

(b) We then determine correspondences between stable expansions for combinations of logic programs with classical theories. Here, besides the shape of the logic program we also take the shape of the classical theory, as well as the types of formulas of interest for correspondence, into account. To this end, we consider different fragments of classical logic that are important for knowledge representation, including Horn, universal, and generalized Horn theories, where the latter are of particular interest for ontologies, since it includes the DL Horn-$\mathcal{SHIQ}$ [Hustadt *et al.*, 2005]. Our main result for embeddings in combinations (Theorem 6.2) gives a complete picture of the correspondences, which reveals that they behave differently in general, and shows the restrictions on the program or theory that give rise to correspondence.

The results of these studies not only deepen the understanding of the individual embeddings, but also have practical implications with respect to the use of the embeddings. They tell us in which situations one embedding may be used instead of another. For example, if the stable expansions of two embeddings of arbitrary logic programs, under combination with propositional theories, correspond with respect to objective ground formulas, the two embeddings may be used interchangeably as long as one is only concerned with the objective ground consequences of such combinations.

The embeddings of logic programs we study in this paper can be seen as building blocks for actual combinations of classical theories and logic programs. The most straightforward combination of a classical theory Φ and a logic program P is $\iota(\Phi, P) = \sigma(\Phi) \cup \tau_\chi(P)$, where σ is the identity mapping and τ_χ is one of the embeddings studied in this paper. One could also imagine adding axioms to, or changing axioms in Φ; similarly, rules could be changed in, or added to P before translating them. If Φ' and P' are the thus obtained classical theory and logic program, the results in this paper are still applicable to the combination $\iota'(\Phi, P) = \Phi' \cup \tau_\chi(P')$. In fact, whenever the combination is of the form $\Phi' \cup \tau_\chi(P')$, no matter what the original Φ and P look like, the correspondences and differences between the embeddings established in this paper hold.

Arguably, no one embedding can a priori be considered to be superior to the others. Our results give useful insight into the properties of the different embeddings, both on their own right and for knowledge combination. They provide helpful guidance for the selection of an embedding in a particular scenario.

The remainder of the paper is structured as follows. We review the definitions of first-order logic and logic programs in Section 2. We proceed to describe first-order autoepistemic logic (FO-AEL) and present a novel characterization of stable expansions for certain kinds of theories in Section 3. The embeddings of normal and disjunctive logic programs and our results about faithfulness of the embeddings are described in Section 4. We investigate the relationships between the embeddings themselves, and under combination with first-order theories, in Sections 5 and 6. We discuss the implications of our results in Section 7, and related work in Section 8. We conclude and outline future work in Section 9.
Proofs of the results in Sections 5 and 6 can be found in the appendix.

2 Preliminaries

Let us briefly recapitulate some basic elements of first-order logic and logic programs as well as some relevant notation.

2.1 First-Order Logic

We consider first-order logic with equality. A language $\mathcal{L}$ is defined over a signature $\Sigma = (\mathcal{F}, \mathcal{P})$, where $\mathcal{F}$ and $\mathcal{P}$ are countable sets of *function* and *predicate symbols*, respectively. Function symbols with arity 0 are also called *constants*. Furthermore, $\mathcal{V}$ is a countably infinite set of *variables*. Terms and atomic formulas (atoms) are constructed as usual. Ground terms are also called *names*; $\mathcal{N}_\Sigma$ denotes the set of names of a given signature Σ. Complex formulas are constructed as usual using the primitive symbols $\neg$, $\wedge$, $\exists$, '(', and ')'. As usual, $\phi \vee \psi$ is short for $\neg(\neg\phi \wedge \neg\psi)$, $\phi \supset \psi$ is short for $\neg\phi \vee \psi$, and $\forall x.\phi(x)$ is short for $\neg\exists x.\neg\phi(x)$. We sometimes write $t_1 \neq t_2$, where t_1 and t_2 are terms, as an abbreviation for $\neg(t_1 = t_2)$. The universal closure of a formula ϕ is denoted by $(\forall)\ \phi$. $\mathcal{L}_g$ is the restriction of $\mathcal{L}$ to ground formulas; $\mathcal{L}_{ga}$ is the restriction of $\mathcal{L}_g$ to atomic formulas. An *FO theory* $\Phi \subseteq \mathcal{L}$ is a set of closed formulas, i.e., every variable is bound by a quantifier.

An *interpretation* of a language $\mathcal{L}$ is a tuple $w = \langle U, \cdot^I \rangle$, where U is a nonempty set, called the *domain*, and $\cdot^I$ is a mapping which assigns to every n-ary function symbol $f \in \mathcal{F}$ a function $f^I : U^n \to U$ and to every n-ary predicate symbol $p \in \mathcal{P}$ a relation $p^I \subseteq U^n$. A *variable assignment* B for w is a mapping that assigns to every variable $x \in \mathcal{V}$ an element $x^B \in U$. A variable assignment B' is an *x-variant* of B if $y^B = y^{B'}$ for every variable $y \in \mathcal{V}$ such that $y \neq x$. The interpretation of a term t, denoted $t^{w,B}$, is defined as usual; if t is ground, we sometimes write t^w.

We call an individual k *named* if there is some name $t \in \mathcal{N}$ such that $t^w = k$, and *unnamed* otherwise. Interpretations are *named* if all individuals are named. The *unique names assumption* applies to an interpretation if all names are interpreted distinctly, and the *standard names assumption* applies if, in addition, the interpretation is named.[2]

A *variable substitution* β is a partial function that assigns variables in $\mathcal{V}$ names from $\mathcal{N}$; we also write $x/\beta(x)$ for $(x, \beta(x))$. As usual, β is *total* if its domain is $\mathcal{V}$. Given a variable assignment B for an interpretation w, we define the set of named variables in B as $V_\mathcal{N}^{w,B} = \{x \mid x^B \text{ is named}\}$. A substitution β is *associated with* B if its domain is $V_\mathcal{N}^{w,B}$ and $x^B = \beta(x)^w$, for each $x \in V_\mathcal{N}^{w,B}$. The *application* of a variable substitution β to some term, formula, or theory χ, denoted by $\chi\beta$, is defined as syntactical replacement, as usual. Clearly, if the unique names assumption applies, each variable assignment has a unique associated substitution; if the standard names assumption applies, each associated substitution is total.

Example 2.1. *Consider a language $\mathcal{L}$ with constants $\mathcal{F} = \{a,\ b, c\}$, and an interpretation $w = \langle U, \cdot^I \rangle$ with $U = \{k, l, m\}$ such that $a^w = k$, $b^w = l$, and $c^w = l$, and the variable assignment B: $x^B = k$, $y^B = l$, and $z^B = m$. B has two associated variable substitutions, $\beta_1 = \{x/a, y/b\}$ and $\beta_2 = \{x/a, y/c\}$, which are both not total.*

2.2 Logic Programs

A *disjunctive logic program* P consists of rules of the form

$$h_1 \mid \ldots \mid h_l \ \leftarrow\ b_1,\ \ldots,\ b_m,\ not\ c_1,\ \ldots,\ not\ c_n \tag{1}$$

where $h_1, \ldots, h_l, b_1, \ldots, b_m, c_1, \ldots, c_n$ are equality-free atoms, with $m, n \geq 0$ and $l \geq 1$. $H(r) = \{h_1, \ldots, h_l\}$ is the set of *head atoms* of r, $B^+(r) = \{b_1, \ldots, b_m\}$ is the set of *positive body atoms* of

[2] We note here that the term "standard names assumption" is used with various slightly different meanings in the literature; see Section 7 for further discussion.

r, and $B^-(r) = \{c_1, \ldots, c_n\}$ is the set of *negated body atoms* of r. If $l = 1$, then r is *normal.* If $B^-(r) = \emptyset$, then r is *positive.* If every variable in r occurs in $B^+(r)$, then r is *safe.* If every rule $r \in P$ is normal (resp., positive, safe), then P is normal (resp., positive, safe).

Each program P has a signature Σ_P, which contains the function and predicate symbols that occur in P. We assume that Σ_P contains some 0-ary function symbol if it has predicate symbols of arity greater than 0. With $\mathcal{L}_P$ we denote the first-order language over Σ_P. As usual, *Herbrand interpretations* M of P are subsets of the set of ground atoms of $\mathcal{L}_P$.

The *grounding* of a logic program P, denoted $gr(P)$, is the union of all possible ground instantiations of P, obtained by replacing each variable in a rule r with a name in $\mathcal{N}_{\Sigma_P}$, for each rule $r \in P$.

Let P be a positive program. A Herbrand interpretation M of P is a *Herbrand model* of P if for every rule $r \in gr(P)$, $B^+(r) \subseteq M$ implies $H(r) \cap M \neq \emptyset$ and for every $t \in \mathcal{N}_{\Sigma_P}$, $t = t \in M$. A Herbrand model M is *minimal* iff for every model M' such that $M' \subseteq M$, $M' = M$.

Following Gelfond and Lifschitz [1991], the *reduct* of a logic program P with respect to an interpretation M, denoted P^M, is obtained from $gr(P)$ by deleting (i) each rule r with $B^-(r) \cap M \neq \emptyset$ and (ii) *not* c from the body of every remaining rule r with $c \in B^-(r)$. If M is a minimal Herbrand model of P^M, then M is a *stable model* of P.

Example 2.2. *Consider the program*

$$P = \{p(a);\ p(b);\ q(x) \mid r(x) \leftarrow p(x), not\ s(x)\}$$

and the interpretation $M_1 = \{p(a), p(b), q(a), r(b)\}$.[3] *The reduct*

$$P^{M_1} = \{p(a); p(b); q(a) \mid r(a) \leftarrow p(a); q(b) \mid r(b) \leftarrow p(b)\}$$

has M_1 *as a minimal model, thus* M_1 *is a stable model of* P*. The other stable models of* P *are* $M_2 = \{p(a),\ p(b),\ q(a),\ q(b)\}$, $M_3 = \{p(a),\ p(b),\ q(b),\ r(a)\}$*, and* $M_4 = \{p(a),\ p(b),\ r(a),\ r(b)\}$.

3 First-order Autoepistemic Logic

We adopt first-order autoepistemic logic (FO-AEL) under the any- and all-name semantics of Konolige [1991]. These semantics allow quantification over arbitrary domains and generalize classical first-order logic with equality, thereby allowing a trivial embedding of first-order theories (with equality). Other approaches like those by Kaminski and Rey [2002] or Levesque and Lakemeyer [2000] require interpretations to follow the unique or standard names assumptions and therefore do not allow such direct embeddings.

An *FO-AEL language* $\mathcal{L}_\mathsf{L}$ is defined relative to a first-order language $\mathcal{L}$ by allowing the unary modal operator L in the construction of formulas—$\mathsf{L}\phi$ is usually read as "ϕ is known" or "ϕ is believed". As usual, closed formulas, i.e., formulas without free variable occurrences, are called *sentences*; formulas of the form $\mathsf{L}\phi$, where ϕ is a formula, are *modal atoms*; and L-free formulas are *objective*. *Standard autoepistemic logic* is variable-free FO-AEL.

To distinguish between semantic notions defined for the any- resp. all-name semantics, we use the symbols E ("*E*xistence of name") and A ("for *A*ll names").

[3] For brevity, we leave out equality atoms in the example.

An *autoepistemic interpretation* is a pair $\langle w, \Gamma\rangle$, where $w = \langle U, \cdot^I\rangle$ is a first-order interpretation and $\Gamma \subseteq \mathcal{L}_{\mathsf{L}}$ is a set of sentences, called a *belief set*. Satisfaction of a formula $\mathsf{L}\phi$ in an interpretation $\langle w, \Gamma\rangle$ with respect to a variable assignment B under the *any-name semantics*, denoted $(w, B) \models^{\mathrm{E}}_{\Gamma} \mathsf{L}\phi$, is defined as

$(w, B) \models^{\mathrm{E}}_{\Gamma} \mathsf{L}\phi$ iff, for *some* variable substitution β associated with B, $\phi\beta$ is closed and $\phi\beta \in \Gamma$.

Satisfaction of arbitrary formulas is then as follows, where $\phi, \psi \in \mathcal{L}_{\mathsf{L}}$:

- $(w, B) \models^{\mathrm{E}}_{\Gamma} p(t_1, \ldots, t_n)$ iff $(t_1^{w,B}, \ldots, t_n^{w,B}) \in p^I$;
- $(w, B) \models^{\mathrm{E}}_{\Gamma} t_1 = t_2$ iff $t_1^{w,B} = t_2^{w,B}$;
- $(w, B) \models^{\mathrm{E}}_{\Gamma} \neg\phi$ iff $(w, B) \not\models^{\mathrm{E}}_{\Gamma} \phi$;
- $(w, B) \models^{\mathrm{E}}_{\Gamma} \phi \wedge \psi$ iff $(w, B) \models^{\mathrm{E}}_{\Gamma} \phi$ and $(w, B) \models^{\mathrm{E}}_{\Gamma} \psi$;
- $(w, B) \models^{\mathrm{E}}_{\Gamma} \exists x.\phi$ iff for some x-variant B' of B, $(w, B') \models^{\mathrm{E}}_{\Gamma} \phi$.

An interpretation $\langle w, \Gamma\rangle$ is a *model* of ϕ, denoted $w \models^{\mathrm{E}}_{\Gamma} \phi$, if $(w, B) \models^{\mathrm{E}}_{\Gamma} \phi$ for every variable assignment B for w. This extends to sets of formulas in the usual way. A set of formulas $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$ *entails* a formula $\phi \in \mathcal{L}_{\mathsf{L}}$ with respect to a belief set Γ, denoted $\Phi \models^{\mathrm{E}}_{\Gamma} \phi$, if for every interpretation w such that $w \models^{\mathrm{E}}_{\Gamma} \Phi$, $w \models^{\mathrm{E}}_{\Gamma} \phi$.

The notions of satisfaction and entailment under the *all-name semantics*, for which we use the symbol $\models^{\mathrm{A}}_{\Gamma}$, are analogous, with the only difference that satisfaction of modal atoms is defined as

$(w, B) \models^{\mathrm{A}}_{\Gamma} \mathsf{L}\phi$ iff, for *all* variable substitutions β associated with B, $\phi\beta$ is closed and $\phi\beta \in \Gamma$.

Note that the any- and all-name semantics always coincide for objective formulas and, if the unique (or standard) names assumption applies, also for arbitrary formulas in $\mathcal{L}_{\mathsf{L}}$; this was also observed by Kaminski and Rey [2002]. In such situations, i.e., where both semantics coincide, we sometimes use $\models_{\Gamma}$ rather than $\models^{\mathrm{E}}_{\Gamma}$ or $\models^{\mathrm{A}}_{\Gamma}$. Furthermore, when talking about entailment $\Phi \models_{\Gamma} \phi$ under the standard names assumption, we mean entailment considering only interpretations for which the standard names assumption holds. That is, $\Phi \models_{\Gamma} \phi$ under the standard names assumption if for every interpretation w such that the standard names assumption applies in w and $w \models_{\Gamma} \Phi$, $w \models_{\Gamma} \phi$.

Example 3.1. *Consider the formula $\phi = \forall x(p(x) \supset \mathsf{L}p(x))$ and some interpretation $\langle w, \Gamma\rangle$. Then, $w \models^{\mathrm{E}}_{\Gamma} \phi$ iff, for every variable assignment B, $(w, B) \models^{\mathrm{E}}_{\Gamma} p(x) \supset \mathsf{L}p(x)$, which in turn holds iff $(w, B) \not\models^{\mathrm{E}}_{\Gamma} p(x)$ or $(w, B) \models^{\mathrm{E}}_{\Gamma} \mathsf{L}p(x)$. Now, $(w, B) \models^{\mathrm{E}}_{\Gamma} \mathsf{L}p(x)$, with $x^B = k$, iff, for some $t \in \mathcal{N}_{\Sigma}$, $t^w = k$, and $p(t) \in \Gamma$. Thus, ϕ is false (unsatisfied) in any interpretation where p^I contains unnamed individuals. Analogous for the all-name semantics.*

The following example illustrates the difference between the any- and all-name semantics.

Example 3.2. *Consider a language with constant symbols a, b and unary predicate symbol p, and an interpretation $\langle w, \Gamma\rangle$ with $w = \langle\{k\}, \cdot^I\rangle$ and $\Gamma = \{p(a)\}$. Then, $w \models^{\mathrm{E}}_{\Gamma} \exists x.\mathsf{L}p(x)$, while $w \not\models^{\mathrm{A}}_{\Gamma} \exists x.\mathsf{L}p(x)$, since $b^w = a^w = k$ but $p(b) \notin \Gamma$.*

A stable expansion is a set of beliefs of an ideally introspective agent (i.e., an agent with perfect reasoning capabilities and with knowledge about its own beliefs), given some theory $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$. Formally, a belief set $T \subseteq \mathcal{L}_{\mathsf{L}}$ is a *stable*$^{\mathrm{E}}$ *expansion* of a theory $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$ iff $T = \{\phi \mid \Phi \models^{\mathrm{E}}_{T} \phi\}$. Similarity, T is a *stable*$^{\mathrm{A}}$ *expansion* of Φ iff $T = \{\phi \mid \Phi \models^{\mathrm{A}}_{T} \phi\}$.

Recall that $\mathcal{L}_g$ and $\mathcal{L}_{ga}$ denote the restrictions of $\mathcal{L}$ to ground and ground atomic formulas, respectively. Given a set of sentences $\Gamma \subseteq \mathcal{L}_{\mathsf{L}}$, Γ_o, Γ_{og}, and Γ_{oga} denote the restrictions of Γ to objective, objective ground, and objective ground atomic formulas, respectively, i.e., $\Gamma_o = \Gamma \cap \mathcal{L}$, $\Gamma_{og} = \Gamma \cap \mathcal{L}_g$, and $\Gamma_{oga} = \Gamma \cap \mathcal{L}_{ga}$.

Every stable expansion T of Φ is a *stable set* [Stalnaker, 1993], which means that it satisfies the following conditions: (a) T is closed under first-order entailment, (b) if $\phi \in T$ then $\mathsf{L}\phi \in T$, and (c) if $\phi \notin T$ then $\neg\mathsf{L}\phi \in T$. Furthermore, if T is consistent, the converse statements of (b) and (c) hold.

Konolige [1991] shows that a stable expansion T of a theory $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$ is determined by its objective subset T_o, also called the *kernel* of T. He further obtained the following result:

Proposition 3.3 ([Konolige, 1991])**.** *Let $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$ be a theory without nested modal operators, $\Gamma \subseteq \mathcal{L}$ a set of objective formulas, and $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$. Then, $\Gamma = \{\phi \in \mathcal{L} \mid \Phi \models^{\mathrm{X}}_{\Gamma} \phi\}$ iff $\Gamma = T_o$, for some stable$^{\mathrm{X}}$ expansion T of Φ.*

We slightly adapt this result as follows:

Proposition 3.4. *Let $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$ be a theory with only objective atomic formulas in the scope of occurrences of L, $\Gamma \subseteq \mathcal{L}$ a set of objective formulas, and $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$. Then, $\Gamma = \{\phi \in \mathcal{L} \mid \Phi \models^{\mathrm{X}}_{\Gamma_{oga}} \phi\}$ iff $\Gamma = T_o$, for some stable$^{\mathrm{X}}$ expansion T of Φ.*

Proof. Since modal atoms in Φ contain only objective atomic formulas, we obtain $\Phi \models^{\mathrm{X}}_{\Gamma_o} \phi$ iff $\Phi \models^{\mathrm{X}}_{\Gamma_{oga}} \phi$, because, by the definition of satisfaction of modal formulas, non-ground and non-atomic formulas in Γ_o do not affect satisfaction of formulas in Φ. Thus,$\{\phi \in \mathcal{L} \mid \Phi \models^{\mathrm{X}}_{\Gamma_{oga}} \phi\} = \{\phi \in \mathcal{L} \mid \Phi \models^{\mathrm{X}}_{\Gamma_o} \phi\}$ follows.

Since there is no nesting of modal operators in Φ, we combine this result with Proposition 3.3 to obtain $\Gamma_o = \{\phi \in \mathcal{L} \mid \Phi \models^{\mathrm{X}}_{\Gamma_o} \phi\} = \{\phi \in \mathcal{L} \mid \Phi \models^{\mathrm{X}}_{\Gamma_{oga}} \phi\}$ iff $\Gamma_o = T \cap \mathcal{L}$ is the kernel of a stable$^{\mathrm{X}}$ expansion T of Φ. □

We note here that, unlike in standard autoepistemic logic, in FO-AEL two different stable expansions may have the same objective subsets, both under the any- and all-name semantics. Consider, for example, the theories $\Phi = \{\forall x.p(x)\}$ and $\Phi' = \{\forall x.\mathsf{L}p(x)\}$ and their respective stable expansions T and T'. We have that $T_o = T'_o$ is the closure under first-order entailment of $\{\forall x.p(x)\}$, but we also have that $\forall x.\mathsf{L}p(x) \in T'$ but $\forall x.\mathsf{L}p(x) \notin T$, because $\forall x.\mathsf{L}p(x)$ is not satisfied in any interpretation that has unnamed individuals.

4 Embedding Non-Ground Logic Programs

We define an embedding τ as a function that takes a logic program P as its argument and returns a set of sentences in the FO-AEL language obtained from Σ_P.

Janhunen [1999] studied translations between nonmonotonic formalisms. He formulated a number of desiderata for such translation functions, namely faithfulness, polynomiality, and modularity (FPM). We adapt these notions to our case of embedding logic programs into FO-AEL.

An embedding τ is *faithful* if, for any logic program P, there is a one-to-one correspondence between stable models of P and consistent stable expansions of $\tau(P)$, with respect to ground atomic formulas.

An embedding τ is *polynomial* if, for any logic program P, $\tau(P)$ can be computed in time polynomial in the size of P.

An embedding τ is *modular* if, for any two logic programs P_1 and P_2, $\tau(P_1 \cup P_2) = \tau(P_1) \cup \tau(P_2)$. Furthermore, we call τ *signature-modular* if, for any two logic programs P_1 and P_2 with the same signature Σ, $\tau(P_1 \cup P_2) = \tau(P_1) \cup \tau(P_2)$.

Since the unique names assumption does not hold in FO-AEL in general, it is necessary to axiomatize default uniqueness of names (as introduced by Konolige [1991]) to assure faithfulness of several of the embeddings. Given a signature Σ, by UNA_Σ we denote the set of axioms

$$\text{UNA} \qquad \neg \mathsf{L}(t_1 = t_2) \supset t_1 \neq t_2, \quad \text{for all distinct } t_1, t_2 \in \mathcal{N}_\Sigma.$$

Default uniqueness, in contrast to rigid uniqueness (i.e., UNA axioms of the form $t_1 \neq t_2$), allows first-order theories that are later combined with the embedding to "override" such inequalities, rather than introducing inconsistency. For example, the theory $\Phi = \{\neg \mathsf{L}(a = b) \supset a \neq b\}$ has a single expansion that includes $a \neq b$; the single expansion of $\Phi \cup \{a = b\}$ is consistent and includes $a = b$.

Observe that the UNA axioms depend on the signature. In addition, the union of the UNA axioms of two signatures is not necessarily the same as the set of UNA axioms of the union of these two signatures: given two signatures Σ_1 and Σ_2 such that $\mathcal{F}_1 \neq \mathcal{F}_2$, $\mathrm{UNA}_{\Sigma_1} \cup \mathrm{UNA}_{\Sigma_2} \neq \mathrm{UNA}_{\Sigma_1 \cup \Sigma_2}$, i.e., the UNA axioms corresponding to different signatures cannot be combined in a modular fashion. This means that embeddings that include such UNA signatures are not modular, but may be signature-modular.

We first present the embeddings of normal programs and then proceed with the embeddings of disjunctive programs.

4.1 Embedding Normal Logic Programs

We consider three embeddings of non-ground logic programs into FO-AEL, called τ_{HP}, τ_{EB}, and τ_{EH}. "*HP*" stands for "*H*orn for *P*ositive rules" (positive rules are translated to objective Horn clauses); "*EB*" stands for "*E*pistemic rule *B*odies" (the body of a rule can only become true if it is *known* to be true); and "*EH*" stands for "*E*pistemic rule *H*eads" (if the body of a rule is true, the head is *known* to be true).

The *HP* embedding is an extension of the one which originally led Gelfond and Lifschitz to the definition of the stable model semantics [Gelfond, 1987; Gelfond and Lifschitz, 1988]. The *EB* and *EH* embeddings are extensions of embeddings by Marek and Truszczynski [1993b]. The *EH* embedding was independently described by Lifschitz and Schwarz [1993] and by Chen [1993]. The original motivation for the *EB* and *EH* embeddings was the possibility to directly embed programs with strong negation and disjunction. Furthermore, Marek and Truszczynski arrived at their embeddings through embeddings of logic programs in *reflexive autoepistemic logic* [Schwarz, 1992], which is equivalent to McDermott's nonmonotonic modal **sw5** [McDermott, 1982], and the subsequent embedding of reflexive autoepistemic logic into standard AEL. Lifschitz and Schwarz arrived at the *EH* embedding through an embedding of logic programs in Lifschitz's nonmonotonic logic of *minimal belief and negation-as-failure* (MBNF) [Lifschitz, 1994] and the subsequent embedding of MBNF into standard AEL. Finally, Chen also arrived at the *EH* embedding via MBNF, but he subsequently embedded MBNF in Levesque's *logic of only knowing* [Levesque, 1990], a subset of which corresponds with standard AEL.

Definition 4.1. *Let r be a normal rule of the form* (1)*. Then,*

$$\tau_{HP}(r) = (\forall)\ \bigwedge_i b_i \wedge \bigwedge_j \neg \mathsf{L} c_j \supset h_1$$
$$\tau_{EB}(r) = (\forall)\ \bigwedge_i (b_i \wedge \mathsf{L} b_i) \wedge \bigwedge_j \neg \mathsf{L} c_j \supset h_1$$
$$\tau_{EH}(r) = (\forall)\ \bigwedge_i (b_i \wedge \mathsf{L} b_i) \wedge \bigwedge_j \neg \mathsf{L} c_j \supset h_1 \wedge \mathsf{L} h_1$$

Furthermore, given a normal logic program P, we define:

$$\tau_\chi(P) = \{\tau_\chi(r) \mid r \in P\} \cup \mathrm{UNA}_{\Sigma_P}, \quad \chi \in \{HP, EB, EH\}.$$

For all three embeddings, we assume $\Sigma_{\tau_\chi(P)} = \Sigma_P$ (here and henceforth "χ" ranges over *HP*, *EB*, and *EH*). Furthermore, by τ_χ^- we denote the embedding τ_χ *without* the UNA axioms: given a normal logic program P, $\tau_\chi^-(P) = \tau_\chi(P) - \mathrm{UNA}_{\Sigma_P}$. The embeddings τ_χ^- are modular and polynomial. The embeddings τ_χ are signature-modular and polynomial, provided $\mathcal{N}_{\Sigma_P}$ is polynomial in the size of P (e.g., if there are no function symbols with arity greater than 0). In the examples of embeddings in the remainder of the paper we do not write the UNA axioms explicitly.

A notable difference between the embedding τ_{HP}, on the one hand, and the embeddings τ_{EB} and τ_{EH}, on the other, is that, given a logic program P, the stable expansions of $\tau_{HP}(P)$ include the "contrapositives" of the rules in P (viewed classically and where $\neg \mathsf{L} a$ is *not* a), which is not true for $\tau_{EB}(P)$ and $\tau_{EH}(P)$ in general.

Example 4.2. *Consider $P = \{p \leftarrow q, not\ r\}$. The stable expansion of $\tau_{HP}(P) = \{q \wedge \neg \mathsf{L} r \supset p\}$ includes $\neg p \supset \neg q \vee \mathsf{L} r$; the expansion of $\tau_{EB}(P) = \{q \wedge \mathsf{L} q \wedge \neg \mathsf{L} r \supset p\}$ includes $\neg p \supset \neg \mathsf{L} q \vee \neg q \vee \mathsf{L} r$, but not $\neg p \supset \neg q \vee \mathsf{L} r$.*

For the case of standard AEL and ground logic programs, the following faithfulness result straightforwardly extends results by Gelfond and Lifschitz [1988] and Marek and Truszczynski [1993b].

Proposition 4.3. *A Herbrand interpretation M of a ground normal logic program P is a stable model of P iff there exists a consistent stable expansion T of $\tau_\chi^-(P)$ in standard AEL such that $M = T \cap \mathcal{L}_{ga}$.*

We now consider the case of non-ground programs. The following example illustrates the embeddings.

Example 4.4. *Consider $P = \{q(a);\ p(x);\ r(x) \leftarrow not\ s(x),\ p(x)\}$, which has single stable model $M = \{q(a), p(a), r(a)\}$. Likewise, each of the embeddings $\tau_\chi(P)$ has a single consistent stable expansion T^χ:*

$$T^{HP} = \{q(a), p(a), \mathsf{L}p(a), \ldots, \forall x(p(x)), \neg\mathsf{L}\forall x(\mathsf{L}p(x)), \forall x(\neg\mathsf{L}s(x) \supset r(x)), \ldots\}$$
$$T^{EB} = \{q(a), p(a), \mathsf{L}p(a), \ldots, \forall x(p(x)), \neg\mathsf{L}\forall x(\mathsf{L}p(x)), \neg\mathsf{L}(\forall x(\neg\mathsf{L}s(x) \supset r(x))), \ldots\}$$
$$T^{EH} = \{q(a), p(a), \mathsf{L}p(a), \ldots, \forall x(p(x)), \forall x(\mathsf{L}p(x)), \forall x(\neg\mathsf{L}s(x) \supset r(x)), \ldots\}$$

The stable expansions in Example 4.4 agree on objective ground atoms, but not on arbitrary formulas. We now extend Proposition 4.3 to the non-ground case. To this end, we use the following two lemma.

Lemma 4.5. *Let P be a normal logic program, let $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$, let T be a stable$^{\mathrm{X}}$ expansion of $\tau_\chi(P)$, and let α be an objective ground atom. Then, $\tau_\chi(P) \models^{\mathrm{X}}_{T_{oga}} \alpha$ iff $\tau_\chi(P) \models_{T_{oga}} \alpha$ under the standard names assumption. Moreover, $\tau^-_{HP}(P) \models^{\mathrm{A}}_{T_{oga}} \alpha$ iff $\tau^-_{HP}(P) \models_{T_{oga}} \alpha$ under the standard names assumption.*

Proof. We start with the first statement.
($\Rightarrow$) This is obvious, as interpretations under the standard names assumption are just special interpretations.

($\Leftarrow$) We start with the case of the any-names assumption. Assume, on the contrary, that $\tau_\chi(P) \models_{T_{oga}} \alpha$ under the standard names assumption, but $\tau_\chi(P) \not\models^{\mathrm{E}}_{T_{oga}} \alpha$. This means that there is some interpretation $w = \langle U, \cdot^I \rangle$ such that $w \models^{\mathrm{E}}_{T_{oga}} \tau_\chi(P)$, but $w \not\models^{\mathrm{E}}_{T_{oga}} \alpha$.

By the fact that the only occurrences of the equality symbol in $\tau_\chi(P)$ are in the UNA axioms, the only atoms in T_{oga} involving equality are of the form $t = t$, for $t \in \mathcal{N}_{\Sigma_P}$. Consider two distinct names $t_1, t_2 \in \mathcal{N}_{\Sigma_P}$ and the UNA axiom $\neg \mathsf{L} t_1 = t_2 \supset t_1 \neq t_2 \in \mathrm{UNA}_{\Sigma_P}$. Since $\langle w, T_{oga} \rangle$ is a model of the axiom and $t_1 = t_2 \notin T_{oga}$, $w \models^{\mathrm{E}}_{T_{oga}} t_1 \neq t_2$. Consequently, it must be the case that $\cdot^I$ maps every name to a distinct individual in U.

We assume that the mapping $\cdot^I$ extends to ground terms in the natural way, i.e., $f(t_1, \ldots, t_m)^I = f^I(t_1^I, \ldots, t_m^I)$. We construct the interpretation $w' = \langle U', \cdot^{I'} \rangle$ as follows: $U' = \mathcal{N}$, $t^{I'} = t$, for $t \in \mathcal{N}$, and $\langle t_1, \ldots, t_n \rangle \in p^{I'}$ if $\langle t_1^I, \ldots, t_n^I \rangle \in p^I$ for n-ary predicate symbol p and every $\langle t_1, \ldots, t_n \rangle \in \mathcal{N}^n$. Clearly, the standard names assumption holds for w', and w and w' agree on objective ground atoms: $w \models \alpha$ iff $w' \models \alpha$ for any $\alpha \in \mathcal{L}_{ga}$. We now show that $w' \models_{T_{oga}} \tau_\chi(P)$.

Clearly, $\langle w', T_{oga} \rangle$ satisfies the UNA axioms since the standard names assumption holds for w' and since T_{oga} contains only the trivial equalities. We first consider the embedding τ_{EH} and some

$$(\forall) \quad \bigwedge_{1 \le i \le m} (b_i \wedge \mathsf{L} b_i) \wedge \bigwedge_{1 \le j \le n} (\neg \mathsf{L} b_j) \supset h_1 \wedge \mathsf{L} h_1 \in \tau_{EH}(P)$$

Since $w \models_{T_{oga}} \tau_{EH}(P)$,

$$(w, B) \models_{T_{oga}} \bigwedge_{1 \le i \le m} (b_i \wedge \mathsf{L} b_i) \wedge \bigwedge_{1 \le j \le n} (\neg \mathsf{L} b_j) \supset h_1 \wedge \mathsf{L} h_1$$

for every variable assignment B of w.

Now, consider a variable assignment B' of w' and the corresponding variable assignment B of w, which we define as follows: $x^B = k$ iff there is a $t \in \mathcal{N}_{\Sigma_P}$ such that $x^{B'} = t$ and $t^I = k$. Observe that B assigns every variable to a named individual. Consider a variable substitution β which is associated with B; since all names are interpreted as distinct individuals (by the UNA axioms), (†) β is unique. Moreover, by construction of B, β is also the only substitution associated with B' .

By construction of w' and since β is the unique substitution associated with B (and B') we have, for every objective atom α such that B is defined for all variables in α, that $(w, B) \models^{\mathrm{E}}_{T_{oga}} \alpha$ iff $(w', B') \models^{\mathrm{E}}_{T_{oga}} \alpha$ and $(w, B) \models^{\mathrm{E}}_{T_{oga}} \mathsf{L}\alpha$ iff $(w', B') \models^{\mathrm{E}}_{T_{oga}} \mathsf{L}\alpha$. Consequently, if $(w, B) \models^{\mathrm{E}}_{T_{oga}} h \wedge \mathsf{L} h$, then $(w', B') \models_{T_{oga}} h \wedge \mathsf{L} h$, and if $(w, B) \not\models^{\mathrm{E}}_{T_{oga}} \bigwedge b_i \wedge \mathsf{L} b_i$, then $(w', B') \not\models_{T_{oga}} \bigwedge b_i \wedge \mathsf{L} b_i$. Now,

$(w, B) \not\models^{\mathrm{E}}_{T_{oga}} \bigwedge \neg \mathsf{L} b_i$ implies $b_1\beta \in T_{oga}, \ldots,$ or $b_n\beta \in T_{oga}$. Hence, $(w', B') \not\models_{T_{oga}} \neg \mathsf{L} b_{m+1} \wedge \cdots \wedge \neg \mathsf{L} b_n$. So, $(w', B') \models_{T_{oga}} b_1 \wedge \mathsf{L} b_1 \wedge \cdots \wedge b_m \wedge \mathsf{L} b_m \wedge \neg \mathsf{L} b_{m+1} \wedge \cdots \wedge \neg \mathsf{L} b_n \supset h$.

Thus, we obtain $w' \models_{T_{oga}} \tau_{EH}(P)$. Since w and w' agree on objective ground atoms, $w' \not\models_{T_{oga}} \alpha$, and thus $\tau_{EH}(P) \not\models_{T_{oga}} \alpha$ under the standard names assumption. This contradicts the initial assumption. Therefore, $\tau_{EH}(P) \models^{\mathrm{E}}_{T_{oga}} \alpha$.

The argument for the embeddings τ_{EB} and τ_{HP} is analogous: simply leave out the positive occurrences of modal atoms in the consequents, respectively consequents and antecedents, in the argument above.

Likewise, the argument for the case of the all-name semantics is analogous. Observe that in the argument about variable assignments (†), β is the only variable substitution associated with B; hence, the any- and all-name semantics coincide, and the subsequent arguments immediately apply also for the all-name semantics.

For the second statement, consider the above argument without the part about the UNA axioms and the following simple adaptation: if $(w, B) \not\models^{\mathrm{A}}_{T_{oga}} \bigwedge \neg \mathsf{L} b_i$, then for all associated variable substitutions β, $b_1\beta \in T, \ldots,$ or $b_n\beta \in T$. One of these variable substitutions is the one associated with B'; the remainder of the argument remains the same. It follows that $\tau^{-}_{HP}(P) \models^{\mathrm{A}}_{T_{oga}} \alpha$ iff $\tau^{-}_{HP}(P) \models_{T_{oga}} \alpha$.

The latter fails for the embeddings τ^{-}_{EB} and τ^{-}_{EH}, as there may be several variable substitutions associated with the assignment B in the "$\Leftarrow$" direction above, while there is a single substitution associated with B' (see also Example 4.8). □

Lemma 4.6. *Let P be a normal logic program and $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$. There exists a stable$^{\mathrm{X}}$ expansion T of $\tau_\chi(P)$ iff there exists a stable$^{\mathrm{X}}$ expansion T' of $\tau_\chi(gr(P))$ such that $T'_{oga} = T_{oga}$. The same result holds for τ^{-}_{HP} and stable$^{\mathrm{A}}$ expansions.*

Proof. We prove the first statement, first for the special case that the standard names assumption applies, and then use Lemma 4.5 to extend it to cases where the standard names assumption does not apply.

Consider a belief set $\Gamma \subseteq \mathcal{L}_{\mathsf{L}}$ and an interpretation w for which the standard names assumption holds. We claim that (*) $w \models^{\mathrm{X}}_{\Gamma} \tau_\chi(gr(P))$ iff $w \models^{\mathrm{X}}_{\Gamma} \tau_\chi(P)$. By the standard names assumption, we have that $w \models^{\mathrm{X}}_{\Gamma} \tau_\chi(P)$ iff for every $\phi \in \tau_\chi(P)$, $w \models^{\mathrm{X}}_{\Gamma} \phi$. In turn, this holds iff for every variable assignment B, $(w, B) \models^{\mathrm{X}}_{\Gamma} \phi$, which in turn holds iff for the variable substitution β associated with B (which is unique and total, by the standard names assumption), $w \models^{\mathrm{X}}_{\Gamma} \phi\beta$. By definition, $\tau_\chi(gr(P))$ contains all (and only) the formulas of the form $\phi\beta$ where $\phi \in \tau_\chi(P)$ and β is a variable substitution associated with some variable assignment B for w; the claim (*) follows immediately from this.

($\Rightarrow$) Let T be a stable expansion of $\tau_\chi(P)$. By Lemma 4.5 and the above we have:

$$\{\phi \in \mathcal{L}_{ga} \,|\, \tau_\chi(P) \models^{\mathrm{X}}_{T_{oga}} \phi\} = \{\phi \in \mathcal{L}_{ga} \,|\, \tau_\chi(gr(P)) \models^{\mathrm{X}}_{T_{oga}} \phi\}$$

Hence by Proposition 3.4,

$$T'_o = \{\phi \in \mathcal{L} \mid \tau_\chi(gr(P)) \models^{\mathrm{X}}_{T_{oga}} \phi\}$$

is the kernel of a stable expansion T' of $\tau_\chi(gr(P))$ and $T' \cap \mathcal{L}_{ga} = T_{ga}$.

The converse is analogous. For the second statement of the lemma, the same proof using Lemma 4.5 works. □

Theorem 4.7. *Let P be a normal logic program and $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$. A Herbrand interpretation M is a stable model of P iff there exists a consistent stable$^{\mathrm{X}}$ expansion T of $\tau_{\chi}(P)$ such that $M = T_{oga}$. The same result holds for τ^{-}_{HP} and stable$^{\mathrm{A}}$ expansions.*

Proof. By Lemma 4.6 we can reduce embeddability of non-ground logic programs to embeddability of ground logic programs.

Consider an embedding $\tau_{\chi}(gr(P))$ and a stable expansion T. Clearly, there is no interaction between the UNA axioms and the axioms resulting from rules in P. Therefore, $\tau^{-}_{\chi}(gr(P))$ has a stable expansion T' such that $T'_{oga} = T_{oga}$, and vice versa. The theorem then follows immediately from Proposition 4.3. □

Note that this result does not extend to the embeddings τ^{-}_{χ} under the any-name semantics, nor does it extend to the embeddings τ^{-}_{EB} and τ^{-}_{EH} under the all-name semantics, as illustrated by the following example.

Example 4.8. *Consider $P = \{p(n_1); r(n_2); q \leftarrow not\ p(x)\}$ such that Σ_P has only two names, n_1 and n_2. P has one stable model, $M = \{p(n_1), r(n_2), q\}$. $\tau^{-}_{HP}(P) = \{p(n_1); r(n_2); \forall x(\neg \mathsf{L}p(x) \supset q)\}$ has one stable$^{\mathrm{E}}$ expansion, $T = \{p(n_1),\ r(n_2),\ \mathsf{L}p(n_1),\ \mathsf{L}r(n_2),\ \neg\mathsf{L}p(n_2), \dots\}$. T does* not *include q. To see why this is the case, consider an interpretation w with only one individual k. $\mathsf{L}p(x)$ is trivially true under the any-name semantics, because there is some name for k such that $p(t) \in T$ (viz. $t = n_1$). In the all-name semantics, this situation does not occur, because for $\mathsf{L}p(x)$ to be true, $p(t)$ must be included in T for* every *name ($t = n_1$ and $t = n_2$) for k. One can similarly verify that the result does not apply to the embeddings τ^{-}_{EB} and τ^{-}_{EH} under the all-name semantics, by the positive modal atoms in the antecedents.*

4.2 Embedding Disjunctive Logic Programs

The embeddings τ_{HP} and τ_{EB} cannot be straightforwardly extended to the case of disjunctive logic programs, even for the propositional case. Consider the program $P = \{a \mid b \leftarrow\}$, which has two stable models: $M_1 = \{a\}$ and $M_2 = \{b\}$. However, a naive extension of τ_{HP}, $\tau_{HP}(P) = \{a \vee b\}$, has one stable expansion $T = \{a \vee b, \mathsf{L}(a \vee b), \neg\mathsf{L}a, \neg\mathsf{L}b, \dots\}$. In contrast, τ_{EH} can be straightforwardly extended because of the modal atoms in the consequent of the implication: $\tau^{\vee}_{EH}(P) = \{(a \wedge \mathsf{L}a) \vee (b \wedge \mathsf{L}b)\}$ has two stable expansions $T_1 = \{a \vee b, a, \mathsf{L}a, \neg\mathsf{L}b, \dots\}$ and $T_2 = \{a \vee b, b, \mathsf{L}b, \neg\mathsf{L}a, \dots\}$.

The so-called *positive introspection axioms* (PIAs) [Przymusinski, 1991a] remedy this situation for $\tau^{\vee}_{HP}$ and $\tau^{\vee}_{EB}$. Let PIA_{Σ} be the set of axioms

PIA $\quad \alpha \supset \mathsf{L}\alpha, \quad$ for every objective ground atom α of $\mathcal{L}_{\Sigma}$.

Each PIA ensures that a consistent stable expansion contains either α or $\neg\alpha$.

It would have been possible to define the PIAs in a different way: $(\forall)\ \phi \supset \mathsf{L}\phi$ for any objective atomic formula ϕ. This would, however, effectively close the domain of the predicates in Σ_P (see Example 3.1). We deem this aspect undesirable in combinations with FO theories.

Definition 4.9. *Let r be a rule of form* (1)*. Then:*

$$\begin{aligned}
\tau^{\vee}_{HP}(r) &= (\forall)\ \textstyle\bigwedge_i b_i \wedge \bigwedge_j \neg\mathsf{L}c_j \supset \bigvee_k h_k\\
\tau^{\vee}_{EB}(r) &= (\forall)\ \textstyle\bigwedge_i (b_i \wedge \mathsf{L}b_i) \wedge \bigwedge_j \neg\mathsf{L}c_j \supset \bigvee_k h_k\\
\tau^{\vee}_{EH}(r) &= (\forall)\ \textstyle\bigwedge_i (b_i \wedge \mathsf{L}b_i) \wedge \bigwedge_j \neg\mathsf{L}c_j \supset \bigvee_k (h_k \wedge \mathsf{L}h_k)
\end{aligned}$$

Given a disjunctive logic program P, we define:

$$\tau^{\vee}_{HP}(P) = \{\tau^{\vee}_{HP}(r) \mid r \in P\} \cup \mathrm{PIA}_{\Sigma_P} \cup \mathrm{UNA}_{\Sigma_P}$$
$$\tau^{\vee}_{EB}(P) = \{\tau^{\vee}_{EB}(r) \mid r \in P\} \cup \mathrm{PIA}_{\Sigma_P} \cup \mathrm{UNA}_{\Sigma_P}$$
$$\tau^{\vee}_{EH}(P) = \{\tau^{\vee}_{EH}(r) \mid r \in P\} \cup \mathrm{UNA}_{\Sigma_P}$$

As before, by $\tau^{\vee-}_{\chi}$ we denote the embedding $\tau^{\vee}_{\chi}$ *without* the UNA axioms. Note that the observations about modularity of the embeddings τ_{χ} extend to the disjunctive embeddings $\tau^{\vee}_{\chi}$; the PIAs do not compromise modularity. However, polynomiality of embeddings with PIAs is lost if the size of $\mathcal{L}_{ga}$ is not polynomial in the size of P. We do not write the UNA and PIA axioms explicitly in the examples below.

For the case of standard AEL and ground disjunctive logic programs, the correspondence between the stable models of P and the stable expansions $\tau^{\vee}_{HP}(P)$ and $\tau^{\vee}_{EH}(P)$, respectively, is due to Przymusinski [1991a] and Marek and Truszczynski [1993b].

Proposition 4.10. *A Herbrand interpretation M of a ground disjunctive logic program P is a stable model of P iff there is a consistent stable expansion T of $\tau^{\vee-}_{HP}(P)$ (resp., $\tau^{\vee-}_{EH}(P)$) in standard AEL such that $M = T_{oga}$.*

We generalize this result to the case of FO-AEL and non-ground programs, and additionally for $\tau^{\vee}_{EB}$, similar to the case of normal programs.

Lemma 4.11. *Let P be a logic program, let $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$, let T be a stable$^{\mathrm{X}}$ expansion of $\tau^{\vee}_{\chi}(P)$, and let α be an objective ground atom. Then, $\tau^{\vee}_{\chi}(P) \models^{\mathrm{X}}_{T_{oga}} \alpha$ iff $\tau^{\vee}_{\chi}(P) \models_{T_{oga}} \alpha$ under the standard names assumption. Moreover, $\tau^{\vee-}_{EH}(P) \models^{\mathrm{A}}_{T_{oga}} \alpha$ iff $\tau^{\vee-}_{EH}(P) \models_{T_{oga}} \alpha$.*

Proof. ($\Rightarrow$) Trivial (cf. the "$\Rightarrow$" direction in Lemma 4.5).
($\Leftarrow$) The argument is a straightforward adaptation of the argument in the "$\Leftarrow$" direction in the proof of Lemma 4.5: simply replace the consequent $h \wedge \mathsf{L}h$ with the disjunction $(h_1 \wedge \mathsf{L}h_1) \vee \cdots \vee (h_l \wedge \mathsf{L}h_l)$. Furthermore, it is also easy to see that, as w and w' agree on ground atomic formulas, if the PIA axioms are satisfied in $\langle w, T_{oga}\rangle$, then they are satisfied in $\langle w', T_{oga}\rangle$. □

Lemma 4.12. *Let P be a logic program and let $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$. There exists a stable$^{\mathrm{X}}$ expansion T of $\tau^{\vee}_{\chi}(P)$ iff there exists a stable$^{\mathrm{X}}$ expansion T' of $\tau^{\vee}_{\chi}(gr(P))$ with $T'_{oga} = T_{oga}$. The same result holds for $\tau^{\vee-}_{EH}$ and stable$^{\mathrm{A}}$ expansions.*

Proof. The proof is obtained from the proof of Lemma 4.6 by replacing occurrences of τ_{χ} with $\tau^{\vee}_{\chi}$ and using Lemma 4.11 in place of Lemma 4.5. □

Theorem 4.13. *Let P be a logic program and let $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$. A Herbrand interpretation M is a stable model of P iff there exists a consistent stable$^{\mathrm{X}}$ expansion T of $\tau^{\vee}_{\chi}(P)$ such that $M = T_{oga}$. The same result holds for $\tau^{\vee-}_{EH}$ and stable$^{\mathrm{A}}$ expansions.*

Proof. The reduction of embeddability of non-ground programs in FO-AEL to ground logic programs in standard AEL follows from Lemma 4.12.

Embeddability of $gr(P)$ using $\tau^{\vee}_{HP}$ and $\tau^{\vee}_{EH}$ follows from Proposition 4.10. Embeddability of $gr(P)$ using $\tau^{\vee}_{EB}$ then follows from the embeddability of $\tau^{\vee}_{HP}$, combined with the PIA axioms ($\alpha \supset \mathsf{L}\alpha$): consider a formula $\bigwedge(b_i \wedge \mathsf{L}b_i) \wedge \bigwedge(\neg \mathsf{L}c_j) \supset \bigvee h_k$ in $\tau^{\vee}_{EB}(gr(P))$ and some b_i. If some model $\langle w, \Gamma\rangle$ of $\tau^{\vee}_{EB}(gr(P))$ satisfies b_i, then $\mathsf{L}b_i$ must also be satisfied in $\langle w, \Gamma\rangle$ (by the PIA axioms). Therefore, the stable expansions of $\tau^{\vee}_{HP}(P)$ and $\tau^{\vee}_{EB}(P)$ must be the same. □

A notable difference between the embeddings $\tau^{\vee}_{HP}$ and $\tau^{\vee}_{EB}$, on the one hand, and $\tau^{\vee}_{EH}$, on the other, is the presence, respectively absence of the PIA axioms, as illustrated in the following example.

Example 4.14. *Consider* $P = \{p \mid q \leftarrow \}$. *Then,* $\tau^{\vee}_{HP}(P) = \{p \vee q\} \cup \mathrm{PIA}_{\Sigma_P}$ *has the stable expansions* $T^{HP}_1 = \{p, \neg q, \mathsf{L}p, \neg\mathsf{L}q, \dots\}$ *and* $T^{HP}_2 = \{q, \neg p, \mathsf{L}p, \neg\mathsf{L}p, \dots\}$, *while* $\tau^{\vee}_{EH}(P) = \{\,(p \wedge \mathsf{L}p) \vee (q \wedge \mathsf{L}q)\,\}$ *has the stable expansions* $T^{EH}_1 = \{p, \mathsf{L}p, \neg\mathsf{L}q, \dots\}$ *and* $T^{EH}_2 = \{q, \mathsf{L}p, \neg\mathsf{L}p, \dots\}$*; the latter include neither* $\neg q$ *nor* $\neg p$.

Note that the embedding τ_{HP} cannot be naively extended to logic programs with strong ("classical") negation $\sim$ [Gelfond and Lifschitz, 1991], even for the propositional case. Take, for example, the logic program $P = \{p \leftarrow\sim p\}$; it has one stable model, namely $M = \emptyset$. The naive extension of τ_{HP} treats strong negation as negation in classical logic and the embedding of P yields $\{\neg p \supset p\}$, which has one stable expansion, which includes p. It was shown by Marek and Truszczynski [1993b] that, for the propositional case, the embeddings τ_{EB} and τ_{EH} *can* be naively extended to the case of logic programs with strong negation: consider a rule of the form (1) such that h_i, b_j, c_k are either atoms or strongly negated atoms, and an extension of the embeddings τ_{EB}, τ_{EH} such that $\sim$ is translated to classical negation $\neg$; then, Proposition 4.3 straightforwardly extends to these extended versions of τ_{EB} and τ_{EH} [Marek and Truszczynski, 1993b]. These results can be straightforwardly extended to the non-ground case. Embedding of logic programs with strong negation using τ_{HP} can be done by rewriting P to a logic program P' without strong negation and subsequently embedding P'; see [Gelfond and Lifschitz, 1991] for such a rewriting.

5 Relationships between the Embeddings

In this section we explore correspondences between the embeddings presented in the previous section. We compare the stable expansions of the individual embeddings and, at the level of inference, we compare the sets of autoepistemic consequences. To this end we introduce the following notation:

Definition 5.1. *Let* $\Phi_1, \Phi_2 \subseteq \mathcal{L}_{\mathsf{L}}$ *be FO-AEL theories and* $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$. *We write* $\Phi_1 \equiv^{\mathrm{X}} \Phi_2$ *if* Φ_1 *and* Φ_2 *have the same stable*$^{\mathrm{X}}$ *expansions. For* $\gamma \in \{o, og, oga\}$ *we write* $\Phi_1 \equiv^{\mathrm{X}}_{\gamma} \Phi_2$ *if, for each stable*$^{\mathrm{X}}$ *expansion* T *of* Φ_1, *there exists some stable*$^{\mathrm{X}}$ *expansion* T' *of* Φ_2 *such that* $T_\gamma = T'_\gamma$, *and vice versa.*

Note the implication chain $\Phi_1 \equiv^{\mathrm{X}} \Phi_2 \Rightarrow \Phi_1 \equiv^{\mathrm{X}}_{o} \Phi_2 \Rightarrow \Phi_1 \equiv^{\mathrm{X}}_{og} \Phi_2 \Rightarrow \Phi_1 \equiv^{\mathrm{X}}_{oga} \Phi_2$.

Definition 5.2. *A formula* ϕ *is an* autoepistemic$^{\mathrm{X}}$ consequence *of a theory* $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$, $\mathrm{X} \in \{\mathrm{E}, \mathrm{A}\}$, *if* ϕ *belongs to every stable*$^{\mathrm{X}}$ *expansion of* Φ. $Cn^{\mathrm{X}}(\Phi)$ *denotes the set of all autoepistemic*$^{\mathrm{X}}$ *consequences of* Φ.

The properties stated in this section holds regardless of whether X = E or X = A is considered. Therefore, we omit the superscript X from $\models^{\mathrm{X}}$, $\equiv^{\mathrm{X}}_{\gamma}$, Cn^{X}, stable$^{\mathrm{X}}$, etc.. Furthermore, we write $Cn_\gamma(\Phi)$ for $Cn(\Phi)_\gamma$ $(= Cn(\Phi) \cap \mathcal{L}_{\mathsf{L}\gamma})$.

In our analysis, we consider different classes of logic programs. With the symbols $\mathcal{LP}$, $s\mathcal{LP}$, and $g\mathcal{LP}$ we denote the classes of arbitrary, safe, and ground disjunctive logic programs, respectively. Observe the following inclusions between the classes:

$$g\mathcal{LP} \subseteq s\mathcal{LP} \subseteq \mathcal{LP}$$

We use the letter n to denote the restriction of the respective classes to the case of normal programs: $n\mathcal{LP}$, $sn\mathcal{LP}$, and $gn\mathcal{LP}$.

We start with an investigation of the correspondences between stable expansions, in Section 5.1, and subsequently consider correspondences between sets of consequences, in Section 5.2. Note that while $\Phi_1 \equiv_\gamma \Phi_2$ implies $Cn(\Phi_1)_\gamma = Cn(\Phi_2)_\gamma$, the converse is not true in general. Thus, for applications based on consequence rather than stable expansions, more flexibility between the choice of equivalent embeddings can be expected as one-to-one correspondence between stable expansions is not required. In order not to interrupt the flow of reading, the proofs of most of the results in this section can be found in the appendix.

5.1 Relationships between Stable Expansions of Embeddings

From Theorems 4.7 and 4.13 we immediately obtain the following result concerning correspondence of stable expansions with respect to ground atomic formulas.

Theorem 5.3. *For every $P \in \mathcal{LP}$, $\tau(P) \equiv_{oga} \tau'(P)$ for all $\tau, \tau' \in \{\tau^\vee_{HP}, \tau^\vee_{EB}, \tau^\vee_{EH}\}$, and if $P \in n\mathcal{LP}$, then $\tau(P) \equiv_{oga} \tau'(P)$ for all $\tau, \tau' \in \{\tau_{HP}, \tau_{EB}, \tau_{EH}, \tau^\vee_{HP}, \tau^\vee_{EB}, \tau^\vee_{EH}\}$.*

Thus, all embeddings may be used interchangeably when concerned with ground atoms. This does not hold for the case of arbitrary objective ground formulas. Consider the logic program $P = \{a \leftarrow b\}$. Then $\tau_{HP}(P) = \{b \supset a\}$ has a single stable expansion, which contains $b \supset a$; also $\tau_{EB}(P) = \{b \wedge \mathsf{L}b \supset a\}$ has a single stable expansion, but it does not contain $b \supset a$. Note that while the latter contains $\mathsf{L}b \supset b$, it does not contain $b \supset \mathsf{L}b$ (which would enable obtaining $b \supset a$). The situation changes for the embeddings $\tau^\vee_{HP}$ and $\tau^\vee_{EB}$, because of the PIA axioms.

Proposition 5.4. *For every $P \in n\mathcal{LP}$, $\tau_{EB}(P) \equiv_{og} \tau_{EH}(P)$, and for every $P \in \mathcal{LP}$, $\tau^\vee_{HP}(P) \equiv_{og} \tau^\vee_{EB}(P)$.*

For non-ground formulas we obtain the following result.

Proposition 5.5. *For every $P \in sn\mathcal{LP}$, $\tau_{EB}(P) \equiv \tau_{EH}(P)$.*

To see that the embeddings τ_{EB} and τ_{EH} differ for arbitrary normal programs, consider $P = \{p(a); p(x); q(x) \leftarrow p(x)\}$. The embedding $\tau_{EH}(P) = \{p(a), \forall x.p(x) \wedge \mathsf{L}p(x), \forall x.p(x) \wedge \mathsf{L}p(x) \supset q(x) \wedge \mathsf{L}q(x)\}$ has one stable expansion, which contains $\forall x.q(x)$, while $\tau_{EB}(P) = \{p(a), \forall x.p(x), \forall x.p(x) \wedge \mathsf{L}p(x) \supset q(x)\}$ has one stable expansion, which does not contain $\forall x.q(x)$, because $\forall x.\mathsf{L}p(x)$ is not necessarily true when $\forall x.p(x)$ is true; in other words, the converse Barcan formula ($\mathsf{L}\forall x.\phi(x) \supset \forall x.\mathsf{L}\phi(x)$) is not universally valid, which is a property of FO-AEL under both the any- and all-name semantics [Konolige, 1991].

Note that the result also does not extend to the embeddings τ_{HP} and $\tau^\vee_{HP}$. Consider the logic program $P = \{q(x) \leftarrow p(x)\}$. Then, $\tau_{HP}(P) = \{\forall x.p(x) \supset q(x)\}$ has one stable expansion, which contains $\forall x.p(x) \supset q(x)\}$, while $\tau_{EB}(P) = \{\forall x.p(x) \wedge \mathsf{L}p(x) \supset q(x)\}$ has one stable expansion which does not contain $\forall x.p(x) \supset q(x)$. This difference is caused by the fact that $\mathsf{L}p(x)$ will be false in case an unnamed individual is assigned to x. Similar for $\tau^\vee_{HP}$; the PIA axioms do not help, since they are only concerned with ground atoms and thus do not apply to unnamed individuals.

Proposition 5.6. *If $P \in g\mathcal{LP}$, then $\tau^\vee_{HP}(P) \equiv \tau^\vee_{EB}(P)$.*

Note that this result does not extend to the embedding $\tau^\vee_{EH}$; it does not include the PIA axioms, and thus the argument used in the proof of Proposition 5.6 does not apply.

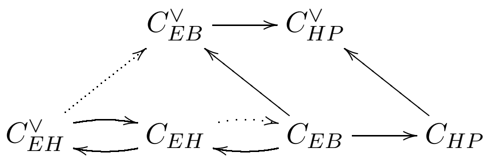


Figure 1: Containment relationships between sets of consequences.

5.2 Relationships between Consequences of Embeddings

In order to investigate the relationships between the embeddings with respect to autoepistemic consequences, we first compare the embeddings with respect to their autoepistemic models. Recall that an autoepistemic interpretation $\langle w, T\rangle$ consists of a first-order interpretation w and a belief set $T \subseteq \mathcal{L}_{\mathsf{L}}$.

Proposition 5.7. *For every $P \in n\mathcal{LP}$ and interpretation $\langle w, T\rangle$, $w \models_T \tau_{EH}(P)$ implies $w \models_T \tau_{EB}(P)$ and $w \models_T \tau_{HP}(P)$ implies $w \models_T \tau_{EB}(P)$.*

Proposition 5.8. *For every $P \in \mathcal{LP}$ and interpretation $\langle w, T\rangle$, $w \models_T \tau^{\vee}_{HP}(P)$ implies $w \models_T \tau^{\vee}_{EB}(P)$; if P is safe, then $w \models_T \tau^{\vee}_{EB}(P)$ implies $w \models_T \tau^{\vee}_{EH}(P)$.*

We now consider the relative behavior of the embeddings with respect to autoepistemic consequences. In order to present our results in a compact and accessible way, we show a small (yet sufficient) number of relationships between the sets of consequences in a graph (Figure 1). Every particular relationship between embeddings can be easily derived from paths in this graph.

Specifically, in Figure 1 $C^{(\vee)}_{\chi}$ is short for $Cn_o(\tau^{(\vee)}_{\chi}(P))$, the straight arrow $\longrightarrow$ represents set inclusion ($\subseteq$), and the dotted arrow $\dashrightarrow$ represents set inclusion in case P is safe. Since $\longrightarrow$ implies $\dashrightarrow$, dotted arrows are only shown if straight arrows are absent. The following lemma states the correctness of Figure 1.

Lemma 5.9. *If $C^{(\vee)}_{\chi} \longrightarrow C^{(\vee)}_{\gamma}$ (resp., $C^{(\vee)}_{\chi} \dashrightarrow C^{(\vee)}_{\gamma}$) in the graph of Figure 1, then $Cn_o(\tau^{(\vee)}_{\chi}(P)) \subseteq Cn_o(\tau^{(\vee)}_{\gamma}(P))$ for every $P \in n\mathcal{LP}$ (resp., for every $P \in sn\mathcal{LP}$).*
Furthermore, if $C^{\vee}_{\chi} \longrightarrow C^{\vee}_{\gamma}$ (resp., $C^{\vee}_{\chi} \dashrightarrow C^{\vee}_{\gamma}$), then $Cn_o(\tau^{\vee}_{\chi}(P)) \subseteq Cn_o(\tau^{\vee}_{\gamma}(P))$ for every $P \in \mathcal{LP}$ (resp., for every $P \in s\mathcal{LP}$).

Note that by transitivity of $\subseteq$, paths in the graph yield further relations; e.g., $Cn_o(\tau_{EB}(P)) \subseteq Cn_o(\tau^{\vee}_{HP}(P))$ since C_{EB} reaches $C^{\vee}_{HP}$ reaches via a path with straight edges. We now show that the graph exactly characterizes the containment relationships via paths. To this end, we first note some negative relationships between embeddings.

Lemma 5.10. *The following inclusion relations do not hold: $Cn_o(\tau^{\vee}_{EH}(P)) \subseteq Cn_o(\tau^{\vee}_{HP}(P))$, for every $P \in \mathcal{LP}$; $Cn_o(\tau^{\vee}_{EB}(P)) \subseteq Cn_o(\tau_{HP}(P))$, for every $P \in sn\mathcal{LP}$; and $Cn_o(\tau_{HP}(P)) \subseteq Cn_o(\tau^{\vee}_{EB}(P))$, for every $P \in sn\mathcal{LP}$.*

From these negative relationships, combined with the positive ones above, we can infer further negative relationships. For example, from $C^{\vee}_{EH} \not\subseteq C^{\vee}_{EB}$ and $C^{\vee}_{EH} \subseteq C_{EH}$, we infer $C_{EH} \not\subseteq C^{\vee}_{EB}$. Exploiting such inferences, we show the following main result of this section.

Theorem 5.11. *For* $P \in n\mathcal{LP}$ *(resp.,* $P \in sn\mathcal{LP}$*),* $Cn_o(\tau_\chi^{(\vee)}(P)) \subseteq Cn_o(\tau_\gamma^{(\vee)}(P))$ *iff* $C_\gamma^{(\vee)}$ *is reachable from* $C_\chi^{(\vee)}$ *in the graph in Figure 1 on a path with* $\longrightarrow$ *arcs (resp., with arbitrary arcs).*
Likewise, for $P \in \mathcal{LP}$ *(resp.,* $P \in s\mathcal{LP}$*),* $Cn_o(\tau_\chi^{\vee}(P)) \subseteq Cn_o(\tau_\gamma^{\vee}(P))$ *iff* $C_\gamma^{\vee}$ *is reachable from* $C_\chi^{\vee}$ *on a path with* $\longrightarrow$ *(resp., with arbitrary arcs).*

Proof. By Lemmas 5.9 and 5.10, the respective containment relationships are correct. Clearly, reflexivity and transitivity are sound inference rules for set inclusion, and thus paths in the graph of Figure 1 are sound with respect to positive containments. Their completeness, for both arbitrary P and safe P, is established using the following basic inference rules for non-inclusion: (i) $A \not\subseteq B$, $C \subseteq B$ implies $A \not\subseteq C$ and (ii) $A \not\subseteq B$, $A \subseteq C$ implies $C \not\subseteq B$.[4] Exhaustive application to the (non-)containments in Lemmas 5.9 and 5.10 (e.g., using a simple logic program) yields one of $C_\chi^{(\vee)} \subseteq C_\gamma^{(\vee)}$ and $C_\chi^{(\vee)} \not\subseteq C_\gamma^{(\vee)}$ for each pair $C_\chi^{(\vee)}$, $C_\gamma^{(\vee)}$. □

Accordingly, $C_{EB} \subseteq C_{EH}^{\vee}$ and $C_{EB} \subseteq C_{HP}^{\vee}$ are the only nontrivial inclusions for arbitrary programs besides those in Figure 1; for safe programs, there are more. We note that the figure is minimal, in the sense that if any of the arcs is removed (or turned from solid into dashed), the theorem no longer holds.

6 Combinations with First-Order Theories

In this section we explore correspondences between the logic program embeddings from Section 4 in combinations with FO theories. To this end, we consider a basic combination of logic programs P and FO theories Φ defined as

$$\iota_\chi^{(\vee)}(\Phi, P) = \Phi \cup \tau_\chi^{(\vee)}(P) \subseteq \mathcal{L}_{\mathsf{L}}$$

where $\Sigma_{\mathcal{L}_{\mathsf{L}}}$ is the union of the signatures Σ_Φ and Σ_P. More involved combinations (e.g., which augment P and Φ with further rules and axioms, respectively) might be recast to such basic combinations.

In the preceding sections we have considered both the any- and all-name semantics, both in the definition of the embeddings and in our analysis of the differences between the embeddings of logic programs. It turned out that the embeddings are faithful for both semantics (cf. Theorems 4.7 and 4.13), implying correspondence with respect to objective ground atoms between the two semantics for all embeddings $\tau_\chi^{(\vee)}$, and the relationships between the embeddings stated in the previous section hold for both semantics. However, in combinations with FO theories the two semantics diverge since names from the first-order part may not be provably identical to or different from other names. The following example illustrates differences between the semantics in the face of positive and negative occurrences of the modal operator.

Example 6.1. *Consider the logic program* P*:*

$$q(a)$$
$$r \leftarrow p(x), not\ q(x)$$
$$s(x) \leftarrow p(x)$$

[4] Note that non-inclusion for normal programs implies non-inclusion for disjunctive programs, since every normal program is a disjunctive program.

and the FO theory $\Phi = \{p(b)\}$. *We note here that the signature of* P *contains only one function symbol, the constant constant* a. *Consequently,* $\mathrm{UNA}_{\Sigma_P} = \emptyset$.

$\iota_{EB}(\Phi, P)$ *has one stable*$^{\mathrm{E}}$ *expansion* T^{E} *and one stable*$^{\mathrm{A}}$ *expansion* T^{A}. T^{E} *contains* $q(a)$, *but not* $q(b)$; *both contain* $p(b)$, *but not* $p(a)$. *Consider an interpretation* $w = \langle U, \cdot^I \rangle$ *such that* $a^w = b^w = k$, $k \in p^I, k \in q^I, k \in s^I$ *and* $w \not\models r$, *and a variable assignment* B *such that* $x^B = k$. *Then,* $\beta = \{x \mapsto a\}$ *is an associated variable substitution and* $q(x)\beta \in T^{\mathrm{E}}$, *and so* $(w, B) \models^{\mathrm{E}}_{T^{\mathrm{E}}} \mathsf{L}q(x)$. *Another associated variable substitution is* $\beta = \{x \mapsto b\}$, *and so* $(w, B) \models^{\mathrm{E}}_{T^{\mathrm{E}}} \mathsf{L}p(x)$. *So,* $w \models^{\mathrm{E}}_{T^{\mathrm{E}}} \iota_{EB}(\Phi, P)$ *and thus* $r \notin T^{\mathrm{E}}$. *One can straightforwardly argue that* $\iota_{EB}(\Phi, P) \models^{\mathrm{E}}_{T} s(b)$, *and so* $s(b) \in T^{\mathrm{E}}$.

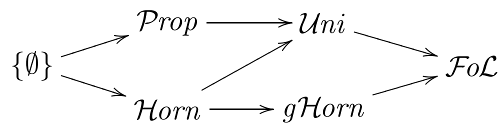
Consider interpretation w' *that is like* w, *except that* $k \notin s^I$ *and* $w' \models r$. *Since* T^{A} *does not contain* $p(a)$, $(w', B) \not\models^{\mathrm{A}}_{T^{\mathrm{A}}} \mathsf{L}p(x)$. T^{A} *does not contain* $q(b)$, *but* $(w, B) \models^{\mathrm{A}}_{T^{\mathrm{A}}} \mathsf{L}q(x)$ *holds only if* $q(x)\beta \in T^{\mathrm{A}}$ *for* every β *associated with* B, *including* $\beta = \{x \mapsto b\}$, *and so* $(w', B) \models^{\mathrm{A}}_{T^{\mathrm{A}}} \neg\mathsf{L}q(x)$ *and* $w' \models^{\mathrm{A}}_{T^{\mathrm{A}}} \iota_{EB}(\Phi, P)$. *Consequently,* $s(b) \notin T^{\mathrm{A}}$. *It is straightforward to verify that* $r \in T^{\mathrm{A}}$.

In order to avoid a proliferation of results, following Konolige [1991] we concentrate in this section on the any-name semantics. In the following section we discuss our results in the light of the standard names assumption, for which the any- and all-name semantics coincide.

In our analysis we consider the same syntactic classes of programs as in the previous section and we consider the following classes of objective theories:

> arbitrary ($\mathcal{F}o\mathcal{L} = 2^{\mathcal{L}}$), universal ($\mathcal{U}ni$), Horn ($\mathcal{H}orn$), generalized Horn ($g\mathcal{H}orn$),[5] propositional ($\mathcal{P}rop$), and empty ($\{\emptyset\}$, i.e., in semantic terms tautological) FO theories.

The order diagram is as follows (arrows stand for set inclusion):



For all pairs of classes of logic programs and FO theories we determine the relationships between stable expansions of different combinations $\iota^{(\vee)}_{\chi}(\Phi, P)$ and $\iota^{(\vee)}_{\gamma}(\Phi, P)$ at different levels of granularity. As in Section 5.1, we concentrate here on correspondences of stable expansions $\equiv^{\mathrm{E}}_{\mathrm{x}}$; they imply that relative to the class of formulas x, the embeddings $\tau^{(\vee)}_{\chi}(P)$ and $\tau^{(\vee)}_{\gamma}(P)$ are interchangeable in combinations.

In Section 6.1 we state our main result on the relationships between stable expansions of combinations and make several observations. In Section 6.2 we establish the partial results necessary for deriving our main result. The proofs of the partial results can be found in the appendix.

6.1 Relationships between Stable Expansions of Combinations

Our results are summarized in Table 1, which gives a complete picture of the correspondences, where each entry represents a *most general correspondence*, i.e., neither the correspondence $\equiv^{\mathrm{E}}_{\mathrm{x}}$ nor the logic program or FO theory class may be relaxed. This is formally stated in the main theorem of this section.

[5]Generalized Horn formulas have the form $(\forall) b_1 \wedge \cdots \wedge b_n \supset \exists \vec{y}.h$ where all b_i and h are atomic, and variables $\vec{y}$ occur only in h, which may be absent. Such formulas are relevant in, e.g., the context of description logics like Horn-$\mathcal{SHIQ}$ [Hustadt *et al.*, 2005] and in reasoning about actions.

$\Phi \setminus P$	$\mathcal{LP}$	$s\mathcal{LP}$	$g\mathcal{LP}$
$\mathcal{FoL}$	$\iota_{EH} \equiv^{\mathrm{E}} \iota^{\vee}_{EH}$		$\iota_{EB} \equiv^{\mathrm{E}} \iota_{EH}$ $\iota^{\vee}_{HP} \equiv^{\mathrm{E}} \iota^{\vee}_{EB}$
$\mathcal{U}ni$			
$g\mathcal{H}orn$			$\iota_{HP} \equiv^{\mathrm{E}}_{oga} \iota_{EH}$ $\iota^{\vee}_{HP} \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{EH}$
$\mathcal{H}orn$			
$\mathcal{P}rop$	$\iota^{\vee}_{HP} \equiv^{\mathrm{E}}_{og} \iota^{\vee}_{EB}$ $\iota_{EB} \equiv^{\mathrm{E}}_{og} \iota_{EH}$	$\iota_{EB} \equiv^{\mathrm{E}} \iota_{EH}$	
$\{\emptyset\}$	$\iota_{HP} \equiv^{\mathrm{E}}_{oga} \iota_{EB}$ $\iota_{HP} \equiv^{\mathrm{E}}_{oga} \iota_{EH}$ $\iota^{\vee}_{HP} \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{EH}$		

Table 1: Correspondence between stable expansions of combinations; $\iota_x^{(\vee)}$ is short for $\iota_x^{(\vee)}(\Phi, P)$.

We call $\Phi_1 \equiv^{\mathrm{E}}_{\mathrm{x}} \Phi_2$ a *trivial inference* from a set of equivalences Q if it is derivable from Q by the fact that $\Phi_1 \equiv^{\mathrm{E}} \Phi_2$ implies $\Phi_1 \equiv^{\mathrm{E}}_{og} \Phi_2$ and $\Phi_1 \equiv^{\mathrm{E}}_{og} \Phi_2$ implies $\Phi_1 \equiv^{\mathrm{E}}_{oga} \Phi_2$ and by reflexivity, transitivity, and symmetry of $\equiv^{\mathrm{E}}_{\mathrm{y}}$, $\mathrm{y} \in \{\epsilon, og, oga\}$.

Theorem 6.2. *Let $\mathcal{X}$ be a class of FO theories, let $\mathcal{Y}$ be a class of programs, and let $\mathrm{x} \in \{\epsilon$, og, $oga\}$. Then $\iota_\chi(\Phi, P) \equiv^{\mathrm{E}}_{\mathrm{x}} \iota_\gamma(\Phi, P)$ holds for all $\Phi \in \mathcal{X}$ and $P \in \mathcal{Y}$ iff $\iota_\chi(\Phi, P) \equiv^{\mathrm{E}}_{\mathrm{x}} \iota_\gamma(\Phi, P)$ follows for cell $(\mathcal{X}, \mathcal{Y})$ in Table 1 by trivial inferences, where $\chi, \gamma \in \{^{\vee}_{HP}, {}^{\vee}_{EB}, {}^{\vee}_{EH}\}$ if $P \in \mathcal{LP}$ and $\chi, \gamma \in \{_{HP}, {}_{EB}, {}_{EH}, {}^{\vee}_{HP}, {}^{\vee}_{EB}, {}^{\vee}_{EH}\}$ if $P \in n\mathcal{LP}$.*

We will establish the results of Table 1 and provide some intuitive explanations about partial results in the next subsection.

Note that removing any statement from Table 1 or modifying any correspondence type invalidates the theorem. We do not explicitly consider correspondence of stable expansions with respect to objective formulas, i.e., $\equiv^{\mathrm{E}}_{o}$. Clearly, $\Phi_1 \equiv^{\mathrm{E}} \Phi_2$ implies $\Phi_1 \equiv^{\mathrm{E}}_{o} \Phi_2$; in addition, all the counterexamples to $\Phi_1 \equiv^{\mathrm{E}} \Phi_2$ presented in the following subsection also apply to $\Phi_1 \equiv^{\mathrm{E}}_{o} \Phi_2$. Hence, $\equiv^{\mathrm{E}}_{o}$ coincides with $\equiv^{\mathrm{E}}$. We can make the following observations about the results.

- The various combinations behave differently in the general case. Only two of them, $_{EH}$ and $^{\vee}_{EH}$, are always equivalent (they coincide on normal programs).
- For combinations with arbitrary FO theories, further correspondences are only present for ground logic programs in some cases. Narrowing to any of the classes that allow predicates of arity > 0 ($\mathcal{H}orn$, $g\mathcal{H}orn$, $\mathcal{U}ni$) does not change the picture.
- For arbitrary logic programs, only in case of propositional theories do some combinations behave equivalently. Requiring safety leads only for propositional theories and in one case ($_{EB}$ and $_{EH}$) to a stronger correspondence
- $\mathcal{U}ni$ and $\mathcal{H}orn$ show no most general correspondences, which means that with respect to more general or more restrictive classes, their change in syntax does not affect equivalence.

- In contrast, the important class $g\mathcal{H}orn$ has maximal correspondences for ground programs. Thus, for combinations with FO theories from the classes $\mathcal{H}orn$, $g\mathcal{H}orn$, and $\mathcal{U}ni$, we have equivalent behavior only for ground programs in some cases (apart from $_{EH}$ and $^{\vee}_{EH}$). However, one could imagine that in a specific combination non-ground programs are grounded before they are embedded, thereby making the results applicable to a large class of combinations.

We illustrate the use of the result in Theorem 6.2 with an example. Note that if the stable expansions of two embeddings or combinations correspond with respect to a certain class formulas, then the embeddings, respectively combinations also agree on autoepistemic consequences for these classes.

Example 6.3. *Consider $P = \{q(a);\ p(x);\ r(x) \leftarrow not\ s(x),\ p(x)\}$ from Example 4.4, which is neither safe nor ground. Hence to determine correspondence between embeddings, we use the first column of Table 1. As P is normal, all equations in the column are applicable. We have that, e.g., $\tau_{EB}(P) \equiv^{\mathrm{E}}_{og} \tau_{EH}(P)$, $\tau^{\vee}_{HP}(P) \equiv^{\mathrm{E}}_{og} \tau^{\vee}_{EB}(P)$, and $\tau_{HP}(P) \equiv^{\mathrm{E}}_{oga} \tau_{EB}(P)$. Let Φ be a propositional theory; then we also have $\iota_{EB}(\Phi, P) \equiv^{\mathrm{E}}_{og} \iota_{EH}(\Phi, P)$ and $\iota^{\vee}_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{og} \iota^{\vee}_{EB}(\Phi, P)$, but not $\iota_{HP}(P) \equiv^{\mathrm{E}}_{oga} \iota_{EB}(P)$. Furthermore, we can conclude that $\iota_{EB}(P)$ and $\iota_{EH}(P)$, and also $\iota^{\vee}_{HP}(P)$ and $\iota^{\vee}_{EB}(P)$, agree on objective ground autoepistemic consequences.*

6.2 Derivation of the Results

We start with the positive results. Trivially, ι_{EH} and $\iota^{\vee}_{EH}$ coincide for arbitrary FO theories, and the the equivalence results for empty Φ in Table 1 carry over from the respective results on embeddings in Section 5.

We show that for ground programs, the τ_{EB} and τ_{EH} embeddings are interchangeable in any combination with an FO theory.

Proposition 6.4. *For every $(\Phi, P) \in \mathcal{FoL}\times gn\mathcal{LP}$, $\iota_{EB}(\Phi, P) \equiv^{\mathrm{E}} \iota_{EH}(\Phi, P)$.*

Intuitively, this holds because only named individuals matter in rules, and hence the modal atoms $\mathsf{L}h$ in embedded rule heads do not matter. However, this does not generalize from ground to safe programs, as the evaluation of literals $\neg\mathsf{L}p(x)$ in the rule bodies does not amount to grounding (see Proposition 6.9(4)).

Also the $\tau^{\vee}_{HP}$ and $\tau^{\vee}_{EH}$ embeddings are interchangeable in combinations with arbitrary FO theories if the logic program is ground.

Proposition 6.5. *For every $(\Phi, P) \in \mathcal{FoL}\times g\mathcal{LP}$, $\iota^{\vee}_{HP}(\Phi, P) \equiv^{\mathrm{E}} \iota^{\vee}_{EB}(\Phi, P)$.*

The reason is that we can eliminate all modal atoms $\mathsf{L}b$ from rule bodies with the PIA axioms in $\iota^{\vee}_{EB}(\Phi, P)$ and obtain $\iota^{\vee}_{HP}(\Phi, P)$. Such elimination is not possible in the non-ground case, since the PIAs only apply to atoms from Σ_P.

Moving now to fragments of $\mathcal{FoL}$, i.e., down the rows in Table 1, we first have:

Proposition 6.6. *For $(\Phi, P) \in g\mathcal{H}orn\times g\mathcal{LP}$, $\iota^{\vee}_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{EH}(\Phi, P)$, and if $P \in gn\mathcal{LP}$, then $\iota_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota_{EH}(\Phi, P)$.*

Intuitively, in the first case we can add modal atoms $\mathsf{L}b$ in the bodies and $\mathsf{L}h$ in the heads of $\tau_{HP}(P)$, by the PIA axioms, thereby obtaining $\iota^{\vee}_{EH}(\Phi, P)$. To go from τ_{EH} to τ_{HP} is possible if Φ is not disjunctive with respect to atoms h. This is the case for a Horn Φ, and similarly for a generalized Horn Φ, as we can apply skolemization. In the second case, there are no PIA axioms, but we can similarly apply skolemization and obtain a disjunction-free theory that is Horn modulo modal atoms. Skolemization does not work for non-ground programs in this case, as previously unnamed individuals are named by skolem terms.

We note that, combined with previous results, we can infer from Proposition 6.6 that Theorem 5.3 generalizes from embeddings to combinations with generalized Horn theories for the case of ground logic programs.

For propositional theories, we obtain a result symmetric to Proposition 6.5 for arbitrary logic programs.

Proposition 6.7. *For every* $(\Phi, P) \in \mathcal{P}rop \times \mathcal{LP}$, $\iota^{\vee}_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{og} \iota^{\vee}_{EB}(\Phi, P)$.

Intuitively, this holds because a propositional Φ can not interfere with names of individuals—it has no names. Therefore, as in the case of the embeddings $\tau^{\vee}_{HP}$ and $\tau^{\vee}_{EB}$, we can eliminate all modal atoms $\mathsf{L}b$ from rule bodies in $\iota^{\vee}_{EB}(\Phi, P)$ to obtain $\iota^{\vee}_{HP}(\Phi, P)$. For similar reasons, also correspondence results for the embeddings τ_{EB} and τ_{EH} extend to combinations with propositional theories.

Proposition 6.8. *For every* $(\Phi, P) \in \mathcal{P}rop \times n\mathcal{LP}$, $\iota_{EB}(\Phi, P) \equiv^{\mathrm{E}}_{og} \iota_{EH}(\Phi, P)$, *and if* $P \in sn\mathcal{LP}$, *then* $\iota_{EB}(\Phi, P) \equiv^{\mathrm{E}} \iota_{EH}(\Phi, P)$.

It turns out that the results in the preceding propositions cannot be extended to more general classes of programs or theories, or larger subsets of stable expansions.

Proposition 6.9. *There are pairs* (Φ, P) *in*

1. $\mathcal{P}rop \times gn\mathcal{LP}$ *such that* $\iota\chi(\Phi, P) \not\equiv^{\mathrm{E}}_{oga} \iota\gamma(\Phi, P)$, *for* $(\chi, \gamma) \in \{(_{HP}, _{EB}), (_{HP}, {}^{\vee}_{HP}), ({}^{\vee}_{HP}, {}^{\vee}_{EH})\}$;
2. $\{\emptyset\} \times gn\mathcal{LP}$ *such that* $\iota\chi(\Phi, P) \not\equiv^{\mathrm{E}}_{og} \iota\gamma(\Phi, P)$, *for* $(\chi, \gamma) \in \{(_{HP}, _{EB}), (_{HP}, {}^{\vee}_{HP}), ({}^{\vee}_{HP}, {}^{\vee}_{EH})\}$;
3. $\{\emptyset\} \times sn\mathcal{LP}$ *such that* $\iota^{\vee}_{HP}(\Phi, P) \not\equiv^{\mathrm{E}} \iota^{\vee}_{EB}(\Phi, P)$;
4. $\mathcal{H}orn \times sn\mathcal{LP}$ *such that* $\iota\chi(\Phi, P) \not\equiv^{\mathrm{E}}_{oga} \iota\gamma(\Phi, P)$, *for* $(\chi, \gamma) \in \{(_{HP}, _{EB}), (_{HP}, _{EH}), (_{EB}, _{EH}), (_{HP}, {}^{\vee}_{HP}), (_{HP}, {}^{\vee}_{EB}), (_{EB}, {}^{\vee}_{EB}), (_{EB}, {}^{\vee}_{HP}), ({}^{\vee}_{HP}, {}^{\vee}_{EB}), (_{EH}, {}^{\vee}_{EB}), (_{EH}, {}^{\vee}_{HP})\}$; *and*
5. $\{\emptyset\} \times n\mathcal{LP}$ *such that* $\iota_{EB}(\Phi, P) \not\equiv^{\mathrm{E}} \iota_{EH}(\Phi, P)$.

Proof of Theorem 6.2. Correctness of the table (i.e., the '$\Leftarrow$' direction of the theorem) follows from the fact that $\tau_{EH}(P)$ and $\tau^{\vee}_{EH}(P)$ are identical for normal programs P (thus $\iota_{EH}(\Phi, P) = \iota^{\vee}_{EH}(\Phi, P)$), Theorem 5.3, Propositions 6.4–6.8, the inclusion relations between the classes of programs and FO theories, and the properties of the $\equiv^{\mathrm{E}}_{\mathrm{x}}$ relation. Completeness (i.e., '$\Rightarrow$' direction) is shown analogously, by exploiting the counterexamples in Proposition 6.9. One can verify with little yet tedious effort that the table is complete and that no entries can be relaxed (e.g., using a simple logic program). □

Most of the counterexamples in the proof of Proposition 6.9 use only positive programs. However, for the case (4) (where $\Phi \in \mathcal{H}orn$), the counterexamples use negation. It turns out that this use is essential in some cases, as we have the following result for positive programs and Horn theories.

Proposition 6.10. *For every* $(\Phi, P) \in g\mathcal{H}orn \times sn\mathcal{LP}$ *such that* P *is positive,* $\iota_{EB}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota_{EH}(\Phi, P)$*, and for every* $(\Phi, P) \in \mathcal{H}orn \times n\mathcal{LP}$ *such that* P *is positive, it holds that* $\iota_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota_{EB}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota_{EH}(\Phi, P)$*.*

This result does not extend to the disjunctive embeddings, because there are no PIA axioms for the atoms involving names not in P, and it does not extend to more general formulas (see, e.g., Example 4.2 on page 9).

7 Discussion

In this section we discuss implications of the results in the previous sections. We first discuss the implications of our results on the relationships between the embeddings, and make a number of observations about those relationships. We then discuss how the results in this paper can be used in the context of combining classical theories (ontologies) with logic programs (rules). Specifically, how the embeddings studied in this paper can be used as building blocks for such combinations. Finally, we discuss our choice of FO-AEL as the underlying formalism, and compare the semantics for quantification (quantifying-in) with other approaches to quantifying-in in autoepistemic logic (e.g., [Levesque, 1990; Levesque and Lakemeyer, 2000; Kaminski and Rey, 2002]).

7.1 Relationships between the Embeddings

Using the results obtained in Sections 5 and 6, we can make a number of observations about the embeddings:

(1) The differences between the embeddings by themselves do not depend on the use of negation in the program. Generally speaking, the differences originate from the positive use of the modal operator in the antecedent and the consequent, and the use of the PIA axioms. However, in combinations with FO theories, the interaction between names in Φ for which there are no UNA axioms and negation in the rules gives rise to different behavior of the embeddings (see Proposition 6.9(4)).

(2) The stable expansions of embeddings with and embeddings without the PIAs generally tend to differ. However, we can note that the former are generally stronger in terms of autoepistemic consequences (cf. Figure 1 and Example 4.14).

(3) The embeddings τ_{HP} and $\tau^{\vee}_{HP}$ are generally the strongest in terms of consequences (see Figure 1), when comparing to other embeddings without and with PIAs, respectively. They allow to derive the contrapositive of rules (cf. Example 4.2) and the bodies of rules are applicable to unnamed individuals, whereas the antecedents of the axioms in the other embeddings are only applicable to named individuals, because of the positive modal atoms in the bodies.

(4) For unsafe programs, the embeddings τ_{EH} and $\tau^{\vee}_{EH}$ are generally not comparable with the others; embeddings of unsafe rules may result in axioms of the form $\forall x.\mathsf{L}p(x)$ (cf. Example 4.4), which result in all individuals being named.

(5) If names in Φ that lack UNA axioms and rules in P (e.g., Φ is propositional or P is ground) do not interact, then τ_{EB} and τ_{EH} are in most cases interchangeable.

Special care needs to be taken if one selects an embedding that includes the PIA axioms (i.e., $\tau^{\vee}_{HP}$ and $\tau^{\vee}_{EB}$). These axioms of the form $\alpha \supset \mathsf{L}\alpha$, since they ensure that α or $\neg\alpha$ is included in every stable expansion, for every ground atom of Σ_P. Note that the PIA axioms have no effect when considering individuals that are not named by ground terms in Σ_P.

The UNA axioms in embeddings, which serve to make individuals different by default, may interact with the FO theory in a combination. For example, consider $P = \{p(a);\ p(b)\}$ and $\Phi = \{a \neq b \supset r,\ a \neq c \supset s\}$. Then, every stable expansion of $\iota_\chi(\Phi, P)$, for any embedding τ_χ we considered, contains r as $a \neq b$ is concluded by default, but not s (as c is unknown in P). To shortcut such (possibly undesired) inequality transfers from P to Φ, the unique names or even the standard names assumption may be adopted *a priori.* Recall that the results on the embeddings in Section 4 were obtained by stepping through the standard names assumption, and thus they also hold under the unique names or standard names assumption, as shown by de Bruijn *et al.* [2008]. On the one hand, this should extend to the positive results about correspondences in Sections 5 and 6, whose proofs rely on named interpretations and no equalities between individuals are enforced. On the other hand, some counterexamples for correspondences fail, including those for the first item in Lemma 5.10 and Proposition 6.9(4), and thus further correspondences may hold. An in-depth study of the effect of unique names and standard names assumptions on the correspondences and differences between the embeddings is an interesting subject for further work.

7.2 Different Embeddings and Combinations

Recall the general setting for combining a first-order theory Φ and a logic program P in a unifying formalism (FO-AEL) that we sketched in the introduction. The combination operator ι takes as arguments the theory Φ and the program P, and returns an FO-AEL theory $\iota(\Phi, P)$. The operator provides two embedding functions: σ and τ map first-order theories, respectively logic programs to FO-AEL theories. We also mentioned that in the simplest case the combination is the union of the two individual embeddings: $\iota(\Phi, P) = \sigma(\Phi) \cup \tau(P)$.

In Section 4 we investigated several candidates for the embedding function for logic programs, τ. All these embedding functions are faithful, in the sense that the stable models of the program P correspond to the sets of objective ground atomic formulas in the stable expansions of the embedding $\tau(P)$. In Section 5 we investigated the relationships between the stable expansions of these embeddings when considering more general formulas. It turned out that there are already significant differences between the expansions when considering non-ground or non-atomic formulas.

Now, in Section 6, we investigated the relationships between the expansions when considering combinations of the embeddings with first-order theories. We have found that, under certain circumstances—namely, the first-order theory and program are of particular shapes and we are interested in a particular kind of formulas (e.g., ground formulas)—certain embeddings can be used interchangeably (cf. Table 1 on page 19). For example, if the program is normal and ground ($P \in gn\mathcal{LP}$), the theory is generalized Horn ($\Phi \in g\mathcal{H}orn$), and we are interested in objective formulas, we can use the embeddings τ_{EB} and τ_{EH} interchangeably: $\iota_{EB} \equiv \iota_{EH}$ for $P \in gn\mathcal{LP}$ and $\Phi \in \mathcal{F}o\mathcal{L}$, according to Table 1, $\mathcal{H}orn \subseteq \mathcal{F}o\mathcal{L}$, and the set of objective formulas is a subset of the set of formulas.

Our results are not limited to combinations of the form $\iota(\Phi, P) = \Phi \cup \tau_{\chi}^{(\vee)}(P)$, where $\tau_{\chi}^{(\vee)}$ is one of the embeddings investigated in this paper. One could imagine adding axioms to Φ or rules to P to achieve the desired interoperation between the two components, or even changing the axioms or rules (e.g., by grounding), obtaining a first-order theory Φ' and program P'. In this more general setting, the combination is defined as

$$\iota(\Phi, P) = \Phi' \cup \tau_{\chi}^{(\vee)}(P')$$

where Φ' and P' are obtained from Φ and P by adding and/or replacing axioms and rules. The results about the relationships between the embeddings obtained in Section 6 can be applied, provided that Φ' and P' are in the respective classes of theories and programs, independent of the shapes of Φ and P.

As discussed in Section 4, embeddings that include the UNA axioms are not modular in general, but only signature-modular. This can be remedied by instead using the single axiom

$$(\forall)\ \mathsf{L}x = x \wedge \mathsf{L}y = y \wedge \neg\mathsf{L}x = y \supset x \neq y$$

which has the same effect for embeddings. However, using this axiom would effect default uniqueness on all names in a combination, not only those from the signature of the program (if desired, such default uniqueness can be easily accomplished by just mentioning respective terms in the logic program). As a consequence, also the combinations behave differently.

7.3 Quantifying-in in First-Order Autoepistemic Logic

We consider here FO-AEL, with the semantics for quantifying-in as defined by Konolige [1991], as an underlying formalism for combinations of first-order theories and logic programs. However, further semantics for quantifying-in have been proposed in the literature.

Levesque [1990] defined the logic of only knowing (see also [Levesque and Lakemeyer, 2000]), which is essentially a superset of FO-AEL. Levesque's semantics for quantifying-in is slightly different from the one of Konolige [1991] that we used in this paper. He adopted a standard names assumption that amounts to a special case of the notion in Section 2.1; there is a countably infinite number of constant symbols in the language, but there are no (other) function symbols. Likewise, the variant of FO-AEL by Kaminski and Rey [2002] also employs a standard names assumption, although under a somewhat different guise: the domain of every interpretation is an extended Herbrand interpretation, i.e., it is a superset of the set of constant symbols in the theory; function symbols not considered. The semantics of Konolige does not impose such restrictions, e.g., the domain may be infinite, while the number of constants is finite, and function symbols are allowed.

It is well known that reasoning in standard first-order logic can be reduced to reasoning in first-order logic with the standard names assumption, as long as there are sufficiently many constant symbols available (cf. [Fitting, 1996]).

Different from Levesque [1990], Kaminski and Rey [2002] did not consider equality in the language. However, equality in first-order logic with standard names behaves quite differently from equality in standard first-order logic. In the former case, two constant symbols may be interpreted as the same element in the domain, whereas in the latter case all constant symbols are interpreted distinctly, e.g., $a = b$ cannot be satisfied if a and b are distinct constant symbols.[6] It is, however, possible to reduce reasoning in standard first-order logic with equality to reasoning in first-order logic with standard names using a special congruence predicate (cf. [Fitting, 1996, Theorem 9.3.9]). Motik and Rosati [2007] use such a predicate in their variant of the logic MKNF [Lifschitz, 1991; Lifschitz, 1994], as do de Bruijn *et al.* [2008] in a variant of FO-AEL with standard names; see Section 8.2 for further discussion about this work.

8 Related Work

We review here two areas of related work: extensions of logic programming and description logic semantics with open domains and nonmonotonicity, respectively, and approaches to combining rules and ontologies.

8.1 Extensions of LP and DL Semantics

We have studied the combination of logic programs and ontologies using embeddings in a unifying formalism (FO-AEL). One could imagine, in contrast, extensions of the semantics of logic programs or ontologies to incorporate (parts of) the other formalism. One such extension of logic programming semantics is that of open domains [Gelfond and Przymusinska, 1993]. Such extended semantics can be used to accommodate incomplete knowledge, an important aspect of ontology languages.

Heymans *et al.* [2006] describe an extension of the stable model semantics with open domains, called *open answer set programming*. They show how an expressive description logic, $\mathcal{SHIQ}$, can be embedded in this language and Heymans *et al.* [2008] show how it can be used for combinations of rules and ontologies, following the $\mathcal{DL}$+*log* semantics ([Rosati, 2006]; see Section 8.2).

Van Belleghem *et al.* [1997] define *open logic programs*, which are combinations of sets of rules and first-order logic formulas; the set of predicate symbols is partitioned into a set of *open* and a set of *closed* predicates. The semantics of the program is the first-order theory consisting of Clark's completion of the closed predicates and the first-order formulas in the open program. They then discuss how description logics can be embedded in such open logic programs and they discuss the correspondence between abduction in open programs and reasoning in description logics.

Several nonmonotonic extensions of description logics have been defined in the literature [Baader and Hollunder, 1995; Donini *et al.*, 1998a; Donini *et al.*, 2002; Bonatti *et al.*, 2006]. These might be further extended to accommodate logic programs by well-known correspondences of the latter to nonmonotonic formalisms. In more detail, extensions of DL semantics with defaults and circumscription have been described by Baader and Hollunder [1995] and Bonatti *et al.* [2006], respectively. Extensions with nonmonotonic modal operators, inspired by the logic MKNF [Lifschitz, 1991], have been described by Donini *et al.* [1998a,2002]. Both works mention a notion of procedural or default rules, which are rules involving description logic concepts. Donini *et al.* [1998a] allow rules of the form $C \Rightarrow D$, where C and D are DL concepts (i.e., unary predicates); such rules are intuitively read "if an individual is proved to be an instance of C, then derive that it

[6] Levesque and Lakemeyer [2000] extend the logic of only knowing by allowing the use of constants and function symbols different from standard names; several ground terms may be associated with one *standard name*, and for any constant symbols a and b with this property, $a = b$ is satisfied.

is also an instance of D." The default rules considered by Donini *et al.* [2002] are a generalization; they are of the form $C_0, not\ C_1, \ldots, not\ C_n \Rightarrow D$, $n \geq 0$, where all C_i and D are DL concepts. Intuitively, "if an individual is proved to be an instance of C_0 and is not proved to be an instance of C_1, ..., or C_n, then derive that it is also an instance of D." The work of Donini *et al.* inspired some more advanced formalisms for combining rules and ontologies, which we consider next.

8.2 Combinations of Rules and Ontologies

Roughly speaking, we can distinguish between three kinds of combinations of rules and ontologies: (1) *uniform combinations* (e.g., CARIN [Levy and Rousset, 1998] and SWRL [Horrocks *et al.*, 2005]), (2) *hybrid combinations* (e.g., dl-programs [Eiter *et al.*, 2008] and $\mathcal{DL}$+*log* [Rosati, 2006]), and (3) *embedding combinations* (e.g., the MKNF combination by Motik and Rosati [2007] and combination based on Quantified Equilibrium Logic [de Bruijn *et al.*, 2007]); for more discussion, see, e.g., [Eiter *et al.*, 2008; de Bruijn *et al.*, 2006]. See [de Bruijn *et al.*, 2008] for embeddings of dl-programs, $\mathcal{DL}$+*log*, and MKNF into FO-AEL.

8.2.1 Uniform Combinations

With uniform combinations we mean combinations of ontologies that are essentially classical first-order theories and of Horn logic formulas that are essentially positive rules. The combined theory, which is the set-theoretic union of the formulas in the ontology and the Horn formulas, is interpreted under the standard first-order logic semantics.

In the CARIN approach [Levy and Rousset, 1998], the ontologies are theories of the description logic $\mathcal{ALCNR}$ and the rules are Datalog rules, i.e., safe positive normal rules as defined in Section 2.2, with the further restriction that predicates which occur in the ontology may not be used in rule heads. Levy and Rousset show that reasoning with these combinations is undecidable in the general case, but becomes decidable when suitably restricting either the ontology or the rules.

SWRL [Horrocks *et al.*, 2005] extends the expressive language OWL DL, which essentially corresponds to description logic $\mathcal{SHOIN}$, with Datalog rules. Horrocks *et al.* [2005] do not impose any restrictions on the shape of the ontology or of the rules, and show that reasoning with the combination is undecidable. Motik *et al.* [2005] show that decidability can be regained by suitably restricting the use of variables in rules to *DL-safeness*: every variable must occur in a non-negated atom in the body whose predicate does not occur in the ontology.

Considering the embedding τ_{HP} (see Definition 4.1 on page 9) we observe that if Φ is a first-order theory and P is a positive logic program, then the combination $\iota_{HP}(\Phi, P) = \Phi \cup \tau_{HP}(P)$ is a first-order theory and is a uniform combination in the sense of CARIN and SWRL.

8.2.2 Hybrid Combinations

Hybrid approaches combine logic programs with nonmonotonic negation (usually, under the stable model semantics or the well-founded semantics) with a description logic knowledge base or, in more abstract terms, theories in first-order logic. The two most prominent such approaches are dl-programs [Eiter *et al.*, 2008] and $\mathcal{DL}$+*log* [Rosati, 2006]. The main difference between the two is the way in which the interaction between the individual components (the logic program and the ontology) is managed. For both approaches, we assume that the ontology component is a DL theory and the logic program is function-free and safe.

In dl-programs, the interoperation between the program and the ontology is achieved by *DL queries*, which are queries to the DL ontology, in the bodies of the rules; in addition, some information from the program may be temporarily added to the ontology for the query. Thus, interoperation is based on the DL entailment relation, resulting in a narrow interface between the program and the ontology. Eiter et al. 2008 show that query answering in dl-programs is decidable as long as reasoning in the individual components (ontology and logic program) is decidable. HEX-programs [Eiter *et al.*, 2005] are a generalization of dl-programs to more general external evaluations that are not limited to queries on DL ontologies.

$\mathcal{DL}$+*log* makes a distinction between ontology and rules predicates; rules predicates may not occur in the ontology, but the ontology predicates may occur in the rules. The combination is interpreted by a single first-order interpretation, but the part of the interpretation concerned with the rules predicates is subject to stability conditions corresponding to the usual restrictions on stable models (cf. Section 2.2 and [Gelfond and Lifschitz, 1988]). Thus, the interoperation is based on single models, resulting in a broad interface between the program and the ontology. Rosati [2005] shows that if the rules are DL-safe and satisfiability checking in the ontology component is decidable, then reasoning with the combination is decidable. Rosati [2006] shows that reasoning is decidable if the problem of containment of conjunctive queries in unions of conjunctive queries is decidable for the underlying DL, provided that the rules are *weakly DL-safe*; this notion dispenses DL-safety for variables that occur only in ontology predicates in rule bodies, which makes it possible to access unnamed individuals in rules. $\mathcal{AL}$-*log* [Donini *et al.*, 1998b] can be seen as a precursor of $\mathcal{DL}$+*log* that considers only positive programs and that allows (unary) ontology predicates only in rule bodies and effectively requires DL-safeness. The differences between and underlying principles of the dl-programs and $\mathcal{DL}$+*log* approaches are discussed in more detail by de Bruijn *et al.* [2006].

Since we did not distinguish between rule and ontology predicates in our embeddings—indeed, in the introduction we claimed this is undesirable—there is no straightforward correspondence between any of the embeddings we considered and the mentioned hybrid approaches. The embeddings we considered in this paper can be used to construct combinations that have a tight integration between the components and that do not have a separation between ontology and rules predicates. In fact, the $\mathcal{DL}$+*log* approach can be reconstructed by an extension of simple combinations $\iota_\chi(\Phi, P) = \Phi \cup P$ with *classical interpretation axioms*, which loosely speaking fix the value of classical predicates for stable expansions; we refer to [de Bruijn *et al.*, 2008] for details.

8.2.3 Embedding Combinations

Motik and Rosati [2007] propose a combination of DL ontologies and nonmonotonic logic programs through an embedding into the bimodal nonmonotonic logic MKNF [Lifschitz, 1991], which uses the modal operators K, which stands for "knowledge", and not, which stands for "negation as failure". The variant of MKNF used by Motik and Rosati employs a standard names assumption similar to [Levesque, 1990]: there is a one-to-one correspondence between the countably many constant symbols in the language and elements in the domains of interpretations (functions symbols are not considered). The equality symbol of first-order logic (=) is embedded using a special binary predicate symbol $\approx$ and the usual congruence axioms (cf. [Fitting, 1996, Chapter 9]) are added. Logic programs are embedded into MKNF using the transformation described by Lifschitz [1994]: a rule r (see (1) on page 4) is embedded as the formula

$$\tau_{MKNF}(r) = \bigwedge_i \mathsf{K} b_i \wedge \bigwedge_j \mathsf{not}\, c_j \supset \bigvee_k \mathsf{K} h_k$$

A classical theory Φ is embedded as a conjunction comprising all the formulas in the theory, preceded by the modal operator K: $\sigma_{MKNF}(\Phi) = \mathsf{K}(\bigwedge \Phi)$. Finally, the combination of the logic program P and the first-order theory Φ is simply $\iota_{MKNF} = \tau_{MKNF}(P) \cup \{\sigma_{MKNF}(\Phi)\}$.

Comparing τ_{MKNF} to the embeddings in Section 4, we can see that it is close in spirit to the embedding $\tau_{EH}^{\vee}$; both embeddings feature modal belief operators in front of positive atoms in both the body and the head of the rule. In fact, it turns out that, when using a variant of FO-AEL with standard names, there is a one-to-one correspondence between the stable expansions of $\tau_{EH}^{\vee-}(P)$ and the MKNF models of $\tau_{MKNF}(P)$ (recall that $\tau_{EH}^{\vee-}(P)$ is $\tau_{EH}^{\vee}(P)$ without the UNA axioms); however, this correspondence does not extend to combinations with FO theories, as shown by de Bruijn *et al.* [2008].

Besides the obvious differences between MKNF and autoepistemic logic—illustrated by the differences between the τ_{MKNF} and $\tau_{EH}^{\vee-}$ embedding functions—there is a difference in the semantics for quantifying-in between the variant of MKNF used by Motik and Rosati [2007] and Konolige's any- and all-name semantics that we used in this paper. Since FO-AEL permits arbitrary interpretations, we needed to utilize UNA axioms. Motik and Rosati employ the standard names assumption and thus do not need such axioms.

Another nonmonotonic logic that has been used for combining ontologies and logic programs is *quantified equilibrium logic* (QEL) [Pearce and Valverde, 2005]. While FO-AEL and MKNF are nonmonotonic modal logics, QEL is based on the nonclassical *logic of here-and-there*, which is an intermediate logic between classical and intuitionistic logic. Negation in QEL is nonmonotonic; however, by axiomatizing the law of the excluded middle (LEM) through $\forall \vec{x}(p(\vec{x}) \vee \neg p(\vec{x}))$, one can enforce that a predicate p is interpreted classically, and negation of this predicates becomes classical. de Bruijn *et al.* [2007] used a slightly generalized version of QEL that does not assume uniqueness of names and includes equality to show that the QEL theory obtained by adding such LEM axioms to the combination $\iota(\Phi, P) = \Phi \cup P$ of a FO theory Φ and a logic program P yields the $\mathcal{DL}$+*log* semantics.

In the discussion so far, we have considered logic programs under the stable model—or answer set—semantics. Another popular semantics for logic programs with negation, particularly in the database context, is the well-founded semantics [Gelder *et al.*, 1991]. Eiter *et al.* [2004] present a well-founded semantics for dl-programs, which is based on a suitable generalization of the notion of an unfounded set. Knorr *et al.* [2008] consider a three-valued variant of the logic MKNF and use it for combining ontologies with normal logic programs under the well-founded semantics. Drabent and Maluszynski [2007] follow the approach of the $\mathcal{DL}$+*log* family and present *hybrid rules* under well-founded semantics, whose implementation integrates a DL-reasoner and a well-founded semantics engine [Drabent *et al.*, 2007]. Extensions of the FO-AEL (e.g., based on [Denecker *et al.*, 2003]) to capture the three-value models of the well-founded semantics is a possible topic for future work.

9 Conclusion

We have defined various embeddings of non-ground programs into first-order autoepistemic logic (FO-AEL) that generalize respective embeddings of propositional logic programs into standard AEL, and we have investigated their semantic properties. We have shown that these embeddings are faithful, in the sense that the stable models (or answer sets) of a given non-ground logic program

P are in one-to-one correspondence to the stable expansions of the embeddings $\tau_\chi^{(\vee)}(P)$ with respect to objective ground atomic formulas. Furthermore, we have analyzed the correspondences between the embeddings at more fine-grained levels, revealing their commonalities and differences.

Our results provide a basis and a stepping stone for the more complex endeavor to combine classical knowledge bases and non-ground logic programs in a uniform logical formalism (which is one of the targets of the Semantic Web architecture), namely the well-known and amply studied formalism of autoepistemic logic.

In this direction, we have investigated correspondences between simple combinations of embeddings of logic programs with FO theories for various classes of logic programs and FO theories. The results of our investigation provide useful insights into the behavior of different embeddings for logic programs with respect to a context, given by a first-order theory, and allows some conclusions about the replaceability of one embedding by another without altering the behavior of the combination. Based on the results in the present paper, more elaborated combinations of logic programs with FO theories are investigated by de Bruijn *et al.* [2008], who show how well-known approaches to combining rules and ontologies in the Semantic Web context can be embedded into FO-AEL, including [Eiter *et al.*, 2008; Rosati, 2006; Motik and Rosati, 2007]. Notably, the $\mathcal{DL}+log$ approach can be embedded into FO-AEL by adding further axioms to the simple combination that we have considered here. The companion paper confirms the value of our results, on which also other novel combinations of logic programs and first-order theories can be built.

Several issues remain for future work. In the present paper, we focused on semantic aspects of embeddings of logic programs, but we did not address computational issues. Since the embeddings are easily computed, they may be exploited to establish decidable fragments of combinations of rules and ontologies, and to craft sound (but possibly incomplete) algorithms for specific reasoning tasks for such combinations. There are several promising starting points for devising algorithms for computing stable expansions and/or autoepistemic consequence in FO-AEL. Niemelä [1992] presents a general procedure for computing stable expansions in FO-AEL without quantifying-in. Levesque and Lakemeyer [2000] present a sound, but incomplete proof theory for the logic of only knowing, which extends FO-AEL with standard names. Finally, Rosati [1999] presents techniques for reasoning with first-order MKNF (with standard names) with a limited form of quantifying-in; the not operator in MKNF is equivalent to $\neg\mathsf{L}$ in autoepistemic logic [Rosati, 1997].

Other issues are extensions of the language used for logic programs. Adding classical negation to the τ_{EB} and τ_{EH} is routine, and has been done in [de Bruijn *et al.*, 2008] for FO-AEL with standard names. Other interesting extensions include nesting [Lifschitz *et al.*, 1999], where the closeness between nesting in logic programs and the logic MKNF suggests that an embedding is straightforward, and aggregates [Faber *et al.*, 2004; Ferraris, 2005; Pelov *et al.*, 2007; Son and Pontelli, 2007].

Furthermore, in the present work we considered embeddings of logic programs interpreted under the stable model semantics, which adopts a two-valued semantics. It would be interesting to consider also embedding of logic programs under many-valued semantics, most importantly under the well-founded semantics, which is a three-valued semantics for logic programs with negation. Three-valued extensions of autoepistemic logic (e.g., [Denecker *et al.*, 2003; Bonatti, 1995; Przymusinski, 1991b]) may be used as a starting point.

A Appendix: Proofs of Sections 5 and 6

This appendix contains the remaining proofs of the results stated in Sections 5 and 6. We remind the reader that we omit the superscript X from $\models^{X}$, $\equiv^{X}_{\gamma}$, Cn^{X}, stable^{X}, etc. if the stated property holds regardless of whether $X = E$ or $X = A$.

Proofs of Section 5

Proof of Proposition 5.4. Let T be a stable expansion of $\tau_{EB}(P)$ or $\tau_{EH}(P)$ (resp., $\tau^{\vee}_{HP}(P)$ or $\tau^{\vee}_{EB}(P)$). We only consider T_{oga} in the following; so, by Theorem 5.3, the choice between τ_{EB} and τ_{EH} (resp., $\tau^{\vee}_{HP}$ and $\tau^{\vee}_{EB}$) is immaterial. Let

$$\Gamma^{\tau} = \{\phi \in \mathcal{L}_g \mid \tau(P) \models_{T_{oga}} \phi\}, \tau \in \{\tau_{EB}, \tau_{EH}, \tau^{\vee}_{HP}, \tau^{\vee}_{EB}\}$$

By Proposition 3.4, $\tau(P)$ has a stable expansion T^{τ} such that $\Gamma^{\tau} = T^{\tau}_{og}$.
First we will show that $T^{\tau_{EB}}_{og} = T^{\tau_{EH}}_{og}$ by establishing

$$T^{\tau}_{og} = \{\phi \in \mathcal{L}_g \mid T_{oga} \models \phi\} \tag{2}$$

$T^{\tau_{EB}}_{og} = T^{\tau_{EH}}_{og}$ follows from this claim, thereby establishing the first part of the proposition, concerning the embeddings τ_{EB} and τ_{EH}.

Every entailed objective ground formula is equivalent to a conjunction of ground clauses $c = l_1 \vee \cdots \vee l_k$, where each l_i is either an atom p_i or a negated atom $\neg p_i$. Clearly, $T_{oga} \models c$ iff $l_i \in T_{oga}$ for some l_i in c. To prove (2), clearly $T^{\tau}_{og} \supseteq \{\phi \in \mathcal{L}_g \mid T_{oga} \models \phi\}$. For the other inclusion, suppose that $\tau(P) \models_{T_{oga}} c$ (hence $c \in T^{\tau}_{og}$), but $T_{oga} \not\models c$. Hence, $l_i \notin T_{oga}$ for every l_i in c. Consider an arbitrary interpretation w such that $w \models_{T_{oga}} \tau(P)$; then, $w \models_{T_{oga}} c$. Let w' result from w by flipping the truth value of the atom p_i of each literal l_i in c such that $w \models_{T_{oga}} l_i$; clearly, $w' \not\models_{T_{oga}} c$. We now show that $w' \models_{T_{oga}} \tau(P)$; this contradicts the assumption $\tau(P) \models_{T_{oga}} c$ and proves (2).

Let $\alpha \in \tau(P)$ be an instance of an axiom that originates from a rule in P. We show that $w' \models_{T_{oga}} \alpha$. Suppose first that some flipped atom p_i occurs in the antecedent α_a of α. If $w' \models_{T_{oga}} \neg p_i$, then clearly $w' \models_{T_{oga}} \alpha$. Otherwise, $w \models_{T_{oga}} \neg p_i$ and thus $\mathsf{L}p_i$, which also occurs in the antecedent of α, is false in w'. Hence, $w' \models_{T_{oga}} \alpha$. Suppose then that α has no flipped p_i in α_a, and that $w' \not\models_{T_{oga}} \alpha$. Hence, $w' \models_{T_{oga}} \alpha_a$ and $w' \models_{T_{oga}} \neg p_i$ for some flipped p_i that occurs in the consequent of α. As $w \models_{T_{oga}} \alpha_a$, it follows $p_i \in T_{oga}$; as $l_i \notin T_{oga}$, we have $l_i = \neg p_i$ and thus $w \models_{T_{oga}} \neg p_i$, which implies $w' \not\models_{T_{oga}} \neg p_i$ by definition; this is a contradiction. This proves $w' \models_{T_{oga}} \alpha$.

As unique names axioms in $\tau(P)$ are clearly satisfied in w', $w' \models_{T_{oga}} \tau(P)$, thereby establishing the claim (2) and thus the first part of the proposition.

For the second part concerning $\tau^{\vee}_{HP}(P)$ and $\tau^{\vee}_{EB}(P)$, we exploit the PIA axioms: thanks to them each objective ground atom α or its negation $\neg\alpha$ is included in the stable expansion T^{τ}, for $\tau \in \{\tau^{\vee}_{HP}, \tau^{\vee}_{EB}\}$. Thus, $T^{\tau}_{og} = \{\phi \in \mathcal{L}_g \mid T_{oga} \models \phi\}$ (by structural induction), from which $T^{\tau^{\vee}_{HP}}_{og} = T^{\tau^{\vee}_{EB}}_{og}$ follows immediately. □

Proof of Proposition 5.5. Let T be a stable expansion of $\tau_{EB}(P)$ (resp., $\tau_{EH}(P)$). By Theorem 5.3 we know that

$$\tau_{EB}(P) \equiv_{oga} \tau_{EH}(P) \tag{3}$$

hence, $T_{oga} = T'_{oga}$ for some stable expansion T' of $\tau_{EH}(P)$ (resp., $\tau_{EB}(P)$).

We proceed as follows. We first claim that (†) given a safe program P and an interpretation w, one can construct a named interpretation w' such that $w \models_{T_{oga}} \tau(P)$ iff $w' \models_{T_{oga}} \tau(P)$, where $\tau \in \{\tau_{EB}, \tau_{EH}\}$. It follows that, for every formula ϕ,

$$\tau(P) \models_{T_{oga}} \phi \text{ iff for every named interpretation } w, \quad w \models_{T_{oga}} \tau(P) \text{ implies } w \models_{T_{oga}} \phi \quad (4)$$

We write the latter—i.e., entailment restricted to named interpretations—as $\tau(P) \models^{\mathcal{N}}_{T_{oga}} \phi$.

We then claim that (‡) for every formula $\phi \in \mathcal{L}_{\mathsf{L}}$, $\tau_{EB}(P) \models^{\mathcal{N}}_{T_{oga}} \phi$ iff $\tau_{EH}(P) \models^{\mathcal{N}}_{T_{oga}} \phi$; combined with (3) and (4), this establishes $\tau_{EB}(P) \equiv \tau_{EH}(P)$. We now proceed to prove the individual claims (†) and (‡).

(†) Let $w = \langle U, \cdot^I \rangle$ be an interpretation. Let the named interpretation $w^{\mathcal{N}} = \langle U^{\mathcal{N}}, \cdot^{I^{\mathcal{N}}} \rangle$ be as follows: $U^{\mathcal{N}} = \{t^I \mid t \in \mathcal{N}\}$; for every n-ary predicate symbol p, $p^{I^{\mathcal{N}}} = p^I \cap (U^{\mathcal{N}})^n$; and for every n-ary function symbol f and $\vec{k} \in (U^{\mathcal{N}})^n$, $f^{I^{\mathcal{N}}}(\vec{k}) = f^I(\vec{k})$.

Consider $\tau_{EB}(P)$, which contains two kinds of formulas: UNA axioms and axioms of the form $(\forall) \bigwedge(b_i \wedge \mathsf{L}b_i) \wedge \bigwedge \neg\mathsf{L}c_j \supset h$. The former are obviously satisfied in $\langle w^{\mathcal{N}}, T_{oga} \rangle$ iff they are satisfied in $\langle w, T_{oga} \rangle$, because they are variable-free. Consider an open formula $\alpha = \bigwedge(b_i \wedge \mathsf{L}b_i) \wedge \bigwedge \neg\mathsf{L}c_j \supset h$ and a variable assignment B of w. Since P is safe, every variable in α occurs in some b_i. Therefore, if B assigns any variable in α to an unnamed individual, $(w, B) \not\models_{T_{oga}} \mathsf{L}b_i$; hence $(w, B) \models_{T_{oga}} \alpha$. If B assigns all variables in α to named individuals, then $(w, B) \models_{T_{oga}} \alpha$ iff $(w^{\mathcal{N}}, B) \models_{T_{oga}} \alpha$, by construction of $w^{\mathcal{N}}$. The proof for the case of $\tau_{EH}(P)$ is analogous; the antecedents of the implications in $\tau_{EB}(P)$ and $\tau_{EH}(P)$ are the same. This proves the claim (†).

It remains to prove (‡). Let T be a stable expansion of $\tau_{EB}(P)$.

(⇒) By the shape of the formulas in $\tau_{EB}(P)$ and $\tau_{EH}(P)$, for every interpretation w clearly $w \models^{\mathcal{N}}_{T_{oga}} \tau_{EH}(P)$ implies $w \models^{\mathcal{N}}_{T_{oga}} \tau_{EB}(P)$. Hence, for every formula $\phi \in \mathcal{L}_{\mathsf{L}}$ it holds that $\tau_{EB}(P) \models^{\mathcal{N}}_{T_{oga}} \phi$ implies $\tau_{EH}(P) \models^{\mathcal{N}}_{T_{oga}} \phi$.

(⇐) We proceed by contradiction. Suppose that $\tau_{EH}(P) \models^{\mathcal{N}}_{T_{oga}} \phi$, but $\tau_{EB}(P) \not\models^{\mathcal{N}}_{T_{oga}} \phi$. Hence there must be a named interpretation w such that $w \models_{T_{oga}} \tau_{EB}(P)$ and $w \not\models_{T_{oga}} \phi$. We will show that $w \models_{T_{oga}} \tau_{EH}(P)$, which contradicts the assumption.

The UNA axioms in $\tau_{EH}(P)$ are obviously satisfied in w, since they are also in $\tau_{EB}(P)$. Consider a formula $(\forall)\, \alpha \supset h \wedge \mathsf{L}h \in \tau_{EH}(P)$ that is not satisfied in $\langle w, T_{oga} \rangle$, where $\alpha = \bigwedge(b_i \wedge \mathsf{L}b_i) \wedge \bigwedge \neg\mathsf{L}c_j$. Since $w \models_{T_{oga}} (\forall)\, \alpha \supset h \in \tau_{EB}(P)$, we have that, for some variable assignment B, $(w, B) \models_{T_{oga}} \alpha$ and $(w, B) \not\models^{\mathrm{E}}_{T_{oga}} \mathsf{L}h$ (resp., $(w, B) \not\models^{\mathrm{A}}_{T_{oga}} \mathsf{L}h$); hence, $h\beta \notin T_{oga}$ for all (resp., for some) variable substitution(s) β associated with B. However, since $(w, B) \models_{T_{oga}} \alpha$, and thus $(w, B) \models^{\mathrm{E}}_{T_{oga}} \mathsf{L}b_i$ and $(w, B) \not\models^{\mathrm{E}}_{T_{oga}} \mathsf{L}c_j$ (resp., $(w, B) \models^{\mathrm{A}}_{T_{oga}} \mathsf{L}b_i$ and $(w, B) \not\models^{\mathrm{A}}_{T_{oga}} \mathsf{L}c_j$), it follows that $b_i\beta_{b_i} \in T_{oga}$ and $c_j\beta_{c_j} \notin T_{oga}$[7] for some/all (resp., for all/some) variable substitutions β_{b_i}, β_{c_j} associated with B. Now, since T is a stable expansion of $\tau_{EB}(P)$, from (4) and the UNA axioms contained in $\tau_{EB}(P)$ it follows that $h\beta \in T_{oga}$ for some (resp., for all) variable substitution(s) β associated with B. Therefore, $w \models_{T_{oga}} (\forall)\, \alpha \supset h \wedge \mathsf{L}h$ and $w \models_{T_{oga}} \tau_{EH}(P)$, establishing the desired contradiction. This proves that $\tau_{EH}(P) \models^{\mathcal{N}}_{T_{oga}} \phi$ implies $\tau_{EB}(P) \models^{\mathcal{N}}_{T_{oga}} \phi$.

The case of T being a stable expansion of $\tau_{EH}(P)$ is analogous. □

[7] Observe that $b_i\beta_{b_i}$ and $c_j\beta_{c_j}$ are well-defined, because all individuals in w are named.

Proof of Proposition 5.6. By the PIA axioms $b_i \supset \mathsf{L}b_i$ we can eliminate the modal atoms of the form $\mathsf{L}b_i$ from the antecedents of the axioms in $\tau^{\vee}_{EB}(P)$ that originate from rules in P. The remaining theory is the same as $\tau^{\vee}_{HP}(P)$ and thus the stable expansions correspond. □

Proof of Proposition 5.7. As $\tau_{EB}(P)$ and $\tau_{EH}(P)$ differ only in that the latter has a conjunction $h \wedge \mathsf{L}h$ in the consequent of embedded rules $\tau_{EH}(r)$ while the former simply has h, the first implication is immediate. The second implication is argued similarly; each $\tau_{EB}(r)$ has a stronger antecedent than $\tau_{HP}(r)$. □

Proof of Proposition 5.8. Similarly as $\tau^{\vee}_{EB}(P)$ and $\tau^{\vee}_{HP}(P)$ in Proposition 5.7, $\tau^{\vee}_{EB}(P)$ and $\tau^{\vee}_{HP}(P)$ only differ in that the embedding $\tau^{\vee}_{EB}(r)$ of each rule r has a stronger antecedent than the corresponding $\tau^{\vee}_{HP}(P)$, from which the first part follows.

For the second part, as in Proposition 5.5 it suffices to consider the case of named individuals. Consider any interpretation $\langle w, T\rangle$ such that w is named and $w \models_T \tau^{\vee}_{EB}(P)$. By the PIA axioms, for every variable assignment B, associated variable substitution β, and $a \in \mathcal{L}_{oga}$, we have that $(w, B) \models_T a$ implies $w \models_T \mathsf{L}a\beta$. Hence, $w \models_T \tau^{\vee}_{EH}(P)$. □

Proof of Lemma 5.9. Recall that a formula $\phi \in \mathcal{L}_{\mathsf{L}}$ is an autoepistemic consequence of a theory $\Phi \subseteq \mathcal{L}_{\mathsf{L}}$ if ϕ is included in every stable expansion of Φ. Recall also that, by Theorem 5.3, the stable expansions of all embeddings $\tau^{(\vee)}_{\chi}(P)$ correspond with respect to $\mathcal{L}_{oga}$. From this and the fact that, by Proposition 3.4, every stable expansion T is determined by T_{oga} and $\tau^{(\vee)}_{\chi}(P)$, we can conclude $C^{(\vee)}_{\chi} \subseteq C^{(\vee)}_{\gamma}$ if $w \models_{T_{oga}} \tau^{(\vee)}_{\gamma}(P)$ implies $w \models_{T_{oga}} \tau^{(\vee)}_{\chi}(P)$ for every interpretation $\langle w, T_{oga}\rangle$.

- $C_{EH} \longrightarrow C^{\vee}_{EH}$, $C^{\vee}_{EH} \longrightarrow C_{EH}$: follow from the fact that, if $P \in n\mathcal{LP}$, $\tau_{EH}(P)$ and $\tau^{\vee}_{EH}(P)$ coincide.
- $C_{EH} \dashrightarrow C_{EB}$: follows from Proposition 5.5.
- $C_{EB} \longrightarrow C_{EH}$, $C_{EB} \rightarrow C_{HP}$: follows from Proposition 5.7.
- $C_{HP} \longrightarrow C^{\vee}_{HP}$, $C_{EB} \rightarrow C^{\vee}_{EB}$: if $P \in n\mathcal{LP}$, then clearly $\tau_{HP}(P) \subseteq \tau^{\vee}_{HP}(P)$, so $w \models_{T_{oga}} \tau^{\vee}_{HP}(P)$ implies $w \models_{T_{oga}} \tau_{HP}(P)$. Likewise for $\tau_{EB}(P)$ and $\tau^{\vee}_{EB}(P)$.
- $C^{\vee}_{EB} \longrightarrow C^{\vee}_{HP}$: $C^{\vee}_{EB} \subseteq C^{\vee}_{HP}$ follows from Proposition 5.8.
- $C^{\vee}_{EH} \dashrightarrow C^{\vee}_{EB}$: follows from Proposition 5.8.

□

Proof of Lemma 5.10.

- $C^{\vee}_{EH} \not\subseteq C^{\vee}_{HP}$: Consider $P = \{p(x);\ q(c) \mid c \in \mathcal{N}\}$, which is not safe. Then $\tau^{\vee}_{EH}(P)$ has a single stable expansion T, and since $\forall x.p(x) \wedge \mathsf{L}p(x)$ is in T, in every model of T all individuals must be named. Hence T, and thus $C^{\vee}_{EH}$, contains $\forall x.q(x)$. Also $\tau^{\vee}_{HP}(P)$ has a single stable expansion, but individuals may be unnamed in its models, and so $\forall x.q(x) \notin C^{\vee}_{HP}$ holds.
- $C^{\vee}_{EB} \not\subseteq C_{HP}$: Consider $P = \{b \leftarrow a\}$, which is clearly safe. Then $\tau^{\vee}_{EB}(P)$ has a single stable expansion, and $a \notin T$, which implies $\neg\mathsf{L}a \in T$. As T contains the contrapositive $\neg\mathsf{L}a \supset \neg a$ of the PIA axiom $a \supset \mathsf{L}a$, T contains also $\neg a$. Consequently, $\neg a \in C^{\vee}_{EB}$. Since $\tau_{HP}(P)$ contains no PIA axioms, $\neg a \notin C_{HP}$.

- $C_{HP} \not\subseteq C^{\vee}_{EB}$: Consider the safe program $P = \{q(x) \leftarrow p(x)\}$. The single stable expansion of $\tau_{HP}(P)$ contains $\forall x(p(x) \supset q(x))$, while the single stable expansion of $\tau^{\vee}_{EB}(P)$ does not. Hence, $\forall x(p(x) \supset q(x)) \in C_{HP}$, but $\forall x(p(x) \supset q(x)) \notin C^{\vee}_{EB}$.

□

Proofs of Section 6

Proof of Proposition 6.4. We show that, given an FO interpretation w and a stable expansion T of $\iota_{\chi}(\Phi, P)$, with $x \in \{EB, EH\}$, $w \models^{\mathrm{E}}_{T} \iota_{EB}(\Phi, P)$ iff $w \models^{\mathrm{E}}_{T} \iota_{EH}(\Phi, P)$. Recall that $\iota_{EB}(\Phi, P)$ and $\iota_{EH}(\Phi, P)$ only differ in embedding of rules r from P, which are mapped to $\tau_{EB}(r) = \alpha \supset h$ in $\iota_{EB}(\Phi, P)$ and to

$$\tau_{EH}(r) = \alpha \supset (h \wedge \mathsf{L}h) \tag{5}$$

in $\iota_{EH}(\Phi, P)$, where $\alpha = \bigwedge(b_i \wedge \mathsf{L}b_i) \wedge \bigwedge\neg\mathsf{L}c_j$. Consider first $x = EB$.

($\Leftarrow$) Clearly, every model of $\iota_{EH}(\Phi, P)$ is a model of $\iota_{EB}(\Phi, P)$.

($\Rightarrow$) Suppose $w \models^{\mathrm{E}}_{T} \iota_{EB}(\Phi, P)$. Consider an axiom (5) such that $(w, B) \models^{\mathrm{E}}_{T} \alpha$. It follows that $b_i \in T$ and $c_j \notin T$, and, since $\alpha \supset h$ is in $\iota_{EB}(\Phi, P)$ and T is a stable expansion thereof (and consequently closed under first-order entailment), that $h \in T$. This proves $w \models^{\mathrm{E}}_{T} \iota_{EH}(\Phi, P)$.

The case $x = EH$, i.e., T is a stable expansion of $\iota_{EH}(\Phi, P)$, is analogous. □

Proof of Proposition 6.5. Follows from the proof of Proposition 5.6. □

Proof of Proposition 6.6. We first prove the result for $\Phi \in \mathcal{H}orn$, and then extend the result to the case $\Phi \in g\mathcal{H}orn$. We use the following Lemma.

Lemma A.1. *Let J be a (countable) index set, let $w_j = \langle U_j, \cdot^{I_j}\rangle$, for $j \in J$, be models of an FO theory Φ in which all individuals are named, and let $w = \langle U, \cdot^{I}\rangle$ be the "intersection" of all w_j, defined as follows:*

- $U = \{[t^{I_1}, \ldots, t^{I_j}, \ldots] \mid t \in \mathcal{N}\}$ *(recall that $\mathcal{N}$ is a set of ground terms);*
- *for n-ary function symbols f, $f^{I}(u_1, \ldots, u_n) = [v_1, \ldots, v_j, \ldots]$ such that $v_j = f^{I_j}(u_{1,j}, \ldots, u_{n,j})$, for all $j \in J$, and $u_i = [u_{i,1}, \ldots, u_{i,j}, \ldots]$, for $i = 1, \ldots, n$;*
- *for n-ary predicate symbols p, $\langle u_1, \ldots, u_n\rangle \in p^{I}$ iff $\langle u_{1,j}, \ldots, u_{n,j}\rangle \in p^{I_j}$, for all $j \in J$, where $u_i = [u_{i,1}, \ldots, u_{i,j}, \ldots]$, for $i = 1, \ldots, n$.*

Then, if Φ is Horn, w is a model of Φ.

This lemma generalizes the well-known intersection property of Herbrand models of Horn theories. Observe that for ground terms $t, t' \in \mathcal{N}$ we have $w \models t = t'$ iff $w_j \models t = t'$ for every $j \in J$.

To show that $\iota^{\vee}_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{EH}(\Phi, P)$, it is by Proposition 3.4 sufficient to show that given any stable expansion T of (1) $\iota^{\vee}_{EH}(\Phi, P)$, resp. (2) $\iota^{\vee}_{HP}(\Phi, P)$, for each $\alpha \in \mathcal{L}_{ga}$ it holds that $\iota^{\vee}_{HP}(\Phi, P) \models^{\mathrm{E}}_{T} \alpha$ iff $\iota^{\vee}_{EH}(\Phi, P) \models^{\mathrm{E}}_{T} \alpha$. We leave out the parameters (Φ, P) in the remainder.

The *if* direction always holds, as every model $\langle w, T\rangle$ of $\iota^{\vee}_{HP}$ is also a model of $\iota^{\vee}_{EH}$: both theories contain Φ and the same UNA axioms; by the PIA axioms we have that (*) for any ground atom α

occurring in P, $w \models \alpha$ iff $\alpha \in T$ and, consequently, for every rule r (which must be ground, as P is ground), $w \models^{\mathrm{E}}_{T} \tau^{\vee}_{HP}(r)$ iff $w \models^{\mathrm{E}}_{T} \tau^{\vee}_{EH}(r)$. Thus, it remains to consider the *only-if* direction.

Case 1) Let T be a stable expansion of $\iota^{\vee}_{EH}$. Suppose $\iota^{\vee}_{HP} \models^{\mathrm{E}}_{T} \alpha$, but $\iota^{\vee}_{EH} \not\models_{T} \alpha$. Let $\{\alpha_j \mid j \in J\} = \{\alpha' \in \mathcal{L}_{ga} \mid \iota^{\vee}_{EH} \not\models_{T} \alpha'\}$ and let for each α_j, w_j be an interpretation such that $w_j \models^{\mathrm{E}}_{T} \iota^{\vee}_{EH} \cup \{\neg\alpha_j\}$; such a w_j must exist and, w.l.o.g., w_j has only named individuals (as $\iota^{\vee}_{EH}$ is a universal theory). Note that $\alpha = \alpha_j$ for some index $j \in J$.

Let w^* be the "intersection" of all w_j, constructed as in Lemma A.1. We first establish $w^* \models^{\mathrm{E}}_{T} \iota^{\vee}_{EH}$: by the lemma, $w^* \models^{\mathrm{E}}_{T} \Phi$; w^* does not satisfy equalities that are not satisfied by the w_j, satisfaction of the UNA axioms follows; finally, $\langle w^*, T\rangle$ does not satisfy antecedents of any $\tau^{\vee}_{EH}(r)$, for $r \in P$, that are not satisfied by some $\langle w_j, T\rangle$, and so $w^* \models^{\mathrm{E}}_{T} \tau^{\vee}_{EH}(r)$.

We now establish $w^* \models^{\mathrm{E}}_{T} \iota^{\vee}_{HP}$. As T is a stable expansion of $\iota^{\vee}_{EH}$, every $\alpha' \in \mathcal{L}_{ga}$ such that $\alpha' \in T$ is satisfied in w^*. The PIA axioms ($\alpha' \supset \mathsf{L}\alpha'$) are satisfied: if $\alpha' \notin T$ (i.e., $\mathsf{L}\alpha'$ is false), then α' is false in w^*, as for each $\alpha' \notin T$, there is a w_j such that $w_j \not\models \alpha'$. Satisfaction of the rule embeddings $\tau^{\vee}_{EH}(r)$ follows by the same argument as in the *if* direction. Therefore, $w^* \models^{\mathrm{E}}_{T} \iota^{\vee}_{HP}$. However, as $\alpha = \alpha_j$ for some j, $w^* \not\models \alpha$. This implies $\iota^{\vee}_{HP} \not\models_{T} \alpha$, a contradiction.

Case 2) Let T be a stable expansion of $\iota^{\vee}_{HP}$. Suppose again that $\iota^{\vee}_{HP} \models^{\mathrm{E}}_{T} \alpha$, but $\iota^{\vee}_{EH} \not\models^{\mathrm{E}}_{T} \alpha$. Hence there is some model $\langle w', T\rangle$ of $\iota^{\vee}_{EH} \cup \{\neg\alpha\}$ such that w' is named. As $w' \not\models_{T} \iota^{\vee}_{HP}$ must hold, the difference of $\iota^{\vee}_{HP}$ and $\iota^{\vee}_{EH}$ (which is only in rule embeddings) together with groundness of P implies that α must occur in some rule head of P. Let $\langle w'', T\rangle$ be a model of $\iota^{\vee}_{HP}$ such that w'' is named. Such a model must exist: if $\iota^{\vee}_{HP}$ would have no such model, then $\mathcal{L}_{ga} \subseteq T$ would hold as T is a stable expansion of $\iota^{\vee}_{HP}$. As P is ground, we can replace all non-negated modal atoms in PIA axioms and in rule embeddings $\tau^{\vee}_{EH}(r)$ with $\top$. Modulo simplifications, we then obtain $\iota^{\vee}_{HP}$. But then $w' \models^{\mathrm{E}}_{T} \iota^{\vee}_{HP}$, which is a contradiction.

Let w be the "intersection" of w' and w'' as in Lemma A.1. We claim that $w \models^{\mathrm{E}}_{T} \iota^{\vee}_{HP}$: $\langle w, T\rangle$ satisfies Φ, the UNA, and the PIA axioms by argument analogous to the case (1). Satisfaction of the rule embeddings $\tau^{\vee}_{HP}(r)$ follows from the property (*).

For establishing $\iota_{HP}(\Phi, P) = \iota_{EH}(\Phi, P)$ for $P \in gn\mathcal{LP}$, the argument is similar but slightly diverging, as there are no PIA axioms. Even if property (*) does not hold, in (2) some w'' exists such that $\langle w'', T\rangle$ satisfies $\bigwedge_i b_i \wedge \mathsf{L}h_i$ resp. $h_{k'} \wedge \mathsf{L}h_{k'}$ in an embedded rule $\tau_{EH}(r)$ iff $w'' \models \bigwedge_i b_i$ resp. $w'' \models h_{k'}$ as EH is, modulo modal atoms $\mathsf{L}\alpha$—which can be eliminated from it, see below—an FO Horn theory, and so Lemma A.1 applies.

This proves the result for $\Phi \in \mathcal{H}orn$. To generalize it to $\Phi \in g\mathcal{H}orn$, we exploit standard skolemization (see, e.g., [Fitting, 1996]), by which we obtain from Φ an equi-satisfiable FO Horn theory Φ'. We show that skolemization of Φ commutes with its combination with the ground program P.

In detail, as already shown, $\iota^{\vee}_{EH}(\Phi', P) \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{HP}(\Phi', P)$ and $\iota_{EH}(\Phi', P) \equiv^{\mathrm{E}}_{oga} \iota_{HP}(\Phi', P)$ hold. Thus it remains to show for each $\chi \in \{^{\vee}_{HP}, ^{\vee}_{EH}, _{HP}, _{EH}\}$ that for every stable expansion T of $\iota\chi(\Phi, P) \subseteq \mathcal{L}_{\mathsf{L}}$, there exists some stable expansion T' of $\iota\chi(\Phi', P) \subseteq \mathcal{L}'_{\mathsf{L}}$ such that $T \cap \mathcal{L}_{ga} = T' \cap \mathcal{L}_{ga}$, and vice versa.

We transform $\iota\chi(\Phi, P)$ to an FO theory $J_\chi(T)$ as follows: we replace each modal atom $\mathsf{L}\alpha$ in $\iota\chi(\Phi, P)$ with $\top$ if $\alpha \in T$, and with $\perp$ if $\alpha \notin T$. Intuitively, in $J_\chi(T)$ we fix the value of $\mathsf{L}\alpha$ according to T. Since P is ground, $J_\chi(T)$ is then indeed an FO theory. Clearly, for every $\alpha \in \mathcal{L}_{ga}$, $\iota\chi(\Phi, P) \models^{\mathrm{E}}_{T} \alpha$ iff $J_\chi(T) \models \alpha$.

We construct $J'_\chi(T)$ from $\iota\chi(\Phi', P)$ in the same way, and obtain that for every $\alpha \in \mathcal{L}_{ga}$,

$\iota\chi(\Phi', P) \models^{\mathrm{E}}_{T} \alpha$ iff $J'_\chi(T) \models \alpha$. Now observe that $J'_\chi(T)$ is a skolemization of $J_\chi(T)$, and thus equi-satisfiable; the same holds for $J_\chi(T) \cup \{\neg\alpha\}$ and $J'_\chi(T) \cup \{\neg\alpha\}$, for arbitrary $\alpha \in \mathcal{L}_{ga}$, and thus $J_\chi(T) \models \alpha$ iff $J'_\chi(T) \models \alpha$; equivalently,

$$\iota\chi(\Phi, P) \models^{\mathrm{E}}_{T_{oga}} \alpha \text{ iff } \iota\chi(\Phi', P) \models^{\mathrm{E}}_{T_{oga}} \alpha.$$

From Proposition 3.4, we can conclude that $\iota\chi(\Phi', P)$ has some stable expansion T' such that $T_{oga} = T' \cap \mathcal{L}_{ga}$. (Note that T' might contain objective ground atoms that are not in T_{oga}, due to the Skolem functions.) The converse is obtained analogously. □

Proof of Proposition 6.7. Follows from the proof of Proposition 5.4 and the fact that there is a PIA axiom for each propositional symbol that occurs in any formula in $\tau^{\vee}_{HP}(P)$ or $\tau^{\vee}_{EB}(P)$. □

Proof of Proposition 6.8. To show the first part, let T be a stable expansion of $\iota_{EB}(\Phi, P)$. We show that for every $\phi \in \mathcal{L}_g$, $\iota_{EB}(\Phi, P) \models^{\mathrm{E}}_{T} \phi$ iff $\iota_{EH}(\Phi, P) \models^{\mathrm{E}}_{T} \phi$.

($\Rightarrow$) Trivial, as every model of $\iota_{EH}(\Phi, P)$ is a model of $\iota_{EB}(\Phi, P)$ (see also Proposition 5.7).

($\Leftarrow$) Let $\phi \in \mathcal{L}_g$. Suppose $\iota_{EH}(\Phi, P) \models^{\mathrm{E}}_{T} \phi$, but $\iota_{EB}(\Phi, P) \not\models_T \phi$, and so there is a $w = \langle U, \cdot^I \rangle$ such that $w \models^{\mathrm{E}}_{T} \iota_{EB}(\Phi, P)$ and $w \not\models_T \phi$. Without loss of generality, we may assume that w is named by straightforward adaptation of Lemma 4.5; note that individuals do not affect satisfaction of propositional formulas, and unnamed individuals to not affect satisfaction of ground formulas. By analogous argument to the proof of the *if* part of (‡) in Proposition 5.5, we can conclude $w \models^{\mathrm{E}}_{T} \iota_{EH}(\Phi, P)$, a contradiction.

The case of T being a stable expansion of $\iota_{EH}(\Phi, P)$ is analogous. The proof of the second part is obtained from the proof of Proposition 5.5 by replacing $\tau_\chi(P)$ with $\iota_\chi(\Phi, P)$ and the observation that the domain of interpretation U has no bearing on the satisfaction of propositional formulas. □

Proof of Proposition 6.9.

(1): Consider the program $P = \{r \leftarrow p; r \leftarrow q; s \leftarrow s\}$ and the theory $\Phi = \{p \vee q, \neg s \supset t\}$. The combinations $\iota\chi(\Phi, P)$, with $\chi \in \{HP, EB, \overset{\vee}{HP}, \overset{\vee}{EH}\}$, all have one stable expansion $T\chi$. Now T_{HP} contains r while T_{EB} does not, due to the modal operators in the antecedents in $\tau_{EB}(P)$. We have that s is not included in any $T\chi$. So, by the PIA axioms, $\neg s$ and, consequently, t are included in $T^{\vee}_{HP}$. Neither is included in T_{HP} or $T^{\vee}_{EH}$.

(2): Consider $P = \{q \leftarrow p\}$. The stable expansions of both $\iota_{HP}(\emptyset, P)$ and $\iota^{\vee}_{HP}(\emptyset, P)$ contain $p \supset q$, whereas the expansions of $\iota_{EB}(\emptyset, P)$ and $\iota^{\vee}_{EH}(\emptyset, P)$ do not; see also Example 4.2. In addition, $\neg p$ and $\neg q$ are in the expansion of $\iota^{\vee}_{HP}(\emptyset, P)$, due to the PIA axioms, but not in the expansions of $\iota_{HP}(\emptyset, P)$ and $\iota^{\vee}_{EH}(\emptyset, P)$.

(3): Consider $P = \{p(a); q(x) \leftarrow p(x)\}$. The single expansion of $\tau^{\vee}_{HP}(P)$ contains $\forall x.p(x) \supset q(x)$, whereas the expansion of $\tau^{\vee}_{EB}(P)$ does not.

(4): Consider $P = \{q' \leftarrow r(x), \mathit{not}\ p(x);\ s \leftarrow q;\ s \leftarrow q';\ r(a)\}$ and the theory $\Phi = \{p(a), p(b) \supset q, r(b)\}$. We have

$$\begin{aligned}\iota_{HP}(\Phi, P) = \{\ & p(a), r(b), r(a),\\ & p(b) \supset q,\ q \supset s,\ q' \supset s\\ & \forall x.r(x) \wedge \neg \mathsf{L}p(x) \supset q'\ \}\end{aligned}$$

A given model $\langle w, T\rangle$ of τ_{HP} either satisfies $a = b$ or $a \neq b$. In the former case, $p(b)$, and thus q and s, must be satisfied. In the latter case, assuming that $p(a) \in T$ and $p(b) \notin T$, $\mathsf{L}p(x)$ is not satisfied in case x is assigned to the individual denoted by b. Consequently, q' and s are satisfied. One can then verify that the single consistent stable expansion T of $\iota_{HP}(\Phi, P)$ is such that $T_{oga} = \{r(a), r(b), p(a), s\}$.

The combination $\iota_{EB}(\Phi, P)$ is like $\iota_{HP}(\Phi, P)$, except that the axioms for s are:

$$\mathsf{L}q \wedge q \supset s,\ \mathsf{L}q' \wedge q' \supset s$$

The single stable expansion of $\iota_{EB}(\Phi, P)$ does not contain s, as neither q nor q' can be derived. So, $q, q' \notin T$, and thus s cannot be derived. This disproves $\iota_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota_{EB}(\Phi, P)$.

Consider the axiom corresponding to the first rule in P in $\iota_{EH}(\Phi, P)$:

$$\forall x.\mathsf{L}r(x) \wedge r(x) \wedge \neg\mathsf{L}p(x) \supset \mathsf{L}q' \wedge q'$$

Let T' be a stable expansion of $\iota_{EH}(\Phi, P)$ such that $q', p(b) \notin T'$; clearly, $r(b), p(a) \in T'$. Consequently, $\mathsf{L}r(x) \wedge r(x) \wedge \neg\mathsf{L}p(x)$ is not satisfied in any model $\langle w, T'\rangle$ for any variable assignment B. Consider B such that $x^B = b^w$. Clearly, $(w, B) \models^{\mathrm{E}}_{T'} \mathsf{L}r(x) \wedge r(x)$, and thus it must be the case that $(w, B) \models^{\mathrm{E}}_{T'} \mathsf{L}p(x)$. Since $p(b) \notin T'$, it must be the case that $w \models a = b$. But then, $\iota_{EH}(\Phi, P) \models^{\mathrm{E}}_{T'} a = b$, and so $p(b) \in T'$, a contradiction. As the expansions of $\iota_{EB}(\Phi, P)$ and $\iota_{HP}(\Phi, P)$ contain neither q' nor $p(b)$, this disproves $\iota_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota_{EH}(\Phi, P)$ and $\iota_{EB}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota_{EH}(\Phi, P)$

The combinations $\iota^{\vee}_{HP}(\Phi, P)$ and $\iota^{\vee}_{EB}(\Phi, P)$ both have a stable expansion that contains q' and not q or $p(b)$: since q is not in the expansion, $\neg q$, and also $\neg p(b)$, must be satisfied in every model, by the PIA axioms. Consequently, every model must satisfy $a \neq b$ and thus also q' and s, which are hence included in the expansion. This disproves $\iota\chi(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota\xi(\Phi, P)$ for $(\chi, \xi) = ({}_{HP}, {}^{\vee}_{HP})$, $({}_{HP}, {}^{\vee}_{EB})$, $({}_{EB}, {}^{\vee}_{EB})$, and $({}_{EB}, {}^{\vee}_{HP})$.

Consider the following modification of (Φ, P): in Φ replace q with $q(b)$ and in P replace q and q' with $q(x)$ and $q'(x)$. Observe that the embeddings $\tau^{\vee}_{EB}(P)$ and $\tau^{\vee}_{HP}(P)$ do not contain PIA axioms involving $q(b)$ or $q'(b)$. By an argument analogous to the case of ι_{HP} and ι_{EB}, we obtain that s is included in the single stable expansion of $\iota^{\vee}_{HP}(\Phi, P)$, but not in the single expansion of $\iota^{\vee}_{EB}(\Phi, P)$ and that neither expansion contains $q'(b)$ or $p(b)$. This disproves $\iota^{\vee}_{HP}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{EB}(\Phi, P)$.

One can then argue, analogously to the case of the original combination above, that there is no stable expansion T of $\iota_{EH}(\Phi, P)$ that contains neither $q'(b)$ nor $p(b)$. This disproves $\iota_{EH}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{EB}(\Phi, P)$ and $\iota_{EH}(\Phi, P) \equiv^{\mathrm{E}}_{oga} \iota^{\vee}_{HP}(\Phi, P)$.

(5): Reconsider $P = \{p(a);\ p(x);\ q(x) \leftarrow p(x)\}$ following Proposition 5.5. □

Proof of Proposition 6.10. We reduce the first part of the proposition to the second part through skolemization.

Let $P \in sn\mathcal{LP}$, $\Phi \in g\mathcal{H}orn$ and let $\Sigma = \Sigma_\Phi \cup \Sigma_P$ be the combined signature. To ease argumentation, we use a fresh unary predicate d as a domain predicate in P: let P' result from P by adding for each variable x occurring in rule $r \in P$ the atom $d(x)$ to the body of r, and let $\Phi' = \Phi \cup \{d(t) \mid t \in \mathcal{N}_\Sigma\}$.

As an easy lemma, the stable expansions of $\iota_\chi(\Phi, P)$ and $\iota_\chi(\Phi', P')$ ($\chi \in \{EB, EH\}$) correspond and agree on formulas in $\mathcal{L} = \mathcal{L}_\Sigma$. Indeed, as every rule $r \in P$ is safe, every variable x in $\tau_\chi(r)$ occurs in a positive modal atom $\mathsf{L}b_i$ in the antecedent; thus, given a model $\langle w, T\rangle$ of $\iota_\chi(\Phi, P)$ and a variable assignment B, $(w, B) \models^{\mathrm{E}}_{\Gamma} \mathsf{L}b_i$ holds only if x^B is a named individual, and thus $(w', B) \models d(x)$, where w' is obtained from w by setting the extension of d to the set of named individuals. On the other hand, as d occurs in $\iota_\chi(\Phi', P')$ only in facts and positively in antecedents of embedded rules $\tau_\chi(r)$, we can, in every model $\langle w, T\rangle$ of $\iota_\chi(\Phi', P')$, decrease the extension of d to the set of named individuals without compromising satisfaction. It follows that if T is a stable expansion of $\iota_\chi(\Phi, P)$, then there exists a stable expansion S of $\iota_\chi(\Phi', P')$ such that $S_{oga} = T_{oga} \cup \{d(t) \mid t \in \mathcal{N}_\Sigma\}$, and vice versa.

Now let Φ'' be a skolemization of Φ', let Σ' (resp., Σ'') be the combined signature of P' and Φ' (resp., P' and Φ''), and let $\mathcal{L}'$ (resp., $\mathcal{L}''$) be the resulting FO language. We now establish correspondence between the stable expansions of $\Delta' = \iota_\chi(\Phi', P')$ and $\Delta'' = \iota_\chi(\Phi'', P')$ with respect to $\mathcal{L}'_{ga}$. As $(\Phi'', P') \in \mathcal{H}orn \times n\mathcal{LP}$ and P' is positive, the first part of the proposition then follows from second.

($\Leftarrow$) Let T'' be a stable expansion of Δ''. We note an important fact: (†) $d(t') \in T''$ such that $t' \notin \mathcal{N}_{\Sigma'}$ (i.e., t' is a Skolem term) implies that $t = t' \in T''$ for some $t \in \mathcal{N}_{\Sigma'}$ (as d occurs in Δ'' only in ground facts and in antecedents of embedded rules). We show that, for every $\alpha \in \mathcal{L}'_{ga}$, it holds that $\Delta' \models^{\mathrm{E}}_{T''_{oga}} \alpha$ iff $\alpha \in T''_{oga}$ (i.e., $\Delta'' \models^{\mathrm{E}}_{T''_{oga}} \alpha$); by Proposition 3.4 this proves that there is a stable expansion T' of Δ' such that $T'_{oga} = T'' \cap \mathcal{L}'_{ga}$.

Suppose first that $w \models^{\mathrm{E}}_{T''_{oga}} \Delta' \cup \{\neg\alpha\}$. Then we can extend w by suitable interpreting the Skolem functions, thereby obtaining some w' such that $w' \models \Phi''$. Clearly, $w' \models^{\mathrm{E}}_{T''_{oga}} \mathrm{UNA}_{\Sigma_P}$. Now suppose that for some variable assignment B and some variable x that occurs in $\tau_\chi(r)$ ($r \in P'$), $(w', B) \models^{\mathrm{E}}_{T''_{oga}} \mathsf{L}d(x)$. Then, by (†), x^B is named by some $t \in \mathcal{N}_{\Sigma'}$ and hence $(w, B) \models^{\mathrm{E}}_{T''_{oga}} \mathsf{L}d(x)$ and the antecedent of $\tau_\chi(r)$ is satisfied by $\langle w, T''_{oga}\rangle$ with respect to B. It follows that $w' \models^{\mathrm{E}}_{T''_{oga}} \tau_\chi(r)$ and thus $w' \models^{\mathrm{E}}_{T''_{oga}} \Delta'' \cup \{\neg\alpha\}$.

Conversely, suppose $w' \models^{\mathrm{E}}_{T''_{oga}} \Delta'' \cup \{\neg\alpha\}$. Let w be the restriction of w' to Σ'. Then clearly $w \models^{\mathrm{E}}_{T''_{oga}} \Phi \cup \mathrm{UNA}_{\Sigma_P}$. Furthermore, if $(w, B) \models^{\mathrm{E}}_{T''_{oga}} \mathsf{L}d(x)$ for some variable x that occurs in $\tau_\chi(r)$, then $(w', B) \models^{\mathrm{E}}_{T''_{oga}} \mathsf{L}d(x)$; as $w' \models^{\mathrm{E}}_{T''_{oga}} \tau_\chi(r)$, it follows that $w \models^{\mathrm{E}}_{T''_{oga}} \tau_\chi(r)$. Hence, $w \models^{\mathrm{E}}_{T''_{oga}} \Delta' \cup \{\neg\alpha\}$.

($\Rightarrow$) Now let T' be a stable expansion of Δ'. We show that Δ'' has a stable expansion T'' such that $T'_{oga} = T'' \cap \mathcal{L}'_{ga}$. Let Γ be the smallest set $\Gamma \subseteq \mathcal{L}''_{ga}$ such that (a) $T'_{oga} \subseteq \Gamma$ and (b) $\Delta'' \models^{\mathrm{E}}_{\Gamma} \Gamma$; the set Γ exists since $\Delta'' \models^{\mathrm{E}}_{T'_{oga}} \beta$ for each $\beta \in T'_{oga}$ and $\Delta'' \models^{\mathrm{E}}_{S} \beta$ implies $\Delta'' \models^{\mathrm{E}}_{S'} \beta$ for each $\beta \in \mathcal{L}''_{oga}$ and $S \subseteq S' \subseteq \mathcal{L}''_{oga}$. Hence, Δ'' has a stable expansion T'' such that $T''_{oga} = \Gamma$. As $T'_{oga} \subseteq \Gamma$, it remains to show that $\Gamma \cap \mathcal{L}'_{ga} \subseteq T'_{oga}$.

We can obtain Γ as $\Gamma = \bigcup_{i \geq 0} \Gamma^i$, where $\Gamma^0 = T'_{oga}$ and $\Gamma^{i+1} = \{\alpha \in \mathcal{L}''_{ga} \mid \Delta'' \models^{\mathrm{E}}_{\Gamma^i} \alpha\}$, $i \geq 0$; furthermore, $d(t') \in \Gamma^i$ implies that $t = t' \in \Gamma^i$ for some $t' \in \mathcal{N}_{\Sigma'}$ by induction on $i \geq 0$ (note that Φ is Horn). Now let i be the least index such that $\Gamma^{i+1} \cap \mathcal{L}'_{ga} \not\subseteq T'_{oga}$. Hence, $\Delta'' \models^{\mathrm{E}}_{\Gamma^i} \alpha$ for some $\alpha \in \mathcal{L}'_{ga} \setminus \Gamma^i$, and for each model (w, Γ^i) of Δ'' there is some variable assignment B and axiom $\tau_\chi(r) \in \Delta''$ such that $(w, B) \models^{\mathrm{E}}_{\Gamma^i} b \wedge \mathsf{L}b$ for each b in the antecedent of $\tau_\chi(r)$, in particular for each $d(x)$, and $\tau_\chi(r)$ has consequent h such that $\alpha = h\beta$ for some variable substitution β associated with B. From minimality of i and safety of r, we may assume that β has only names from $\mathcal{N}_{\Sigma'}$. But then $\Delta' \models^{\mathrm{E}}_{T'_{oga}} \tau_\chi(r)$ implies that $\alpha \in T'_{oga}$, which is a contradiction. Hence $\Gamma^{i+1} \cap \mathcal{L}'_{ga} \subseteq T'_{oga}$ holds for all $i \geq 0$ and thus $\Gamma \cap \mathcal{L}'_{ga} \subseteq T'_{oga}$.

For the second part, we show that for any stable expansion T of $\Delta\chi = \iota_\chi(\Phi, P)$, $\chi \in \{{\scriptstyle HP, EB, EH}\}$, there exists a stable expansion T' of $\Delta'\chi = \iota_\chi(\Phi, gr_{\mathcal{L}}(P))$ such that $T'_{oga} = T_{oga}$, and vice versa. The result then follows by Propositions 6.6 and 6.4.

In what follows, let Σ_P be the signature of P, let Σ be the combined signature of Φ and P, and let $\Delta^{(')}\chi[\Sigma_P]$ (resp., $\Delta^{(')}\chi[\Sigma]$) result from $\Delta^{(')}\chi$ by removing the UNA_{Σ_P} axioms and adding, for each $a, b \in \mathcal{N}_{\Sigma_P}$ (resp., $a, b \in \mathcal{N}_{\Sigma}$), the atom $a = b$ if $a = b \in T$ and the negated atom $\neg(a = b)$ if $a = b \notin T$.

For the case $x = HP$, we note that $\Delta_{HP}[\Sigma_P]$ and $\Delta_{HP}[\Sigma]$ are classical FO Horn theories. Let $\alpha \in \mathcal{L}_{ga}$. Given that T is a stable expansion of Δ_{HP}, Proposition 3.4 implies $\alpha \in T_{oga}$ iff (a) $\Delta_{HP} \models^{\mathrm{E}}_{T_{oga}} \alpha$; the latter is equivalent to (b) $\Delta_{HP}[\Sigma_P] \models^{\mathrm{E}}_{T_{oga}} \alpha$. By well-known minimal model properties of Horn theories (cf. Lemma A.1), (b) iff (c) $\Delta_{HP}[\Sigma] \models^{\mathrm{E}}_{T_{oga}} \alpha$; furthermore, if we replace all embedded rules $\tau_{HP}(r)$, $r \in P$, in $\Delta_{HP}[\Sigma]$ with their groundings with respect to $\mathcal{N}_{\Sigma}$, we obtain $\Delta'_{HP}[\Sigma]$, and we that (c) iff (d) $\Delta'_{HP}[\Sigma] \models^{\mathrm{E}}_{T_{oga}} \alpha$. The latter is clearly equivalent to $\Delta'_{HP} \models^{\mathrm{E}}_{T_{oga}} \alpha$, which implies that Δ'_{HP} has some stable expansion T' such that $T_{oga} = T'_{oga}$. The converse direction is argued similarly.

In the cases $x = EB$ and $x = EH$, $\Delta^{(')}\chi[\Sigma_P]$ and $\Delta^{(')}\chi[\Sigma]$ are not classical FO Horn theories. However, we can turn them into such theories: for each modal atom $\mathsf{L}p(\vec{x})$ in a rule, we view $\mathsf{L}p$ as a fresh classical predicate, and add, for each ground atomic formula α with predicate p, the formula $\mathsf{L}\alpha$ to Φ if $\alpha \in T$ and we add the negated atom $\neg\mathsf{L}\alpha$ to Φ if $\alpha \notin T$. One can then straightforwardly verify that, for any $\alpha \in \mathcal{L}_{ga}$, the resulting theory entails α iff the original theory entails α with respect to T, i.e., if Ψ is the original theory, $\Psi \models^{\mathrm{E}}_{T} \alpha$. The argument then proceeds analogously to the case $x = HP$. $\square$

References

[Baader and Hollunder, 1995] Franz Baader and Bernard Hollunder. Embedding defaults into terminological knowledge representation formalisms. *Journal of Automated Reasoning*, 14:149–180, 1995.

[Bonatti *et al.*, 2006] Piero Bonatti, Carsten Lutz, and Frank Wolter. Expressive non-monotonic description logics based on circumscription. In *Proc. 10th International Conference on Principles of Knowledge Representation and Reasoning (KR'06)*, pages 400–410. AAAI Press, 2006.

[Bonatti, 1995] Piero A. Bonatti. Autoepistemic logics as a unifying framework for the semantics of logic programs. *Journal of Logic Programming*, 22(2):91–149, 1995.

[Borgida, 1996] Alex Borgida. On the relative expressiveness of description logics and predicate logics. *Artificial Intelligence*, 82(1–2):353–367, 1996.

[Chen, 1993] Jianhua Chen. Minimal knowledge + negation as failure = only knowing (sometimes). In *Proc. 2nd International Workshop on Logic Programming and Nonmonotonic Reasoning (LPNMR'93)*, pages 132–150. MIT Press, 1993.

[de Bruijn *et al.*, 2006] Jos de Bruijn, Thomas Eiter, Axel Polleres, and Hans Tompits. On representational issues about combinations of classical theories with nonmonotonic rules. In *Proc. 1st International Conference on Knowledge Science, Engineering and Management (KSEM'06)*, pages 1–22. Springer, 2006.

[de Bruijn *et al.*, 2007] Jos de Bruijn, David Pearce, Axel Polleres, and Agustín Valverde. Quantified equilibrium logic and hybrid rules. In *Proc. 1st International Conference on Web Reasoning and Rule Systems (RR'07)*, pages 58–72. Springer, 2007.

[de Bruijn *et al.*, 2008] Jos de Bruijn, Thomas Eiter, and Hans Tompits. Embedding approaches to combining rules and ontologies into autoepistemic logic. In *Proc. 11th International Conference on Principles of Knowledge Representation and Reasoning (KR'08)*, pages 485–495. AAAI Press, 2008.

[Denecker *et al.*, 2003] Marc Denecker, V. Wiktor Marek, and Miroslaw Truszczynski. Uniform semantic treatment of default and autoepistemic logics. *Artificial Intelligence*, 143(1):79–122, 2003.

[Donini *et al.*, 1998a] Francesco M. Donini, Maurizio Lenzerini, Daniele Nardi, Werner Nutt, and Andrea Schaerf. An epistemic operator for description logics. *Artificial Intelligence*, 100(1–2):225–274, 1998.

[Donini *et al.*, 1998b] Francesco M. Donini, Maurizio Lenzerini, Daniele Nardi, and Andrea Schaerf. $\mathcal{AL}$-log: integrating datalog and description logics. *Journal of Intelligent Information Systems*, 10:227–252, 1998.

[Donini *et al.*, 2002] Francesco M. Donini, Daniele Nardi, and Riccardo Rosati. Description logics of minimal knowledge and negation as failure. *ACM Transactions on Computational Logic*, 3(2):177–225, 2002.

[Drabent and Maluszynski, 2007] Wlodzimierz Drabent and Jan Maluszynski. Well-founded semantics for hybrid rules. In *Proc. 1st International Conference on Web Reasoning and Rule Systems (RR'07)*, pages 1–15. Springer, 2007.

[Drabent *et al.*, 2007] Wlodzimierz Drabent, Jakob Henriksson, and Jan Maluszynski. Hybrid reasoning with rules and constraints under well-founded semantics. In *Proc. 1st International Conference on Web Reasoning and Rule Systems (RR'07)*, pages 348–357. Springer, 2007.

[Eiter *et al.*, 2004] Thomas Eiter, Thomas Lukasiewicz, Roman Schindlauer, and Hans Tompits. Well-founded semantics for description logic programs in the semantic web. In *Proc. Workshop on Rules and Rule Markup Languages for the Semantic Web* (*RuleML04*), pages 81–97. Springer, 2004.

[Eiter *et al.*, 2005] Thomas Eiter, Giovambattista Ianni, Roman Schindlauer, and Hans Tompits. A uniform integration of higher-order reasoning and external evaluations in answer-set programming. In *Proc. 19th International Joint Conference on Artificial Intelligence (IJCAI'05)*, pages 90–96. Morgan Kaufmann, 2005.

[Eiter *et al.*, 2008] Thomas Eiter, Giovambattista Ianni, Thomas Lukasiewicz, Roman Schindlauer, and Hans Tompits. Combining answer set programming with description logics for the semantic web. *Artificial Intelligence*, 172(12-13):1495–1539, 2008.

[Faber *et al.*, 2004] W. Faber, N. Leone, and G. Pfeifer. Recursive aggregates in disjunctive logic programs: Semantics and complexity. In *Proc. 9th European Conference on Logics in Artificial Intelligence (JELIA'04)*, pages 200–212. Springer, 2004. Full paper to appear in *Artificial Intelligence*.

[Ferraris, 2005] Paolo Ferraris. Answer sets for propositional theories. In *Proc. 8th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'05)*, pages 119–131. Springer, 2005.

[Fitting, 1996] Melvin Fitting. *First Order Logic and Automated Theorem Proving (2nd ed.)*. Springer, New York, 1996.

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

[Gelfond and Lifschitz, 1988] Michael Gelfond and Vladimir Lifschitz. The stable model semantics for logic programming. In *Proc. 5th International Conference on Logic Programming (ICLP'88)*, pages 1070–1080. MIT Press, 1988.

[Gelfond and Lifschitz, 1991] Michael Gelfond and Vladimir Lifschitz. Classical negation in logic programs and disjunctive databases. *New Generation Computing*, 9(3–4):365–386, 1991.

[Gelfond and Przymusinska, 1993] Michael Gelfond and Halina Przymusinska. Reasoning on open domains. In *Proc. 2nd International Workshop on Logic Programming and Nonmonotonic Reasoning (LPNMR'93)*, pages 397–413. MIT Press, 1993.

[Gelfond, 1987] Michael Gelfond. On stratified autoepistemic theories. In *Proc. 6th National Conference on Artificial Intelligence (AAAI'87)*, pages 207–211. Morgan Kaufmann, 1987.

[Heymans *et al.*, 2006] Stijn Heymans, Davy van Nieuwenborgh, and Dirk Vermeir. Conceptual logic programs. *Annals of Mathematics and Artificial Intelligence*, 47(1–2):103–137, 2006.

[Heymans *et al.*, 2008] Stijn Heymans, Jos de Bruijn, Livia Predoiu, Cristina Feier, and Davy Van Nieuwenborgh. Guarded hybrid knowledge bases. *Theory and Practice of Logic Programming*, 8(3):411–429, 2008.

[Horrocks *et al.*, 2005] Ian Horrocks, Peter F. Patel-Schneider, Sean Bechhofer, and Dmitry Tsarkov. OWL rules: A proposal and prototype implementation. *Journal of Web Semantics*, 3(1):23–40, 2005.

[Hustadt *et al.*, 2005] Ullrich Hustadt, Boris Motik, and Ulrike Sattler. Data complexity of reasoning in very expressive description logics. In *Proc. 19th International Joint Conference on Artificial Intelligence (IJCAI'05)*, pages 466–471. AAAI Press, 2005.

[Janhunen, 1999] Tomi Janhunen. On the intertranslatability of non-monotonic logics. *Annals of Mathematics and Artificial Intelligence*, 27(1-4):79–128, 1999.

[Kaminski and Rey, 2002] Michael Kaminski and Guy Rey. Revisiting quantification in autoepistemic logic. *ACM Transactions on Computational Logic*, 3(4):542–561, 2002.

[Knorr *et al.*, 2008] Matthias Knorr, Jose Julio Alferes, and Pascal Hitzler. A coherent well-founded model for hybrid MKNF knowledge bases. In *Proc. 18th European Conference on Artificial Intelligence (ECAI'08)*. IOS Press, 2008.

[Konolige, 1991] Kurt Konolige. Quantification in autoepistemic logic. *Fundamenta Informaticae*, 15(3–4):275–300, 1991.

[Levesque and Lakemeyer, 2000] Hector J. Levesque and Gerhard Lakemeyer. *The Logic of Knowledge Bases*. MIT Press, Cambridge, MA, USA, 2000.

[Levesque, 1990] Hector J. Levesque. All I know: A study in autoepistemic logic. *Artificial Intelligence*, 42(2–3):263–309, 1990.

[Levy and Rousset, 1998] Alon Y. Levy and Marie-Christine Rousset. Combining horn rules and description logics in CARIN. *Artificial Intelligence*, 104:165–209, 1998.

[Lifschitz and Schwarz, 1993] Vladimir Lifschitz and Grigori Schwarz. Extended logic programs as autoepistemic theories. In *Proc. 2nd International Workshop on Logic Programming and Nonmonotonic Reasoning (LPNMR'93)*, pages 101–114. MIT Press, 1993.

[Lifschitz *et al.*, 1999] Vladimir Lifschitz, Lappoon R. Tang, and Hudson Turner. Nested expressions in logic programs. *Annals of Mathematics and Artificial Intelligence*, 25(3–4):369–389, 1999.

[Lifschitz, 1991] Vladimir Lifschitz. Nonmonotonic databases and epistemic queries. In *Proc. 12th International Joint Conference on Artificial Intelligence (IJCAI'91)*, pages 381–386. Morgan Kaufmann, 1991.

[Lifschitz, 1994] Vladimir Lifschitz. Minimal belief and negation as failure. *Artificial Intelligence*, 70(1–2):53–72, 1994.

[Marek and Truszczyński, 1993a] V. Wiktor Marek and Miroslaw Truszczyński. *Nonmonotonic Logic: Context-dependent Reasoning*. Springer-Verlag, Berlin Heidelberg, 1993.

[Marek and Truszczynski, 1993b] V. Wiktor Marek and Miroslaw Truszczynski. Reflexive autoepistemic logic and logic programming. In *Proc. 2nd International Workshop on Logic Programming and Nonmonotonic Reasoning (LPNMR'93)*, pages 115–131. MIT Press, 1993.

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

[McDermott, 1982] Drew McDermott. Nonmonotonic logic II: Nonmonotonic modal theories. *Journal of the ACM*, 29(1):33–57, 1982.

[Moore, 1985] Robert C. Moore. Semantical considerations on nonmonotonic logic. *Artificial Intelligence*, 25(1):75–94, 1985.

[Motik and Rosati, 2007] Boris Motik and Riccardo Rosati. A faithful integration of description logics with logic programming. In *Proc. 20th International Joint Conference on Artificial Intelligence (IJCAI'07)*, pages 477–482. AAAI Press, 2007.

[Motik *et al.*, 2005] Boris Motik, Ulrike Sattler, and Rudi Studer. Query answering for OWL-DL with rules. *Journal of Web Semantics*, 3(1):41–60, 2005.

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

[Pearce and Valverde, 2005] David Pearce and Agustín Valverde. A first-order nonmonotonic extension of constructive logic. *Studia Logica*, 80:321–246, 2005.

[Pelov *et al.*, 2007] Nikolay Pelov, Marc Denecker, and Maurice Bruynooghe. Well-founded and stable semantics of logic programs with aggregates. *Theory and Practice of Logic Programming*, 7(3):301–353, 2007.

[Przymusinski, 1991a] Teodor C. Przymusinski. Stable semantics for disjunctive programs. *New Generation Computing*, 9(3–4):401–424, 1991.

[Przymusinski, 1991b] Teodor C. Przymusinski. Three-valued nonmonotonic formalisms and semantics of logic programs. *Artificial Intelligence*, 49(1–3):309–343, 1991.

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

[Rosati, 1997] Riccardo Rosati. Embedding minimal knowledge into autoepistemic logic. In *Proc. 6th Conference of the Italian Association for AI (AI*IA-97)*, pages 231–241. Springer, 1997.

[Rosati, 1999] Riccardo Rosati. Towards first-order nonmonotonic reasoning. In *Proc. 5th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR'99)*, pages 332–346. Springer, 1999.

[Rosati, 2005] Riccardo Rosati. On the decidability and complexity of integrating ontologies and rules. *Journal of Web Semantics*, 3(1):61–73, 2005.

[Rosati, 2006] Riccardo Rosati. $\mathcal{DL}+log$: Tight integration of description logics and disjunctive datalog. In *Proc. 10th International Conference on Principles of Knowledge Representation and Reasoning (KR'06)*, pages 68–78. AAAI Press, 2006.

[Sattler *et al.*, 2003] Ulrike Sattler, Diego Calvanese, and Ralf Molitor. Relationships with other formalisms. In Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors, *The Description Logic Handbook*, chapter 4, pages 142–183. Cambridge University Press, Cambridge, UK, 2003.

[Schwarz, 1992] Grigori Schwarz. Reflexive autoepistemic logic. *Fundamenta Informaticae*, 17(1–2):157–173, 1992.

[Shvarts, 1990] Grigori F. Shvarts. Autoepistemic modal logics. In *Proc. 3rd Conference on Theoretical Aspects of Reasoning about Knowledge (TARK'90)*, pages 97–109. Morgan Kaufmann, 1990.

[Son and Pontelli, 2007] Tran Cao Son and Enrico Pontelli. A constructive semantic characterization of aggregates in answer set programming. *Theory and Practice of Logic Programming*, 7(3):355–375, 2007.

[Stalnaker, 1993] Robert Stalnaker. A note on non-monotonic modal logic. *Artificial Intelligence*, 64(2):183–196, 1993.

[Van Belleghem *et al.*, 1997] Kristof Van Belleghem, Marc Denecker, and Danny De Schreye. A strong correspondence between description logics and open logic programming. In *Proc. 14th International Conference on Logic Programming (ICLP97)*, pages 346–360. MIT Press, 1997.