

OI-implication: Soundness and Refutation Completeness

Floriana Esposito, Nicola Fanizzi, Stefano Ferilli, Giovanni Semeraro

Dipartimento di Informatica, Università di Bari, Via Orabona, 4 I-70125 Bari, Italy
{esposito, fanizzi, ferilli, semeraro}@di.uniba.it

Abstract

Weakening implication by assuming the *object identity* bias allows for both a model-theoretical and a proof-theoretical definition of a novel and more manageable ordering relationship over clausal spaces. In this paper, we give two important results, namely the soundness and the refutation completeness (through a subsumption theorem) of the underlying derivation procedure, that make this relationship particularly appealing for inducing a generalization model for clausal search spaces.

1 Introduction

Logical implication, as a relationship on clausal spaces, is particularly hard to handle due to many negative results descending from its non-decidability [Schmidt-Schauss, 1988] and intrinsic complexity [Gottlob and Fermüller, 1993]. This is the reason why subsumption, being more tractable and efficiently implementable than stronger relationships (i.e., relationships that entail it), is the generalization model commonly exploited in *Inductive Logic Programming* (ILP).

Attempts have been made to weaken subsumption and implication in order to obtain more manageable ordering relationships between clauses. Consider, for instance, forms of *non-decreasing subsumption* such as the *weak subsumption* [Badea and Stanciu, 1999], or weaker but decidable forms of implication such as *T-implication* [Idestam-Almqvist, 1995]. Indeed, the algebraic structure of the search space, and hence the refinement operators that can be defined, are affected by the choice of the underlying ordering relationship.

Weakening implication by assuming the *object identity* bias allowed for the definition of *OI-implication*, a novel relationship between clauses which makes the search space more manageable [Esposito *et al.*, 2000]. Specifically, it has been possible to prove the existence of *ideal* refinement operators for the search spaces ordered by the resulting form of logical implication, while this is not possible in standard clausal spaces [Nienhuys-Cheng and de Wolf, 1997]. This is particularly important for the efficiency of the refinement in search spaces with *dense solutions* [Badea and Stanciu, 1999]. Indeed, it has been shown as the computational complexity of the resulting operators is much more appealing than the one of the refinement operators defined for other search spaces.

Furthermore, the object identity framework for machine learning has been also exploited in the definition of operators performing inductive learning as well as abstraction and abduction, so to have a unique knowledge representation and reasoning framework for *multistrategy learning*.

In this paper, a proof-theoretic definition is adapted from [Esposito *et al.*, 2000] and a new model-theoretic definition of the implication under object identity is introduced. These definitions are shown to be equivalent. Indeed, two main results are demonstrated, namely the soundness of the derivation procedure compliant with the assumption of the mentioned bias and a subsumption theorem which allows for the demonstration that this procedure is also refutation-complete.

These results open new directions for the investigation of other interesting properties of this form of implication, such as the proof of a compactness theorem, the decidability and all the other algebraic properties of the search space, ordered by the relationship induced by this form of implication.

This paper is organized as follows. In Section 2, a semantics is defined, complying with the object identity assumption; hence, the definition of OI-implication is presented together with the proof of the soundness of the related proof procedure. Refutation completeness through the Subsumption Theorem, are proven in Section 3. Section 4 draws the conclusions of this research, outlining the ongoing developments.

2 OI-implication

In our framework, we adopt a representation language $\mathcal{L}$ expressing theories as *logic programs* made up of *clauses*. The basic notions about clausal representation can be found in [Lloyd, 1987; Nienhuys-Cheng and de Wolf, 1997].

The framework we propose relies essentially on the following assumption, which is commonly employed. For instance, defining the concept of bicycle, one would mention the two wheels among its components. Then, in order to rule out an instance of a different concept such as a monocycle, the identification of the two wheels should be avoided (a bicycle is two-wheeled by definition).

A thorough treatment of these issues for the Datalog clausal language is offered in [Semeraro *et al.*, 1998]. We want to extend this to a more expressive clausal language, without the restriction to function-free terms. In our setting we want to avoid the identification of terms.

Assumption 2.1 (Object Identity) *In a clause, terms denoted with different symbols must be distinct, i.e. they represent different entities of the domain.*

This notion can be viewed as an extension of Reiter's *unique names* assumption [Reiter, 1980]. Object identity is the basis for the definition of both an equational theory for the language clauses and a quasi-order upon them.

This assumption brings to a modified equational theory consisting in the addition of one axiom to an Equality Theory [Lloyd, 1987], namely:

$$\forall \neg(t_1 = t_2)$$

for all pairs (t_1, t_2) of distinct terms occurring in a clause.

Nienhuys-Cheng and de Wolf have embedded an equality theory in the semantics, for demonstrating the completeness of the SLDNF resolution, when dealing with the completion of normal programs, yielding the so called *E-Herbrand interpretations* [Nienhuys-Cheng and de Wolf, 1997]. Badea and Stanciu, in the definition of the weak subsumption [1999], deal with substitutions which do not identify literals. In this paper, we adopt a different approach, founded on a semantic basis. A similar approach to capturing the object identity assumption can be considered that of the *unique substitution semantics* [Khardon, 1999].

2.1 Semantics and Model-Theoretic Definition

Some definitions follow, specifying how the object identity assumption changes the semantics for the clausal representation. We start off by defining the relation between the terms of the language and the domain we deal with.

A *pre-interpretation* J under object identity of a first-order language is exactly a standard pre-interpretation:

- A non-empty set $\mathcal{D}$ called the *domain* of the pre-interpretation;
- Each constant in $\mathcal{L}$ is assigned to an element of $\mathcal{D}$;
- Each n -ary function symbol f is assigned a mapping J_f from $\mathcal{D}^n$ to $\mathcal{D}$.

Then, we need to map variables (and terms) to the objects of the domain:

Definition 2.1 *Let J be a pre-interpretation of $\mathcal{L}$ with domain $\mathcal{D}$. A variable assignment under object identity V w.r.t. $\mathcal{L}$ is a one-to-one mapping from the variables of $\mathcal{L}$ to the domain $\mathcal{D}$ of J . We use $V(X/d)$ to denote that the variable X is mapped to $d \in \mathcal{D}$ in the variable assignment V , and the other variables are mapped according to V . The term assignment under object identity Z in a formula ϕ w.r.t. J and V of terms in $\mathcal{L}$ is the following mapping from the terms in $\mathcal{L}$ to $\mathcal{D}$:*

1. J maps each constant to an element of $\mathcal{D}$;
2. V maps each variable to a distinct element of $\mathcal{D}$;
3. If $d_1, \dots, d_n$ are the elements of $\mathcal{D}$ to which the terms $t_1, \dots, t_n$ are mapped, then the term $f(t_1, \dots, t_n)$ is mapped to $J_f(d_1, \dots, d_n)$, where J_f is the function from $\mathcal{D}^n$ to $\mathcal{D}$ assigned to the function symbol f by J , s.t. $\forall t_1, t_2 \in \text{terms}(\phi) : Z(t_1/d_1), Z(t_2/d_2)$ implies $d_1 \neq d_2$.

By combining pre-interpretations and variable (term) assignments, we can specify the notion of interpretation under object identity:

Definition 2.2 *Given J a pre-interpretation of $\mathcal{L}$ with domain $\mathcal{D}$, an interpretation I under object identity (OI-interpretation) based on J assigns a mapping I_P from $\mathcal{D}^n$ to $\{T, F\}$ to each predicate symbol P .*

Given a variable assignment under object identity V , be Z the corresponding term assignment w.r.t. J and V . Then a formula ϕ in $\mathcal{L}$ has a truth value under I defined as follows:

1. *For each atom $A = P(t_1, \dots, t_n)$, if $\forall i = 1, \dots, n : d_i \in \mathcal{D}$ is assigned to t_i by Z , then the truth value of A under I and V is $I_P(d_1, \dots, d_n)$;*
2. *The truth values of the formula ϕ obtained by using connectives is determined using the standard truth tables.*
3. *If $\phi = \exists X \psi$ then ϕ has truth value T under I and V if there exists $d \in \mathcal{D}$ for which ψ has truth value T under I and $V(X/d)$, where $V(X/d)$ is a legal variable assignment under object identity. Otherwise, ϕ has truth value F under I and V .*
4. *If $\phi = \forall X \psi$ then ϕ has truth value T under I and V if for all elements $d \in \mathcal{D}$ ψ has truth value T under I and $V(X/d)$, where $V(X/d)$ is a legal variable assignment under object identity. Otherwise, ϕ has truth value F under I and V .*

The standard notions of *tautology*, *contradiction*, *satisfiability* and *consistency* can be straightforwardly extended to the new semantics.

In ILP we focus our interest on clauses. Since they are closed formulas, the truth values do not depend on variable assignments. We can represent an interpretation as a set of ground literals, i.e. those that are true according with the interpretation.

Definition 2.3 *Given a closed formula ϕ , an OI-interpretation I is called a model under object identity (OI-model) for ϕ iff I satisfies ϕ . Given a set of closed formulas Σ , I is an OI-model for Σ iff it is an OI-model for all formulas $\phi \in \Sigma$.*

OI-interpretations are always standard interpretations. Conversely, we show with an example a case when an OI-model for a formula is not a model.

Example 2.1 *Let $C = p(X) \vee p(a)$. An OI-model for C , assumed the domain is $\mathcal{D} = \{a, b\}$, is represented by the Herbrand OI-interpretation $I = \{p(b)\}$.*

I is not a model for C in the standard meaning. Indeed, since $p(a)$ is stated as false, the truth of the clause depends on $p(X)$, where variable X is universally quantified. Under object identity, it is sufficient that $p(b)$ be true for $\forall X p(X)$ to hold. The same does not apply in the standard semantics where both $p(a)$ and $p(b)$ should be true.

We are now ready to define the form of implication that is compliant with the object identity assumption:

Definition 2.4 *Given a set of formulas Σ and a formula ϕ , Σ OI-implies ϕ or ϕ is a logical consequence under object identity of Σ (denoted with $\Sigma \models_{\text{OI}} \phi$) iff all OI-models I of Σ are OI-also models of ϕ .*

Thus, useful results hold also in this semantics:

Theorem 2.1 (Deduction Theorem) *Let Σ be a set of formulas and ϕ and ψ be formulas. Then $\Sigma \cup \{\phi\} \models_{\text{OI}} \psi$ iff $\Sigma \models_{\text{OI}} (\phi \rightarrow \psi)$.*

Proof. $\Sigma \cup \{\phi\} \models_{\text{OI}} \psi$ iff
all OI-models of $\Sigma \cup \{\phi\}$ are OI-models of ψ iff
all OI-models of Σ are OI-models of $\neg\phi$ or of ψ iff
all OI-models of Σ are OI-models of $\neg\phi \vee \psi$ iff
 $\Sigma \models_{\text{OI}} (\phi \rightarrow \psi)$.

As a consequence, it is possible to prove also this result:

Theorem 2.2 *Let Σ be a set of formulas and ϕ be a formula. Then $\Sigma \models_{\text{OI}} \phi$ iff $\Sigma \cup \{\neg\phi\}$ is unsatisfiable.*

Proof. $\Sigma \models_{\text{OI}} \phi$ iff (by Theorem 2.1) $\Sigma \cup \{\neg\phi\} \models_{\text{OI}} \square$ iff
all OI-models of $\Sigma \cup \{\neg\phi\}$ are OI-models of $\square$ iff
 $\Sigma \cup \{\neg\phi\}$ is unsatisfiable under object identity.

As concerns clauses, in the standard semantics it is possible concentrate the interest on Herbrand models only [Lloyd, 1987], This can be extended to our setting.

Proposition 2.3 *Let Σ be a set of clauses in a first order language $\mathcal{L}$. Then Σ has an OI-model iff Σ has an Herbrand OI-model.*

Proof.

($\Rightarrow$) Suppose Σ has an OI-model I . Then we define the Herbrand OI-interpretation $\bar{I}$ with:

1. the same pre-interpretation of I and having as domain the Herbrand universe $\mathcal{U}_{\mathcal{L}}$;
2. Let P an n -ary predicate symbol occurring in Σ . Then we define the function $\bar{I}_P$ from $\mathcal{U}_{\mathcal{L}}$ to $\{F, T\}$ as follows: $\bar{I}_P(t_1, \dots, t_n) = T$ if $P(t_1, \dots, t_n)$ is true under I and $\bar{I}_P(t_1, \dots, t_n) = F$ otherwise.

($\Leftarrow$) Obvious.

2.2 Proof-Theoretic Definition and Soundness

Given a notion of subsumption under object identity, and using a subsumption theorem, we define implication under object identity proof-theoretically. The goal here is to define this relationship in a constructive way.

First, we have to define a form of resolution coping with the object identity assumption. To this purpose, we discuss some further properties required to substitutions in order to fulfill this assumption. In fact, a substitution can be regarded as a function mapping variables to terms. Hence, we require these functions to satisfy certain additional properties in order to preserve object identity.

Definition 2.5 *Given a set of terms T , a substitution σ is an OI-substitution w.r.t. T iff $\forall t_1, t_2 \in T : t_1 \neq t_2$ implies $t_1\sigma \neq t_2\sigma$.*

In the following, we will omit the set of terms T when it is obvious. Based on OI-substitutions, it is possible to define other related notions such as the *composition* of OI-substitutions, *ground* and *renaming* OI-substitutions, as well as *instance* clauses, *ground* clauses and *variants*.

In order to cope with the object identity assumption, a new relationship has been derived [Semeraro et al., 1998] from the classic θ -subsumption. Its definition can be given by exploiting the notion of OI-substitution [Esposito et al., 2000].

Definition 2.6 *Let C and D be two clauses. Then, C θ_{OI} -subsumes D iff there exists σ , an OI-substitution w.r.t. $\text{terms}(C)$, such that $C\sigma \subseteq D$.*

The notion of θ_{OI} -subsumption induces a quasi-order upon the space of clauses [Semeraro et al., 1998]. Besides, a notion of equivalence for θ_{OI} -subsumption can be defined, that partitions the space of clauses in equivalence classes. Since θ_{OI} -subsumption maps each literal of the subsuming clause onto a single literal in the subsumed one, equivalent clauses under θ_{OI} -subsumption must have the same number of literals. Thus, a search space ordered by θ_{OI} -subsumption is made up of non-redundant clauses. This yields smaller equivalence classes than those in a space ordered by θ -subsumption.

Now, as we have done for the substitutions, we can specify a notion of a unifier fulfilling Assumption 2.1:

Definition 2.7 *Given a finite set of simple expressions S , we say that θ is an OI-unifier iff $\exists E \forall E_i \in S : E_i\theta = E$ and θ is an OI-substitution w.r.t. $\text{terms}(E_i)$. An OI-unifier θ for S is called a most general OI-unifier (mgu_{OI}) for S iff, for each OI-unifier σ of S , there exists an OI-substitution τ such that $\sigma = \theta\tau$.*

In this setting, clauses cannot have proper subsets as factors. Hence, this notion can be used only for standardizing apart the variables of the parent clauses before a step of resolution [Esposito et al., 2000].

Definition 2.8 *Given the clauses C and D , suppose they are standardized apart. A clause R is an OI-resolvent of C and D iff, there exist $M \subseteq C$ and $N \subseteq D$ such that $\{M, N\}$ is unifiable through the mgu_{OI} and*

$$R = ((C \setminus M) \cup (D \setminus N))\theta$$

We write $R = \mathcal{R}_{\text{OI}}(\{C, D\})$.

This definition of OI-resolution follows Robinson's original definition [Robinson, 1965]. However, a different definition has been given, based on the one given in [Chang and Lee, 1973], taking into account one complementary pair of literals for each resolution step [Esposito et al., 2000].

Definition 2.9 *Let T be a set of clauses. The OI-resolution closure of T , denoted by $\mathcal{R}_{\text{OI}}^*(\Sigma)$, is defined inductively:*

$$\begin{aligned} (n = 0) \quad \mathcal{R}_{\text{OI}}^0(\Sigma) &= \Sigma \\ (n > 0) \quad \mathcal{R}_{\text{OI}}^n(\Sigma) &= \mathcal{R}_{\text{OI}}^{n-1}(\Sigma) \cup \\ &\quad \{R = \mathcal{R}_{\text{OI}}(\{C, D\}) \mid C, D \in \mathcal{R}_{\text{OI}}^{n-1}(\Sigma)\} \\ \mathcal{R}_{\text{OI}}^*(\Sigma) &= \mathcal{R}_{\text{OI}}^0(\Sigma) \cup \mathcal{R}_{\text{OI}}^1(\Sigma) \cup \dots \cup \mathcal{R}_{\text{OI}}^n(\Sigma) \cup \dots \end{aligned}$$

If there exists an OI-derivation of C from Σ , this will be denoted with $\Sigma \vdash_{\text{OI}} C$ which is equivalent to $C \in \mathcal{R}_{\text{OI}}^*(\Sigma)$.

It is interesting to consider also the case of *linear resolution* [Chang and Lee, 1973], which was originally exploited for the proof-theoretic definition of OI-implication [Esposito et al., 2000]. If C can be derived by means of zero or more linear OI-resolution steps from the set of clauses Σ , we denote

this with $C \in \mathcal{L}_{OI}^n(\Sigma)$, $n > 0$, where $\mathcal{L}_{OI}$ denotes the linear OI-resolution operator and $\mathcal{L}_{OI}^*$ stands for its closure.

In order to prove the soundness of the proof-procedure defined above, some preliminary results are needed.

Lemma 2.4 *Let C_1 and C_2 be clauses. If there exists R that is the OI-resolvent of C_1 and C_2 then it holds $\{C_1, C_2\} \models_{OI} R$.*

Proof. Assume that the two clauses are standardized apart. Suppose R is an OI-resolvent of C_1 and C_2 , resolved upon the subsets of literals M and N , and let θ be the $mg_{OI}(M, \bar{N})$ employed. Suppose the OI-interpretation I with domain $\mathcal{D}$ is an OI-model for $\{C_1, C_2\}$.

Let $\{X_1, \dots, X_n\} = \text{vars}(C_1) \cup \text{vars}(C_2)$ and V be a variable assignment under object identity. Then, $\forall i = 1, \dots, n, d_i \in \mathcal{D}$ if $M\theta$ is false under I and $V(X_1/d_1) \cdots (X_n/d_n)$, at least one of the other literals in C_1 is true under I and $V(X_1/d_1) \cdots (X_n/d_n)$. Similarly for $\bar{N}$. Note that R is made up of all literals in $C_1\theta$ and $C_2\theta$ except $M\theta$ or $N\theta$.

Since either $M\theta$ or $N\theta$ is false under I and $V(X_1/d_1) \cdots (X_n/d_n)$, at least one of the literals in R is true under I and $V(X_1/d_1) \cdots (X_n/d_n)$, hence I is an OI-model of R . Therefore, $\{C_1, C_2\} \models_{OI} R$.

This result can be inductively extended to OI-derivations, proving the soundness of OI-resolution:

Theorem 2.5 (Soundness) *Let Σ be a set of clauses, and C be a clause. If $\Sigma \vdash_{OI} C$ then $\Sigma \models_{OI} C$.*

Proof. Suppose $\Sigma \vdash_{OI} C$. Then there exists an OI-derivation $R_1, \dots, R_k = C$ from Σ . By induction on k , it holds:

$(k = 1)$ $R_1 = C \in \Sigma$, thus obviously $\Sigma \models_{OI} C$.

$(k > 1)$ Suppose the thesis holds for $m \leq k$.

Let $R_1, \dots, R_{k+1} = C$ be an OI-derivation of C from Σ . If $R_{k+1} \in \Sigma$ then the theorem is obvious. Otherwise, R_{k+1} is an OI-resolvent of some R_i and R_j ($i, j \leq k$). By the inductive hypothesis, we have $\Sigma \models_{OI} R_i$ and $\Sigma \models_{OI} R_j$. From Lemma 2.4, it follows that $\{R_i, R_j\} \models_{OI} R_{k+1} = C$.

Now we can give a proof-theoretic definition of the notion of OI-implication. In the next section, we will justify this formally, showing the equivalence of the two versions.

Definition 2.10 *Let C and D be two clauses. Then C implies D under object identity (C OI-implies D), denoted $C \Rightarrow_{OI} D$ iff either D is a tautological clause or there exists a clause $E \in \mathcal{R}_{OI}^*(\{C\})$ such that $E \theta_{OI}$ -subsumes D . Equivalence under OI-implication is denoted by $\Leftrightarrow_{OI}$.*

From the definition above, it is easy to see that OI-implication is a strictly a stronger ordering relationship than θ_{OI} -subsumption. Moreover the set of most specific clauses $\perp$ will be made up by tautologies.

We will prove in the next section that Definitions 2.4 and 2.10 are equivalent, by means of a subsumption theorem.

3 Refutation Completeness

The soundness theorem for OI-implication bridges the gap from the proof-theoretic to the model-theoretic definition of

this notion. In this section we give a proof of a completeness result to make the same step in the other direction. Following the proof-schema given in [Nienhuys-Cheng and de Wolf, 1996] for the standard case, first we prove this result for ground clauses, then we lift it to the general case.

3.1 Preliminaries

Let us start from the simplest case when both Σ and C are ground.

Lemma 3.1 *Let Σ be a logic program of ground clauses and C be a ground clause. If $\Sigma \models_{OI} C$ then there exists D such that $\Sigma \vdash_{OI} D$ and $D \subseteq C$.*

The proof is almost the same as in the case of standard implication (see [Nienhuys-Cheng and de Wolf, 1996]) since no variable assignment is made, for the clauses are ground. Yet in our case we considered finite sets of clauses. This can be extended by means of a compactness theorem.

In order to prove the Subsumption Theorem when only clause C is required to be ground, we need two theorems that are valid for the notion of unsatisfiability for standard interpretations but can be proven also when the object identity is assumed.

Theorem 3.2 *A set of clauses Σ is unsatisfiable under object identity iff there exists a finite unsatisfiable set Γ of ground instances of clauses from Σ .*

See [Nienhuys-Cheng and de Wolf, 1997] for a proof for standard interpretations. In this setting, it comes by using Proposition 2.3.

From this result and the Deduction Theorem another result follows:

Theorem 3.3 *Let Σ be a non-empty set of clauses and C be a ground clause. Then $\Sigma \models_{OI} C$ iff there exists a finite set Γ of ground instances of clauses from Σ such that $\Gamma \models_{OI} C$.*

Again, see [Nienhuys-Cheng and de Wolf, 1997] for a proof in the standard case. In this setting, Theorems 2.1 and 3.2 are to be exploited.

Now we have to show how to lift an OI-resolution step from the case of ground parent clauses to that of unrestricted clauses.

Lemma 3.4 *Let C_1 and C_2 be two clauses and C'_1 and C'_2 , respectively, two instances of them. If R' is an OI-resolvent of C'_1 and C'_2 , then there exists a clause R , that is an OI-resolvent of C_1 and C_2 , such that R' is an instance of R .*

Proof. Suppose that C_1, C_2, C'_1 and C'_2 are standardized apart. Be $\sigma_{1,2}$ the OI-substitutions such that $C'_1 = C_1\sigma_1$ and $C'_2 = C_2\sigma_2$.

Let $N'_1 \subseteq C'_1$ and $N'_2 \subseteq C'_2$ be the subsets of literals resolved upon for obtaining R' , then there exist $N_1 \subseteq C_1$ and $N_2 \subseteq C_2$ such that $N'_1 = N_1\sigma_1$ and $N'_2 = N_2\sigma_2$. Suppose that μ be an $mg_{OI}(N'_1, \bar{N}'_2)$.

Now let $D'_1 = C'_1 \setminus N'_1$ and $D'_2 = C'_2 \setminus N'_2$ for some $D_1 \subseteq N_1$ and $D_2 \subseteq N_2$. Hence, $R' = (D'_1 \cup D'_2)\mu = (D_1\sigma_1 \cup D_2\sigma_2)\mu$. Since the clauses are standardized apart, we can write: $R' = (D_1 \cup D_2)\sigma_1\sigma_2\mu$.

Consider N_1 and N_2 . They can be unified through the OI-unifier $\sigma_1\sigma_2\mu$. Then, let θ be an $mg_{OI}(N_1, \bar{N}_2)$, there exists

an OI-substitution δ such that $\sigma_1\sigma_2\mu = \theta\delta$.

Thus, C_1 and C_2 can be resolved under object identity upon N_1 and N_2 by using θ : $R = ((C_1 \setminus N_1) \cup (C_2 \setminus N_2))\theta = (D_1 \cup D_2)\theta$. Then $R\delta = (D_1 \cup D_2)\theta\delta = (D_1 \cup D_2)\sigma_1\sigma_2\mu = R'$. Hence $R' = R\delta$.

We can generalize the lifting technique to the case of OI-derivations, i.e. to multiple OI-resolution steps.

Lemma 3.5 Let Σ be a finite set of clauses and Σ' a set of instances of clauses from Σ . Suppose that $R'_1, \dots, R'_k$ be an OI-derivation of clause R'_k from Σ' . Then there exists an OI-derivation $R_1, \dots, R_k$ of clause R_k from Σ such that R'_i is an instance of R_i , for each $i = 1, \dots, k$.

Proof. By induction on the length of the derivation k .

($k = 1$) In this case $R'_1 \in \Sigma'$. But Σ' is made up of instances of the clauses in Σ . Hence, $\exists R_1 \in \Sigma$ such that R'_1 is an instance of R_1 .

($k > 1$) By hypothesis, let us suppose the thesis proven up to a derivation of length m . We have to prove it for $k = m + 1$.

Let $R'_1, \dots, R'_{m+1}$ be an OI-derivation of R'_{m+1} from Σ' , so that R'_{m+1} is the resolvent of two clauses in $\Sigma' \cup \{R'_1, \dots, R'_m\}$.

By hypothesis, there exists an OI-derivation $R_1, \dots, R_m$ of R_m from Σ such that $\forall i = 1, \dots, m : R'_i$ is an instance of R_i . Hence, by Lemma 3.4, there exists R_{m+1} such that R'_{m+1} is one of its instances.

We have now all the instruments for proving the following result:

Lemma 3.6 Let Σ be a finite set of clauses and C be a ground clause. If $\Sigma \models_{OI} C$ then there exists D such that $\Sigma \vdash_{OI} D$ and $D \subseteq C$.

Proof. Assume that C is not a tautology. We look for a clause D fulfilling the thesis. From $\Sigma \models_{OI} C$ and Theorem 3.3, then there exists a finite set Σ_g such that each clause in Σ_g is a ground instance of a clause in Σ and $\Sigma_g \models_{OI} C$.

Then, from Lemma 3.1 there exists a clause D' such that $\Sigma_g \vdash_{OI} D'$ and $D' \subseteq C$. Let $R'_1, \dots, R'_k = D'$ be an OI-derivation of D' from Σ_g .

By means of Lemma 3.5, the OI-derivation can be lifted to an OI-derivation $R_1, \dots, R_k$ of R_k from Σ , where $D' = R'_k$ is an instance of R_k . Let $D = R_k$. Then $\Sigma \vdash_{OI} D$ and $D \theta_{OI}$ -subsumes C (since $D' = D\delta$ and $D' \subseteq C$).

3.2 Subsumption Theorem and Consequences

Preliminarily, we have to recall the notion of a particular substitution that will be used in the next proofs:

Definition 3.1 Let Σ be a set of clauses and C be a clause. Let $\{X_1, \dots, X_n\} = \text{vars}(C)$ and suppose we have a set of constants $\{a_1, \dots, a_n\} \cap \text{consts}(\Sigma \cup C) = \emptyset$. Then $\sigma = \{X_1/a_1, \dots, X_n/a_n\}$ is a Skolem substitution for C w.r.t. Σ .

This notion is easily extensible to this setting, it is easy to see that Skolem substitutions are also OI-substitutions.

Now, given the results of the previous sections, we can prove the main result and some consequences.

Theorem 3.7 (Subsumption Theorem) Let Σ be a finite set of clauses and C be a clause. Then $\Sigma \models_{OI} C$ iff there exists a clause D such that $\Sigma \vdash_{OI} D$ and $D \theta_{OI}$ -subsumes C .

Proof.

($\Rightarrow$) Assume that C is not a tautology. Let θ be a Skolem OI-substitution for C w.r.t. Σ . Then $C\theta$ is a ground clause which is not a tautology and $\Sigma \models_{OI} C\theta$.

By Lemma 3.6, there is a clause D such that $\Sigma \vdash_{OI} D$ and $D \theta_{OI}$ -subsumes $C\theta$. Since D is derived from Σ , D cannot contain constants yielded by θ . Therefore θ maps variables of C to constants that are not in D .

Let δ be an OI-substitution such that $D \theta_{OI}$ -subsumes $C\theta$ and σ be the OI-substitution obtained by replacing each binding X_i/t_i in δ with X_i/a_i . Then $\delta = \sigma\theta$. Since θ only replaces the variables X_i by a_i ($1 \leq i \leq n$), it follows that $D\sigma \subseteq C$. Then $D \theta_{OI}$ -subsumes C . Hence, the thesis holds.

($\Leftarrow$) For the soundness of the OI-derivation (and θ_{OI} -subsumption).

As a consequence of this theorem, similarly to the standard case, it is nearly straightforward to demonstrate some important results originally due to Gottlob [1987].

Definition 3.2 A clause C is ambivalent when there are two literals $L_1, L_2 \in C$ with the same predicate symbol but opposite sign. Besides, if they unify then the clause C is called recursive.

Of course only unification through OI-substitutions is taken into account here. Given these notions, it is possible to show when OI-implication and θ_{OI} -subsumption coincide:

Proposition 3.8 Let C and D be clauses. If C is not recursive and D is not tautological, then $C \models_{OI} D$ iff $C \theta_{OI}$ -subsumes D .

Proof. Assume that $C \models_{OI} D$. By Theorem 3.7 it holds that $\exists E \in \mathcal{R}_{OI}^*(\{C\}) : E \theta_{OI}$ -subsumes D . Since C is not recursive $\{C\} = \mathcal{R}_{OI}^*(\{C\})$. Hence, $C \theta_{OI}$ -subsumes D .

The ($\Leftarrow$) part is obvious.

Now, given a clause C , we denote with C^+ and C^- , respectively, the set of its positive and negative literals.

Proposition 3.9 Let C and D be clauses. If $C \models_{OI} D$ then C^+ θ_{OI} -subsumes D^+ and C^- θ_{OI} -subsumes D^- .

Proof. Observe that C^+ θ_{OI} -subsumes C then $C^+ \models_{OI} C$. By the hypothesis, it holds: $C \models_{OI} D$, thus $C^+ \models_{OI} D$. Note that by construction C^+ and D^+ contain only positive literals, then C^+ cannot be OI-resolved with itself, and D^+ is not a tautology. By Proposition 3.8, it must hold that C^+ θ_{OI} -subsumes D .

Again, observing that C^+ is made up of positive literals only, one can conclude that C^+ θ_{OI} -subsumes D^+ .

Analogously C^- θ_{OI} -subsumes D^- .

This property can be used for the proof of the decidability of the implication under object identity. Yet, this subject goes beyond the scope of this paper.

Finally, we state a sufficient condition for the equivalence to hold between OI-implication and θ_{OI} -subsumption:

Proposition 3.10 Let C and D be clauses. If D is not ambivalent, then $C \models_{OI} D$ iff C θ_{OI} -subsumes D .

Proof. Assume $C \models_{OI} D$ and D non ambivalent. Thus D cannot be a tautology.

By Proposition 3.9 it holds $C^+ \theta \subseteq D^+$ and $C^- \delta \subseteq D^-$ for some OI-substitutions θ and δ .

Now, if C were recursive then there would be a literal $L = L\sigma_1 = L\sigma_2$ such that $L\sigma_1 \in C^+$ and $\neg L\sigma_2 \in C^-$, with σ_1 and σ_2 OI-unifiers. But then $L\sigma_1 \theta \in D^+$ and $\neg L\sigma_2 \delta \in D^-$, which does not hold since D is not ambivalent.

Thus C is not recursive and D is not tautological. By Proposition 3.8, it follows that C θ_{OI} -subsumes D .

The Subsumption Theorem showed that OI-resolution, together with θ_{OI} -subsumption, can derive the same conclusions drawn model-theoretically by implication under object identity. In standard implication, resolution is not deduction complete (some θ -subsumption steps could be needed), yet it might be complete with respect to an unsatisfiable set of clauses. We prove this result in the object identity setting.

Theorem 3.11 (Refutation Completeness) Let Σ be a finite set of clauses. Then Σ is unsatisfiable under object identity iff $\Sigma \vdash_{OI} \square$.

Proof.

($\Rightarrow$) Be Σ a finite set of clauses such that $\Sigma \models_{OI} \square$. For the Subsumption Theorem 3.7, there exists D such that $\Sigma \vdash_{OI} D$ and D θ_{OI} -subsumes $\square$. Then D must be the empty clause itself. Hence $\Sigma \vdash_{OI} \square$

($\Leftarrow$) For the soundness of the OI-derivation (Theorem 2.5).

All the theorems and propositions of the previous section can be extended to the case of an infinite set of formulas Σ if a Compactness Theorem like that for standard models [Boolos and Jeffrey, 1989] can be demonstrated. Yet, here we are interested in these issues concerning logic programs, hence finite sets of clauses only.

4 Conclusions and Further Work

A framework has been presented for the novel notion of implication under object identity. By assuming this bias over a clausal representation, we were able to define OI-implication both model-theoretically and proof-theoretically. The two definitions can be unified by exploiting the soundness of the OI-derivations and a subsumption theorem which is the basis for proving the refutation completeness of our procedure. This relationship, that is stronger than θ_{OI} -subsumption, induces a quasi-order over the clausal search space conferring different algebraic properties than standard implication.

We expect encouraging results on the complexity of learning/revising logic theories in such a search space. Hence, from a theoretical viewpoint, we are now concerned with the demonstration of the compactness and decidability of this new relationship. Another research stream will concern the algebraic properties of the resulting search space. In particular, algorithms for calculating minimal generalizations and specializations are to be developed.

In a more long term perspective, we aim at building an integrated framework for multistrategy learning that will benefit by these theoretical findings.

References

- [Badea and Stanciu, 1999] L. Badea and M. Stanciu. Refinement operators can be (weakly) perfect. In S. Džeroski and P. Flach, editors, *Proceedings of the 9th International Workshop on Inductive Logic Programming - ILP99*, volume 1634 of *LNAI*, pages 21–32. Springer, 1999.
- [Boolos and Jeffrey, 1989] G.S. Boolos and R.C. Jeffrey. *Computability and Logic*. Cambridge University Press, Cambridge, U.K., 3rd edition, 1989.
- [Chang and Lee, 1973] C.L. Chang and R.C.T. Lee. *Symbolic Logic and Mechanical Theorem Proving*. Academic Press, San Diego, CA, 1973.
- [Esposito et al., 2000] F. Esposito, N. Fanizzi, S. Ferilli, and G. Semeraro. Ideal refinement under object identity. In P. Langley, editor, *Proceedings of the 17th International Conference on Machine Learning - ICML2000*, pages 263–270, Palo Alto, CA, 2000. Morgan Kaufmann.
- [Gottlob and Fermüller, 1993] G. Gottlob and C.G. Fermüller. Removing redundancy from a clause. *Artificial Intelligence*, 61:263–289, 1993.
- [Gottlob, 1987] G. Gottlob. Subsumption and implication. *Information Processing Letters*, 24(2):109–111, 1987.
- [Idestam-Almqvist, 1995] P. Idestam-Almqvist. Generalization of clauses under implication. *Journal of Artificial Intelligence Research*, 3:467–489, 1995.
- [Kharon, 1999] R. Kharon. Learning function-free Horn expressions. *Machine Learning*, 37(3):241–275, December 1999.
- [Lloyd, 1987] J.W. Lloyd. *Foundations of Logic Programming*. Springer, 2nd edition, 1987.
- [Nienhuys-Cheng and de Wolf, 1996] S.-H. Nienhuys-Cheng and R. de Wolf. The subsumption theorem in inductive logic programming: Facts and fallacies. In L. De Raedt, editor, *Advances in Inductive Logic Programming*, pages 265–276. IOS Press, 1996.
- [Nienhuys-Cheng and de Wolf, 1997] S.-H. Nienhuys-Cheng and R. de Wolf. *Foundations of Inductive Logic Programming*, volume 1228 of *LNAI*. Springer, 1997.
- [Reiter, 1980] R. Reiter. Equality and domain closure in first order databases. *Journal of ACM*, 27:235–249, 1980.
- [Robinson, 1965] J.A. Robinson. A machine-oriented logic based on the resolution principle. *Journal of the ACM*, 12(1):23–41, January 1965.
- [Schmidt-Schauss, 1988] M. Schmidt-Schauss. Implication of clauses is undecidable. *Theoretical Computer Science*, 59:287–296, 1988.
- [Semeraro et al., 1998] G. Semeraro, F. Esposito, D. Malerba, N. Fanizzi, and S. Ferilli. A logic framework for the incremental inductive synthesis of Datalog theories. In N.E. Fuchs, editor, *Proceedings of LOPSTR97*, volume 1463 of *LNCS*, pages 300–321. Springer, 1998.