Spaces of Theories with Ideal Refinement Operators

Content Areas: Machine Learning, Inductive Logic Programming

Abstract

Refinement operators for theories avoid the problems related to the myopia of many relational learning algorithms based on the operators that refine single clauses. However, the non-existence of ideal refinement operators has been proven for the standard clausal search spaces based on θ-subsumption or logical implication, which scales up to the spaces of theories. By adopting different generalization models constrained by the assumption of object identity, we extend the theoretical results on the existence of ideal refinement operators for spaces of clauses to the case of spaces of theories.

1 Motivation

The investigation of learning through refinement operators allows to decouple the search from the heuristics in the study of relational learning algorithms. Therefore, the choice of the generalization model for a search space plays a key role since it affects both its algebraic structure and the definition of refinement operators for that space.

Logical implication and θ-subsumption are the relationships that are commonly employed for inducing generalization models in relational learning (the latter turning out to be more tractable with respect to the former). Yet, they are not fully satisfactory because of the complexity issues that the resulting search spaces present, although subspaces have been found where the generalization model is more manageable.

Indeed, the effectiveness and efficiency of learning as a refinement process depends on the properties of the search space and, as a consequence, of the operators. In some cases the important property to be required to operators is *flexibility* [Badea, 2001], meaning that they should be capable of focussing dynamically on certain zones of the search space that may be more promising. Conversely, the property of *ideality* [Nienhuys-Cheng and de Wolf, 1997] has been recognized as particularly important for the efficiency of incremental algorithms in search spaces with *dense solutions*. It is also possible to derive *non-redundant* operators from ideal ones, since the former are recognized to be more suitable for spaces with *rare* solutions [Badea and Stanciu, 1999].

Weakening implication by assuming *object identity*, an extension of the *unique names assumption* [Reiter, 1980], as a semantic bias has led to the definition of θ_{OI}-*subsumption* and *OI-implication* [Esposito *et al.*, 2001a], clausal relationships which induce more manageable search spaces. The existence of ideal refinement operators in these generalization models is possible [Esposito *et al.*, 2001b], while this does not hold in clausal spaces ordered by θ-subsumption or implication [Nienhuys-Cheng and de Wolf, 1997]. The objective of this work is to extend their result on spaces of clauses and prove the existence of ideal refinement operators for spaces of theories in those generalization models.

Indeed, most algorithms for relational learning, such as those employed in FOIL [Quinlan, 1990] and PROGOL [Muggleton, 1995], adopt greedy iterative covering strategies based of the refinement of clauses. Although these refinements may turn out to be optimal for a single clause, the result of assembling them in a theory is not guaranteed to be globally effective, since the interdependence of the clauses with respect to covering may lead to better theories made up of locally non-optimal clauses [Bratko, 1999].

This urges more complex refinement operators to be adopted in algorithms obeying to a more global strategy that are able to take into account the possible interactions between the single clausal refinements. Hence, the new problem is defining operators that refine whole theories rather than single clauses [Midelfart, 1999; Badea, 2001]. The resulting extended setting would also take into account background knowledge that may be available, and then it is also comparable to *generalized* and *relative subsumption* [Buntine, 1988; Plotkin, 1971] or *implication* [Nienhuys-Cheng and de Wolf, 1997]. However, these subjects concern the heuristics of refinement, which is no covered in this paper.

This paper is organized as follows. In Section 2, we present the semantics and proof-theory adopted in the framework. Section 3 deals with refinement operators and their properties. Then, in Section 4, refinement operators for the search space considered are defined and proven to be ideal. Section 5 summarizes the paper outlining possible developments.

2 Generalization Models and Object Identity

The representation language adopted in the proposed framework concerns logic theories (whose space is denoted $2^{\mathcal{C}}$) made up of clauses (space $\mathcal{C}$). For the basic notions about clausal representations in Inductive Logic Programming, the reader can refer to [Nienhuys-Cheng and de Wolf, 1997].

The framework relies essentially on the following bias proposed in [Esposito *et al.*, 2001a]:

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

The intuition for this bias is the following: considering the two clauses $C = q(x) \leftarrow p(X,X)$ and $D = q(x) \leftarrow p(X,X), p(X,Y), p(Y,Z), p(Z,X)$ in spaces where θ-subsumption (or implication) is adopted for inducing the generalization model, they are equivalent (in fact C is the *reduced* clause of D); this is not so natural as it may appear, since more elements of the domain can be accounted for in D than in C (indeed in this framework C is more general than D).

The expressive power is not diminished by this bias, since it is always possible to convey the same meaning of a clause, yet it may be necessary to employ more clauses, e.g. $q(X) \leftarrow p(X,Y)$ is equivalent to the couple of clauses (a theory) $\{(q(X) \leftarrow p(X,X)),\ (q(X) \leftarrow p(X,Y))\}$ when object identity is assumed.

The following definitions specify how the assumption above can be captured in the syntax and in the semantics of a clausal representation.

2.1 Semantics and Proof Theory

In citeEsposito01 substitutions that fulfill the assumption made have been defined. Since a substitution can be regarded as a mapping from the variables to the terms of a language, we require these functions to satisfy additional properties to avoid the identification of terms:

Definition 2.1 *Given a set of terms T (omitted when obvious) 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* $t_1\sigma \neq t_2\sigma$.

Based on OI-substitutions, it is possible to define related notions such as *ground* and *renaming* OI-substitutions and their *composition*, *instance* clauses, *ground instances* and *alphabetic variants*.

When the object identity assumption biases the interpretations, the resulting semantics can be defined as follows:

Definition 2.2 *Given a non-empty* domain $\mathcal{D}$, *a* pre-interpretation J *of the language* $\mathcal{C}$ *assigns each constant to an element of* $\mathcal{D}$ *and each n-ary function symbol f to a mapping from* $\mathcal{D}^n$ *to* $\mathcal{D}$.
An OI-interpretation I *based on* J *is a set of ground instances of atoms with arguments mapped in* $\mathcal{D}$ *through* J.
Given a ground OI-substitution γ *mapping* $vars(\mathcal{C})$ *to* $\mathcal{D}$, *an instance* $A\gamma$ *of an atom* A *is* true *in* I *iff* $A\gamma \in I$ *otherwise it is* false *in* I. *A negative literal* $\neg A\gamma$ *is* true *in* I *iff* $A\gamma$ *is not, otherwise it is* false *in* I.
I *is an* OI-model *for the clause* C *iff for all ground OI-substitutions* γ *there exists at least a literal in* $C\gamma$ *that is true in* I, *otherwise the clause is* false *in* I.

The standard notions of *tautology, contradiction, satisfiability* and *consistency* can be straightforwardly transposed to this semantics. Hence, they defined the form of implication that is compliant with this semantics [Esposito *et al.*, 2001a]. Besides, this relationship induces a quasi-order on spaces of clauses and theories.

Definition 2.3 *Let* C, D *be two clauses.* C implies D under object identity *(and then* C *is* more general *than* D *w.r.t. OI-implication) iff all OI-models for* C *are also OI-models for* D. *This relationship is denoted with* $C \models_{\mathrm{OI}} D$.
Analogously, a theory T implies C under object identity, *denoted with* $T \models_{\mathrm{OI}} C$, *iff all OI-models for* T *are also OI-models for* C. *Finally, a theory* T *is* more general *than a theory* T' *w.r.t. OI-implication iff* $\forall C' \in T'$: $T \models_{\mathrm{OI}} C'$.

OI-implication is a constrained form of logical implication biassed by the object identity assumption, as shown in the following example:

Example 2.1 *Given the two theories* $T = \{(p(X) \leftarrow q(f(X),Y), q(Y,f(X))),\ (q(f(X),f(X)) \leftarrow r(Z))\}$ *and* $T' = \{p(X) \leftarrow r(Z)\}$, *observe that* $T \models T'$ *but* $T \not\models_{\mathrm{OI}} T'$. *This depends on the disallowed binding* $Y/f(X)$ *that would identify terms within the same clause.*

The definition of the proof-theory is now briefly recalled:

Definition 2.4 *Given a finite set of clauses* S, *we say that* θ *is an* OI-unifier *iff* $\exists E$ *such that* $\forall E_i \in S$: $E_i\theta = E$ *and* θ *is an OI-substitution w.r.t.* $terms(E_i)$.
An OI-unifier θ *for* S *is a* most general OI-unifier *for* S *iff for each OI-unifier* σ *of* S *there exists an OI-substitution* τ *such that* $\sigma = \theta\tau$. *This is denoted with* $\mathrm{mgu}_{\mathrm{OI}}(S)$.

The following notions represent resolution and derivation when exclusively OI-unifiers are used.

Definition 2.5 *Given the clauses* C *and* D *that are supposed 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* $\{M, \overline{N}\}$ *is unifiable through the* $\mathrm{mgu}_{\mathrm{OI}}$ θ *and* $R \in ((C \setminus M) \cup (D \setminus N))\theta$.
$\mathcal{R}_{\mathrm{OI}}(C,D)$ *is the set of the OI-resolvents of* C *and* D.

An *OI-derivation* is obtained by successively chaining OI-resolutions.

Definition 2.6 *For any theory* T, *the* closure of OI-resolution *is defined*

$$\mathcal{R}^*_{\mathrm{OI}}(T) = \bigcup_{n \geq 0} \mathcal{R}^n_{\mathrm{OI}}(T)$$

where $\mathcal{R}^0_{\mathrm{OI}}(T) = T$ *and* $\mathcal{R}^n_{\mathrm{OI}}(T) = \mathcal{R}^{n-1}_{\mathrm{OI}}(T) \cup \{R = \mathcal{R}_{\mathrm{OI}}(C,D) \mid C, D \in \mathcal{R}^{n-1}_{\mathrm{OI}}(T)\}$.
If $\exists C \in \mathcal{R}^n_{\mathrm{OI}}(T)$ *then there is an* OI-derivation *of* C *from* T *of length* n.

This proof-procedure was proven sound in [Esposito *et al.*, 2001a], thus bridging the gap from the proof-theory to the model-theoretic definition of OI-implication.

2.2 θ_{OI}-subsumption and OI-implication

A syntactic relationship, similar to θ-subsumption but biased by the object identity assumption, has been defined based the notion of OI-substitution.

Definition 2.7 *Given two clauses* C *and* D, C θ_{OI}-subsumes D *iff there exists an OI-substitution* σ *w.r.t.* $terms(C)$ *such that* $C\sigma \subseteq D$. *In this case,* C *is* more general *than* D *w.r.t.* θ_{OI}*-subsumption, denoted* $C \geq_{\mathrm{OI}} D$. *If also* $D \geq_{\mathrm{OI}} C$ *then they are* equivalent *w.r.t.* θ_{OI}*-subsumption, denoted* $C \sim_{\mathrm{OI}} D$.
Analogously, given the theories T, T', T *is* more general *than* T' *w.r.t.* θ_{OI}*-subsumption iff* $\forall D \in T'$ $\exists C \in T$: $C \geq_{\mathrm{OI}} D$, *denoted* $T \geq_{\mathrm{OI}} T'$.

This relationship induces a quasi-order on spaces of clauses (and theories) which is weaker than OI-implication. Indeed, the following result [Esposito *et al.*, 2001a] shows that exactly:

Theorem 2.1 *Given a theory T and a non-tautological clause C, $T \models_{\mathrm{OI}} C$ iff there exists $D \in \mathcal{R}^*_{\mathrm{OI}}(T)$ such that D θ_{OI}-subsumes C.*

This result bridges the gap from model-theory to proof-theory in this framework. It also suggests the way to decompose OI-implication that is exploited for defining complete refinement operators.

Similarly to standard implication, it is nearly straightforward to demonstrate some consequences of Theorem 2.1 originally due to Gottlob [Gottlob, 1987]. Given a clause C, let C^+ and C^- denote, respectively, the sets of its positive and negative literals. Then, it holds:

Proposition 2.1 *Let C and D be clauses. If $C \models_{\mathrm{OI}} D$ then C^+ θ_{OI}-subsumes D^+ and C^- θ_{OI}-subsumes D^-.*

Since OI-substitutions map different literals of the subsuming clause onto different literals in the subsumed one, equivalent clauses under θ_{OI}-subsumption have the same number of literals. Thus, a space ordered by θ_{OI}-subsumption is made up of non-redundant clauses. Indeed, it holds:

Proposition 2.2 *Let C and D be two clauses. If C θ_{OI}-subsumes D then $|C| \leq |D|$. Moreover, $C \sim_{\mathrm{OI}} D$ iff they are alphabetic variants.*

As a consequence of the propositions above, it is possible to prove the following results on the depth and cardinality of clauses [Fanizzi and Ferilli, 2002], giving lower bounds for some measures definable on clauses, when implication under object identity holds:

Definition 2.8 *The* depth *of a term t is 1 when t is a variable or a constant. If $t = f(t_1, \ldots, t_n)$, then $depth(t) = 1 + max_{i=1}^{n}(depth(t_i))$. The* depth *of a clause C, denoted $depth(C)$, is the maximum depth among its terms.*

Proposition 2.3 *Given the clauses C and D, if $C \models_{\mathrm{OI}} D$ then it holds that $depth(C) \leq depth(D)$ and $|C| \leq |D|$.*

3 Theory Refinement and Object Identity

A learning problem can be cast as a search problem [Mitchell, 1982] where theory refinement is triggered when new evidence made available is to be assimilated. The canonical inductive paradigm requires the fulfillment of the properties of *completeness* and *consistency* for the synthesized theory with respect to a set of input examples.

When an inconsistent (respectively, incomplete) hypothesis is detected, a specialization (resp., generalization) of the hypothesis is required in order to restore this property of the theory. In the former case the refinement operators must search the space looking for more specific theories (downward refinements); in the latter, more general theories (upward refinements) are required.

The formal definition of the refinement operators for generic search spaces, is based on the algebraic notion of quasi-ordered set:

Definition 3.1 *A set $\mathcal{S}$ endowed with a relationship $\preceq$ that is reflexive and transitive is a* quasi-ordered set $(\mathcal{S}, \preceq)$

Then, the following definitions specify the notion of a function for computing refinements in generic quasi-ordered search spaces.

Definition 3.2 *Given a quasi-ordered set $(\mathcal{S}, \preceq)$, a* refinement operator *is a mapping from $\mathcal{S}$ to $2^{\mathcal{S}}$ such that:*

- $\forall C \in \mathcal{S} : \rho(C) \subseteq \{D \in \mathcal{S} \mid D \preceq C\}$ (downward refinement operator)
- $\forall C \in \mathcal{S} : \delta(C) \subseteq \{D \in \mathcal{S} \mid C \preceq D\}$ (upward refinement operator)

A notion of closure upon refinement operators is required when proving the completeness of the operators.

Definition 3.3 *In a quasi-ordered set $(\mathcal{S}, \preceq)$, let τ be a refinement operator. The* closure *of τ for $C \in \mathcal{S}$ is defined*

$$\tau^*(C) = \bigcup_{n \geq 0} \tau^n(C)$$

where $\tau^0(C) = \{C\}$ and
$\tau^n(C) = \{D \mid \exists E \in \tau^{n-1}(C) \colon D \in \tau(E)\}$.

Ultimately, refinement operators should construct chains of refinements from the starting elements (theories in this case) to target ones.

3.1 Properties of the Refinement Operators

As mentioned above, the properties of the refinement operators depend on the algebraic structure of the search space. A refinement operator induces a *refinement graph* [Nienhuys-Cheng and de Wolf, 1997], that is a directed graph containing an edge from T to T' in $\mathcal{S}$ in case the operator τ is such that $T' \in \tau(T)$. Refinement operators compute such steps.

A major source of inefficiency may come from refinements that turn out to be equivalent to the starting ones. Depending on the search algorithm adopted, computing refinements that are equivalent to some element that has been already discarded may introduce a lot of useless computation. As to the effectiveness of the search, a refinement operator should be able to find a path between any two comparable elements of the search space (or their equivalent representatives). It is desirable that at least one path in the graph can lead to target elements. This means that a complete refinement operator can derive any comparable element in a finite number of steps. The following properties formally define these concepts:

Definition 3.4 *In a quasi-ordered set $(\mathcal{S}, \preceq)$, a refinement operator τ is* locally finite *iff $\forall C \in \mathcal{S} : \tau(C)$ is finite and computable.*
A downward (resp. upward) refinement operator ρ (resp. δ) is proper *iff $\forall C \in \mathcal{S}$: $D \in \rho(C)$ implies $D \prec C$ (resp. $D \in \delta(C)$ implies $C \prec D$).*
A downward (resp. upward) refinement operator ρ (resp. δ) is complete *iff $\forall C, D \in \mathcal{S}$, $D \prec C$ implies $\exists E \in \mathcal{S} : E \in \rho^*(C)$ and $E \sim D$ (resp. $C \prec D$ implies $\exists E \in \mathcal{S} : E \in \delta^*(C)$ and $E \sim D$).*

Let us observe that local finiteness and completeness ensure the existence of a computable refinement chain to a target element, and properness ensure a more efficient refinement process, by avoiding the search of equivalent clauses. Then, the combination of these properties confers more effectiveness and efficiency to an operator:

Definition 3.5 *In a quasi-ordered set $(\mathcal{S}, \preceq)$, a refinement operator is* ideal *iff it is locally finite, proper and complete.*

As mentioned in the introduction, other important properties of refinement operators have been defined, yet they go beyond the scope of this paper which focusses on ideality.

3.2 Minimal Refinements of Clauses

The existence of maximal specializations and minimal generalizations of clauses was proven for both the $\theta_{\rm OI}$-subsumption and the OI-implication generalization model [Fanizzi and Ferilli, 2002]. These results are briefly recalled here for being used in the construction of ideal refinement operators presented in the following section.

As a consequence of Theorem 2.1, some limitations are provable as concerns depth and cardinality for a clause that implies (subsumes) another clause under object identity. This yields a bound to the proliferation of possible generalizations:

Proposition 3.1 *Let C and D be two clauses. The set of generalizations of C and D w.r.t. OI-implication is finite.*

The proof is straightforward since the depths and cardinalities of the generalizations are limited, by Proposition 2.3. Now, given two clauses C and D, let us denote with G the set of generalizations of $\{C, D\}$ w.r.t. OI-implication. Observe that $G \neq \emptyset$ since $\Box \in G$. Proposition 3.1 yields that G is finite. Thus, since the test of OI-implication between clauses is decidable [Fanizzi and Ferilli, 2002], it is theoretically possible to determine the minimal elements of G by comparing the clauses in G and eliminating those that are overly general.

For computing theories that are proper generalizations of the starting ones, an operator for inverting OI-resolutions is needed which is similar to the V-operator [Muggleton and Buntine, 1988]. Indeed, it is possible to define a theoretic operator in the following way:

Definition 3.6 *Let T be a theory. The* operator for the inversion of the OI-resolution *is defined:*

$$\mathcal{V}_{\rm OI}(T) = \{D \in \mathcal{C} \mid \exists C \in T, D' \in \mathcal{C}\colon\ C \in \mathcal{R}_{\rm OI}(D, D')\}$$

Of course there is a lot of indeterminacy in this definition. Yet it suffices to our theoretical purposes. In fact, the definition of an actual operator to be implemented in a learning system should consider also other information (such as examples, background knowledge, etc.), to specify the underlying heuristic component.

As regards maximal specializations, the major difficulty comes from the fact that under standard implication, $C \cup D$ is a clause that preserves the models of either clause, hence turning out to be a maximal specialization. In this setting, as expected, more clauses are needed than a single one; indeed the following operator has been defined:

Definition 3.7 *Let C_1 and C_2 be two clauses such that C_1 and C_2 are standardized apart and K a set of new constants such that: $|K| \geq |vars(C_1 \cup C_2)|$. A new set of clauses is defined $\mathcal{U}_{\rm OI}(C_1, C_2) = \{C \mid C = (C_1\sigma_1 \cup C_2\sigma_2)\sigma_1^{-1}\sigma_2^{-1}\}$ where σ_1 and σ_2 are Skolem substitutions for, respectively, C_1 and C_2 with K as their term set*[1].

Example 3.1 *Given two clauses $C_1 = \{p(X, Y), q(X)\}$ and $C_2 = \{p(X', Y'), r(X')\}$, the OI-substitutions $\sigma_1 = \{X/a, Y/b\}$ and $\sigma_2 = \{X'/a, Y'/b\}$ yield the following clause: $F_1 = \{p(X, Y), q(X), r(X)\}$. Similarly $\sigma_3 = \{X/a, Y/b\}$ and $\sigma_4 = \{X'/b, Y'/a\}$ yield the clause: $F_2 = \{p(X, Y), p(Y, X), q(X), r(X)\}$ and so on.*

It is easy to see that, clauses $(C_1\sigma_1 \cup C_2\sigma_2)\sigma_1^{-1}\sigma_2^{-1}$ are equivalent to those in $(C_1\sigma_1 \cup C_2\sigma_2)\sigma_2^{-1}\sigma_1^{-1}$. Besides, the clauses in $\mathcal{U}_{\rm OI}(C, D)$ preserve the OI-models of C and D:

Proposition 3.2 *Let C, D and E be clauses such that C and D are standardized apart. If $C \models_{\rm OI} E$ and $D \models_{\rm OI} E$ then $\forall F \in \mathcal{U}_{\rm OI}(C, D)\colon F \models_{\rm OI} E$.*

This result implies that $\mathcal{U}_{\rm OI}(C, D)$ contains maximal specializations of the two clauses w.r.t. OI-implication. For the ideality of the operator proved in the next section, it is also important to note that this set of specializations is finite. Moreover the definition of $\mathcal{U}_{\rm OI}$ and also Proposition 3.2 can be extended to the case of multiple clauses [Fanizzi and Ferilli, 2002].

4 Ideal Operators for Theories

Nonexistence conditions for ideal refinement operators for generic spaces are given in [van der Laag, 1995; Nienhuys-Cheng and de Wolf, 1997]. A close relationship is detected between ideality and the *covers* of elements of $(\mathcal{S}, \preceq)$, a cover of C being a $D \in S$ such that $D \prec C$ and $\nexists E\colon\ D \prec E \prec C$ (resp. $C \prec D$ and $\nexists E\colon\ C \prec E \prec D$). A condition that is necessary for the ideality of refinement operators is that they return supersets of the sets of covers (up to equivalence).

Theorem 4.1 *In the space $(\mathcal{C}, \leq_\theta)$, where $\leq_\theta$ denotes the order induced by θ-subsumption, on a language with at least a binary predicate, an ideal refinement operator does not exist.*

The non-existence of ideal refinement operators for spaces of clauses ordered by implication can be proven as a consequence of this result [Nienhuys-Cheng and de Wolf, 1997], since θ-subsumption is weaker than logical implication. Besides, this can be extended, proving the non-existence of refinement operators for search spaces of theories endowed with the ordering relationship induced by θ-subsumption [Midelfart, 1999] or the one induced by logical implication [Nienhuys-Cheng and de Wolf, 1997].

Conversely, in this framework it is possible to exploit the properties of the refinement operators for clausal spaces.

4.1 Ideal Operators for $\theta_{\rm OI}$-subsumption

As regards the spaces of clauses in the generalization model induced by $\theta_{\rm OI}$-subsumption, we exploit the ideality of the operators given in [Esposito *et al.*, 2001b].

[1]The *term set* of a set of clauses T by the Skolem substitution σ is the set of all terms occurring in $T\sigma$.

Definition 4.1 *In the quasi-ordered space* $(2^{\mathcal{C}}, \geq_{\mathrm{OI}})$*, given a theory* T*, let* T_{nr} *be a non redundant theory equivalent to* T*. The* downward refinement operator ρ_{OI} *is defined as follows:*

- $[(T_{nr} \setminus S) \cup \{D \in \rho_{\mathrm{OI}}(C) \mid C \in S\}] \in \rho_{\mathrm{OI}}(T)$ *where* $S \subseteq T_{nr}$
- $T_{nr} \setminus \{C\} \in \rho_{\mathrm{OI}}(T)$ *if* $C \in T_{nr}$

The upward refinement operator δ_{OI} *is defined as follows:*

- $[(T_{nr} \setminus S) \cup \{D \in \delta_{\mathrm{OI}}(C) \mid C \in S\}] \in \delta_{\mathrm{OI}}(T)$ *where* $S \subseteq T_{nr}$
- $T_{nr} \cup \{C\} \in \delta_{\mathrm{OI}}(T)$ *if* $C \notin T_{nr}$

The ideality of these operators is proven as follows:

Theorem 4.2 *In the search space* $(2^{\mathcal{C}}, \geq_{\mathrm{OI}})$*, the refinement operators* ρ_{OI} *and* δ_{OI} *are ideal.*
Proof:

ρ_{OI}: *(locally finite) obvious.*
(proper) by the properness of ρ_{OI} *for clauses.*
(complete) Suppose $T \geq_{\mathrm{OI}} T'$ *and* $T' \not\geq_{\mathrm{OI}} T$*. Let* $T' = \{D_i \mid i = 1, \ldots, n\}$*. The theories can be supposed to be non redundant, otherwise the reduced equivalent theory can be computed by removing clauses by means of the second item of the operator.*
By definition $T \geq_{\mathrm{OI}} T'$ *means that* $\forall D_i \in T'$, $i \in \{1, \ldots, n\}, \exists C_i \in T\colon\ C_i \geq_{\mathrm{OI}} D_i$.
By the completeness of the operator ρ_{OI} *for clauses, it holds that* $\forall i \in \{1, \ldots, n\} \exists k_i\colon\ D_i \in \rho_{\mathrm{OI}}^{k_i}(C_i)$.
Starting from $T_1 = T$*, the first component of the operator is iterated, obtaining for each* T_j *a refinement* T_{j+1}*, by choosing* S_j *as the subset of* T_j *made up of the clauses that are not in the target theory* T' *while being strictly more general than clauses in* T'*, that is* $S_j = \{C \in T_j \setminus T' \mid \exists D \in T'\colon\ C >_{\mathrm{OI}} D\}$.
Eventually it holds that $\exists T_k \in \rho_{\mathrm{OI}}^{k}(T)$ *for some* $k \leq \max_{i=1}^{n}(k_i)$ *such that* $\forall i\ D_i \in T_k$.
T_k *may be larger than* T'*. Thus the second component of* ρ_{OI} *for theories can be employed for deleting the exceeding clauses from* T_k *yielding* T'.
Finally we have that $T' \in \rho_{\mathrm{OI}}^{*}(T)$.

δ_{OI}: *Analogously.*

These operators will be employed in the definition of the refinement operators for the stronger order induced by OI-implication.

4.2 Ideal Operators for OI-implication

As regards the generalization model induced by OI-implication, some notions and results given in Section 3.2 are exploited. In particular, these operator should be able to compute specializations and generalizations that are able to reach those clauses involved in OI-resolution steps (and their inverse).

Definition 4.2 *In the quasi-ordered space* $(2^{\mathcal{C}}, \models_{\mathrm{OI}})$*, let* $T \in 2^{\mathcal{C}}$*. The* downward refinement operator ρ_{OI} *is defined as follows:*

- $\mathcal{U}_{\mathrm{OI}}(C)_{C \in S} C \in \rho_{\mathrm{OI}}(T)$ *where* $S \subseteq T$
- $\bigcup_{C \in S}((T \setminus S) \cup \rho_{\mathrm{OI}}'(C)) \in \rho_{\mathrm{OI}}(T)$ *where* $S \subseteq T$ *and* ρ_{OI}' *denotes the downward refinement operator for clauses wrt* $\models_{\mathrm{OI}}$
- $T' \in \rho_{\mathrm{OI}}(T)$ *if* $T' \in \rho_{\mathrm{OI}}''(T)$ *where* ρ_{OI}'' *denotes the downward refinement operator for theories wrt* $\geq_{\mathrm{OI}}$

The upward refinement operator δ_{OI} *is defined as follows:*

- $T \cup \{C\} \in \delta_{\mathrm{OI}}(T)$ *where* $\exists R \subseteq T$ *and* $C \in \mathcal{V}_{\mathrm{OI}}(T)$
- $\bigcup_{C \in S}((T \setminus S) \cup \delta_{\mathrm{OI}}'(C)) \in \delta_{\mathrm{OI}}(T)$ *where* $S \subseteq T$ *and* δ_{OI}' *denotes the upward refinement operator for clauses wrt* $\models_{\mathrm{OI}}$
- $T' \in \delta_{\mathrm{OI}}''(T)$ *if* $T' \in \delta_{\mathrm{OI}}(T)$ *where* δ_{OI}'' *denotes the upward refinement operator for theories wrt* $\geq_{\mathrm{OI}}$

The ideality of these operators is stated by the following result (Figure 1 depicts the related refinement graph):

Theorem 4.3 *In the search space* $(2^{\mathcal{C}}, \models_{\mathrm{OI}})$*, the refinement operators* ρ_{OI} *and* δ_{OI} *are ideal.*
Proof:

ρ_{OI}: *(locally finite) by definition of the various operators and the finiteness of the theories.*
(proper) by the properness of the refinement operators employed in the various items.
(complete) Suppose $T \models_{\mathrm{OI}} T'$ *and* $T' \not\models_{\mathrm{OI}} T$ *with* $n = |T'|$*. Since redundancy can be eliminated by removing redundant clauses (and tautologies) through the last item of the operator, we will consider the theories as non redundant.*
Let us observe that $T \models_{\mathrm{OI}} T'$ *is equivalent to* $\forall C_i \in T'\colon$ $T \models_{\mathrm{OI}} C_i$*, for* $i = 1, \ldots, n$.
By Theorem 2.1, for each C_i*, there exists a* D_i *such that* $D_i \in \mathcal{R}_{\mathrm{OI}}^{k_i}(T)$*, for some* k_i*, and* D_i θ_{OI}*-subsumes* C_i.
Observe that, by using the first item of the refinement operator, it is possible to produce the theory T_S *containing the maximal specializations of the clauses in* $S \subseteq T$ *employed in the OI-derivation the of* D_i*'s (by the extension of Lemma 3.2 to the case of multiple clauses where each maximal specialization is more general wrt OI-implication than an OI-resolvent* D_i*):* $T_S \in \rho_{\mathrm{OI}}(T)$.

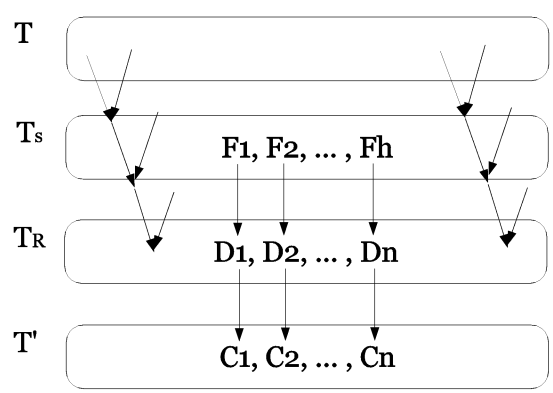

Figure 1: The Refinement Graph

Observe that $\forall F \in T_S\ F \models_{\mathrm{OI}} D_i$. Now, it is possible to iterate (at most n times for the D_i that are properly by some F) the second item of the operator, in order to compute the theory $T_R = \{D_i \mid i = 1, \ldots, n\}$, thus we can write: $T_R \in \rho^k_{\mathrm{OI}}(T_S), k \leq n$.
By construction $\forall i\ D_i \geq_{\mathrm{OI}} C_i$, then $T_R \geq_{\mathrm{OI}} T'$, then we can exploit the ideal operator for theories wrt θ_{OI}-subsumption (third item of the operator for theories wrt OI-implication) writing: $T' \in \rho^*_{\mathrm{OI}}(T_R)$.
Finally, by chaining these steps, it is possible to conclude that: $T' \in \rho^*_{\mathrm{OI}}(T)$.

δ_{OI}: *Analogously to the previous proof and in the same hypotheses, it is possible to invert the θ_{OI}-subsumption of T_R wrt T' using the last item of the definition of δ_{OI} for theories wrt OI-implication: $T_R \in \delta^*_{\mathrm{OI}}(T')$.*
*Then, a number of **OI**-resolutions are to be inverted by using the first item of δ_{OI}. This number is finite due to Proposition 2.3 and then can be done tentatively in a finite number of steps. Then $T \in \delta^*_{\mathrm{OI}}(T_R)$.*
Finally, by chaining these steps, it is possible to conclude that: $T \in \delta^*_{\mathrm{OI}}(T')$.

Differently from the standard generalization models, in this framework the number of OI-resolution steps is bounded because, during OI-implication or θ_{OI}-subsumption steps, the sizes of the clauses increase (decrease) monotonically, as a consequence of Propositions 2.2 and 2.3.

5 Conclusions

In this work the existence of ideal refinement operators was proved in the search space of theories ordered by generalization models based on object identity. Coupled with some heuristics, this allows for the definition of efficient refinement algorithms that avoid the myopia of the traditional relational learning approaches.

We focussed on the effectiveness of the refinement operators, that is related to their static properties. In general this is not sufficient for defining a learning algorithm: efficiency plays a key role when dealing with first order logics. The successive step is to investigate the dynamic properties of these operators when they are to be guided by means of heuristics based on the available examples and/or other criteria.

We have also mentioned that in spaces with rare solutions it is more suitable to have an operator that is non redundant, because almost all of the paths that could be constructed would not lead to a target theory. It should be investigated how to define non redundant operators in this framework.

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*, volume 1634 of *LNAI*, pages 21–32. Springer, 1999.

[Badea, 2001] L. Badea. A refinement operator for theories. In C. Rouveirol and M. Sebag, editors, *Proceedings of the 11th International Workshop on Inductive Logic Programming*, volume 2157 of *LNCS*, pages 1–14. Springer, 2001.

[Bratko, 1999] I. Bratko. Refining complete hypotheses in ILP. In S. Džeroski and P. Flach, editors, *Proceedings of the 9th International Workshop on Inductive Logic Programming*, volume 1634 of *LNAI*, pages 44–55. Springer, 1999.

[Buntine, 1988] W. Buntine. Generalized subsumption and its application to induction and redundancy. *Artificial Intelligence*, 36(2):149–176, 1988.

[Esposito *et al.*, 2001a] F. Esposito, N. Fanizzi, S. Ferilli, and G. Semeraro. OI-implication: Soundness and refutation completeness. In B. Nebel, editor, *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, pages 847–852, Seattle, WA, 2001.

[Esposito *et al.*, 2001b] F. Esposito, N. Fanizzi, S. Ferilli, and G. Semeraro. A generalization model based on OI-implication for ideal theory refinement. *Fundamenta Informaticae*, 47:15–33, 2001.

[Fanizzi and Ferilli, 2002] N. Fanizzi and S. Ferilli. Minimal generalizations under OI-implication. In M.-S. Hacid, Z. Raś, D. Zighed, and Y. Kodratoff, editors, *Proceedings of the 13th International Symposium on Methodologies for Intelligent Systems*, volume 2366 of *LNAI*, Lyon, France, 2002. Springer.

[Gottlob, 1987] G. Gottlob. Subsumption and implication. *Information Processing Letters*, 24(2):109–111, 1987.

[Midelfart, 1999] H. Midelfart. A bounded search space of clausal theories. In S. Džeroski and P. Flach, editors, *Proceedings of the 9th International Workshop on Inductive Logic Programming*, volume 1634 of *LNAI*, pages 210–221. Springer, 1999.

[Mitchell, 1982] T.M. Mitchell. Generalization as search. *Artificial Intelligence*, 18:203–226, 1982.

[Muggleton and Buntine, 1988] S. Muggleton and W. Buntine. Machine invention of first-order predicates by inverting resolution. In J. Laird, editor, *Proceedings of the 5th International Conference on Machine Learning*, pages 339–352, San Mateo, CA, 1988. Morgan Kaufmann.

[Muggleton, 1995] S.H. Muggleton. Inverse entailment and Progol. *New Generation Computing*, 13:245–286, 1995.

[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.

[Plotkin, 1971] G.D. Plotkin. A further note on inductive generalization. *Machine Intelligence*, 6:101–121, 1971.

[Quinlan, 1990] J.R. Quinlan. Learning logical definitions from relations. *Machine Learning*, 5:239–266, 1990.

[Reiter, 1980] R. Reiter. Equality and domain closure in first order databases. *Journal of ACM*, 27:235–249, 1980.

[van der Laag, 1995] P.R.J. van der Laag. *An Analysis of Refinement Operators in Inductive Logic Programming*. Ph.D. Thesis, Erasmus University, Rotterdam, The Netherlands, 1995.