From Inverse Entailment to Inverse Subsumption

Yoshitaka Yamamoto[1,2], Katsumi Inoue[2,3], Koji Iwanuma[1]

[1] University of Yamanashi
4-3-11 Takeda, Kofu-shi, Yamanashi 400-8511, Japan.
[2] Graduate University for Advanced Studies
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan.
[3] National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan.

Abstract. Modern explanatory ILP methods like Progol, Residue procedure, CF-induction, HAIL and Imparo use the principle of Inverse Entailment (IE). It is based on the fact that the negation of a hypothesis is derived from a prior background theory and the negation of examples. IE-based methods commonly compute a hypothesis in two steps: by first constructing an intermediate theory and next by generalizing its negation into the hypothesis with the inverse relation of entailment. In this paper, we focus on the sequence of intermediate theories that constructs a derivation from the background theory and the negation of examples to the negation of the hypothesis. We then show the negations of those derived theories in a sequence are represented with inverse subsumption. Using our result, inverse entailment can be reduced into inverse subsumption, while it preserves the completeness for finding hypotheses.
Keywords: inverse entailment, generalization, inverse subsumption

1 Introduction

Given a background theory B and observations E, the task of explanatory induction is to find a hypothesis H such that $B \wedge H \models E$ and $B \wedge H$ is consistent [1]. By the principle of Inverse Entailment (IE) [4], this is logically equivalent to finding a consistent hypothesis H such that $B \wedge \neg E \models \neg H$. This equivalence means that the inductive hypothesis H can be computed by deducing its negation $\neg H$ from B and $\neg E$. We represent this derivation process as follows:

$$B \wedge \neg E \models F_1 \models \cdots \models F_i \models \cdots \models F_n \models \neg H \tag{1}$$

where each F_i $(1 \leq i \leq n)$ denotes a clausal theory. IE-based methods [1, 2, 4, 6–8, 10] compute a hypothesis H in two ways: by first constructing an intermediate theory F_i in Formula (1) and next generalizing its negation $\neg F_i$ into the hypothesis H with the inverse relation of entailment. The logical relation between $\neg F_i$ and H is obtained from the contrapositive of Formula (1) as follows:

$$\neg(B \wedge \neg E) \dashv \neg F_1 \dashv \cdots \dashv \neg F_i \dashv \cdots \dashv \neg F_n \dashv H \tag{2}$$

where $\dashv$ denotes the inverse relation of entailment. Hereafter, we call it the *generalization relation*. In brief, IE-based methods first use the entailment relation to construct F_i with Formula (1), and then switch to the generalization relation to generate the hypothesis H with Formula (2) (See Fig. 1).

$$\overline{B \wedge \neg E \models F_1 \models \cdots \models F_i} \searrow \text{(Generalization relation)}$$
$$\neg F_i \dashv \cdots \dashv \neg F_n \dashv H$$

Fig. 1. Hypothesis finding based on inverse entailment

The inverse relation of entailment ensures the completeness in finding hypotheses from an intermediate theory by Formula (2). However, the generalization procedures with this relation need a variety of different operators such as *inverse resolution* [3]. There are several such operators and each one can be applied in many different ways, which lead to a large number of choice points. For this reason, systems like Progol [4, 8] and HAIL [6, 7] use subsumption due to computational efficiency, though their generalization procedures may become incomplete. On the other hand, systems like CF-induction [1] use entailment to find any hypothesis, though they need to handle a huge search space.

In this paper, we show inverse subsumption is an alternative generalization relation to ensure the completeness for finding hypotheses (See Fig. 2).

$$\overline{B \wedge \neg E \models F_1 \models \cdots \models F_i} \searrow \text{(Generalization relation)}$$
$$\neg F_i \preceq \cdots \preceq \neg F_n \preceq H$$

Fig. 2. Hypothesis finding based on inverse subsumption

Our result is used to reduce the non-deteriminism in generalization without losing the completeness. The key idea lies in that for two ground clausal theories S and T such that $S \models T$, $\neg S$ and $\neg T$ translated into CNF are represented with inverse subsumption. This feature is applied to the logical relation $F_i \models \neg H$ where F_i is an intermediate theory and H is a hypothesis in Formula (1). We then obtain its alternative relation $\neg F_i \preceq H$ represented by inverse subsumption. Since F_i is a CNF formula, there are several ways to represent $\neg F_i$ in CNF.

This paper focuses on two CNF formulas translated $\neg F_i$ into CNF, called *residue* and *minimal* complements, respectively. After Section 2 describes the theoretical background as well as some related work, we show inverse subsumption with residue and minimal complements ensure completeness of generalization in Section 3 and 4, respectively. In Section 5, we conclude. Due to space limitations, full proofs are given in the supplementary material.

2 Background

2.1 Notation and terminology

Here, we review the notation and terminology used in the paper. A *clause* is a finite disjunction of literals which is often identified with the set of its literals. A clause $\{\neg A_1, \ldots, \neg A_n, B_1, \ldots, B_m\}$, where each A_i, B_j is an atom, is also written as $A_1 \wedge \cdots \wedge A_n \supset B_1 \vee \cdots \vee B_m$. A clause is *Horn* if it contains at most one positive literal. A clause is *tautology* if it has complementary literals A and $\neg A$.

A *clausal theory* is a finite set of clauses. A clausal theory is *full* if it contains at least one non-Horn clause. A *conjunctive normal form* (CNF) formula is a conjunction of clauses, and a *disjunctive normal form* (DNF) is a disjunction of conjunctions of literals. A clausal theory is often identified with the conjunction of its clauses. Let S and T be clausal theories. We write $S_1 \models S_2$ if S_2 is a logical consequence of S_1. Especially, S_1 and S_2 are said to be *equivalent*, denoted by $S_1 \equiv S_2$ if $S_1 \models S_2$ and $S_2 \models S_1$. Note that a clausal theory S can include tautological clauses. Then, τS denotes the set of non-tautological clauses in S.

Let C and D be clauses. C *subsumes* D, denoted by $C \succeq D$, if there is a substitution θ such that $C\theta \subseteq D$. C *properly* subsumes D if $C \succeq D$ but $D \not\succeq C$. Let S be a clausal theory. μS denotes the set of clauses in S not properly subsumed by any clause in S. Let S_1 and S_2 be clausal theories. S_1 (*theory-*) *subsumes* S_2, denoted by $S_1 \succeq S_2$, if for any clause $D \in S_2$, there is a clause $C \in S_1$ such that $C \succeq D$.

Let S be a ground clausal theory $\{C_1, C_2, \ldots, C_n\}$ where C_i $(1 \leq i \leq n)$ $= l_{i,1} \vee l_{i,2} \vee \cdots \vee l_{i,m_i}$. The *complement* of S, denoted by $\overline{S}$, is defined as follows:

$$\overline{S} = \left\{ \neg l_{1,k_1} \vee \neg l_{2,k_2} \vee \cdots \vee \neg l_{n,k_n} \;\middle|\; \begin{array}{c} 1 \leq k_1 \leq m_1,\ 1 \leq k_2 \leq m_2, \\ \ldots,\ 1 \leq k_n \leq m_n \end{array} \right\}.$$

Note that $\overline{S}$ is a CNF formula such that $\overline{S} \equiv \neg S$. Accordingly, $\tau\overline{S}$ and $\mu\overline{S}$ are also CNF formulas logically equivalent to $\neg S$. We often denote $\tau\overline{S}$ and $\mu\overline{S}$ as the functions $R(S)$ and $M(S)$, called the *residue* and *minimal* complement of S, respectively. $R^2(S)$ and $M^2(S)$ denote $R(R(S))$ and $M(M(S))$, respectively.

2.2 Hypothesis finding based on inverse entailment

We give the definition of a hypothesis H in the logical setting of ILP as follows:

Definition 1 (Hypothesis). Let B and E be clausal theories, representing a background theory and (positive) examples/observations, respectively. Let H be a clausal theory. Then H is a *hypothesis* wrt B and E if H satisfies that $B \wedge H \models E$ and $B \wedge H$ is consistent. To ensure that the inductive task is not trivial, we assume $B \not\models E$ throughout this paper. We refer to a hypothesis instead of a hypothesis wrt B and E if no confusion arises.

Hypothesis finding in Definition 1 is logically equivalent to seeking a consistent hypothesis H such that $B \wedge \neg E \models \neg H$. Using this alternative condition, IE-based procedures [1, 2, 4, 6–8, 10] compute a hypothesis H in two steps. First, they construct an intermediate theory F such that F is ground and $B \wedge \neg E \models F$. Hereafter, we call F a *bridge theory* wrt B and E as follows.

Definition 2 (Bridge theory). Let B and E be a background theory and observations, respectively. Let F be a ground clausal theory. Then F is a *bridge theory* wrt B and E if $B \wedge \neg E \models F$ holds. If no confusion arises, a bridge theory wrt B and E will simply be called a bridge theory.

Every IE-based procedure constructs a bridge theory and generalizes into a hypothesis in its own way. Progol [4, 8], one of the state of the art ILP systems, uses the technique of *Bottom Generalization.* Its bridge theory F corresponds to the conjunction of ground literals each of which is derived from $B \wedge \neg E$. After constructing $\neg F$, called the *bottom clause*, Progol generalizes it with the inverse relation of subsumption, instead of entailment.

HAIL [6, 7] constructs so-called *Kernel Sets* to overcome some limitation on Bottom Generalization. Each clause C_i in a Kernel Set $\{C_1, \ldots, C_n\}$ is given by $B^i_1 \wedge \cdots \wedge B^i_{m_i} \supset A^i$, where $B \cup \{A^1, \ldots, A^n\} \models E$ and $B \models \{B^1_1, \ldots, B^n_{m_n}\}$. After constructing a Kernel Set, HAIL generalizes it using the inverse relation of subsumption like Progol. Note that a Kernel Set is regarded as the negation of a certain bridge theory F. In other words, they directly construct $\neg F$ by separately computing head and body literals of each clause in $\neg F$. We remark there is a recent work to extend Kernel Sets with an iterative procedure [2].

The residue procedure [10], which has been firstly proposed to find hypotheses in full clausal theories, constructs a bridge theory F consisting of ground instances from $B \wedge \neg E$. It then computes the residue complement $R(F)$, and generalizes it with inverse subsumption. In contrast, CF-induction [1], which is sound and complete for finding hypotheses in full clausal theories, constructs a bridge theory F consisting of ground instances from so-called *characteristic clauses* of $B \wedge \neg E$. Each characteristic clause is a subsume-minimal consequence of $B \wedge \neg E$ that satisfies a given language bias. Then CF-induction translates $\neg F$ into a CNF formula and generalizes it with inverse entailment.

Given a bridge theory F, inverse entailment ensures the completeness in generalizing $\neg F$ into hypotheses. There are several generalization operators within this relation such as *inverse resolution* [3] applying the inverse of the resolution principle, *anti-weakening* adding some clauses, and *anti-subsumption* dropping some literals from a clause. These operators can be soundly applied. However, there are many ways to apply each of them. Besides, any combination of them can be applied as another generalization operator. This fact makes generalization procedures with inverse entailment highly non-deterministic.

Our motivation is to reduce the non-determinism in generalization. In the following, we show inverse subsumption is an alternative generalization relation to ensure the completeness. In other words, given a bridge theory F, any hypothesis H such that $F \models \neg H$ can be obtained from $\neg F$ translated into CNF using inverse subsumption, instead of entailment.

3 Inverse Subsumption with Residue Complements

Our approach is based on the fact that for two ground clausal theories S and T such that $S \models T$, the logical relation between $\neg S$ and $\neg T$ translated in CNF is represented by inverse subsumption. We intend to apply this feature to Formula (1), later. Since $\neg S$ and $\neg T$ are DNF formulas, there are several ways to represent $\neg S$ and $\neg T$ in CNF. In this section, we use the residue complement and consider the logical relation between $R(S)$ and $R(T)$, which is represented primarily by the following theorem:

Theorem 1 ([10]). Let S and T be two clausal theories such that T is ground and both S and T do not include tautological clauses. If $S \models T$, there is a finite subset U of ground instances from S such that $R(T) \succeq R(U)$.

By Theorem 1, the following, which deals with the case that S is ground, holds:

Proposition 1. Let S and T be two ground clausal theories such that S and T do not include tautological clauses. If $S \models T$, then $R(T) \succeq R(S)$.

We apply this proposition to the logical relation $F \models \neg H$ where F is a bridge theory and H is a ground hypothesis. We represent $\neg H$ using the residue complement $R(H)$. Suppose that F does not include any tautological clauses. Then, by Propositon 1, $R^2(H) \succeq R(F)$ holds. In other words, $R^2(H)$, which is logically equivalent to H, can be obtained from $R(F)$ using inverse subsumption.

Theorem 2. Let F be a bridge theory such that F do not include tautological clauses, and H a hypothesis such that $F \models \neg H$. Then, there is a hypothesis H^* such that $H^* \equiv H$ and $H^* \succeq R(F)$.

However, every target hypothesis is not necessarily obtained from the residue complement by inverse subsumption. The below example describes such a case.

Example 1. Let B, E and H be a background theory, observations and a target hypothesis as follows:

$$B = \{p(a)\}. \ E = \{p(f(f(a))\}.$$
$$H = \{\{p(a) \supset p(f(a)),\ p(f(a)) \supset p(f(f(a)))\}\}.$$

Let F be the clausal theory $\{p(a), \neg p(f(f(a)))\}$. Since $F = B \cup \neg E$, F is a bridge theory wrt B and E such that $F \models \neg H$. $R(F)$ is $\{p(a) \supset p(f(f(a)))\}$. Then we notice that $R(F)$ is not subsumed by H. Indeed, $R(F)$ is the resolvent of two clauses in H. Then, we may need to apply an inverse resolution operator to $R(F)$ for obtaining the target hypothesis H.

This problem is caused by the fact that $R^2(H) = H$ cannot necessarily hold for a ground clausal theory H. The key idea in Theorem 2 lies in the logical relation $R^2(H) \succeq R(F)$. Hence, if $R^2(H) = H$ should not hold, H cannot be obtained from $R(F)$ using inverse subsumption. We thus need some CNF formula $F(H)$ for representing the negation of a hypothesis H such that $F(F(H)) = H$.

4 Inverse Subsumption with Minimal Complements

We here investigate minimal complements which satisfy the following theorem:

Theorem 3 ([11]). Let S be a ground clausal theory. Then, $M^2(S) = \mu S$ holds.

This theorem can be regarded as a fix point theorem on the function M computing the minimal complement. Unlike residue complements, $M^2(S)$ corresponds with S in case that S is subsume-minimal. Thus, minimal complements may not cause the problem in residue complements that they cannot necessarily obtain a target hypothesis using inverse subsumption, as described in Section 3.

Example 2. Let S be the clausal theory $\{a \vee b,\ b \vee c,\ \neg c\}$. Then, $\overline{S}$, $R(S)$, $R^2(S)$, $M(S)$ and $M^2(S)$ are as follows. In fact, $M^2(S) = S$ holds, whereas $R^2(S)$ does not. Note that $M(S)$ contains a tautological clause.

$$\overline{S} = \{\neg a \vee \neg b \vee c,\ \neg a \vee \neg c \vee c,\ \neg b \vee c,\ \neg b \vee \neg c \vee c\}.$$
$$R(S) = \{\neg a \vee \neg b \vee c,\ \neg b \vee c\}. \quad R^2(S) = \{a \vee b,\ \neg c \vee a,\ b,\ \neg c \vee b,\ \neg c\}.$$
$$M(S) = \{\neg a \vee \neg c \vee c,\ \neg b \vee c\}.\ M^2(S) = \{a \vee b,\ b \vee c,\ \neg c\}.$$

On the other hand, minimal complements do not necessarily satisfy the logical relation that $M(T) \succeq M(S)$ for ground clausal theories S and T such that $S \models T$. We recall Example 1. Whereas $F \models M(H)$ holds, $M^2(H)$, which is equal to H by Theorem 3, does not subsume $M(F)$. This is because minimal complements can include tautological clauses that residue complements never have. Indeed, Proposition 1, which shows the logical relation between $R(T)$ and $R(S)$, does not allow tautological clauses to be included in S and T. We then extend Proposition 1 so as to deal with tautological clauses as follows:

Theorem 4. Let S and T be ground clausal theories such that $S \models T$ and for every tautological clause $D \in T$, there is a clause $C \in S$ such that $C \succeq D$. Then,

$$\tau M(T) \succeq \tau M(S).$$

Theorem 4 enables us to construct an alternative generalization procedure using minimal complements. To describe the hypotheses that can be found by this, we first introduce the following language bias, called an *induction field*:

Definition 3 (Induction field). An *induction field*, denoted by $\mathcal{I}_{\mathcal{H}}$, is a finite set of literals to be appeared in ground hypotheses. A ground hypothesis H_g *belongs to* $\mathcal{I}_{\mathcal{H}}$ if every literal in H_g is included in $\mathcal{I}_{\mathcal{H}}$.

We next define the target hypotheses using the notion of an induction field $\mathcal{I}_{\mathcal{H}}$, together with a bridge theory F as follows:

Definition 4 (Hypothesis wrt $\mathcal{I}_{\mathcal{H}}$ and F). Let H be a hypothesis. H is a *hypothesis* wrt $\mathcal{I}_{\mathcal{H}}$ and F if there is a ground hypothesis H_g such that H_g consists of instances from H, $F \models \neg H_g$ and H_g belongs to $\mathcal{I}_{\mathcal{H}}$.

Now, the generalization procedure based on inverse subsumption with minimal complements is as follows:

Definition 5. Let B, E and $\mathcal{I}_{\mathcal{H}}$ be a background theory, observations and an induction field, respectively. Let F be a bridge theory wrt B and E. A clausal theory H is derived by *inverse subsumption with minimal complements* from F wrt $\mathcal{I}_{\mathcal{H}}$ if H is constructed as follows.

Step 1. $Taut(\mathcal{I_H}) := \{\neg A \vee A \mid A \in \mathcal{I_H} \ and\ \neg A \in \mathcal{I_H}\}$;
Step 2. Compute $\tau M(F \cup Taut(\mathcal{I_H}))$;
Step 3. Construct a clausal theory H satisfying the condition:

$$H \succeq \tau M(F \cup Taut(\mathcal{I_H})). \qquad (3)$$

Inverse subsumption with minimal complements ensures the completeness for finding hypotheses wrt $\mathcal{I_H}$ and F, by way of (3).

Main Theorem. Let B, E and $\mathcal{I_H}$ be a background theory, observations and an induction field, respectively. Let F be a bridge theory wrt B and E. For every hypothesis H wrt $\mathcal{I_H}$ and F, H is derived by inverse subsumption with minimal complements from F wrt $\mathcal{I_H}$.

Example 3. We show how a target hypothesis is derived by inverse subsumption with minimal complements using the below example on pathway completion:

$$B = \{arc(a,b),\ arc(X,Y) \wedge path(Y,Z) \supset path(X,Z)\}.\ E = \{path(a,c)\}.$$
$$\mathcal{I_H} = \{arc(b,c),\ \neg arc(b,c),\ path(b,c),\ \neg path(b,c)\}.$$
$$H = \{arc(b,c),\ arc(X,Y) \supset path(X,Y)\}.$$

One arc from b to c and one rule on pathways are missing in B. The task is to find the hypothesis H that completes these missing fact and rule. To complete H, both abduction and induction must involve, but most current ILP systems cannot compute it. This advanced inference has a possibility to be effectively applied to systems biology [12]. Let F be the clausal theory $\{arc(a,b),\ arc(a,b) \wedge path(b,c) \supset path(a,c),\ \neg path(a,c)\}$. Since F is the set of ground instances from $B \wedge \neg E$, F is a bridge theory wrt B and E. Since there is a ground hypothesis $H_g = \{arc(b,c),\ arc(b,c) \supset path(b,c)\}$ such that H_g consists of instances from H, $F \models \neg H_g$ and H_g belongs to $\mathcal{I_H}$, H is a hypothesis wrt $\mathcal{I_H}$ and F. Then, H could be derived by inverse subsumption with minimal complements. We first compute $Taut(\mathcal{I_H})$. Then, $Taut(\mathcal{I_H})$ is the set $\{\neg arc(b,c) \vee arc(b,c),\ \neg path(b,c) \vee path(b,c)\}$. After adding $Taut(\mathcal{I_H})$ to F, we compute $\tau M(F \cup Taut(\mathcal{I_H}))$ represented as follows. We then notice that H subsumes

$$\{\ \neg arc(a,b) \vee path(b,c) \vee \boxed{arc(b,c)} \vee path(a,c),$$
$$\neg arc(a,b) \vee \boxed{\neg arc(b,c) \vee path(b,c)} \vee path(a,c)\ \}.$$

$\tau M(F \cup Taut(\mathcal{I_H}))$ (See the dotted surrounding parts). Therefore, H can be derived by inverse subsumption with minimal complements.

In contrast, Since $R(F)$ is $\{\neg arc(a,b) \vee path(b,c) \vee path(a,c)\}$, H does not subsume the residue $R(F)$. Hence, H cannot be obtained from the residue complement, whereas the minimal complement can do with inverse subsumption.

5 Related Work and Conclusion

This paper has shown that inverse subsumption is an alternative generalization relation to ensure completeness for finding hypotheses. This result can be applied to each IE-based procedures. Generalization in Progol [4, 8], HAIL [6, 7] and Imparo [2] described in Section 2, are based on inverse subsumption, instead of entailment, whereas it has not been clarified so far whether or not this logical reduction makes generalization incomplete. For this open problem, we have showed that inverse subsumption can ensure the completeness only by adding tautological clauses associated with a language bias to a bridge theory. The generalization of Residue procedure [10] corresponds to inverse subsumption with residue complements, which has been studied in Section 3. CF-induction uses inverse entailment as the generalization relation. This can be reduced into inverse subsumption with minimal complements shown in Section 4. The issue on efficient implementation needs to be addressed in future work. For computation of minimal complements, we expect an efficient algorithm for enumerating the minimal hitting sets [9], which is equivalent to computing the minimal complement. We also intend to develop heuristics for searching relevant hypotheses on the subsumption lattice, which has been used in the state of the art systems.

References

1. K. Inoue, Induction as consequence finding. *Machine Learning*, 55(2), 109–135 (2004)
2. T. Kimber, K. Broda and A. Russo, Induction on failure: learning connected Horn theories. *Proc. of the 10th Int. Conf. on LPNMR.* LNCS, 5753, 169–181 (2009)
3. S. H. Muggleton and W. L. Buntine. Machine invention of first order predicates by inverting resolution. *Proc. of the 5th Int. Conf. on ML.* pages 339–352 (1988).
4. S. H. Muggleton, Inverse entailment and Progol. *New Generation Computing*, 13, 245–286 (1995)
5. S. Nienhuys-Cheng and R. de Wolf, *Foundations of inductive logic programming.* LNCS, 1228 (1997)
6. O. Ray and K. Broda and A. M. Russo, Hybrid abductive inductive learning. *Proc. of the 13th Int. Conf. on ILP.* LNCS, 2835, 311–328 (2003)
7. O. Ray and K. Inoue, Mode directed inverse entailment for full clausal theories. *Proc. of the 17th Int. Conf. on ILP.* LNCS, 4894, 225–238 (2008)
8. A. Tamaddoni-Nezhad and S. M. Muggleton, The lattice structure and refinement operators for the hypothesis space bounded by a bottom clause. *Machine Learning*, 76, 37–72 (2009)
9. T. Uno. A practical fast algorithm for enumerating minimal set coverings. *IPSJ SIG Notes*, 2002 (29), 9–16 (2002)
10. A. Yamamoto, Hypothesis finding based on upward refinement of residue hypotheses. *Theoretical Computer Science*, 298, 5–19 (2003)
11. Y. Yamamoto, K. Inoue and K. Iwanuma, Hypothesis enumeration by CF-induction. *Proc. of the 6th Workshop on LLLL*, 80–87 (2009)
12. Y. Yamamoto, K. Inoue and Andrei Doncescu, Integrating abduction and induction in biological inference using CF-induction. In H. Lodhi and S. Muggleton (Eds.), *Elements of Computational Systems Biology*, Chapter 9, 213–234 (2009)