Completing Inverse Entailment

Stephen Muggleton
Department of Computer Science,
University of York,
Heslington, York, YO1 5DD,
United Kingdom.

No Institute Given

Abstract. Yamamoto has shown that the *Inverse Entailment* (IE) mechanism described previously by the author is complete for Plotkin's relative subsumption but incomplete for entailment. That is to say, an hypothesised clause H can be derived from an example E under a background theory B using IE if and only if H subsumes E relative to B in Plotkin's sense. Yamamoto gives examples of H for which $B \cup H \models E$ but H cannot be constructed using IE from B and E. The main result of the present paper is a theorem to show that by enlarging the bottom set used within IE, it is possible to make a revised version of IE complete with respect to entailment for Horn theories. Furthermore, it is shown for function-free definite clauses that given a bound k on the arity of predicates used in B and E, the cardinality of the enlarged bottom set is bounded above by the polynomial function $p(c+1)^k$, where p is the number of predicates in B, E and c is the number of constants in $B \cup \overline{E}$.

1 Introduction

In [5] Yamamoto shows that the mechanism of *Inverse Entailment* (IE) introduced in [2] is complete for Plotkin's relative subsumption, but incomplete for entailment. In this paper it is shown that enlarging the bottom set leads to completeness of IE with respect to entailment for Horn theories.

This paper is organised as follows. The next section gives formal definitions used in the rest of the paper. Some useful properties of Herbrand models of Horn theories are proved in Section 2.3. Section 3 introduces the definitions of IE and and the revised bottom set. A useful model intersection property of Horn theories is also proved. The definitions are applied to the example Yamamoto uses to show the incompleteness of IE. The main completeness theorem is proved in Section 4. The results are summarised and an open problem is discussed in Section 5.

2 Preliminaries

It is assumed the reader is familiar with first-order logic and logic programming (see [4]).

2.1 Clauses and clausal theories

A positive literal is an atom, a negative literal is the negation of an atom. A clause is a finite set of literals, and is treated as a universally quantified disjunction of those literals. A finite set of clauses is called a clausal theory and is treated as a conjunction of those clauses. A Horn clause is a clause containing at most one positive literal. Non-definite Horn clauses are called goals. A Horn theory is a clausal theory containing only Horn clauses. A definite clause is a clause containing exactly one positive literal. A definite clause program is a clausal theory containing only definite clauses. A clause C can be written as

$$a_0; ..; a_m \leftarrow b_1, .., b_n$$

where the a_i are the positive literals and the b_j are the negative literals. In this case C^+, C^- denote the clauses consisting of the positive and negative literals of C respectively.

A clause is said to be function-free whenever it does not contain functions of arity 1 or more. A clausal theory is said to be function-free whenever all its clauses are function-free.

A unique Skolem constant c_v will be assumed to be associated with every variable v. If F is a first-order formula then $N(F)$ is formed by replacing each variable x in F by the Skolem constant c_x and each Skolem constant c_y by the associated variable y. If l is a literal then $\overline{l}$ is formed by removing all occurrences of double negation ($\neg\neg$) in $\neg N(l)$. Thus $\overline{\overline{l}} = l$ for all literals l. If $C = \{l_1, ..l_n\}$ is a clause then $\overline{C}$ is the Horn theory $\{\{\overline{l_1}\}, .., \{\overline{l_n}\}\}$.

Definition 1. Subsumption. *Let C and D be clauses. C subsumes D, denoted $C \succeq D$, if and only if there exists a substitution θ such that $C\theta \subseteq D$.*

2.2 Herbrand models

Suppose T is a clausal theory. $B(T)$ will denote the Herbrand Base of T (see [4] for a definition of $B(T)$). As usual, Herbrand interpretations and Herbrand models of T are represented by subsets of $B(T)$. $M \subseteq B(T)$ is a Herbrand model of T if and only if it is a model of all the clauses in T. M is a Herbrand model of a clause $C \in T$ if and only if there does not exist a ground substitution θ and a clause $D = C\theta$ such that $D^+ \subseteq B(T) \setminus M$ and $D^- \subseteq M$. $\mathcal{M}(F) = \{M_1, M_2, ..\}$ will be used to denote the Herbrand models of T. When $\mathcal{M}(T)$ is non-empty $M(T)$ denotes $\bigcap_i M_i$. If $M(T)$ is a model of T then it will be known as the least Herbrand model of T. According to Herbrand's theorem T is satisfiable if $\mathcal{M}(T)$ is non-empty and unsatisfiable otherwise.
The following is a restatement of Proposition 6.1 in [1] Theorem 2.14 in [4] and Proposition 7.13 in [3].

Theorem 1. Model intersection property. *Let P be a definite clause program and $\mathcal{M} = \{M_1, M_2, ..\}$ be a non-empty set of Herbrand models of P. Then $\bigcap_i M_i$ is a Herbrand model of P.*

2.3 Herbrand models of Horn theories

The following Lemma is a generalisation of Theorem 1 to Horn theories[1].

Lemma 1. Horn theory model intersection property. *Let $T = T_0 \cup T_1$ be a satisfiable Horn theory in which T_0, T_1 are the non-definite and definite subsets of T respectively. Let $\mathcal{M} = \{M_1, M_2, ..\}$ be a non-empty set of Herbrand models of T. Then $M = \bigcap_i M_i$ is a Herbrand model of T.*
Proof. *Assume false. Then there exists T for which $\mathcal{M}$ is non-empty and M is not a model of T. M must either not be a model of T_0 or not be a model of T_1. However, every model in $\mathcal{M}$ is a model of T_1 since $T \models T_1$ and thus according to Theorem 1 M is a model of T_1. It follows that M must not be a model of T_0. Thus there must exist a ground instance of a clause in T_0:*

$$\leftarrow a_1, .., a_n$$

for which $a_1, .., a_n$ are all in M. Then $a_1, .., a_n$ are in every model in $\mathcal{M}$. But this means that no element of $\mathcal{M}$ is a model of T, which contradicts the assumption that $\mathcal{M}$ is non-empty. This completes the proof.

Corollary 1. Least Herbrand model of Horn theories. *Let T be a satisfiable Horn theory. $M(T)$ is the least Herbrand model of T.*
Proof. *Follows directly from Lemma 1 and the definitions of satisfiability and least Herbrand model.*

3 IE and the enlarged bottom set

The following is a variant on the definition of the bottom set used by Yamamoto in describing IE.

Definition 2. Enlarged bottom set. *Let B be a Horn theory, E be a clause, such that $F = (B \cup \overline{E})$ is satisfiable (ie. $B \not\models E$)). The enlarged bottom set of E under B is denoted $BOT(B, E)$ and is defined as follows.*

$$\begin{aligned} BOT^+(B,E) &= \{a | a \in B(F) \setminus M(F)\} \\ BOT^-(B,E) &= \{\neg a | a \in M(F)\} \\ BOT(B,E) &= BOT^+(B,E) \cup BOT^-(B,E) \end{aligned}$$

Note that the cardinality of BOT(B, E) can be infinite. However, this will not be the case when B, E are function-free.

Remark 1. **Cardinality of BOT(B, E) (function-free case).** Let B be a Horn theory and E be a clause, such that $F = (B \cup \overline{E})$ is satisfiable. Suppose F contains p predicate symbols with maximum arity k and c is the number of constants in F. The cardinality of BOT(B, E) has an upper bound of $p(c + 1)^k$.
Proof. Follows from the fact that for each of the p predicate symbols there are at most $(c + 1)^k$ different atoms that can be constructed from the at most $c + 1$ constants in $B(F)$.

[1] Lemma 1 is not found in logic programming theory texts such as [1, 4, 3].

Below we give a revised definition of IE similar to that in [5], but based on the enlarged bottom set.

Definition 3. IE based on enlarged bottom set. *Let B be a Horn theory and E be a clause, such that $B \not\models E$. A clause H is derived by IE from E under B if and only if there exists a clause $H' \subseteq BOT(B, E)$ such that $H \succeq H'$.*

The following example demonstrates the construction of $\mathrm{BOT}(B, E)$ for a function-free version of the example Yamamoto uses to demonstrate the incompleteness of IE.

Example 1. **Yamamoto's example.**

$$B = \left\{ \begin{array}{l} even(0) \leftarrow \\ even(y) \leftarrow s(x,y), odd(x) \end{array} \right\}$$
$$E = odd(z) \leftarrow s(y,z), s(x,y), s(0,x)$$
$$\overline{E} = \left\{ \begin{array}{l} \leftarrow odd(c_z) \\ s(c_y, c_z) \leftarrow \\ s(c_x, c_y) \leftarrow \\ s(0, c_x) \leftarrow \end{array} \right\}$$
$$B(B \cup \overline{E}) = \left\{ \begin{array}{l} even(0), even(c_x), even(c_y), even(c_z), odd(0), \\ odd(c_x), odd(c_y), odd(c_z), s(0,0), s(0,c_x), .., s(c_z, c_z) \end{array} \right\}$$
$$M(B \cup \overline{E}) = \{ even(0), s(c_y, c_z), s(c_x, c_y), s(0, c_x) \}$$
$$\mathrm{BOT}(B, E) = even(c_x); even(c_y); ..; odd(c_x) \leftarrow even(0), s(c_y, c_z), s(c_x, c_y), s(0, c_x)$$

Note that $\mathrm{BOT}(B, E)$ is subsumed by the following clause H.

$$odd(u) \leftarrow s(v, u), even(v)$$

Thus H can be derived by IE from E under B. H is a function-free version of the one which Yamamoto showed could not be derived by the original definitions of IE.

4 Completeness

Below is the main completeness result for clauses derived using IE.

Theorem 2. Horn theory completeness of IE. *Let B be a Horn theory, E be a clause, such that $F = (B \cup \overline{E})$ is satisfiable (ie. $B \not\models E$)). $G = (B \cup H \cup \overline{E})$ is unsatisfiable (ie. $B \cup H \models E$) and $B(G) = B(F)$ only if H is derived by IE from E under B.*

Proof. *Assume false. Then G is unsatisfiable and $B(G) = B(F)$ and (according to Definition 3) there does not exist $H' \subseteq BOT(B, E)$ such that $H \succeq H'$. Since F is a satisfiable Horn theory according to Corollary 1 $M(F)$ is its least Herbrand model. But $M(F)$ cannot be a Herbrand model of G since G is unsatisfiable, and therefore there must be a ground substitution θ such that $H^+\theta \subseteq (B(F) \setminus M(F))$ and $H^-\theta \subseteq M(F)$. Therefore letting $H' = H\theta$ clearly $H \succeq H'$ and $H' \subseteq BOT(B, E)$. This contradicts the assumption and completes the proof.*

It is worth drawing attention to the "only if" in the theorem above. For some clauses H derived by IE from E under B it is not the case that $B \cup H \models E$. In example 1, the following is such a clause.

$$odd(u) \leftarrow s(v,u), s(w,v), even(w)$$

5 Discussion

Yamamoto [5] demonstrated the incompleteness of the mechanism of IE described in [2]. This paper defines an enlargement of the bottom set definitions of [2, 5]. In fact this enlargement is a generalisation of the use of sub-saturants in [2]. Although the enlarged bottom set is infinite in the non-function-free case, a polynomial bound to the cardinality of the enlarged bottom set is provided in the case of function-free Horn background theories. Example 1 shows how the enlarged bottom set deals with a function-free form of the example Yamamoto used to show the incompleteness of IE. Finally a completeness result is given for definite clauses for the revised version of IE.

It is an open question as to whether a further generalisation of this approach would be complete for arbitrary clausal background theories.

Acknowledgements

The author would like to thank his wife Thirza and daughter Clare for their support during the writing of this paper. Thanks also to David Page, James Cussens, Alan Frisch, Akihiro Yamamoto and Koichi Furukawa for relevant discussions on this topic. This work was supported partly by the Esprit Long Term Research Action ILP II (project 20237), EPSRC grant GR/K57985 on Experiments with Distribution-based Machine Learning and an EPSRC Advanced Research Fellowship held by the author.

References

1. J.W. Lloyd. *Foundations of Logic Programming.* Springer-Verlag, Berlin, 1987. Second edition.
2. S. Muggleton. Inverse entailment and Progol. *New Generation Computing*, 13:245–286, 1995.
3. S-H. Nienhuys-Cheng and R. de Wolf. *Foundations of Inductive Logic Programming.* Springer-Verlag, Berlin, 1997. LNAI 1228.
4. U. Nilsson and J. Maluszynski. *Logic, Programming and Prolog.* Wiley, New York, 1995. Second edition.
5. A. Yamamoto. Which hypotheses can be found with inverse entailment? In N. Lavrac and S. Dzeroski, editors, *Proceedings of the Seventh International Workshop on Inductive Logic Programming*, pages 296–308. Springer-Verlag, Berlin, 1997. LNAI 1297.