Disjunctive Bottom Set and Its Computation

Wenjin Lu and **Ross King**
Department of Computer Science, University of Wales, Aberystwyth
Ceredigion, SY23 3DB, Wales, UK
e-mail:{wwl, rdk}@aber.ac.uk

Abstract

This paper presents the concept of the disjunctive bottom set and discusses its computation. The disjunctive bottom set differs from existing extensions of the bottom set, such as kernel sets(Ray, Broda, & Russo 2003), by being the weakest minimal single hypothesis for the whole hypothesis space. The disjunctive bottom set may be characterized in terms of minimal models. Therefore, as minimal models can be computed in polynomial space complexity, so can the disjunctive bottom set. We outline a flexible inductive logic programming framework based on the disjunctive bottom set. Compared with existing systems based on bottom set, such as Progol (Muggleton 1995), it can probe an enlarged hypothesis space without increasing space complexity. Another novelty of the framework is that it provides an avenue, via hypothesis selection function, for the integration of more advanced hypothesis selection mechanisms.

Introduction

Inverse Entailment(IE) (Muggleton 1995) is one of the most important inference mechanisms in inductive logic programming (ILP). It is an inverse process of deductive reasoning. Formally, given a background knowledge B and an example E and $B \not\models E$, IE will work out a set of rules H such that

$$B \wedge H \models E$$

In practice, a typical implementation of IE will include the following modules:

1. **Bottom set computation**: The bottom set of E under B, is defined as a specific (ground) clause set whose negation is derivable from $B \wedge \bar{E}$.
2. **Bottom set generalisation**: This constructs a clausal theory H such that every clause in the bottom set is θ-subsumed by a clause in H.

Therefore the concept of a bottom set is a key idea in the implementation of Inverse Entailment(IE) (Muggleton 1995). The composition of the bottom set has a critical effect on what kind of hypotheses can be induced by the system.

Example 1 *Given background knowledge B and an example E as follows,*

$$\begin{aligned} B = \ & \{\, b \rightarrow a, \\ & c \wedge d \rightarrow a, \\ & e \rightarrow c, \\ & f \rightarrow c, \\ & g \rightarrow d, \\ & h \rightarrow d \,\} \\ E = \ & a \end{aligned}$$

the possible (minimal) inductive hypotheses are:

(1) $\{a\}$,
(2) $\{b\}$,
(3) $\{c,\ d\}$,
(4) $\{c,\ g\}$,
(5) $\{c,\ h\}$,
(6) $\{d,\ e\}$,
(7) $\{d,\ f\}$,
(8) $\{e,\ g\}$,
(9) $\{e,\ h\}$,
(10) $\{f,\ g\}$,
(11) $\{f,\ h\}$,
(12) $\{a \vee b\}$,
(13) $\{a \vee b \vee c \vee e \vee f,\ \ a \vee b \vee d \vee g \vee h\}$

Depending on the selection of bottom set, existing ILP systems may deliver different solutions. As it is limited to single Horn clause hypotheses, the Progol family takes (12) as the bottom set and may deliver (1) or (2) as hypotheses. The HAIL system (Ray, Broda, & Russo 2003) allows hypotheses consisting of many Horn clauses and takes (1), (2), ..., (11) all together as the bottom set. It may deliver as hypothesis (1) to (11) but neither (12) nor (13). Hypothesis (13), however, does possess some desirable properties as a bottom set:

- it is a minimal hypothesis in the sense that no proper subset of (13) is a hypothesis.
- it is the weakest hypothesis in the sense that it is subsumed by all other hypotheses.
- it is complete in the sense that all other hypothesis can be obtained from (13) by selecting some literals from each clause in it. Therefore it represents the multi-solution in a compact way.

This observation has led us to introduce the concept of the disjunctive bottom set which is defined as the weakest minimal ground hypothesis[1] for given background knowledge B and an example E. In addition to the properties listed above, the disjunctive bottom set also has the following advantages:

- it can be characterised by the minimal models of a simple duality transformation of B and E.
- With some restriction on the syntax of B, the disjunctive bottom set can be computed in polynomial space complexity, as this is possible for minimal model computation.

The rest of the paper is organized as follows. After introducing some preliminaries in the next section, in section of Disjunctive Bottom Set, we present the definition of the disjunctive bottom set. Section On Computation of the Disjunctive Bottom Set discusses the issues of computing the disjunctive bottom set. The comparison with related work is presented in the section of Related work. We conclude the paper in the last section by discussing some future work

Preliminaries

In this section, based on the assumption of familiarity with first order logic and logic programming (Lloyd 1987), we give a brief review on the inverse entailment and its variants.

Given a first order language $\mathcal{L}$, here are the necessary notation and terminology. A positive literal is an atom and a negative literal is the negation of an atom. A ground literal is a literal without variables. We denote $HB(\mathcal{L})$ the Herbrand base of $\mathcal{L}$, the set of all ground atoms formed from $\mathcal{L}$. The disjunctive Herbrand base, denoted as $dHB(\mathcal{L})$, is the set of all (finite) positive ground disjunctions formed from the elements of the Herbrand base $HB(\mathcal{L})$. The set of all ground literals of $\mathcal{L}$ is denoted by $GL(\mathcal{L})$. A clause is a disjunction of literals where all variables in the clause are (implicitly) universally quantified. Conventionally, a clause is also represented as a set of literals which means a disjunction of the literals in the set. In logic programming setting, a clause C is written as

$$B_1 \wedge ... \wedge B_n \rightarrow A_1 \vee ... \vee A_m$$

where $m, n \geq 0$ and A_i, B_i are atoms. A Horn clause is a clause containing at most one positive literal, that is, $m \leq 1$. A (Horn) clausal theory is a conjunction of (Horn) clauses. Given C as above, $\overline{C} = (B_1 \wedge ... \wedge B_n \wedge \neg A_1 \wedge ... \wedge \neg A_m)\sigma$ is called the complement of C, where σ is a Skolemising substitution for C.

Given a clausal theory B, an (Herbrand) interpretation of B is a subset of the Herbrand base. Given an interpretation I, a ground clause $C = B_1 \wedge ... \wedge B_k \rightarrow A_1 \vee ... \vee A_l$ is true in the I iff $\{B_1, ..., B_k\} \subseteq I$ implies $\{A_1, ..., A_l\} \cap I \neq \emptyset$, denoted as $I \models C$. I is a model of B iff all clauses in B are true in I. A model M of B is minimal model iff there is no model M_1 of B such that $M_1 \subset M$. The set of all minimal models of B is denoted by $\mathcal{MM}(B)$.

[1] see definitions in the section Disjunctive Bottom Set

The central task of ILP is to find a hypothesis H from given background knowledge B and examples E such that

$$B \wedge H \models E$$

where H, B and E are all finite clausal theories. Inverse Entailment fulfills this task by so-called bottom generalisation, which is, in turn, based on bottom set (Muggleton 1995). The following definitions and notations are taken from (Yamamoto 1997) with B and E are limited to a Horn theory and a Horn clause, respectively.

Definition 1 (Muggleton's bottom set) *Let B be a Horn theory and E be a Horn clause. Then the bottom Set of B and E is the clause*

$$bot(B, E) = \{L \mid L \in GL(\mathcal{L}) \text{ and } B \wedge \overline{E} \models \neg L\}$$

Denoting $bot^+(B, E)$ the set of atoms in $bot(B, E)$ and $bot^-(B, E)$ the set of atoms whose negation is in $bot(B, E)$, then we have

$$bot(B, E) \equiv \bigwedge bot^-(B, E) \rightarrow \bigvee bot^+(B, E)$$

Definition 2 (Bottom generalisation) *Let B be a Horn theory and E be a Horn clause. A Horn clause H is said to be derivable by bottom generalization from B and E iff H θ-subsumes $bot(B, E)$.*

For computational purpose, Bottom set has been rephrased in (Yamamoto 1997) in terms of deductive and abductive reasoning. In the following, without loss of generality, we assume that example E is a ground atom, as in the case E is a Horn clause, normalisation process can be applied [2].

Proposition 1 *Given Horn theory B and ground atom E with $B \not\models E$. Then*

$$bot^-(B, E) = \{a \mid a \in HB(\mathcal{L}) \text{ and } B \models a\}$$
$$bot^+(B, E) = \{b \mid b \in HB(\mathcal{L}) \text{ and } B \wedge \{b\} \models E\}$$

The interesting point with this reformulation is that it explicitly reveals the relationship between inductive logic programming and abductive logic programming, that is, $bot^+(B, E)$ can be generated by employing an abductive procedure to abduce all single atom hypotheses (assuming that all atoms are abducible). As indicated in (Ray, Broda, & Russo 2003), however, Muggleton's bottom set is incomplete due to its restriction to single clause hypotheses. This has led to a further generalisation of the bottom set by allowing abductive hypotheses with multiple atoms (Ray, Broda, & Russo 2003; 2004), which provides a semantic underpinning to a larger hypothesis space than that computed using Muggleton's bottom set.

Definition 3 (Kernel, Kernel generalisation) *Let B be a Horn theory and E a ground atom with $B \not\models E$. Then the Kernel of B and E, written as $\mathcal{K}er(B, E)$, is the formula defined as follows:*

$$\mathcal{K}er(B, E) \equiv \bigwedge \mathcal{K}er^-(B, E) \rightarrow \bigvee \mathcal{K}er^+(B, E)$$

[2] Given a Horn theory B and Horn clause $E = a_1 \wedge ... \wedge a_n \rightarrow b$, $\mathcal{B} = B \wedge a_1\sigma \wedge ... \wedge a_m\sigma$ and $\epsilon = b\sigma$ is called a normalisation of B and E, where σ is a Skolemising substitution for E (Ray, Broda, & Russo 2003)

where

$$\begin{array}{ll} \mathcal{K}er^-(B,E) = & \{a \mid a \in HB(\mathcal{L}) \text{ and } B \models a\} \\ \mathcal{K}er^+(B,E) = & \{\Delta \mid \Delta \subseteq HB(\mathcal{L}) \text{ and } B \wedge \Delta \models E\} \end{array}$$

A Horn theory H is said to be derivable by Kernel Generalisation iff $H \models \mathcal{K}er(B,E)$.

It has been shown that kernel generalisation is sound in the sense that give B and E as above, for any Horn theory H, $H \models \mathcal{K}er(B,E)$ only if $B \wedge H \models E$.

The Disjunctive Bottom Set

This section presents the formal definition of the disjunctive bottom set. We show that for a given background knowledge B and a ground atom E such that $B \not\models E$, there exist an unique weakest hypothesis H such that $B \wedge H \models E$. The disjunctive bottom set is then defined to be this weakest hypothesis. We start with the following simple facts.

Proposition 2 *Let B be a Horn theory and E be a ground atom. Then for $C = c_1 \vee ... \vee c_n \in dHB$, $B \wedge C \models E$ iff $B \wedge c_i \models E$ for all $i = 1, ..., n$.*

Proposition 3 *Let B be a Horn theory and E a ground atom with $B \not\models E$. For any $H \in dHB$, if*
$B \wedge H \models E$, then $H \models \bigvee bot^+(B,E)$.

Proposition 2 and proposition 3 together establish that $bot^+(B,E)$ is simply the weakest positive ground hypothesis consisting of single clause for B and E. For example, the hypothesis (12) in example 1. Considering the fact that Muggleton's bottom set is incomplete due to this limitation; by the above propositions, it would be natural to select the weakest ground hypothesis in the whole hypothesis space as the bottom set. *This is the idea behind the definition of the disjunctive bottom set.* In the following we give a formal account of "the weakest" ground hypothesis.

Definition 4 (Positive ground hypothesis (PGH)) *Let B be a Horn theory and E be a ground atom where $B \not\models E$. A positive ground hypothesis of B and E is a set of positive ground clauses of the form*

$$PH = \{\mathcal{D}_i \mid \mathcal{D}_i \in dHB,\ i = 0, 1, ..., m\}$$

satisfying

$$B \wedge \mathcal{D}_1 \wedge ... \wedge \mathcal{D}_m \models E$$

A positive ground hypothesis PH is called minimal if there is no positive ground hypothesis PH' such that $PH' \subset PH$.

A clausal theory S is said to clausally subsume a clausal theory T, written as $S \sqsupseteq T$, if every clause in T is θ-subsumed by at least one clause in S. If $S \sqsupseteq T$, then we say T is weaker than S.

Definition 5 (Weakest PGH) *Let PH be a minimal positive ground hypothesis of a Horn theory B and a ground atom E where $B \not\models E$. PH is called weakest iff there is no minimal positive ground hypothesis PH' of B and E such that $PH \sqsupseteq PH'$ and $PH \neq PH'$.*

The following lemma shows that all weakest positive ground hypotheses are logically equivalent.

Lemma 1 (Uniqueness of weakest PGH) *Let B be a Horn theory and E be a ground atom where $B \not\models E$. If both H_1 and H_2 are weakest positive ground hypotheses, then $H = H'$*[3].

Proof: Let $H = H_1 \vee H_2$, then $B \wedge H \models E$. Convert H into a conjunctive normal form (CNF) and remove all clauses which are subsumed by others. Let the resulting CNF be H_c, then H_c is a positive ground hypothesis and is weaker than H_1 and H_2. But if H_1 and H_2 both are weakest, we have $H_c \sqsupseteq H_i$ (i = 1, 2). As H_1, H_2 and H_c are all positive ground, we have $H_1 = H_c = H_2$.

For a given Horn theory B and an example E satisfying $B \not\models E$, we still need to show the existence of the weakest positive ground hypothesis. To fulfill this task, we borrow the approach and results from (Yahya 2002) which discusses the duality for goal-driven query processing in disjunctive deductive databases. The interesting point for us is that it shows that the weakest minimal hypothesis can be obtained by computing the minimal models of a duality transformation of B and E. The following result taken from (Yahya 2002) has been tailored and rephrased according to our needs. A more general version and its proof can be found in (Yahya 2002).

Definition 6 (Dual clause (Yahya 2002)) *Let $C = B_1 \wedge ... \wedge B_k \rightarrow A_1 \vee ... \vee A_l$ be a clause, the dual clause of C, denoted by C^d, is a clause of the form*

$$C^d = A_1 \wedge ... \wedge A_l \rightarrow B_1 \vee ... \vee B_m$$

The dual of a set of clauses S is the set S^d of duals of each of the members of S.

Theorem 1 ((Yahya 2002)) *Let B be a Horn theory and E be a ground atom. Let $B^d_E = B^d \cup \{E\}$. If $\mathcal{MM}(B^d_E)$ is non empty, then*

- *$B \not\models E$*
- *E becomes derivable from the updated clause theory B' achieved by adding to B the set of clauses S such that $S \sqsupseteq \mathcal{MM}(B^d_E)$.*
- *$S = \mathcal{MM}(B^d_E)$ is the minimal and weakest such set that can be added to B to guarantee the derivablility of E from B'.*

The following corollary clarifies the relationship between minimal models and positive ground disjunctive hypotheses.

Corollary 1 (Existence of weakest PGH) *Let B be a Horn theory and E be a ground atom with $B \not\models E$. Then $S = \mathcal{MM}(B^d \cup \{E\})$ is the weakest minimal positive ground hypothesis.*

Example 2 *Let B and E be as in example 1, then*

$$\begin{array}{ll} B^d = \{ & a \rightarrow b, \\ & a \rightarrow c \vee d, \\ & c \rightarrow e, \\ & c \rightarrow f, \\ & d \rightarrow g, \\ & d \rightarrow h\} \end{array}$$

[3] here we read a ground hypothesis as a set of clauses, which, in turn, are sets of ground atoms.

$\mathcal{MM}(B^d \cup \{a\}) = \{\{a,b,c,e,f\},\{a,b,d,g,h\}\}$. *As* $bot^-(B,E) = \emptyset$*, we have*

$$B \cup \{a \vee b \vee c \vee \vee e \vee f,\ a \vee b \vee d \vee \vee g \vee h\} \models a$$

With lemma 1 and corollary 1, we have the following theorem.

Theorem 2 *Let B be a Horn theory and E be a ground atom with $B \not\models E$. Then there exists an unique weakest minimal positive ground hypothesis.*

With these results, we are now in a position to present the definition of the disjunctive bottom set.

Definition 7 (Disjunctive bottom set) *Let B be a Horn theory and E a ground atom with $B \not\models E$. Let WPH be the weakest minimal positive ground hypothesis of B and E. The disjunctive bottom set of B and E is a clausal theory of the form*

$$dBot(B,E) = \{bot^-(B,E) \rightarrow D | D \in WPH\}$$

Example 3 *Let B and E be as in example 1, By example 2 and $bot^-(B,E) = \emptyset$, we have*

$$dBot(B,E) = \{a \vee b \vee c \vee e \vee f,\ a \vee b \vee d \vee g \vee h\}$$

In the following, we show that the disjunctive bottom set is a real extension of bottom set (theorem 3). The next lemma follows the fact that for any derivation $\mathcal{D}$ of $\neg a$ from $B \wedge \neg E$, we have a derivation $\mathcal{D}^d$ of a from $B^d \wedge E$ obtained by replacing each C in the $\mathcal{D}$ with C^d which is a clause in $B^d \wedge E$.

Lemma 2 *Let B be a Horn theory and E be a ground atom with $B \not\models E$. Then for any ground atom a, $B \wedge \neg E \models \neg a$ iff $B^d \wedge E \models a$.*

Theorem 3 *Let B be a Horn theory and E be a Horn clause. Then $bot(B,E) \sqsupseteq dBot(B,E)$.*

Proof: By lemma 2, for any $a \in bot^+(B,E)$, $B^d \wedge E \models a$. That is, a is true in every minimal model of $B^d \wedge E$. Therefore for every minimal model $M \in \mathcal{MM}(B^d \wedge E)$, $bot^+(B,E) \subseteq M$. Thus the theorem follows the definitions of the bottom set and the disjunctive bottom set.

On Computation of the Disjunctive Bottom Set

In this section we discuss the issues of computing the disjunctive bottom set. By theorem 1 and the definition of the disjunctive bottom set, for given background knowledge B and an example E where $B \not\models E$, the computation of $dBot(B,E)$ turns out to be the generation of minimal models of $B^d \cup E$.

Minimal model computation has been intensively studied in the community of disjunctive logic programming and theorem proving. Many minimal model generation approaches have been proposed in the literature (Niemela 1996; Lu 1997; 1999; Bry & Yahya 2000). Among them the methods based on hyper tableaux seem to offer a promising basis for minimal model reasoning (Niemela 1996; Lu 1997; 1999; Bry & Yahya 2000). The hyper tableau calculus combines the idea from hyper resolution and from analytic tableaux. When applied to minimal model generation, hyper tableaux are defined as a special kind of literal trees. The tree is generated in such a way, that in any step an open branch is a candidate for a partial model.

While it is true that there are many algorithms for minimal model generation, most of them are defined for ground theories or theories with restricted syntax. One such a restriction is that of range restriction clauses (Bry & Yahya 2000).

Definition 8 (Range restricted clause) *A clause is said to be range restricted if every variable occurring in a positive literal also appears in a negative literal. A clause theory is range restricted if every clause in it is range restricted.*

As discussed in (Bry & Yahya 2000), for a non-range restricted clausal theory, a range-restricted transformation can be applied to it to produce a range-restricted clauses theory (Bry & Yahya 2000).

Specifically, for a range restricted clause theory, there exist minimal model generation procedures with a polynomial space complexity. One such procedure is reported in (Niemela 1996). The basic idea is to generate models with a hyper tableau proof procedure and to include an additional test for ruling out those branches in the tableau that do not represent minimal models. This groundedness test is done *locally*, i.e. there is no need to compare a branch with other branches computed previously; hence there is no need to store models. In the following discussion, we will rely on this fact and assume that the minimal model generation procedure provides an API $next\_minimal\_model(B)$, which takes a range-restricted clause theory B and always returns the next minimal model if any without repeating.

Next, under the assumption that for a given background Horn theory B, B^d is range-restricted, we outline an ILP framework based on the disjunctive bottom set. To make the framework more flexible, we introduce the concept of a hypothesis selection function, which will be used to select a ground hypothesis from the disjunctive bottom set.

Definition 9 (Hypothesis selection function) *A hypothesis selection function is a mapping*

$$f : 2^{HB} \rightarrow 2^{HB}$$

such that

- $f(\emptyset) = \emptyset$
- *if $M \neq \emptyset$, then $f(M) \neq \emptyset$ and $f(M) \subseteq M$*

f is called a Horn hypothesis selection function if $f(M)$ contains only one atom.

Algorithm 1 presents a computational procedure to compute inductive hypotheses. The basic idea behind the procedure is as follows: for a given Horn theory B, a ground atom E, as B^d is assumed to be range-restricted, a minimal model generation procedure can be applied to generate all minimal models of $B^d \cup \{E\}$. For each minimal model, apply the hypothesis selection function to produce a partial hypothesis. This partial hypothesis is then generalised by a hypothesis generalising procedure. Once the algorithm terminates, it will produce an inductive hypothesis for E.

The following theorem shows that algorithm 1 is sound.

Algorithm 1 : Computing Inductive hypotheses

Input: A Horn theory B,
A ground atom E,
A hypothesis selection function $\mathcal{F}$
Output: A hypothesis $\mathcal{H}$
begin
$\mathcal{H} = \emptyset$
repeat
$M = next\_minimal\_model(B^d_E)$
if $M \neq$ "no"
let H be a generalisation of
$bot(B,E)^- \rightarrow \mathcal{F}(M)$
$\mathcal{H} = \mathcal{H} \cup \{H\}$
until $M =$ "no"
return $\mathcal{H}$
end

Theorem 4 (Soundness) *Let B be a Horn theory and E a ground atom. If $\mathcal{H}$ is the output of algorithm 1 with the input of the Horn theory B, the ground atom E and the hypothesis selection function f, then $B \wedge \mathcal{H} \models E$.*

Proof: By algorithm 1, a clause H is in $\mathcal{H}$ iff there is a minimal model M such that H is a generalisation of $bot(B,E)^- \rightarrow f(M)$. Then by the definition of disjunctive bottom set, we have $\mathcal{H} \models dBot(B,E)$, which implies that $B \wedge \mathcal{H} \models E$.

The algorithm, in general, is not complete as shown by the example [4]: given $B = \{n(0)\}$ and $E = n(s(s(0))$, $n(X) \rightarrow n(s(X))$ is a hypothesis but can not be produced by algorithm 1. However, the following "weaker" completeness is provable.

Theorem 5 *Let B be a Horn theory and E a ground atom. If a set of Horn clauses $\mathcal{H}$ is a minimal hypothesis of E given B, then there exists a Horn hypothesis selection function f such that for each clause $H = a_1 \wedge ... \wedge a_l \rightarrow b \in \mathcal{H}$, there is a minimal model $M \in \mathcal{MM}(B^d \cup \{E\})$ such that b is a generalisation of $f(M)$.*

Proof: Let $HH = \bigwedge\{b \mid a_1 \wedge ... \wedge a_l \rightarrow b \in \mathcal{H}\}$, then HH is a minimal positive hypothesis. As $\mathcal{MM}(B^d \cup \{E\})$ is the weakest minimal hypothesis, for each minimal model $M \in \mathcal{MM}(B^d \cup \{E\})$, there must be a $b \in HH$ such that b θ-subsumes M, that is $b\theta \in M$. Defining $f(M) = b\theta$, then we get the desired Horn hypothesis selection function f

Related work

The work presented here has been influenced by several existing work. The bottom set concept was first introduced in (Muggleton 1995). As rephrased in (Yamamoto 1997; Ray, Broda, & Russo 2003; 2004), given a background knowledge B and ground atom E, the bottom set $bot(B,E)$ can be represented in two parts, $bot^-(B,E)$ and $bot^+(B,E)$, where $bot^-(B,E)$ is the least Herbrand model of B and $bot^+(B,E)$ is the set of atoms abducible from B and E. By proposition 2 and 3, $bot^+(B,E)$ is nothing but the weakest *single* clause which is an hypothesis for E given B. In this sense, the disjunctive bottom set is a natural extension of the bottom set, as it is the weakest set of clauses which, combined, form a hypothesis for E given B.

The disjunctive bottom set concept has been influenced by the kernel set approach (Ray, Broda, & Russo 2003), which is a generalisation on bottom set. Given B and E, the kernel can be represented as

$$\mathcal{K}er(B,E) \equiv \bigwedge \mathcal{K}er^-(B,E) \rightarrow \bigvee \mathcal{K}er^+(B,E)$$

where

$$\begin{array}{ll} \mathcal{K}er^-(B,E) = & \{a \mid a \in HB(\mathcal{L}) \text{ and } B \models a\} \\ \mathcal{K}er^+(B,E) = & \{\Delta \mid \Delta \subseteq HB(\mathcal{L}) \text{ and } B \wedge \Delta \models E\} \end{array}$$

As the Kernel set is a complete extension of the bottom set, it it is not surprising that, the disjunctive bottom set and the Kernel are semantically equivalent; in the sense that they represented each other in a dual way. More precisely we have the following result.

Theorem 6 (The disjunctive bottom set and the kernel) *Let B be a Horn theory and E be a ground atom where $B \not\models E$. Then*

$$\bigvee \mathcal{K}er^+(B,E) \leftrightarrow \bigwedge \mathcal{MM}(B^d \cup \{E\})$$

Proof: Let $WPH = \bigwedge \mathcal{MM}(B^d \cup \{E\})$, then WPH is a ground clause theory consisting of only positive ground clauses. Let Δ be a model of WPH, then Δ subsumes WPH. As WPH is the weakest hypothesis, we have $B \wedge \Delta \models E$, therefore, $\Delta \in \mathcal{K}er^+(B,E)$. That is, Δ is a model of $\bigvee \mathcal{K}er^+(B,E)$.

Now let δ be a model of $\bigvee \mathcal{K}er^+(B,E)$, then δ must be a hypothesis of E under B. As WPH is the weakest hypothesis, we have δ subsumes WPH. Therefore δ is a model of WPH. This completes our proof.

While it is true that the disjunctive bottom set and the kernel are semantically equivalent, the differences between the two are also clear. The $\mathcal{K}er^+(B,E)$ is defined as a set of hypotheses consisting of ground atoms as each Δ is a hypothesis. The disjunctive bottom set is a single hypothesis. The difference in representation has an impact on their implementation. The kernel set approach has its implementation based on abductive reasoning, while the ILP framework presented here will be implemented on top of minimal model reasoning and thereby share its advantage of lower space complexity.

Another interesting ILP framework is CF-induction (Inoue 2004). It is also sound and complete for finding hypotheses from full clausal theories, and can be used for inducing not only definite clauses but also non-Horn clauses and integrity constraints. The main difference between CF-induction and our framework is the way in which the hypotheses are computed. CF-induction computes hypotheses using a resolution method via consequence finding. Our framework is based on minimal model generation. Another

[4]The example was mentioned by O. Ray via private communication

difference is in dealing with bias. While it is modelled in CF-induction by a production field, inductive bias can be represented in our framework via more general hypothesis selection function.

Conclusions and Future Work

This paper presents the disjunctive bottom set which is a natural extension of Muggleton's bottom set. Different from existing extensions, the disjunctive bottom set is the weakest minimal hypothesis and can be represented by the minimal models of a duality transformation of background knowledge B and an example E. In addition, the disjunctive bottom set can be computed in polynomial space complexity. An ILP framework based on the disjunctive bottom set is also outlined. The main novelty of the new framework is that it can explore an enlarged hypothesis space without increasing space complexity. In addition the hypothesis selection function in the framework leaves an opening to integrate more advanced hypothesis selection mechanism in hypothesis construction.

Much work remains to be done. Firstly we will prototype the framework for experiment and compare the results with existing work. The other point we want to exploit further is to cooperate statistical methods into the hypothesis selection. An interesting application area will be bioinformatics, where ILP has shown great success (King 2004).

Acknowledgements

We would like to thank the three anonymous reviewers for their helpful comments. This work was partially funded by the European Union IST Programme, contract no. FP6-516169 IQ project and BBSRC Project BLID.

References

Bry, F., and Yahya, A. 2000. Positive unit hyperresolution tableaux and their application to minimal model generation. *J. Autom. Reason.* 25(1):35–82.

Inoue, K. 2004. Induction as consequence finding. *Machine Learning* 55(2):109–135.

King, R. D. 2004. Applying inductive logic programming to predicting gene function. *AI Magazine* 25(1):57–68.

Lloyd, J. 1987. *Foundations of Logic Programming*. Berlin: Springer-Verlag. Second edition.

Lu, W. 1997. Minimal model generation based on e-hyper tableaux. In Brewka, G.; Habel, C.; and Nebel, B., eds., *KI-97: Advances in Artificial Intelligence, 21st Annual German Conference on Artificial Intelligence, Freiburg, Germany, September 9-12, 1997, Proceedings*, volume 1303 of *Lecture Notes in Computer Science*, 99–110. Springer.

Lu, W. 1999. Nonmonotonic Reasoning Based on Minimal Models and Its Implementation. Koblenz: Fölbach Verlag. Koblenzer Schriften zur Informatik 11, Dissertation.

Muggleton, S. 1995. Inverse entailment and Progol. *New Generation Computing, Special issue on Inductive Logic Programming* 13(3-4):245–286.

Niemela, I. 1996. A tableau calculus for minimal model reasoning. In *Analytic Tableaux and Related Methods*, 278–294.

Ray, O.; Broda, K.; and Russo, A. 2003. Hybrid abductive inductive learning: a generalisation of progol. In Horváth, T., and Yamamoto, A., eds., *13th International Conference on Inductive Logic Programming*, volume 2835 of *Lecture Notes in AI*, 311–328. Springer Verlag.

Ray, O.; Broda, K.; and Russo, A. 2004. Generalised Kernel Sets for Inverse Entailment. In Demoen, B., and Lifschitz, V., eds., *Proceedings of the 20th International Conference on Logic Programming*, volume 3132 of *Lecture Notes in Computer Science*, 165–179. Springer Verlag.

Yahya, A. H. 2002. Duality for goal-driven query processing in disjunctive deductive databases. *Journal of Automated Reasoning* 28(1):1–34.

Yamamoto, A. 1997. Which hypotheses can be found with inverse entailment? In Džeroski, S., and Lavrač, N., eds., *Proceedings of the 7th International Workshop on Inductive Logic Programming*, volume 1297, 296–308. Springer-Verlag.