Inverse Resolution as Belief Change

Maurice Pagnucco
ARC Centre of Excel. for Autonomous Sys.
School of Comp. Science and Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia.
Email: morri@cse.unsw.edu.au

David Rajaratnam
School of Comp. Science and Engineering
The University of New South Wales
Sydney, NSW, 2052, Australia.
Email: daver@cse.unsw.edu.au

Abstract

Belief change is concerned with modelling the way in which a rational reasoner maintains its beliefs as it acquires new information. Of particular interest is the way in which new beliefs are acquired and determined and old beliefs are retained or discarded. A parallel can be drawn to symbolic machine learning approaches where examples to be categorised are presented to the learning system and a theory is subsequently derived, usually over a number of iterations. It is therefore not surprising that the term '*theory revision*' is used to describe this process [Ourston and Mooney, 1994]. Viewing a machine learning system as a rational reasoner allows us to begin seeing these seemingly disparate mechanisms in a similar light.

In this paper we are concerned with characterising the well known *inverse resolution* operations [Muggleton, 1987; 1992] (and more recently, *inverse entailment* [Muggleton, 1995]) as AGM-style belief change operations. In particular, our account is based on the abductive expansion operation [Pagnucco *et al.*, 1994; Pagnucco, 1996] and characterised by using the notion of epistemic entrenchment [Gärdenfors and Makinson, 1988] extended for this operation. This work provides a basis for reconciling work in symbolic machine learning and belief revision. Moreover, it allows machine learning techniques to be understood as forms of non-monotonic reasoning.

The study of belief change is concerned with how a rational reasoner maintains its beliefs in the face of new information. As this new information is acquired it must be assimilated into the reasoner's stock of beliefs. This requires that decisions be made as to which beliefs are retained, which are abandoned and which are incorporated by the reasoner. In this paper we shall consider belief change that adheres to the popular AGM [Gärdenfors, 1988] paradigm.

Symbolic machine learning deals with the generalisation of data for the purpose of classification and categorisation. Here we find the closely related notion of *theory revision* [Ourston and Mooney, 1994]. It too can be viewed as a form of belief change; one in which the aim is to assimilate the new information by generalising the reasoner's belief corpus.

In this paper we investigate inductive logic programming as a form of belief change. In particular, we consider how the inverse resolution operators introduced by Muggleton [Muggleton, 1987; Muggleton and Buntine, 1988; Muggleton, 1992] can be re-constructed as belief change operators. Our strategy is to view the machine learning process as one in which the agent's belief corpus represents its current theory. Examples are considered consecutively as new information that causes the agent to revise its belief corpus according to one of the operations. This is essentially the approach adopted in theory revision.

Since the inductive process that inverse resolution is trying to capture is an unsound form of logical inference (i.e., the conclusions do not *necessarily* follow from the premises) we require a form of belief change that allows for this type of inference. Unfortunately, the AGM operators for belief change do not cater for this inferential process. The problem stems from the fact that the inductive process requires the reasoner to formulate and accept hypotheses; hypotheses that go beyond the content carried by the new information (examples) and current belief corpus. As a result we use the notion of *abductive expansion* introduced by Pagnucco [Pagnucco *et al.*, 1994; Pagnucco, 1996] and formulate the required conditions to be satisfied by the epistemic entrenchment relation in order for the inverse resolution operators to be carried out.

The aim of this paper is to take a step towards reconciling the areas of symbolic machine learning and belief change. To that end, we show how the various inverse resolution operations—absorption, identification, interconstruction, intra-construction and truncation—can be viewed as AGM-style belief change operations. In particular, we show how the inverse resolution operations can be captured as forms of abductive expansion by providing the required restrictions on the epistemic entrenchment relation.

In the next section we briefly survey AGM belief change, abductive expansion and the inverse resolution operators. In Section 2 we provide the necessary conditions for each inverse resolution operation. Our approach is discussed in Section 3 and suggestions for future work provided in Section 4.

1 Background

We consider an underlying classical propositional logic with a propositional language $\mathcal{L}$, over a set of atoms, or propositional letters, $\mathbf{P} = \{a, b, c, \ldots\}$, and truth-functional connectives $\neg$, $\wedge$, $\vee$, $\rightarrow$, and $\leftrightarrow$. $\mathcal{L}$ also includes the truth-functional constants $\top$ and $\bot$. The closure operator Cn denotes classical propositional consequence. We use the term *belief set* to refer to a set of formulas closed under Cn (i.e., $K = Cn(K)$). $\mathcal{K}$ represents the set of all belief sets. The distinguished element $K_\perp \in \mathcal{K}$ denotes the inconsistent belief set.

We shall also adopt the following linguistic conventions to simplify the presentation. Lower case Greek characters $\phi,\ \psi,\ \chi,\ \ldots$ will represent arbitrary logical formulas. Lower case Roman characters $k,\ l,\ \ldots$ will denote propositional literals. Upper case Roman characters $A,\ B,\ C,\ \ldots$ denote propositional Horn clauses (disjunctions of literals containing at most one positive literal).

1.1 Belief Change

The mostly widely accepted approach to belief change is the one proposed by Alchourrón, Gärdenfors and Makinson [Alchourrón *et al.*, 1985; Gärdenfors, 1988; Hansson, 1999]. This approach is motivated by the concern to characterise *rational* belief change operators. That is, operations of belief change that are guided by principles of rationality. The AGM approach introduces three belief change operations: *expansion* $(+)$, *contraction* $(-)$, and *revision* $(*)$. Expansion deals with the addition of new information without retraction of existing beliefs; contraction takes care of the removal of belief; and, revision deals with the addition of beliefs with the possibility of retracting current beliefs in order to maintain consistency. Each of these operations takes a belief set representing the reasoner's original belief corpus and a sentence representing the new information, returning the modified belief corpus: $(+,-,* : \mathcal{K} \times \mathcal{L} \to \mathcal{K})$. The belief change operations are characterised by rationality postulates and various constructions. The interested reader is referred to [Gärdenfors, 1988; Hansson, 1999] for further details.

In this paper we are concerned with the process of belief expansion as this is the typical setting under which inverse resolution is applied. This should not be seen as a significant restriction because belief revision proper can be obtained by combining contraction and expansion. However, AGM expansion is very limited in scope. It can be shown that AGM expansion corresponds to closure under logical deduction: $K+\phi = Cn(K \cup \{\phi\})$. What we require is a form of belief change that elaborates upon the newly acquired information. Such an operation has been introduced by Pagnucco [Pagnucco, 1996] in the form of an operation called *abductive expansion* $(\oplus : \mathcal{K} \times \mathcal{L} \to \mathcal{K})$. This operation looks to amplify the newly acquired information by trying to find an explanation for it (via abduction) and is reminiscent of the way that inverse entailment is defined [Muggleton, 1992]. Formally, abductive expansion may be defined as follows:

Definition 1.1 *$K \oplus \phi$ is an abductive expansion of K with respect to ϕ iff*

$$K \oplus \phi = \begin{cases} Cn(K \cup \{\psi\}) & \text{for some } \psi \in \mathcal{L} \text{ such that:} \\ & \quad \text{(i) } K \cup \{\psi\} \vdash \phi \\ & \quad \text{(ii) } K \cup \{\psi\} \nvdash \perp \\ K & \text{if no such } \psi \text{ exists} \end{cases}$$

That is, when confronted with new information ϕ the reasoner seeks an hypothesis ψ that explains ϕ with respect to its current beliefs and incorporates this explanation into its corpus. It can be shown that the operation of abductive expansion satisfies the following rationality postulates:

(K$\oplus$1) For any sentence ϕ and any belief set K, $K \oplus \phi$ is a belief set
(K$\oplus$2) If $\neg\phi \notin K$, then $\phi \in K \oplus \phi$
(K$\oplus$3) $K \subseteq K \oplus \phi$
(K$\oplus$4) If $\neg\phi \in K$, then $K \oplus \phi = K$
(K$\oplus$5) If $\neg\phi \notin K$, then $\neg\phi \notin K \oplus \phi$
(K$\oplus$6) If $K \vdash \phi \leftrightarrow \psi$, then $K \oplus \phi = K \oplus \psi$
(K$\oplus$7) $K \oplus \phi \subseteq Cn(K \oplus (\phi \vee \psi) \cup \{\phi\})$
(K$\oplus$8) If $\neg\phi \notin K \oplus (\phi \vee \psi)$, then $K \oplus (\phi \vee \psi) \subseteq K \oplus \phi$

Postulate (K$\oplus$1) maintains the expanded belief corpus as a belief set. (K$\oplus$2) states that the new information should be included in the expanded belief state provided that it is consistent with the original corpus while (K$\oplus$3) requires that the expanded corpus be at least as large as the original. When the new information is inconsistent with the corpus, (K$\oplus$4) prevents change. Postulate (K$\oplus$5) ensures that the corpus does not expand into inconsistency. (K$\oplus$6) states that the expansion process is syntax insensitive. These postulates suffice to characterise abductive expansion as defined above. Often postulates (K$\oplus$7) and (K$\oplus$8) are added and restrict the mechanism that selects hypotheses to one that satisfies transitivity.

An elegant way of constructing an abductive expansion operation is by providing an ordering of *epistemic entrenchment* over the sentences in the language that indicates their plausibility or importance. For AGM operations, entrenchment orders the beliefs while the non-beliefs are relegated together as the least important sentences. For abductive expansion however, we need to be able to discriminate between potential hypotheses and so the traditional epistemic entrenchment ordering is modified so that beliefs are lumped together as the most important while the non-beliefs are ordered by importance in what we term an *abductive entrenchment ordering*.

Definition 1.2 *An ordering $\leq$ over $\mathcal{L}$ is an abductive entrenchment ordering iff it satisfies (SEE1)–(SEE3) and condition (AE4):*

(SEE1) *For any $\phi,\ \psi,\ \chi \in \mathcal{L}$, if $\phi \leq \psi$ and $\psi \leq \chi$ then $\phi \leq \chi$*
(SEE2) *For any $\phi,\ \psi \in \mathcal{L}$, if $\{\phi\} \vdash \psi$ then $\phi \leq \psi$*
(SEE3) *For any $\phi,\ \psi \in \mathcal{L}$, $\phi \leq \phi \wedge \psi$ or $\psi \leq \phi \wedge \psi$*
(AE4) *When $K \neq K_\perp$, $\phi \in K$ iff $\psi \leq \phi$ for all $\psi \in \mathcal{L}$*

This ordering is a total preorder (SEE1)–(SEE3) in which the beliefs are maximally entrenched (AE4); it effectively *ranks* the reasoner's non-beliefs.

The following properties of abductive epistemic entrenchment will be useful in proving the results in this paper.

Lemma 1.1 *Suppose $\leq$ satisfies (SEE1)–(SEE3), then*[1]

1. *For all $\phi \in \mathcal{L}$ either $\phi \leq \psi$ or $\neg\phi \leq \psi$ for all $\psi \in \mathcal{L}$.*
2. *For all $\phi \in \mathcal{L}$, $\{\psi : \phi \leq \psi\} = Cn(\{\psi : \phi \leq \psi\})$*
3. *$\phi \wedge \psi = min(\phi, \psi)$*
4. *$\phi \vee \psi \geq max(\phi, \psi)$*

[1] Note that $\phi < \psi$ iff $\phi \leq \psi$ and $\psi \not\leq \psi$ and $\phi = \psi$ iff $\phi \leq \psi$ and $\psi \leq \psi$.

The first property states that for every sentence ϕ, either ϕ or $\neg\phi$ (possibly both) is minimally entrenched. The second property notes that taking all sentences more entrenched than a certain rank gives a deductively closed set of sentences. The third property indicates that a conjunction is entrenched at the same level as the minimum of the conjuncts while the fourth property states that a disjunction is at least as entrenched as the maximally entrenched disjunct. These are well known properties of orderings satisfying conditions (SEE1)–(SEE3) which are termed *expectation orderings* by Gärdenfors and Makinson [Gärdenfors and Makinson, 1994].

Pagnucco [Pagnucco, 1996] gives a condition that allows us to determine an abductive expansion function $\oplus_{\leq}$ for a particular epistemic state K given an abductive entrenchment ordering $\leq$.

$$(C\oplus) \qquad \psi \in K \oplus_{\leq} \phi \textit{ iff either } \psi \in K \textit{ or } \textit{both } \neg\phi \notin K \textit{ and } \phi \rightarrow \neg\psi < \phi \rightarrow \psi$$

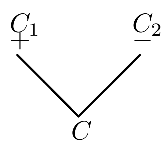
We omit the subscript $\leq$ unless necessary. This condition states that we can accept sentence ψ in the abductively expanded belief state whenever it is believed in the original corpus or when ψ is more plausible given ϕ than $\neg\psi$ given ϕ.

1.2 Inverse Resolution

Resolution is a valid inference procedure which deduces a clause C from two clauses C_1 and C_2. Given a clause C_1 containing a literal l and a clause C_2 containing the literal $\neg l$, the resolvent of C_1 and C_2, denoted $C = C_1.C_2$, is

$$C = (C_1 \setminus \{l\}) \cup (C_2 \setminus \{\neg l\}) \qquad (1)$$

This process may be visualised in the following diagram [Muggleton and Buntine, 1988]:[2]



Inverse resolution (and more recently, inverse entailment), on the other hand, is a machine learning technique based upon the following characterisation of inductive inference [Muggleton, 1987; 1989]. Given a partial domain theory Γ and a positive example E that is not a consequence of the domain theory ($\Gamma \not\vdash E$) we attempt to determine a new domain theory Γ', using Γ and E, that will account for the example ($\Gamma' \vdash E$) and the original domain theory ($\Gamma' \vdash \Gamma$). If we think of Γ' as $\Gamma \cup I$ where I represents the result of inverse resolution, then the relationship with abduction should become much clearer. In practice, the domain theory and example are usually represented as Horn clauses. This technique is based on inverting the resolution process and consists of five operators: two V-operators, two W-operators and the truncation operator.

So as not to countenance invalid inference, the notion of an *oracle* is adopted. An oracle is an entity that accepts a clause, constructed using one of the inverse resolution operators, if it is valid in the intended model. In our framework abductive entrenchment is used to discriminate between potential hypotheses and takes the place of the oracle.

[2]The plus (+) (respectively minus (−)) sign in the diagram denotes that the literal resolved upon appears positive (respectively negative) in that clause.

V-Operators

Previously, a single resolution step was presented in terms of a "V"-shaped diagram. The two V-operators can derive one of the clauses at the top of the V given the other clause at the top of the V and the clause at the bottom of the V. The *absorption* operator derives C_2 given C_1 and C while the *identification* operator derives C_1 given C_2 and C.

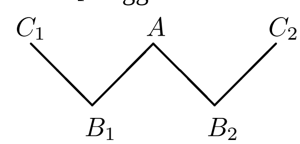
Since the new clause is constructed by finding the inverse of a resolved product, the notion of a resolved quotient[3] of C and C_1 is defined [Muggleton and Buntine, 1988] as $C_2 = C/C_1$. Rearranging equation (1) for resolution can obtain $C_2 = (C \setminus (C_1 \setminus \{l\})) \cup \{\neg l\}$ under the following assumption [Muggleton and Buntine, 1988]:

- *Separability Assumption* — Clauses $C_1 \setminus \{l\}$ and $C_2 \setminus \{\neg l\}$ contain no common literals.

This assumption also simplifies the calculation of resolved quotients (i.e., absorptions or identifications).

W-Operators

Joining two resolution "V"s we obtain a form analogous to that for the V-operators [Muggleton and Buntine, 1988]:



In this situation a common literal l, contained in A, resolves with clauses C_1 and C_2 to produce B_1 and B_2. Clauses B_1 and B_2 represent the new information and clauses A, C_1 and C_2 the constructed clauses. Interestingly, since l is resolved away, the constructed clauses A, C_1 and C_2 will contain a literal whose propositional symbol does not appear in either B_1 or B_2. If l occurs negative in A then the operator is referred to as *intra-construction* and if it occurs positive in A the operator is called *inter-construction*.

Truncation

The *truncation* operator results from the special case where the empty clause occurs at the base of a V or W schemata. In a propositional system, this corresponds to dropping negative literals from a clause. In the first-order case Muggleton and Buntine [Muggleton and Buntine, 1988] show that two literals may be truncated by taking their least-general-generalisation. Rouveriol and Puget [Rouveirol and Puget, 1990] generalise this to a truncation operator which replaces terms by variables and drops literals from clauses.

In the case of a propositional language, a number of schema may be used to compute the required inverse resolution. These are displayed in Table 1 (see [Muggleton, 1987; 1989]).[4] In each instance, the top line represents the initial clauses and the bottom line represents the constructed clause(s). We shall use these directly in our approach to characterise the inverse resolution operators as belief change operators (abductive belief expansion operators in fact).

[3]Absorption is considered here. Identification is similar.

[4]We adopt a slight renaming of the terms to those presented in [Muggleton, 1989], having found them more amenable to study. In the case of absorption and identification, the first clause on the top

Name	*Rule*
Absorption	$\frac{A\to k,\ A\wedge B\to j}{B\wedge k\to j}$
Identification	$\frac{B\wedge k\to j,\ A\wedge B\to j}{A\to k}$
Inter-Construction	$\frac{A\wedge B\to j,\ A\wedge C\to k}{B\wedge l\to j,\ C\wedge l\to k,\ A\to l}$
Intra-Construction	$\frac{A\wedge B\to j,\ A\wedge C\to j}{A\wedge l\to j,\ B\to l,\ C\to l}$
Truncation	$\frac{A\wedge B\to j}{A\to j}$

Table 1: Propositional inverse resolution operators.[5]

2 Approach

Our aim here is to render the individual inverse resolution operations as belief change operators. In particular, we shall give conditions on the abductive epistemic entrenchment that will guarantee that the results of an inverse resolution operation are included in the revised (or, more precisely, abductively expanded) belief corpus.

To make this more precise, we take the reasoner's belief corpus K and an example ϕ as the newly acquired information. We want to construct abductive expansion operations $\oplus$ for each of the five inverse resolution operations such that $K \oplus \phi$ represents the result of applying that particular inverse resolution operation on the given example and elements of the current belief corpus. This construction is achieved by specifying a condition(s) on abductive entrenchment that, when added to conditions (SEE1)–(SEE3) and (AE4), guarantees the result of the inverse resolution operation.

As noted above, abductive entrenchment encodes the 'choices' that are made by the oracle. In other words, entrenchment will be used to determine which potential hypotheses are accepted and which are rejected. Furthermore, it may be the case that in expanding the reasoner's belief corpus in this way, more than one potential hypothesis is accepted for a given example ϕ. This could be easily restricted to one hypothesis by imposing additional restrictions on the entrenchment ordering however, intuitively, this seems procrustean. Multiple generalisations allow for a more flexible characterisation of the inverse resolution process. Finally note that we will only be concerned with examples (new information) that are presented as Horn clauses in keeping with the way in which inverse resolution is traditionally realised.

2.1 Absorption

The typical case for Absorption occurs when $A \to k$ is believed and $A \wedge B \to j$ is presented as an example (newly acquired information). The Absorption operator looks for a sentence $B \wedge k \to j$ to add to the current theory in this case. By $(C\oplus)$, the condition we need to guarantee is therefore:

$$(A\wedge B \to j) \to \neg(B\wedge k \to j) < (A\wedge B \to j) \to (B\wedge k \to j)$$

line of the schemata is taken from the domain theory while the second represents the new data.

[5]Here $A,\ B,\ C$ represent conjunctions of atoms while $j,\ k,\ l$ represent atoms.

What we require here is a restriction on abductive entrenchment that, given $A \to k \in K$, guarantees this condition. Consider the following restriction:

(Abs) $\qquad \bot < B \wedge k \to j$

It states that the sentence $B \wedge k \to j$ is not minimally entrenched (by Lemma 1.1(1), $\bot$ is minimally entrenched). In other words, when an example $A \wedge B \to j$ is presented, any hypothesis of the form $B \wedge k \to j$ that is not minimally entrenched and where $A \to k$ is believed, will be accepted. Now, by Lemma 1.1, condition *(Abs)* is equivalent to $\neg(B \wedge k \to j) < B \wedge k \to j$ which says that this hypothesis is preferred to its negation. The following theorem indicates the appropriateness of this condition.

Theorem 2.1 *Given a belief set $K \neq K_\bot$, a sentence $A \to k \in K$ and an abductive epistemic entrenchment relation $\leq$ satisfying (Abs), the abductive expansion operator $\oplus_\leq$ obtained from $\leq$ via condition $(C\oplus)$ gives the same result as Absorption (i.e., $B \wedge k \to j \in K \oplus (A \wedge B \to j)$).*

Proof: Now $(A \to k) \wedge (B \wedge k \to j) \vdash A \wedge B \to j$ so $(A \to k) \wedge (B \wedge k \to j) \leq A \wedge B \to j$ by (SEE2). But $B \wedge k \to j \leq A \to k$ by (AE4). It follows by Lemma 1.1(3) that $B \wedge k \to j = A \wedge B \to j$. Furthermore $(B \wedge k \to j) \wedge (A \wedge B \to j) = B \wedge k \to j$. But by (Abs) $\bot < (B \wedge k \to j) \wedge (A \wedge B \to j)$. Therefore $\neg((B \wedge k \to j) \wedge (A \wedge B \to j)) = \bot$ by Lemma 1.1(1) and so $\neg((B \wedge k \to j) \wedge (A \wedge B \to j)) < B \wedge k \to j$. It follows by Lemma 1.1(4) $\neg((B \wedge k \to j) \wedge (A \wedge B \to j)) < \neg(A \wedge B \to j) \vee (B \wedge K \to j)$. Hence by logical equivalence $(A \wedge B \to j) \to \neg(B \wedge k \to j) < (A \wedge B \to j) \to (B \wedge k \to j)$ as required. $\square$

Now it may seem strange that condition (Abs) does not mention the new example $A \wedge B \to j$. By Lemma 1.1(2) it can be shown that whenever $A \to k \in K$, it is always the case that $B \wedge k \to j \leq A \wedge B \to j$. That is, the new example cannot be less entrenched than any potential hypothesis.

However, we might re-consider the typical Absorption case and look at what happens when we assume that $A \wedge B \to j \in K$ and $A \to k$ is the new example. In this case we require condition (Abs) together with the following condition:

(Abs') $\qquad B \wedge k \to j \leq A \to k$

The following result indicates this formally:

Theorem 2.2 *Given a belief set $K \neq K_\bot$, a sentence $A \wedge B \to j \in K$ and an abductive epistemic entrenchment relation $\leq$ satisfying (Abs) and (Abs'), the abductive expansion operator $\oplus_\leq$ obtained from $\leq$ via condition $(C\oplus)$ gives the same result as Absorption (i.e., $B \wedge k \to j \in K \oplus (A \to k)$).*

2.2 Identification

The typical case for Identification occurs when $B \wedge k \to j$ is believed and $A \wedge B \to j$ is presented as the new example. Identification returns $A \to k$ to be added to the belief corpus. The condition that we need to guarantee is, by $(C\oplus)$:

$$(A \wedge B \to j) \to \neg(A \to k) < (A \wedge B \to j) \to (A \to k)$$

Analogously to Absorption we require the following restriction on entrenchment to guarantee this condition.

$(Ident)$ $\qquad \bot < A \to k$

It states that any potential hypothesis $A \to k$ is not minimally entrenched and by Lemma 1.1 is equivalent to saying that any potential hypothesis be preferred to its negation. This result is stated formally in the following theorem:

Theorem 2.3 *Given a belief set $K \neq K_\perp$, a sentence $B \wedge k \rightarrow j \in K$ and an abductive epistemic entrenchment relation $\leq$ satisfying $(Ident)$, the abductive expansion operator $\oplus_\leq$ obtained from $\leq$ via condition $(C\oplus)$ gives the same result as Identification (i.e., $A \rightarrow k \in K \oplus (A \wedge B \rightarrow j)$).*

The proof follows a similar idea to that for Absorption and so, due to lack of space, we omit it here.

Again, if we re-consider Identification and suppose that $A \wedge B \rightarrow j \in K$ while $B \wedge k \rightarrow j$ is the new example, then we also require the following condition on entrenchment:

$$(Ident') \qquad A \rightarrow k \leq B \wedge k \rightarrow j$$

2.3 Inter-Construction

In the case of Inter-Construction there is only ever one case to consider because, without loss of generality, we may assume that sentence $A \wedge B \rightarrow j$ is believed and $A \wedge C \rightarrow k$ is presented as the new example. In this case three new clauses are derived ($B \wedge l \rightarrow j$, $C \wedge l \rightarrow k$, and $A \rightarrow l$) and we need to guarantee three conditions which are, by $(C\oplus)$:

$$(A\wedge C \rightarrow k) \rightarrow \neg(B\wedge l \rightarrow j) < (A\wedge C \rightarrow k) \rightarrow (B\wedge l \rightarrow j)$$

$$(A\wedge C \rightarrow k) \rightarrow \neg(C\wedge l \rightarrow k) < (A\wedge C \rightarrow k) \rightarrow (C\wedge l \rightarrow k)$$

$$(A \wedge C \rightarrow k) \rightarrow \neg(A \rightarrow l) < (A \wedge C \rightarrow k) \rightarrow (A \rightarrow l)$$

Given that $A \wedge B \rightarrow j \in K$, the condition we require is:

$$(Inter) \qquad \perp < (B \wedge l \rightarrow j) \wedge (C \wedge l \rightarrow k) \wedge (A \rightarrow l)$$

This condition can be read in a number of ways. Literally it says, in analogy to (Abs) and $(Ident)$, that the conjunction of the potential hypotheses is not minimally entrenched. However, by Lemma 1.1(3), this also means that each of the individual hypotheses themselves is not minimally entrenched and so by Lemma 1.1(1) that they are each preferred to their negations. We now obtain the following result.

Theorem 2.4 *Given a belief set $K \neq K_\perp$, a sentence $A \wedge B \rightarrow j \in K$ and an abductive epistemic entrenchment relation $\leq$ satisfying $(Inter)$, the abductive expansion operator $\oplus_\leq$ obtained from $\leq$ via condition $(C\oplus)$ gives the same result as Inter-construction (i.e., $B \wedge l \rightarrow j,\ C \wedge l \rightarrow k,\ A \rightarrow l \in K \oplus (A \wedge C \rightarrow k)$).*

In inverse resolution, the literal l in the sentences above is newly introduced. It represents a concept that was not present in the original theory. The terms '*predicate invention*' or '*constructive induction*' are often used to describe what the W-operators are achieving by introducing a new element into the object language. Current theories of belief change do not allow for the expansion of the object language so we assume that the language is suitably equipped with a supply of propositions that can be used to create literals like l here. Of course, l will be present in any belief set; tautologies like $l \rightarrow l$ are present in every belief set. However, we can require additional conditions like $l \notin K$ if we want to ensure that l is newly generated. This is not an ideal solution as one is difficult to derive using belief sets but should go some way to achieving the desired results.

2.4 Intra-Construction

In Intra-construction, again without loss of generality, we may assume that $A \wedge B \rightarrow j \in K$ and $A \wedge C \rightarrow j$ is the new example. In this case, by $(C\oplus)$, we need to guarantee the following three conditions:

$$(A\wedge C \rightarrow j) \rightarrow \neg(A\wedge l \rightarrow j) < (A\wedge C \rightarrow j) \rightarrow (A\wedge l \rightarrow j)$$

$$(A \wedge C \rightarrow j) \rightarrow \neg(B \rightarrow l) < (A \wedge C \rightarrow j) \rightarrow (B \rightarrow l)$$

$$(A \wedge C \rightarrow j) \rightarrow \neg(C \rightarrow l) < (A \wedge C \rightarrow j) \rightarrow (C \rightarrow l)$$

Again, the desired result is achieved by ensuring that none of the potential hypotheses is minimally entrenched.

$$(Intra) \qquad \perp < (A \wedge l \rightarrow j) \wedge (B \rightarrow l) \wedge (C \rightarrow l)$$

The following theorem shows that this condition is correct.

Theorem 2.5 *Given a belief set $K \neq K_\perp$, a sentence $A \wedge B \rightarrow j \in K$ and an abductive epistemic entrenchment relation $\leq$ satisfying $(Intra)$, the abductive expansion operator $\oplus_\leq$ obtained from $\leq$ via condition $(C\oplus)$ gives the same result as Intra-construction (i.e., $A \wedge l \rightarrow j,\ B \rightarrow l,\ C \rightarrow l \in K \oplus (A \wedge C \rightarrow j)$).*

2.5 Truncation

For truncation, there are no requirements on the original belief corpus and all we assume is that $A \wedge B \rightarrow j$ is the new information. Truncation generalises this by adding $A \rightarrow j$ to the belief corpus. The condition we need guarantee is:

$$(A \wedge B \rightarrow j) \rightarrow \neg(A \rightarrow j) < (A \wedge B \rightarrow j) \rightarrow (A \rightarrow j)$$

This can be achieved by the following condition stating that $A \rightarrow j$ is not minimally entrenched.

$$(Trunc) \qquad \perp < A \rightarrow j$$

The following theorem indicates that this is correct.

Theorem 2.6 *Given a belief set $K \neq K_\perp$ and an abductive epistemic entrenchment relation $\leq$ satisfying $(Trunc)$, the abductive expansion operator $\oplus_\leq$ obtained from $\leq$ via condition $(C\oplus)$ gives the same result as Truncation (i.e., $A \rightarrow j \in K \oplus (A \wedge B \rightarrow j)$).*

3 Discussion

The idea of using a form of belief expansion capable of introducing new hypotheses has also been suggested by Levi [Levi, 1991] who introduces an operation known as *deliberate expansion*. While his method of determining and selecting hypotheses differs from ours, the notion that belief expansion should allow for the addition of sentences that do not necessarily follow from the new information and the belief corpus is the same. The level to which the reasoner wishes to accept such hypotheses will be determined by their aversion to courting erroneous beliefs and this is reflected in the abductive entrenchment ordering in our framework. Such errors, when realised later, may of course be fixed by subsequent revision. We follow Levi's *commensurability thesis* in maintaining that any form of belief revision can be obtained via a sequence of (abductive) expansions and contractions.

Forms of symbolic machine learning such as inverse resolution are often equated with *inductive* inference (or what might be termed '*enumerative induction*' [Harman, 1968]). However, here we have preferred the notion of abductive inference. There has been a substantial amount of debate in the artificial intelligence community as to what constitutes induction and abduction and whether they are essentially the same form of inference. We do not address this at any great length here. Levi [Levi, 1991] adopts Peirce's later ideas on this subject where abduction is concerned with formulating potential hypotheses and induction selects the best of these. Abductive entrenchment to a large extent serves both purposes here.

One issue that is not strictly addressed by the inverse resolution operators but is important in the wider literature on symbolic machine learning is that of positive and negative examples. Clearly our framework and the inverse resolution operators deal quite well with positive examples. When it comes to negative examples this issue is not so clear. On the one hand new information like $\neg\phi$ can be dealt with quite easily. However, negative information of the form $\phi \notin K$, cannot be dealt with adequately here. It is possible to place restrictions on the entrenchment relation that will not permit ϕ to be believed but it is not clear how this information should be presented. On first sight it would appear that this is more an issue for iterated revision.

4 Conclusions and Future Work

In this paper we have developed an account of how the machine learning technique of inverse resolution, can be viewed as a form of belief change. In particular, we have given restrictions on epistemic entrenchment that will guarantee that the resulting belief change operations behave exactly in the manner of the inverse resolution operations. This is the first step in unifying the study of machine learning and belief change and has important repercussions for the links between machine learning and nonmonotonic reasoning.

This paper opens many avenues for future research in a number of interesting directions. Firstly, the development of belief change postulates that characterise the various inverse resolution operators. Such rationality conditions would allow another way of comparing these operations with belief revision. Perhaps more importantly, at this point it would be relatively straightforward to characterise each of the inverse resolution operators as a nonmonotonic consequence relation in the way of Kraus, Lehman and Magidor [Kraus *et al.*, 1990] allowing comparison with another large body of research.

In this work we have been able to capture each of the inverse resolution operators as a belief change operation. Given an example, we need to state which operation to apply in order to generalise the reasoner's beliefs. A more general form of this theory would not require that the operation be stated but that a single belief change operation represent all inverse resolution operations and that it determine which are applied when new information is given. It is not clear however, how feasible such a *super-selecting* belief change operation is.

One important principle that guides belief change is that of *minimal change*; the reasoner's belief corpus should be modified minimally to accept the new information. It would be interesting to see how this notion in belief change is manifested in work on inverse resolution and symbolic machine learning and the heuristics and metrics applied there to select hypotheses for generalisation. Extending this work to the first-order case and investigating forms of iterated (abductive) belief change are also interesting and important possibilities.

References

[Alchourrón *et al.*, 1985] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[Gärdenfors and Makinson, 1988] P. Gärdenfors and D. Makinson. Revisions of knowledge systems using epistemic entrenchment. In *Proc. 2nd Conf. on Theor. Aspects of Reasoning About Know.*, pages 83–96, 1988.

[Gärdenfors and Makinson, 1994] P. Gärdenfors and D. Makinson. Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65:197–245, 1994.

[Gärdenfors, 1988] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, 1988.

[Hansson, 1999] S. O. Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer, 1999.

[Harman, 1968] G. H. Harman. Enumerative induction as inference to the best explanation. *The Journal of Philosophy*, 65(18):529–533, 1968.

[Kraus *et al.*, 1990] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial intelligence*, 44:167–207, 1990.

[Levi, 1991] I. Levi. *The Fixation of Belief and its Undoing*. Cambridge University Press, 1991.

[Muggleton and Buntine, 1988] S. Muggleton and W. Buntine. Machine invention of first-order predicates by inverting resolution. In *Proc. of the 5th Int. Machine Learning Workshop*, pages 339–352, 1988.

[Muggleton, 1987] S. Muggleton. Duce, an oracle based approach to constructive induction. In *Proc. of the 10th Int. Joint Conf. on AI*, pages 287–292, 1987.

[Muggleton, 1989] S. Muggleton. Inverting the resolution principle. In *Machine Intelligence 12*. Oxford, 1989.

[Muggleton, 1992] S. Muggleton. *Inductive Logic Programming*. Academic Press, 1992.

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

[Ourston and Mooney, 1994] D. Ourston and R. J. Mooney. Theory refinement combining analytical and empirical methods. *Artificial Intelligence*, 66:311–344, 1994.

[Pagnucco *et al.*, 1994] M. Pagnucco, A. C. Nayak, and N. Y. Foo. Abductive expansion: The application of abductive inference to the process of belief change. In *Proc. 7th Aust. Joint Conf. on AI*, pages 267–274, 1994.

[Pagnucco, 1996] M. Pagnucco. *The Role of Abductive Reasoning within the Process of Belief Revision*. PhD thesis, Basser Department of Computer Science, 1996.

[Rouveirol and Puget, 1990] C. Rouveirol and J. F. Puget. Beyond inversion of resolution. In *Proc. of the 7th Int. Machine Learning Workshop*, pages 122–130, 1990.