Brave Induction Revisited

Jianmin Ji

School of Computer Science and Technology
University of Science and Technology of China
Hefei, China
jianmin@ustc.edu.cn

Abstract. Sakama and Inoue introduced brave induction as a novel logic framework for concept-learning. They showed that brave induction has potential applications for problem solving in many domains. In this paper, motivated from Shapiro's definition of model inference problems, we provide an optimization of brave induction called proper brave induction, which prefers hypotheses resulting fewer minimal models. We first propose formal definitions of proper brave induction for clausal theories and nonmonotonic logic programs, then investigate corresponding properties and develop an optimization procedure. At last, we analyze computational complexity of decision problems for proper brave induction in propositional case.

Keywords: Proper brave induction · Brave induction · Inductive logic programming · Nonmonotonic logic programming

1 Introduction

The problem of concept-learning [8] is to construct a general description of a class of objects given a set of examples and non-examples using background knowledge. There are many different logical frameworks for concept-learning, including *explanatory induction* [5] or *learning from entailment* [3] which is considered as a normal setting in *inductive logic programming* (ILP) [9], *learning from satisfiability* (LFS) [3], and *learning from interpretations* (LFI) [3]. Besides, Sakama and Inoue [14] introduced a novel logic framework for concept-learning called brave induction, which allows more hypotheses than explanatory induction and fewer hypotheses than LFS. They showed that brave induction has potential applications for problem solving in systems biology, requirement engineering, and multiagent negotiation.

Example 1. There are 1 teacher and 30 students in a class, of which 20 are European, 7 are Asian, and 3 are American. The situation is represented by background knowledge B and the observation O:

$$B : \mathit{teacher}(0),\ \mathit{student}(1), \ldots, \mathit{student}(30),$$
$$O : \mathit{euro}(1), \ldots, \mathit{euro}(20),\ \mathit{asia}(21), \ldots, \mathit{asia}(27),\ \mathit{usa}(28), \ldots, \mathit{usa}(30),$$

where each number represents a teacher or an individual student. Here are some hypotheses:

$$H_1 : euro(X) \vee asia(X) \vee usa(X) \leftarrow student(X),$$
$$H_2 : euro(X) \vee asia(X) \vee usa(X) \vee teacher(X),$$
$$H_3 : euro(X) \vee asia(X) \vee usa(X) \vee teacher(X) \leftarrow student(X).$$

All of them are allowed by brave induction, while H_1 appears a good hypothesis.

Shapiro [15] discussed logical foundations of inductive learning and defined model inference problems. The intuition behind the definition is that, the "world" is governed by some model M of the language and the inductive learning process is to gather information and correct hypotheses in order to converge to theories that could capture the model M.

If a hypothesis captures more models, then it has more "uncertainties" to capture the "world" model. Motivated from the above intuition, we would prefer hypotheses allowed by brave induction with fewer "uncertainties" to capture the "world" model. In specific, we introduce an optimization of brave induction called proper brave induction, which allows a hypothesis that is allowed by brave induction and there does not exist another such hypothesis whose set of minimal models or answer sets is a proper subset of its. In Example 1, H_1 is allowed by proper brave induction while H_2 and H_3 are not. In the paper, we provide formal definitions of proper brave induction for clausal theories and nonmonotonic logic programs, then investigate corresponding properties and develop an optimization procedure. At last, we analyze computational complexity of decision problems for proper brave induction in propositional case.

2 Preliminaries

In this paper, we assume a function-free first-order language $\mathcal{L}$ with finite sets of constant symbols and predicate symbols, and a countable set of variables. A *term* is either a variable or a constant. A *ground term* is a constant. An *atom* is an expression $p(t_1, \ldots, t_n)$, where p is a predicate symbol with arity $n \geq 1$ and $t_1, \ldots, t_n$ are terms. A *literal* is an atom or the negation of an atom. A *formula* is a propositional combination of atoms. A *ground atom* (resp. *ground formula*) is an atom (resp. formula) that contains no variables.

The *Herbrand universe* of $\mathcal{L}$ is the set of constants and the *Herbrand base* of $\mathcal{L}$ is the set of ground atoms. A *ground instance* of an expression (atom, formula, etc.) is obtained by uniformly instantiating the variables in it with ground terms in the Herbrand universe.

An *interpretation* I is defined as a subset of the Herbrand base of $\mathcal{L}$. I *satisfies* a ground formula F, if I entails F in the sense of classical logic. Moreover, I *satisfies* a (non-ground) formula F, written $I \models F$, if I satisfies every ground instance of F. I *satisfies* a set T of formulas, written $I \models T$, if I satisfies every formula in T. Given sets T_1 and T_2 of formulas, T_1 *entails* T_2, written $T_1 \models T_2$, if every interpretation I satisfies T_1 implies I satisfies T_2.

In the following, we recall the basic notions about clausal theories with the minimal model semantics, answer set programming [1], and brave induction [14].

2.1 Clausal Theories and the Minimal Model Semantics

A *clausal theory* is a finite set of *clauses* of the form:

$$A_1 \vee \cdots \vee A_m \vee \neg A_{m+1} \vee \cdots \vee \neg A_n, \tag{1}$$

where $n \geq m \geq 0$ and $A_1, \ldots, A_n$ are atoms. If $m \leq 1$, it is a *Horn clause*; if $m = n$, it is a *positive clause*; if $A_1, \ldots, A_n$ are ground atoms, it is a *ground clause*. A *Horn clausal theory* is a finite set of Horn clauses and a *ground clausal theory* is a finite set of ground clauses.

An interpretation I *satisfies* a ground clause of form (1) if $\{A_{m+1}, \ldots, A_n\} \subseteq I$ implies $\{A_1, \ldots, A_m\} \cap I \neq \emptyset$. A clause with variables is considered as a shorthand for the set of its ground instantiations. In specific, an interpretation I *satisfies* a (non-ground) clause if I satisfies every ground instance of it. An interpretation I is a *model* of a clausal theory T if I satisfies every clause in T. A model I of a clausal theory T is *minimal* if there does not exist another model I' of T such that I' is a proper subset of I. We use $MM(T)$ to denote the set of minimal models of a clausal theory T. A clausal theory T is *consistent* if $MM(T) \neq \emptyset$; otherwise, T is *inconsistent*.

2.2 Answer Set Programming

Answer set programming (ASP) is one of the most popular rule-based nonmonotonic formalisms [1]. An *ASP program* (or simply a *program*, DLP) is a finite set of *(disjunctive) rules* of the form:

$$A_1 \vee \cdots \vee A_k \leftarrow A_{k+1}, \ldots, A_m, \mathit{not}\ A_{m+1}, \ldots, \mathit{not}\ A_n, \tag{2}$$

where $n \geq m \geq k \geq 0$, $n \geq 1$ and $A_1, \ldots, A_n$ are atoms. If $k \leq 1$, it is a *normal rule*; if $m = n$, it is a *positive rule*. A *normal logic program* (NLP) is a finite set of normal rules and a *positive program* is a finite set of positive rules.

We will also write rule r of form (2) as

$$\mathit{head}(r) \leftarrow \mathit{body}(r),$$

where $\mathit{head}(r)$ is $A_1 \vee \cdots \vee A_k$, $\mathit{body}(r) = \mathit{body}^+(r) \wedge \mathit{body}^-(r)$, $\mathit{body}^+(r)$ is $A_{k+1} \wedge \cdots \wedge A_m$, and $\mathit{body}^-(r)$ is $\neg A_{m+1} \wedge \cdots \wedge \neg A_n$. With a slight abuse of notion, we identify $\mathit{head}(r)$, $\mathit{body}^+(r)$, $\mathit{body}^-(r)$ with their corresponding sets of atoms. We can omit $\leftarrow$ if $\mathit{body}(r)$ is empty.

An interpretation S *satisfies* a ground rule r if $\mathit{body}^+(r) \subseteq S$ and $\mathit{body}^-(r) \cap S = \emptyset$ implies $\mathit{head}(r) \cap S \neq \emptyset$. Similar to clausal theories, a program with variables is considered as a shorthand for the set of its ground instantiations. In specific, an interpretation S *satisfies* a (non-ground) rule if S satisfies every ground instance of it. An interpretation S is a *model* a program P if S satisfies every rule in P, in this case, P is *satisfiable*. A model S of a program P is *minimal* if there does not exist another model S' of P such that $S' \subset S$. We also use $MM(P)$ to denote the set of minimal models of a program P.

Given an ASP program P and an interpretation S, the Gelfond-Lifschitz reduct of P on S, written P^S, is obtained from P by deleting:

- each rule that has a formula $not\, p$ in its body with $p \in S$,
- all formulas of the form $not\, p$ in the bodies of the remaining rules.

Then P^S is a positive program. An interpretation S is an *answer set* of P if $S \in MM(P^S)$. In the following, we use $AS(P)$ to denote the set of answer sets of a program P. P is *consistent* if there exists an answer set of P; otherwise, P is *inconsistent*. P *AS-entails* a formula F (resp. a set T of formulas), written $P \models_{AS} F$ (resp. $P \models_{AS} T$), if for every $S \in AS(P)$, $S \models F$ (resp. $S \models T$).

Notice that, given a clausal theory T, a positive program P can be obtained from T by replacing each clause of form (1) with the positive rule of the form:

$$A_1 \vee \cdots \vee A_m \leftarrow A_{m+1}, \ldots, A_n$$

and $MM(T) = AS(P)$. In the following, we identify a clausal theory T with the corresponding positive program P and a minimal model of T with the corresponding answer set of P.

2.3 Brave Induction

A typical induction task is to construct hypotheses to explain an observation using background knowledge. In specific, one is given a triple $\langle L_b, L_o, L_h \rangle$, where L_b is the language of background knowledge, L_o is the language of observations, and L_h is the language of hypotheses. Here we identify a language with the set of sentences allowed in the language and require that L_b, L_o, L_h are subsets of sentences of $\mathcal{L}$. Given an observation O in L_o and background knowledge B in L_b, a formalization of concept-learning is to define a *cover-relation* from O and B to a hypothesis H in L_h.

Sakama and Inoue [14] proposed a framework of concept-learning called *brave induction* based on the triple $\langle L_{CT}, L_{CT}, L_{CT} \rangle$, where L_{CT} denotes the set of clausal theories in $\mathcal{L}$, and extended the notion to the triple $\langle L_{ASP}, L_{GA}, L_{ASP} \rangle$[1], where L_{ASP} denotes the set of ASP programs and L_{GA} denotes the set of ground atoms in $\mathcal{L}$. Different from existing frameworks, brave induction allows more hypotheses than explanatory induction and fewer hypotheses than LFS. It has been shown that brave induction has potential applications for problem solving in many domains. We recall the basic notions of brave induction here.

Definition 1 (Brave induction). *Let B be background knowledge and O an observation. A hypothesis H covers O under B in* brave induction *if*

[1] In [14], classical negation is allowed in ASP programs and the language of observations is the set of ground literals. However, from [1], an ASP program with classical negation can be equivalently translated to one without classical negation. Without loss of generality, we only consider ASP programs without classical negation here.

– *$B \cup H$ has a minimal model satisfying O, given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$;*
– *$B \cup H$ has an answer set S such that $O \subseteq S$, given the triple $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$.*

In this case, H is called a solution *of brave induction.*

Besides brave induction, there are other frameworks for concept-learning, including *explanatory induction* [5], LFS [3], and *cautious induction* [14].

Definition 2. *Let B be background knowledge and O an observation.*

– *A hypothesis H covers O under B in* explanatory induction *if*
  • *$B \cup H$ is consistent and $B \cup H \models O$, given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$.*
– *A hypothesis H covers O under B in* LFS *if*
  • *$B \cup H$ has a model satisfying O, given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$.*
– *A hypothesis H covers O under B in* cautious induction *if*
  • *$B \cup H$ is consistent and every minimal model of $B \cup H$ satisfies O, given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$;*
  • *$B \cup H$ is consistent and $O \subseteq S$ for every answer set S of $B \cup H$, given the triple $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$.*

H is called a solution *of explanatory induction, LFS, or cautious induction.*

Following propositions summarize the relations between these frameworks.

Proposition 1 (Proposition 2.1 and 6.1 in [14]). *Given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$, let B be background knowledge and O an observation.*

– *If H is a solution of explanatory induction, then H is a solution of cautious induction. The converse implication holds when O is a set of positive clauses.*
– *If H is a solution of cautious induction, then H is a solution of brave induction. The converse implication holds when $B \cup H$ is a Horn clausal theory.*
– *If H is a solution of brave induction, then H is a solution of LFS.*

Proposition 2 (Proposition 3.4 in [14]). *Given the triple $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$, let B be background knowledge and O an observation.*

– *If H is a solution of cautious induction, then H is a solution of brave induction. The converse implication holds when $B \cup H$ is a positive NLP.*

[14] provided computational complexity results of brave induction.

Proposition 3 (Theorem 4.1 and 4.3 in [14]). *Let B be background knowledge and O an observation, the following complexity results hold:*

– *Deciding whether O has a solution of brave or cautious induction under B is NP-complete, given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$ or $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$.*
– *Deciding whether a given hypothesis is a solution of brave induction is Σ_2^P-complete, given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$ or $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$.*
– *Deciding whether a given hypothesis is a solution of cautious induction is coNP-complete, given the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$.*
– *Deciding whether a given hypothesis is a solution of cautious induction is Π_2^P-complete, given the triple $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$.*

3 Motivations

An inductive learning problem often assumes that the "world" is governed by some model M of the language and the learning process is to gather information and correct hypotheses in order to converge to theories that could capture the model M. In this section, we address the problem of brave induction from the perspective of the assumption, which motivates our optimization of brave induction proposed in the next section.

Following Shapiro's [15] definition of model inference problems, we define problems of explanatory induction, LFS, brave induction, and cautious induction. Given a triple $\langle L_b, L_o, L_h \rangle$ and an interpretation M as the model of the "world", we use L_o^M to denote the set of observational sentences satisfied by M.

Definition 3. *Given a triple $\langle L_b, L_o, L_h \rangle$, an interpretation M of $\mathcal{L}$, and background knowledge $B \subseteq L_b$ satisfied by M,*

- *the* explanatory induction problem *is to find a hypothesis $T \subseteq L_h$ such that M satisfies T and $T \cup B \models L_o^M$;*
- *the* LFS problem *is to find a hypothesis $T \subseteq L_h$ such that M satisfies T and there exists a model S of $T \cup B$ with $S \models L_o^M$;*
- *the* brave induction problem *is to find a hypothesis $T \subseteq L_h$ such that M satisfies T and there exists an answer set S of $T \cup B$ with $S \models L_o^M$;*
- *the* cautious induction problem *is to find a hypothesis $T \subseteq L_h$ such that M satisfies T, $T \cup B$ is consistent, and $T \cup B \models_{AS} L_o^M$.*

In this case, T is respectively called an observation complete axiomatization *of the explanatory induction, LFS, brave induction, or cautious induction problem.*

Notice that, if the set of all ground literals in $\mathcal{L}$, denoted by L_{GL}, is a subset of L_o, then the condition "M satisfies T" has been implied and can be omitted in the definition.

We can extend the definition of brave and cautious induction given the triple $\langle L_{ASP}, L_{GL}, L_{ASP} \rangle$.

Definition 4. *Given the triple $\langle L_{ASP}, L_{GL}, L_{ASP} \rangle$, let B be background knowledge and O an observation.*

- *A hypothesis H covers O under B in* brave induction *if $B \cup H$ has an answer set satisfying O.*
- *A hypothesis H covers O under B in* cautious induction *if $B \cup H$ is consistent and every answer set of $B \cup H$ satisfies O.*

Recall that, the language $\mathcal{L}$ is function-free and contains finite sets of constants and predicates, then both L_{GL} and interpretations are finite.

Proposition 4. *Let M be an interpretation of $\mathcal{L}$, background knowledge B satisfied by M, and an observation $O = \{l \mid l \in L_{GL} \cap L_o$ and $M \models l\}$ where L_o is the language of observation.*

– *Given the triple* $\langle L_{CT}, L_{CT}, L_{CT}\rangle$*, a hypothesis* H *is an observation complete axiomatization of the explanatory induction, LFS, brave induction, or cautious induction problem if and only if* H *is a solution of explanatory induction, LFS, brave induction, or cautious induction respectively.*
– *Given the triple* $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$*, a hypothesis* H *is an observation complete axiomatization of the brave induction or cautious induction problem implies* H *is a solution of brave or cautious induction respectively, but not vice versa in general.*
– *Given the triple* $\langle L_{ASP}, L_{GL}, L_{ASP}\rangle$*, a hypothesis* H *is an observation complete axiomatization of the brave induction or cautious induction problem if and only if* H *is a solution of brave or cautious induction respectively.*

Example 2. Consider Example 1, the "world" model M implies that the sets of European, Asian, and American are pairwise disjoint. Given the triple $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$ and the observation $O \cup B$, the hypothesis $\{euro(X),\ asia(X),\ usa(X)\}$ is a solution of brave and cautious induction, but it is not an observation complete axiomatization of the brave or cautious induction problem.

If $L_{GL} \subseteq L_o$ and $L_{GL} \subseteq L_h$, then the set $\{l \mid l \in L_{GL} \text{ and } M \models l\}$ is an observation complete axiomatization of explanatory induction, LFS, brave induction, and cautious induction problems. However, the set is not a good solution for many concept-learning problems. Normally, L_h has a set of restrictions, *i.e.*, language bias [9], which specifies the space of acceptable hypotheses. For example, some set of constants may not be allowed to appear in L_h. To simplify the discussion, we consider the set of clausal theories with no appearance of any constants, denoted by L_{CT^V}, and the set of ASP programs without any constants, denoted by L_{ASP^V}, as examples of restricted language of hypotheses.

Given the triple $\langle L_{CT}, L_{CT}, L_{CT^V}\rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^V}\rangle$, solutions of explanatory induction, LFS, brave induction, and cautious induction are defined the same for the triple $\langle L_{CT}, L_{CT}, L_{CT}\rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP}\rangle$ respectively.

Example 3. Consider Example 1, given the triple $\langle L_{CT}, L_{CT}, L_{CT^V}\rangle$, we could not find out a hypothesis that covers O under B in explanatory induction and cautious induction. On the other hand, hypotheses H_1, H_2, H_3, and

$$H_4 : euro(X) \vee student(X),$$
$$H_5 : \neg asia(X) \vee \neg usa(X)$$

are solutions of LFS. H_1, H_2, and H_3 are solutions of brave induction.

Although the hypothesis space of brave induction is smaller than the space of LFS, candidate solutions of brave induction also need to be optimized. For instance, we would prefer H_1 to H_2 and H_3, as $AS(B \cup H_1) \subset AS(B \cup H_2)$ and $AS(B \cup H_1) \subset AS(B \cup H_3)$. The intuition is that, a solution of brave induction is intended to capture the "world" model with background knowledge, and solutions with fewer "uncertainties" are preferred. Following the intuition, we specify our optimization of brave induction in the next section.

4 Proper Brave Induction

Motivated from previous discussions, we introduce two frameworks of induction.

Definition 5 (Proper brave and cautious induction). *Given the triple $\langle L_{CT}, L_{CT}, L_{CT^{\vee}} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^{\vee}} \rangle$, let B be background knowledge and O an observation.*

- *A hypothesis H covers O under B in* proper brave induction *if*
  - *H covers O under B in brave induction, and*
  - *there does not exist another such hypothesis H' such that $AS(H' \cup B) \subset AS(H \cup B)$.*
- *A hypothesis H covers O under B in* proper cautious induction *if*
  - *H covers O under B in cautious induction, and*
  - *there does not exist another such hypothesis H' such that $AS(H' \cup B) \subset AS(H \cup B)$.*

H is called a solution *of proper brave or cautious induction respectively .*

Example 4. Consider Example 3, H_1, H_2 and H_3 are solutions of brave induction and only H_1 is a solution of proper brave induction.

Relations to brave and cautious induction are as follows.

Proposition 5. *Given the triple $\langle L_{CT}, L_{CT}, L_{CT^{\vee}} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^{\vee}} \rangle$, let B be background knowledge and O an observation.*

- *If H is a solution of proper cautious induction, then H is a solution of proper brave induction.*
- *If H is a solution of proper cautious induction, then H is a solution of cautious induction.*
- *If H is a solution of proper brave induction, then H is a solution of brave induction.*

When $B \cup H$ has only one answer set, the converse implication holds respectively.

Proof. H is a solution of proper cautious induction, then H is a solution of brave induction. If there exists another solution H' of brave induction such that $AS(H') \subset AS(H)$, then H' is also a solution of cautious induction, which conflicts to the precondition that H is a solution of proper cautious induction.

Proper brave (resp. cautious) induction has a solution if and only if brave (resp. cautious) induction has a solution. The conditions for the existence of solutions for triples $\langle L_{CT}, L_{CT}, L_{CT} \rangle$ and $\langle L_{ASP}, L_{GA}, L_{ASP} \rangle$ have been discussed in [14]. We extend the results here.

Proposition 6 (Necessary conditions for the existence of solutions). *Let B be background knowledge and O an observation.*

– *Given the triple $\langle L_{CT}, L_{CT}, L_{CT^{\vee}} \rangle$, proper brave induction (resp. brave induction, proper cautious induction, cautious induction) has a solution, only if $B \cup O$ is consistent.*
– *Given the triple $\langle L_{ASP}, L_{GL}, L_{ASP^{\vee}} \rangle$, proper brave induction (resp. brave induction, proper cautious induction, cautious induction) has a solution, only if $B \cup O$ is satisfiable.*

Proof. (1) If proper brave induction has a solution H, then $B \cup H$ has a minimal model satisfying O, thus $B \cup O$ is consistent. (2) If proper brave induction has a solution H, then $B \cup H$ has an answer set satisfying O, thus $B \cup O$ is satisfiable. The proofs for brave induction, proper cautious induction, and cautious induction are similar.

Given the triple $\langle L_{CT}, L_{CT}, L_{CT^{\vee}} \rangle$, $B \cup O$ is consistent is not a sufficient condition for the existence of solutions. For example, given $B = \emptyset$ and $O = \{p(a), \neg p(b)\}$, $MM(B \cup O) \neq \emptyset$. However, proper brave induction (resp. brave induction, proper cautious induction, cautious induction) does not have a solution. Given the triple $\langle L_{ASP}, L_{GL}, L_{ASP^{\vee}} \rangle$, $B \cup O$ is consistent is not a necessary condition for the existence of solutions. For example, given $B = \{p(a) \leftarrow not\, p(a)\}$ and $O = \{q(a)\}$, $AS(B \cup O) = \emptyset$. However, $H = \{p(X), q(X)\}$ is a solution of proper brave induction.

Corollary 1 (Necessary condition of solutions). *Given the triple $\langle L_{CT}, L_{CT}, L_{CT^{\vee}} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^{\vee}} \rangle$, let B be background knowledge and O an observation. H is a solution of proper brave induction (resp. brave induction, proper cautious induction, cautious induction), only if $B \cup H \cup O$ is consistent.*

Now we provide some properties of proper brave and cautious induction. Given two clausal theories T_1 and T_2, we denote $T_1 \vee T_2$ to be a clausal theory that is logically equivalent to the formula $\bigwedge_{C \in T_1} C \vee \bigwedge_{C \in T_2} C$.

Proposition 7. *Given the triple $\langle L_{CT}, L_{CT}, L_{CT^{\vee}} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^{\vee}} \rangle$, both H_1 and H_2 are solutions of proper brave or cautious induction does not imply that $H_1 \cup H_2$ is a solution of proper brave or cautious induction.*

Example 5. Let $B = \{q(a) \vee r(a), \neg q(a) \vee \neg r(a)\}$ and $O = \{p(a)\}$. Both $H_1 = \{q(X),\, p(X) \vee \neg q(X)\}$ and $H_2 = \{r(X),\, p(X) \vee \neg r(X)\}$ cover O under B in proper brave and cautious induction, but $H_1 \cup H_2$ does not.

Proposition 8. *Given the triple $\langle L_{CT}, L_{CT}, L_{CT^{\vee}} \rangle$, both H_1 and H_2 are solutions of proper brave or cautious induction does not imply that $H_1 \vee H_2$ is a solution of proper brave or cautious induction.*

Note that, the result is different from Proposition 2.5 in [14] for brave and cautious induction. Though $H_1 \vee H_2$ is a solution of brave induction, it is possible that $MM(H_1) \subset MM(H_1 \vee H_2)$. Example 5 is such an example.

Proposition 9. *Given the triple* $\langle L_{CT}, L_{CT}, L_{CT^\vee}\rangle$ *or* $\langle L_{ASP}, L_{GL}, L_{ASP^\vee}\rangle$, *H covers both* O_1 *and* O_2 *under B in proper cautious induction implies that H covers* $O_1 \cup O_2$ *under B in proper cautious induction. But this is not the case for proper brave induction.*

Example 6. Let $B = \{p(X) \vee q(X)\}$, $O_1 = \{p(a)\}$, and $O_2 = \{q(a)\}$. Then $H = \emptyset$ covers both O_1 and O_2 under B in proper brave induction, but H does not cover $O_1 \cup O_2$ under B.

Proposition 10. *Given the triple* $\langle L_{CT}, L_{CT}, L_{CT^\vee}\rangle$ *or* $\langle L_{ASP}, L_{GL}, L_{ASP^\vee}\rangle$, *H covers O under both* B_1 *and* B_2 *in proper brave or cautious induction does not imply that H covers O under* $B_1 \cup B_2$ *in proper brave or cautious induction.*

Example 7. Let $B_1 = \{p(a)\}$, $B_2 = \{\neg f(x) \leftarrow p(x)\}$, $O = \{f(a)\}$. $H = \{f(X)\}$ covers O under both B_1 and B_2 in proper brave and cautious induction, but H does not cover O under $B_1 \cup B_2$.

Sakama and Inoue [14] provided algorithms to compute solutions of brave induction for triples $\langle L_{CT}, L_{CT}, L_{CT}\rangle$ and $\langle L_{ASP}, L_{GA}, L_{ASP}\rangle$. From previous discussions, the idea of proper brave induction can be used to optimize the solutions of brave induction. One such optimization procedure is as follows:

1. for each rule r in the hypothesis H and each atom $A \in head(r)$, let $r' = head(r) \setminus \{A\} \leftarrow body(r)$ and $H' = (H \setminus \{r\}) \cup \{r'\}$;
2. if H' is still a solution of brave induction and $AS(H' \cup B) \subseteq AS(H \cup B)$, then replace r by r'.

In Example 1, H_1 can be obtained from H_3 by the optimization procedure.

Proposition 11. *Given the triple* $\langle L_{CT}, L_{CT}, L_{CT^\vee}\rangle$ *or* $\langle L_{ASP}, L_{GL}, L_{ASP^\vee}\rangle$, *let B be background knowledge, O an observation, H a solution of brave induction, and* H' *a hypothesis obtained from H by the above optimization procedure.* $AS(H' \cup B) \subseteq AS(H \cup B)$.

Proof. Note that, in the optimization procedure, in each step H' is required to satisfy the condition that $AS(H' \cup B) \subseteq AS(H \cup B)$. Then after finite steps, $AS(H' \cup B) \subseteq AS(H \cup B)$.

Notice that, the condition $AS(H' \cup B) \subseteq AS(H \cup B)$ in the optimization procedure is necessary for the validity of Proposition 11.

Example 8. Let $B = \{p(a) \leftarrow q(a)\}$, $O = \{p(a)\})$, $H = \{p(x) \vee q(x)\}$, and $H' = \{q(x)\}$. Both H and H' cover O under B. However, $AS(H \cup B) = \{\{p(a)\}\}$, $AS(H' \cup B) = \{\{p(a), q(a)\}\}$, and $AS(H' \cup B) \not\subseteq AS(H \cup B)$.

There are some syntactic conditions for guaranteeing the condition $AS(H' \cup B) \subseteq AS(H \cup B)$ in the optimization procedure. For instance, if there does not exist another rule $r^* \in H' \cup B$ such that $A \in head(r^*)$, then $AS(H' \cup B) \subseteq AS(H \cup B)$ in the procedure. Moreover, we can improve the optimization procedure for the triple $\langle L_{ASP}, L_{GL}, L_{ASP^\vee}\rangle$.

An improved optimization procedure for the triple $\langle L_{ASP}, L_{GL}, L_{ASP^\vee}\rangle$ is as follows:

1. for each rule r in the hypothesis H and each atom $A \in head(r)$, let $r' = head(r) \setminus \{A\} \leftarrow body(r) \cup \{not\, A\}$ and $H' = (H \setminus \{r\}) \cup \{r'\}$;
2. if H' is still a solution of brave induction, then replace r by r'.

Proposition 12. *Given the triple $\langle L_{ASP}, L_{GL}, L_{ASP^{\vee}} \rangle$, let B be background knowledge, O an observation, H a solution of brave induction, and H' a hypothesis obtained from H by the above optimization procedure. $AS(H' \cup B) \subseteq AS(H \cup B)$.*

Proof. S is an answer set of $H' \cup B$ implies that S is a model of $H \cup B$. If S is not an answer set of $H \cup B$, then there exists a model S' of $(H \cup B)^S$ such that $S' \subset S$.

Let a be a ground atom corresponding to the atom A in the procedure and r^* be the grounding rule corresponding to the rule r' and a. If $a \in S$, then $\{r^{*S}\} = \emptyset$ and S' satisfies r^{*S}. If $a \notin S$, then $a \notin S'$ and S' satisfies r^{*S}. So S' is still a model of $(H' \cup B)^S$, which conflicts to the condition that S is an answer of $H' \cup B$.

5 Computational Complexity

In this section, we consider computational complexity of proper brave induction. The classes Σ_K^P, Π_k^P, Δ_k^P of the Polynomial Hierarchy [11] are defined as follows:

$\Delta_0^P = \Sigma_0^P = \Pi_0^P = P$ and for all k ≥ 1, $\Delta_k^P = P^{\Sigma_{k-1}^P}$, $\Sigma_k^P = NP^{\Sigma_{k-1}^P}$, $\Pi_k^P = \text{co-}\Sigma_k^P$.

The class D_k^P is defined as the class of problems that consist of the conjunction of two independent problems from Σ_k^P and Π_k^P. In particular, NP $= \Sigma_1^P$, co-NP $= \Pi_1^P$, and DP $= D_1^P$. For all $k \geq 1$,

$$\Sigma_k^P \subseteq D_k^P \subseteq \Delta_{k+1}^P \subseteq \Sigma_{k+1}^P \subseteq \text{PSPACE}.$$

First, we provide some computational complexity results about clausal theories and ASP programs. Notice that, we view a program with variables as a shorthand of its ground instantiations. We only consider ground clausal theories and ground ASP programs in this section.

Lemma 1. *Let H_1 and H_2 be (ground) clausal theories or DLPs.*

- *Deciding whether $AS(H_1) \subseteq AS(H_2)$ is Π_2^P-complete.*
- *Deciding whether $AS(H_1) = AS(H_2)$ is Π_2^P-complete.*
- *Deciding whether $AS(H_1) \subset AS(H_2)$ is D_2^P-complete.*

Proof. (1) The complement of the problem is in Σ_2^P, as we can guess a set S such that $S \in AS(H_1)$ and $S \notin AS(H_2)$ and the problem of deciding whether $S \in AS(H_1)$ is co-NP-complete [4]. The hardness can be proved by reducing the problem of deciding whether $AS(H_1) = \emptyset$ to whether $AS(H_1) \subseteq AS(\{p \leftarrow not\, p\})$. Note that, deciding whether $AS(H_1) = \emptyset$ is Π_2^P-complete [4].

(2) can be proved in the same manner.

(3) The problem is equivalent to deciding whether $AS(H_1) \subseteq AS(H_2)$ and $AS(H_1) \neq AS(H_2)$, so it is in $\mathrm{D}_2^{\mathrm{P}}$. The hardness can be proved by reducing the problem of deciding whether the conjunction $\exists \boldsymbol{X} \forall \boldsymbol{Y} E \wedge \forall \boldsymbol{X'} \exists \boldsymbol{Y'} E'$ is satisfiable to whether $AS(H_1) \subset AS(H_2)$, where H_2 has an answer set iff $\exists \boldsymbol{X} \forall \boldsymbol{Y} E$ is satisfiable and H_1 has an answer set iff $\forall \boldsymbol{X'} \exists \boldsymbol{Y'} E'$ is satisfiable.

Similar to the proof for Lemma 1, we have the following lemma.

Lemma 2. *Let H_1 and H_2 be (ground) NLPs.*

- *Deciding whether $AS(H_1) \subseteq AS(H_2)$ is* co-NP*-complete.*
- *Deciding whether $AS(H_1) = AS(H_2)$ is* co-NP*-complete.*
- *Deciding whether $AS(H_1) \subset AS(H_2)$ is* DP*-complete.*

Now we provide computational complexity results of proper brave induction. We use L_{NLP} and L_{NLP^V} to denote the set of NLPs and the set of NLPs without any constants.

Theorem 1. *The following computational complexity results hold:*

- *Given the triple $\langle L_{NLP}, L_{GL}, L_{NLP^V} \rangle$,*
  - *deciding whether a given hypothesis is a solution of brave induction is* NP*-complete;*
  - *deciding the existence of solutions in brave induction or proper brave induction is in Σ_2^{P} and* NP*-hard;*
  - *deciding whether a given hypothesis is a solution of proper brave induction is in Π_2^{P} and* co-NP*-hard.*
- *Given the triple $\langle L_{CT}, L_{CT}, L_{CT^V} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^V} \rangle$,*
  - *deciding whether a given hypothesis is a solution of brave induction is Σ_2^{P}-complete;*
  - *deciding the existence of solutions in brave induction or proper brave induction is in Σ_3^{P} and Σ_2^{P}-hard;*
  - *deciding whether a given hypothesis is a solution of proper brave induction is in Π_3^{P} and Π_2^{P}-hard.*

Proof. (1) The problem is equivalent to checking whether a set O of ground literals is satisfied by some answer set of an NLP $H \cup B$, which is known to be NP-complete.

(2) The problem is in Σ_2^{P}, as we can guess a hypothesis H such that H is a solution of brave induction, which can be verified in polynomial time by an NP oracle. The NP-hardness can be proved by reducing the problem of deciding whether a ground NLP P has an answer set to whether there exists a hypothesis covers an observation O under background knowledge B, where B is obtained from P by replacing each occurrence of a ground atom A in P by an atom $f_A(c_A)$ and adding rules $f_A(c'_A) \leftarrow$ and $\leftarrow f_A(c''_A)$ for each ground atom A in P, and $O = \{f(c)\}$ such that both f and c do not appear in B. It is easy to verify that, there exists a hypothesis (*i.e.*, $\{f(X)\}$) covers O under B iff P has an answer set. Note that, deciding whether a ground NLP has an answer set is NP-complete.

(3) The complement of the problem is in Σ_2^{P}, as we can guess another solution H' of brave induction such that $AS(H' \cup B) \subset AS(H \cup B)$, which can be verified in polynomial time by an NP oracle. The co-NP-hardness can be proved from the fact that H covers an observation O under background knowledge B in brave induction iff H' does not cover O under $B \cup B'$ in proper brave induction, where H' is obtained from H by adding $not\, f(X)$ in $body(r)$ of each $r \in H$ and adding $f'(X) \leftarrow not\, f(X)$ and $f(X) \leftarrow not\, f'(X)$ with new symbols f and f', and $B' = \{A \leftarrow f(X) \mid A \in O\} \cup \{\leftarrow A, f(X) \mid \neg A \in O\}$.

(4) can be proved in the same manner of the proof for Proposition 3.

(5) and (6) can be proved in the same manner for (2) and (3) respectively.

From Proposition 3, given the triple $\langle L_{CT}, L_{CT}, L_{CT} \rangle$ or $\langle L_{ASP}, L_{GA}, L_{ASP} \rangle$, deciding the existence of solutions in brave induction is NP-complete. However, the problem is Σ_2^{P}-hard given the triple $\langle L_{CT}, L_{CT}, L_{CT^V} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^V} \rangle$. The reason is that, the hypothesis space is restricted in the latter case and $B \cup O$ is satisfiable would no longer be the sufficient condition for the existence of solutions in brave induction.

6 Related Work

In previous sections, we provided an optimization of brave induction called proper brave induction and compared it with brave induction in Proposition 5. From the discussion in [14], a solution of explanatory induction is always a solution of brave induction and a solution of brave induction is a solution of LFS. Then a solution of proper brave induction is always a solution of LFS. Similar to brave induction, proper brave induction is neither stronger nor weaker than LFI [3] or confirmatory induction [6]. Notice that, the idea of proper brave induction can also be extended to these logic frameworks of induction.

On the other hand, there has been much work on induction in nonmonotonic logic programs. Otero [10] and Sakama [13] extended the definition of ILP to the stable model semantics and introduced frameworks for learning positive/negative examples in NLPs. Ray [12] developed a nonmonotonic ILP system, called XHAIL, which combines abduction and induction for constructing hypotheses. ASPAL [2], another nonmonotonic ILP system, uses ASP as a solver to compute a solution to a standard ILP task. Later, Law, Russo, and Broda [7] presented a new paradigm for ILP that allows the learning of ASP programs. Note that, the optimization based on fewer stable models or answer sets can also be applied to these nonmonotonic ILP systems.

7 Conclusion

Motivated from Shapiro's definition of model inference problems, we provide an optimization of Sakama and Inoue's brave induction, called proper brave induction, for causal theories and ASP programs. A hypothesis is a solution of proper brave induction, if it is a solution of brave induction and there does not exist another solution whose set of answer sets is a proper subset of its. We investigate formal properties of proper brave induction and develop an optimization procedure. At last, we analyze computational complexity of decision problems for proper brave induction in propositional case. We expect that the idea of the optimization will be extended to other logical frameworks for concept-learning.

Acknowledgments. We thank reviewers for their helpful comments. The work was supported by NSFC under grant 61573386 and 61175057, NSFC for the Youth under grant 61403359, as well as the USTC Key Direction Project and the USTC 985 Project.

References

1. Baral, C.: Knowledge representation, reasoning and declarative problem solving. Cambridge university press (2003)
2. Corapi, D., Russo, A., Lupu, E.: Inductive logic programming in answer set programming. In: Proceedings of the 22nd International Conference on Inductive Logic Programming (ILP-12). pp. 91–97 (2012)
3. De Raedt, L.: Logical settings for concept-learning. Artificial Intelligence 95(1), 187–201 (1997)
4. Eiter, T., Gottlob, G.: On the computational cost of disjunctive logic programming: Propositional case. Annals of Mathematics and Artificial Intelligence 15(3-4), 289–323 (1995)
5. Flach, P.A.: Rationality postulates for induction. In: Proceedings of the 6th conference on Theoretical aspects of rationality and knowledge. pp. 267–281 (1996)
6. Lachiche, N.: Abduction and induction from a non-monotonic reasoning perspective. In: Abduction and Induction: Essays on their Relation and Integration, pp. 107–116. Kluwer (2000)
7. Law, M., Russo, A., Broda, K.: Inductive learning of answer set programs. In: Proceedings of the 14th European Conference on Logics in Artificial Intelligence (JELIA-14). pp. 311–325 (2014)
8. Mitchell, T.M.: Generalization as search. Artificial intelligence 18(2), 203–226 (1982)
9. Muggleton, S., De Raedt, L.: Inductive logic programming: Theory and methods. The Journal of Logic Programming 19, 629–679 (1994)
10. Otero, R.P.: Induction of stable models. In: Proceedings of the 11th International Conference on Inductive Logic Programming (ILP-01). pp. 193–205 (2001)
11. Papadimitriou, C.H.: Computational complexity. John Wiley and Sons Ltd. (2003)
12. Ray, O.: Nonmonotonic abductive inductive learning. Journal of Applied Logic 7(3), 329–340 (2009)
13. Sakama, C.: Induction from answer sets in nonmonotonic logic programs. ACM Transactions on Computational Logic (TOCL) 6(2), 203–231 (2005)
14. Sakama, C., Inoue, K.: Brave induction: a logical framework for learning from incomplete information. Machine Learning 76(1), 3–35 (2009)
15. Shapiro, E.Y.: Inductive inference of theories from facts. Yale University, Department of Computer Science (1981)