Brave Induction Revisited

Jianmin Ji

University of Science and Technology of China

August 20, ILP 2015

Table of Contents

Motivation

- Sakama and Inoue introduced brave induction as a novel logic framework for concept-learning.
  - A *hypothesis* H covers an *observation* O under a *background knowledge* B in *brave induction* if $B \cup H$ has an answer set S such that $O \subseteq S$.
  - Brave induction allows more hypotheses than *explanatory induction* and fewer hypotheses than *learning from satisfiability* (LFS).
  - Sakama and Inoue showed that brave induction has potential applications for problem solving in systems biology, requirement engineering, and multiagent negotiation.
- In order to choose hypothesises produced by brave induction, we introduce an optimization of brave induction called proper brave induction.

Example

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 : $\mathit{euro}(X) \vee \mathit{asia}(X) \vee \mathit{usa}(X) \leftarrow \mathit{student}(X)$,

H_2 : $\mathit{euro}(X) \vee \mathit{asia}(X) \vee \mathit{usa}(X) \vee \mathit{teacher}(X)$,

H_3 : $\mathit{euro}(X) \vee \mathit{asia}(X) \vee \mathit{usa}(X) \vee \mathit{teacher}(X) \leftarrow \mathit{student}(X)$.

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

Intuition

- The intuition behind Shapiro's definition of model inference problems:
  - 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.
- We would prefer hypotheses allowed by brave induction with fewer "uncertainties".
- 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.

Table of Contents

Proper Brave Induction

A triple $\langle L_b, L_o, L_h \rangle$: L_b is the language of background knowledge, L_o for observations, and L_h for hypotheses.
L_{CT}: clausal theories; L_{ASP}: ASP programs; L_{GA}: ground atoms; L_{*V}: language * with variables.

Definition (Proper brave induction)

Given the triple $\langle L_{CT}, L_{CT}, L_{CT^V} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^V} \rangle$, 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 $AS(H' \cup B) \subset AS(H \cup B)$.

H is called a *solution* of proper brave induction.

Proper Cautious Induction

Definition (Proper cautious induction)

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

- 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 cautious induction.

Relation to Brave and Cautious Induction

Proposition

Given the triple $\langle L_{CT}, L_{CT}, L_{CT^V} \rangle$ *or* $\langle L_{ASP}, L_{GL}, L_{ASP^V} \rangle$, *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.*

Necessary Conditions for the Existence of Solutions

Proposition

Let B be background knowledge and O an observation.

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

Corollary (Necessary condition of solutions)

Given the triple $\langle L_{CT}, L_{CT}, L_{CT^V} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^V} \rangle$, 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.

Some Properties

Given the triple $\langle L_{CT}, L_{CT}, L_{CT^V} \rangle$ or $\langle L_{ASP}, L_{GL}, L_{ASP^V} \rangle$.

Proposition

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.

Proposition

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.

Proposition

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.

Optimization Procedure

- Sakama and Inoue provided an algorithm to compute solutions of brave induction.
- Based on Sakama and Inoue's algorithm, an optimization procedure can be added:
  1. for each rule r in the hypothesis H and each atom $A \in head(r)$, let $r' = head(r) \setminus \{A\} \leftarrow body(r)$;
  2. if $(H \setminus \{r\}) \cup \{r'\}$ is still a solution of brave induction, then replace r by r'.

Proposition

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. $AS(H' \cup B) \subseteq AS(H \cup B)$.

Table of Contents

Lemmas

Lemma

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.*

Lemma

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.*

Computational Complexity

Theorem

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.*

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.