Comparing Abductive Theories

Katsumi Inoue
National Institute of Informatics

Chiaki Sakama
Wakayama University

ECAI 2008

Computational issues on abductive reasoning

- Abduction is used in many AI applications, e.g., diagnosis, design, discovery.
- Abduction is an important paradigm for problem solving, and is incorporated in programming technologies, viz, **abductive logic programming** (**ALP**).
- Automated abduction is also studied in the literature as an extension of deductive methods or a part of inductive systems.

Comparing non-deductive capabilities between programs

- Intelligent agents perform non-deductive commonsense reasoning as well as deductive reasoning.
- Comparing capabilities of non-deductive reasoning such as abduction and induction is meaningful to measure intelligence of agents.
- Inoue and Sakama [IJCAI-05] defined two notions of **abductive equivalence** --- explainable/explanatory equivalence.

Comparison of abductive theories

- Evaluation of **abductive power** in ALP.
- **Refinement** and **revision** in ALP.
- **Equivalence** of abductive theories [IS05].
  - optimization, debugging, simplification, verification

- Relations between abductive agents.
- Software development of abductive theories.

- Comparison of **inductive theories** also involves similar computational issues.

When is an abductive theory stronger than another abductive theory?

- No definition in the literature of ALP.
  - In what circumstances, can we say that abduction by agent A is stronger than abduction by agent B?
  - When can we regard that abduction with knowledge P is stronger than abduction with knowledge Q?
  - When can we regard that abduction with hypotheses M is stronger than abduction with hypotheses N?
  - How about the relation between abductive theories (P,M) and (Q,N)?

Considerable parameters

- World
  - background knowledge
  - observations
- Agent who performs abduction
  - her logic of background knowledge
    - language, syntax
    - semantics
    - axioms, inference procedure
  - her logic of hypotheses/observation
    - language, syntax
    - logic of explanation entailment
    - criteria of best explanations

Abductive framework

- $(\mathcal{L}, B, H)$
  - $\mathcal{L}$: language / logic
  - B : background knowledge
  - H : candidate hypotheses

- Given an observation O, E is an **<u>explanation</u> of** O **in** $(\mathcal{L}, B, H)$ iff E belongs to H ($E \subseteq H$) and
  - $B \cup E \vdash_{\mathcal{L}} O$
  - $B \cup E$ is consistent.

- When O has an explanation in $(\mathcal{L}, B, H)$, O is **<u>explainable in</u>** $(\mathcal{L}, B, H)$.

Abductive generality:
First Definition

- Let (L, B_1, H_1) and (L, B_2, H_2) be abductive frameworks.
- (L, B_1, H_1) is **more (or equally) explainable than** (L, B_2, H_2) iff, for any observation O, if O is explainable in (L, B_1, H_1) then O is explainable in (L, B_2, H_2).
- Explainable generality requires that one abductive framework has more explainability than another abductive framework for any observation.

Note: L must be common when comparing frameworks.

Abductive generality: Second Definition

- Let ($\mathcal{L}$, B_1, H_1) and ($\mathcal{L}$, B_2, H_2) be abductive frameworks.

- ($\mathcal{L}$, B_1, H_1) is **more (or equally) explanatory than** ($\mathcal{L}$, B_2, H_2) iff, for any observation O, any explanation of O in ($\mathcal{L}$, B_1, H_1) is also an explanation of O in ($\mathcal{L}$, B_2, H_2).

- Explanatory generality assures that one abductive framework has more explanation power (explanation contents) than another for any observation.

- Explanatory generality implies explainable generality.

📖 Note: $\mathcal{L}$ must be common.

Example

- $A_1 = (\text{FOL}, B_1, \{s,r\})$ and $A_2 = (\text{FOL}, B_2, \{s,r\})$ where

$$B_1: \quad s \rightarrow g$$

$$B_2: \quad s \rightarrow g, \quad r \rightarrow g$$

- A_1 and A_2 are explainably equivalent. That is, A_1 is more or equally explainable than A_2 , and vice versa.
- A_1 and A_2 are not explanatorily equivalent.
- A_2 is more explanatory than A_1 , but not vice versa.
- $A_3 = (\text{FOL}, B_1, \{r\})$ and $A_4 = (\text{FOL}, B_2, \{r\})$ are not explainably equivalent.

Example

- $A_1 = (LP, B_1, \{a,b\})$, $A_2 = (LP, B_2, \{a,b\})$,

  $B_1 : \quad p \leftarrow a, \quad a \leftarrow b$

  $B_2 : \quad p \leftarrow a, \quad p \leftarrow b$

- A_1 and A_2 are explainably equivalent. That is, A_1 is more or equally explainable than A_2, and vice versa.
- A_1 and A_2 are not explanatorily equivalent.

  In fact, $\{b\}$ is an explanation of a in A_1 but is not in A_2.
- A_1 is more explanatory than A_2, but not vice versa.

Results in first-order logic

- **Definition** [Reiter, Poole]**:** An **extension** of (FOL, B, H) is *Th*(B ∪ S), where S is a maximal subset of H such that B ∪ S is consistent.
- The set of all extensions of A is denoted as Ext(A).
- **Lemma** [Poole]**:** O is explainable in (FOL, B, H) iff there is an extension of (FOL, B, H) in which O is true.
- **Theorem:** A_1 is more explainable than A_2 iff, for any $X_2 \in \mathrm{Ext}(A_2)$, there is $X_1 \in \mathrm{Ext}(A_1)$ such that $X_1 \supseteq X_2$.
- **Theorem:** $A_1 = (\mathrm{FOL}, B_1, H_1)$ is more explanatory than $A_2 = (\mathrm{FOL}, B_2, H_2)$ iff $B_1 \models B_2$ and A_1 is more explainable than A_2 .

Results in first-order logic

- **<u>Theorem</u>:** $A_1 = (\text{FOL}, B_1, H_1)$ and $A_2 = (\text{FOL}, B_2, H_2)$ are explainably equivalent iff $Ext(A_1) = Ext(A_2)$.
- **<u>Corollary</u>:** If $B_1 \equiv B_2$ then (FOL, B_1, H) and (FOL, B_2, H) are explainably equivalent.

- **<u>Theorem</u>:** $A_1 = (\text{FOL}, B_1, H)$ and $A_2 = (\text{FOL}, B_2, H)$ are explanatorily equivalent iff $B_1 \equiv B_2$ and A_1 and A_2 are explainably equivalent.

Subclasses in first-order logic

- **Theorem (Assumption-freeness):** (FOL, B_1, O) is more explainable than (FOL, B_2, O) iff $B_2 \models B_1$.

- **Theorem (Semi-monotonicity):** Suppose two abductive frameworks with the same background knowledge, $A_1 = (\mathrm{FOL}, B, H_1)$ and $A_2 = (\mathrm{FOL}, B, H_2)$. If $H_1 \supseteq H_2$ then A_1 is more explainable than A_2 and is more explanatory than A_2 .

Note: $B_1 \models B_2$ does not imply that (FOL, B_1, H) is more explainable than (FOL, B_2, H).

Abductive Logic Programs (ALP)

- (LP, B, H) : abductive framework where
  - B : logic program (GEDP)
  - H : set of **abducibles** (literals)

- *G* : **observation** (a conjunction of ground literals)
- $E \subseteq$ H is a **(credulous) explanation of** *G* **in** (LP, B, H) if all literals in *G* are true in a consistent answer set of B $\cup E$.

Results in abductive logic programs

- **Definition** [IS]**:** A **belief set of (**LP, B, H) **(wrt** *E* **)** is a consistent answer set of $B \cup E$ where $E \subseteq H$.
- When a belief set S is an answer set of $B \cup E$, S is also denoted as S_E.
- The set of all belief sets of A is denoted as BS(A).

- **Theorem**: $A_1 = (LP, B_1, H_1)$ is more explainable than $A_2 = (LP, B_2, H_2)$ iff, for any $S_2 \in BS(A_2)$, there is $S_1 \in BS(A_1)$ such that $S_1 \supseteq S_2$.
- **Theorem:** $A_1 = (LP, B_1, H_1)$ is more explanatory than $A_2 = (LP, B_2, H_2)$ iff, for any $E \subseteq H_2$ and $S_E \in BS(A_2)$, there is $T_E \in BS(A_1)$ such that $E \subseteq H_1$ and $T_E \supseteq S_E$.

Results in abductive logic programs

- **Theorem**: $A_1 = (LP, B_1, H_1)$ and $A_2 = (LP, B_2, H_2)$ are explainably equivalent iff $\max(BS(A_1)) = \max(BS(A_2))$.

- **Theorem:** $A_1 = (LP, B_1, H_1)$ and $A_2 = (LP, B_2, H_2)$ are explanatorily equivalent iff $C_1 = C_2$ and
  $\max(AS(B_1 \cup E)) = \max(AS(B_2 \cup E))$ for any $E \in C_i$,
  where $C_i = \{ E \subseteq H_i \mid B_i \cup E$ is consistent $\}$ for $i=1,2$.

Results in abductive logic programs

- **<u>Definition</u>** [IS04]**:** Let $\mathcal{R}$ be a set of rules. Two programs P_1 and P_2 are **<u>strongly equivalent with respect to $\mathcal{R}$</u>** if $AS(P_1 \cup R) = AS(P_2 \cup R)$ for any $R \subseteq \mathcal{R}$.
- *<u>Note:</u>* If there is no restriction on $\mathcal{R}$, the notion reduces to strong equivalence [Lifschitz, Pearce & Valverde, 2001].
- **<u>Theorem</u>**: Let B_1 and B_2 be EDPs, that is, logic programs without NAF in heads.

  (LP, B_1, H) and (LP, B_2, H) are explanatorily equivalent iff B_1^+ and B_2^+ are strongly equivalent with respect to H, where $B_i^+ = B_i \cup \{ \leftarrow L, \neg L \mid L \in \mathit{Lit} \}$.

Summary of Results

necessary and sufficient conditions

Logic	$A_1=(\mathcal{L}, B_1, H_1)$ is more than $A_2=(\mathcal{L}, B_2, H_2)$			
	explainable	**explanatory**		
FOL	$\forall X_2 \in Ext(A_2)$, $\exists X_1 \in Ext(A_1)$ s.t. $X_1 \supseteq X_2$	$B_1 \models B_2$ and A_1 is more explainable than A_2		
LP	$\forall S_2 \in BS(A_2)$, $\exists S_1 \in BS(A_1)$ s.t. $S_1 \supseteq S_2$	$\forall E \in H_2$ $\forall S_E \in BS(A_2)$, $\exists T_E \in BS(A_1)$ s.t. $E \in H_1$ and $T_E \supseteq S_E$		

Summary of Results

Computational Complexities (Propositional Case)

Logic	$A_1=(\mathcal{L}, B_1, H_1)$ is more than $A_2=(\mathcal{L}, B_2, H_2)$			
	explainable	explanatory		
FOL	Π^P_3 -complete	Π^P_3 -complete		
LP (general)	Π^P_3 -complete	Π^P_3 -complete		
LP (ELPs)	Π^P_2 -complete	Π^P_2 -complete		

Abductive Equivalence

- All generality relations are defined to be anti-symmetric, that is, two abductive frameworks are explainably/explanatorily equivalent in the sense of Inoue & Sakama [IJCAI-05] iff one is both more (or equally) and less (or equally) explainable/explanatory than another at the same time.
- With this correspondence, abductive equivalence can be more easily characterized by combining both directions of properties for abductive generality.

Discussion

- Explainable generality and explanatory generality have the same complexity, and are more complex in general than abductive equivalence.
- Abductive generality can be further characterized with the generality notions in default logic [Inoue & Sakama ICLP'06] and answer set programming [Inoue & Sakama AAAI-07].
- In future work, further parameters can be considered.