Equivalence in Abductive Logic

Katsumi Inoue
National Institute of Informatics
Chiaki Sakama
Wakayama University

IJCAI-2005

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, i.e., **abductive logic programming** (**ALP**).
- Automated abduction is also studied in the literature as an extension of deductive methods or a part of inductive systems.

Issues not fully understood yet

- Evaluation of **abductive power** in ALP.
- Understanding of the semantics of ALP with respect to **modularity** and **contexts**.
- Efficiency in abductive reasoning, e.g., **simplification**, **optimization**.
- Debugging and **verification** in ALP.
- Standardization in ALP.

➢ All these issues are related to different notions of **equivalence** in ALP.

When two ALPs are equivalent?

- No definition in the literature of ALP.
- No such concepts in philosophy, either.
  - When can we consider that an **explanation** E is equivalent to an **explanation** F for an observation?
  - When can we say that an **observation** G is equivalent to an **observation** H in an abductive framework?
  - In what circumstances, can we say that abduction **by** person A is equivalent to abduction **by** person B?
  - When can we regard that abduction **with** knowledge P is equivalent to abduction **with** knowledge Q?

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 and logic
  - B : background knowledge
  - H : candidate hypotheses

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

Abductive equivalence:
First Definition

- Two abductive frameworks ($\mathcal{L}$, B_1, H_1) and ($\mathcal{L}$, B_2, H_2) are **explainably equivalent** iff, for any observation O, there is an explanation of O in ($\mathcal{L}$, B_1, H_1) iff there is an explanation of O in ($\mathcal{L}$, B_2, H_2).

- Explainable equivalence requires that two abductive frameworks have the same explainability for any observation.

Note: $\mathcal{L}$ must be common.

Abductive equivalence:
Second Definition

- Two abductive frameworks ($\mathcal{L}$, B_1, H) and ($\mathcal{L}$, B_2, H) are **explanatorily equivalent** iff, for any observation O,
  E is an explanation of O in ($\mathcal{L}$, B_1, H)
  iff E is an explanation of O in ($\mathcal{L}$, B_2, H).

- Explanatory equivalence assures that two abductive frameworks have the same explanation power (explanation contents) for any observation.

- Explanatory equivalence implies explainable equivalence.

Note: $\mathcal{L}$ and H are common.

Example

- A_1 = (FOL, B_1, {a,b}) and A_2 = (FOL, B_2, {a,b}) where

  B_1 : $a \rightarrow p$

  B_2 : $b \rightarrow p$

- A_1 and A_2 are explainably equivalent.
- A_1 and A_2 are not explanatorily equivalent.
- A_3 = (FOL, B_1, {b}) and A_4 = (FOL, B_2, {b}) are not explainably equivalent.

Example

- ⌘ $A_1=(\text{FOL}, B_1, \{a,b\})$, $A_2=(\text{FOL}, B_2, \{a,b\})$, $A_3=(\text{FOL}, B_3, \{a,b\})$,

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

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

$$B_3: \quad b \rightarrow p, \quad b \rightarrow a$$

- ⌘ A_1, A_2 and A_3 are explainably equivalent.
- ⌘ A_1 and A_2 are explanatorily equivalent.
- ⌘ A_1 and A_3 are not explanatorily equivalent.

  In fact, $\{a\}$ is an explanation of p in A_1 but is not in A_3.

Results in first-order logic

- **Definition** [Reiter]**:** An **extension** of (B,H) is $Th(B \cup S)$ where S is a maximal subset of H s.t. $B \cup S$ is consistent.
- **Theorem:** (FOL, B_1, H_1) and (FOL, B_2, H_2) are explainably equivalent iff the extensions of (B_1, H_1) coincide with the extensions of (B_2, H_2).
- **Corollary:** If $B_1 \equiv B_2$ then (FOL, B_1, H) and (FOL, B_2, H) are explainably equivalent.
- **Theorem:** (FOL, B_1, H) and (FOL, B_2, H) are explanatorily equivalent iff $B_1 \equiv B_2$.

Subclasses in first-order logic

- **<u>Theorem</u>:** For any (FOL, B, H), there is a (FOL, Φ, H') which is explainably equivalent to (FOL, B, H).

- **<u>Theorem</u>:** (FOL, B_1, $\mathscr{L}$) and (FOL, B_2, $\mathscr{L}$) are explainably equivalent iff $B_1 \equiv B_2$.

- **<u>Theorem</u>:** Suppose that both $B_1 \cup H_1$ and $B_2 \cup H_2$ are consistent. Then, (FOL, B_1, H_1) and (FOL, B_2, H_2) are explainably equivalent iff $B_1 \equiv B_2$.

- Equivalence of the minimal explanations for any observation reduces to explanatory equivalence.

Complexity for first-order logic

- **Lemma** [Dix]**:** A formula ψ is true in all extensions of (B,H) iff the extensions of (B $\cup \{\psi\}$, H) are the same as the extensions of (B,H).

- **Theorem:** Deciding explainable equivalence in propositional logic is Π_2^P-complete.

- **Theorem:** Deciding explanatory equivalence in propositional logic is coNP-complete.

Abductive Logic Programs (ALP)

- ⌘ ⟨B, H⟩ : abductive framework (ALP, B, H), where
  - B : logic program (GEDP)
  - H : set of **abducibles** (literals)

- Let G be a conjunction of ground literals.
- $E \subseteq$ H is a **(credulous) explanation of** G **in** ⟨B, H⟩ if every literal in G is true in a consistent answer set of B $\cup E$.

Results in abductive logic programs

- **Definition** [IS]**:** A **belief set of** $\langle B, H \rangle$ **(wrt** E**)** is a consistent answer set of $B \cup E$ where $E \subseteq H$.

- **Theorem**: $\langle B_1,H_1 \rangle$ and $\langle B_2,H_2 \rangle$ are explainably equivalent iff the literals in the belief sets of $\langle B_1,H_1 \rangle$ coincide with those in the belief sets of $\langle B_2,H_2 \rangle$.
- **Corollary:** If the belief sets of $\langle B_1,H_1 \rangle$ coincide with the belief sets of $\langle B_2,H_2 \rangle$, then $\langle B_1,H_1 \rangle$ and $\langle B_2,H_2 \rangle$ are explainably equivalent. The converse does not hold.

⌘ **Example**: If H = {a,b}, and

B_1 : p ← a, ***not*** b, q ← b, ***not*** a,

B_2 : p ← b, ***not*** a, q ← a, ***not*** b,

then $\langle B_1,H \rangle$ and $\langle B_2,H \rangle$ are explainably equivalent.

Results in abductive logic programs

- **Definition** [IS]**:** Let $\mathcal{R}$ be a set of rules. Two programs P_1 and P_2 are **strongly equivalent with respect to $\mathcal{R}$** if $AS(P_1 \cup R) = AS(P_2 \cup R)$ for any $R \subseteq \mathcal{R}$.
- *Note:* If there is no restriction on $\mathcal{R}$, the notion reduces to strong equivalence [Lifschitz, Pearce & Valverde, 2001].
- **Theorem**: $\langle B_1,H \rangle$ and $\langle B_2,H \rangle$ are explanatorily equivalent iff B_1 and B_2 are strongly equivalent with respect to H.

⌘ **Corollary**: Unfold/fold transformation does not preserve explanatory equivalence [Sakama & Inoue, ICLP 1995].

Complexity for ALP

- **<u>Theorem</u>:** Deciding explainable equivalence in propositional ALP is in Δ_3^P, and is Π_2^P-hard.

- **<u>Lemma</u>:** Let $\mathscr{A}$ be a set of literals, and

  $\mu(\mathscr{A}) = \{\ \delta(l);\ \textbf{\textit{not}}\ \delta(l) \leftarrow,\ \ l \leftarrow \delta(l)\ \mid\ l \in \mathscr{A}\ \}.$

  Two programs P_1 and P_2 are strongly equivalent with respect to $\mathscr{A}$ if $AS(P_1 \cup \mu(\mathscr{A})) = AS(P_2 \cup \mu(\mathscr{A}))$.

- **<u>Theorem</u>:** Deciding explanatory equivalence in propositional ALP is Π_2^P-complete.

Extension

- Further generalization of abductive equivalence in ALP where removal of hypotheses is allowed in **extended abduction** [Inoue and Sakama, IJCAI 1995] can be characterized by the notion of **update equivalence** [Inoue and Sakama, JELIA 2004].

- The concepts of abductive equivalence can be applied to **inductive equivalence** [Sakama & Inoue, ILP 2005].
--- relative least generalization, inverse entailment (Progol, CF-induction), descriptive induction (CLAUDIEN), induction of stable models, etc.

Discussion

- Equivalence of abductive theories is the basis for measuring abduction capability.
- Many variations exist for abductive equivalence.
- Logical equivalence (or weak notion of equivalence) of background theories does not necessarily implies abductive equivalence (except for some simple cases).
- Explanatory equivalence is not computationally harder than explainable equivalence.
- In future work, further parameters should be considered.