A Structuralist Theory of Belief Revision

Holger Andreas

Seminar for Philosophy, Logic, and Theory of Science

LMU Munich

September 3, 2010

What is a Structuralist Theory of Belief Revision?

A synthesis of the Sneed formalism
(known as structuralist theory of science)

with belief revision theory
(as initiated by Alchourrón, Gärdenfors, and Makinson)

in the medium of prioritised default logic
(a particular system of nonmonotic reasoning)

Objectives of the Synthesis

- A simpler and allegedly superior conception of epistemic ranking
- Some progress in terms of cognitive adequacy of belief revision theory
- The representation of belief changes in science

Outline

Operations on Sets of Beliefs

1. Expansions: $A + \phi$
2. Contractions: $A \div \phi$
3. Revisions: $A * \phi$

A - a set of beliefs
ϕ - a meaningful sentence

Belief Sets vs. Belief Bases I

Explanation

Belief sets K are logically closed: $Cn(K) = K$.

Belief bases H represent beliefs in that $Cn(H) = K$ or $Inf(H) = K$.

K - set of beliefs
Cn - monotonic consequence operation
Inf - nonmonotonic (inference) operation

Belief Sets vs. Belief Bases II

Arguments in favour of base revisions:

- Finite memory argument
- Less complex epistemic ranking
- Base revisions respect *justifications* more properly

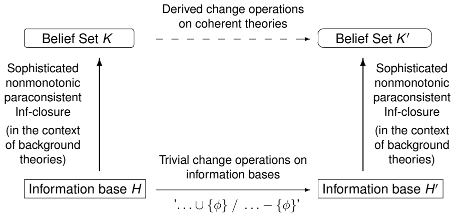
Connections to Epistemology

Foundationalist assumptions in the study of *base* revisions:

- Beliefs can be divided into non-derived and derived ones.
- The belief base may contain all and only non-derived beliefs.
- Other beliefs must be inferable from the belief base plus some optional set of axioms of background theories.

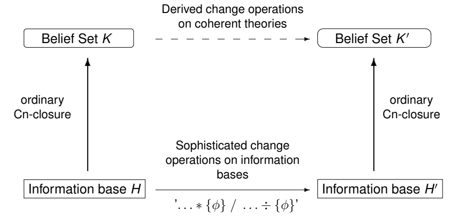
But no claim is made about certainty, truth, or even irrevisability of basic beliefs. Nor are scientific and everyday theories considered derivable from basic beliefs.

Two Kinds of Belief Base Changes: The Direct Mode



Cf. H. Rott (2001): Change, Choice, and Inference.

The Coherence-Constrained Mode



Cf. H. Rott (2001): Change, Choice, and Inference.

More on the Direct Mode

$K = Inf(\mathcal{H}, \mathcal{E})$ — Generation of the belief set from a set $\mathcal{H}$ of prioritised basic beliefs and a set $\mathcal{E}$ of prioritised *expectations*

$K * \alpha = Inf(\mathcal{H} \circ \langle\alpha\rangle, \mathcal{E})$ — Revisions

$K \div \alpha = Inf(\mathcal{H} \setminus \langle\alpha\rangle, \mathcal{E})$ — Contractions

Inf is a nonmonotonic and paraconsistent inference formalism.

Expectations are elements of (defeasibly valid) background theories.

See H. Rott (2001): Change, Choice, and Inference.

Note on Epistemic Rankings

Explanation

An epistemic ranking introduces some sort of ordering on the elements of K, H, or on possible worlds of $\mathcal{L}$. It is necessary for a unique definition of contractions and revisions.

Examples:

- Epistemic entrenchment orderings - Gärdenfors/ Makinson (1988)
- Ranking functions - Spohn (1988)
- System of spheres - Grove (1988)

Epistemic Entrenchment

Requirement by Gärdenfors (1988):

"It is possible to determine the relative epistemic entrenchment of the sentences in a belief set K *independently* of what happens to K in contractions and revisions."

Structuralist Theory of Science

Origin: Sneed (1971): The Logical Structure of Mathematical Physics

Emerges from a combination of:

1. Use of set-theoretical predicates along the lines of P. Suppes
2. Ramsey account of scientific theories

Mature account: Balzer, Moulines, and Sneed (1987): An Architectonic for Science.

Frames in the sense of Minsky I

Motivation:

To explain the effectiveness and the speed of mental activities better than first-order representations of reasoning, memory, and language.

Important objective:

Inter-propositional knowledge representation, i.e., the association of propositions with information how to use these propositions.

History of ideas:

OOP (object oriented programming) is deeply inspired by Minsky-frames.

Minsky (1974): A framework for representing knowledge.

Frames in the sense of Minsky II

Cognitive characterisation:

A structure that we select from memory when we encounter a new situation of which type we are familiar with.

Formal characterisation:

A frame is a data-structure for representing a stereo-typed situation. Such a structure consists of *slots*, i.e., constants whose admissible values are specified by simple and complex conditions. Complex conditions specify *relations* among the values assigned to slots.

Example: A Frame For a Birthday Party

Slots	Simple and complex conditions (complex conditions in italics)
Host	Has birthday. *Receives presents.*
Guests	*Are friends of the host. Give presents to the host.*
Presents	Must please the host.
Games	*Host and guests participate.* Being fun.

Cf. M. Minsky (1974): A Framework for Representing Knowledge

Frame Concepts in Structuralism I

Explanation

A *structure* is a sequence of sets $\langle D_1, \ldots, D_k, R_1, \ldots, R_n \rangle$, where $D_1, \ldots, D_k$ are sets of empirical or mathematical objects and $R_1, \ldots, R_n$ relations on these sets.

A *set-theoretical predicate* applies to structures of type $\langle D_1, \ldots, D_k, R_1, \ldots, R_n \rangle$ and imposes certain conditions on the sets $D_1, \ldots, D_k$ and the relations $R_1, \ldots, R_n$.

What is Default Logic?

A classical approach to nonmonotonic reasoning

Explanation

A logical system with an inference operation *Inf* is nonmonotonic if and only if it does not hold in general that, for all sets A and B of sentences, if $A \subseteq B$, then $Inf(A) \subseteq Inf(B)$.

Motivation for nonmonotonic logical system:

Numerous of our inferential patterns are not strictly truth-preserving. Rather, scientific and everyday reasoning is ampliative, risky and adventurous.

Syntax of Default Logic

A default theory is a pair $\langle W, D \rangle$, where W is a set of closed formulas and D a set of *default rules*.

General syntax of a default rule:

$$\frac{\phi : \psi_1, \ldots, \psi_n}{\chi}$$

ϕ - prerequisite
$\psi_1, \ldots, \psi_n$ - consistency conditions
χ - consequent

Meaning of a Default Rule

$$\frac{\phi : \psi_1, \ldots, \psi_n}{\chi}$$

If ϕ and it is consistent to assume that $\psi_1, \ldots, \psi_n$, then χ.

A default is called *normal* if and only if it has the form

$$\frac{\phi : \psi}{\psi}$$

Cf. G. Antoniou (1997): Nonmonotonic Reasoning

Prioritised Default Logic (PDL)

Explanation

A prioritised default theory is a triple $T = (W, D, <)$, where (W, D) is a default theory and $<$ a strict partial order among the defaults.

Any computation of inferences from $T = (W, D, <)$ must respect the priority ordering $<$.

Example: Federal laws override state laws.

Abductive Inferences in Structuralism

Abductions in the context of some background theory:

$$\frac{y \in \mathbf{I}(\mathbf{T}), y = \mathbf{r}(\mathbf{T})(x), x \in \mathbf{M}(\mathbf{T})}{(x, y) \in \mathbf{AE}(\mathbf{T})}$$

$y \in \mathbf{I}(\mathbf{T})$ - y represents a phenomenon to which **T** is applied.
$y = \mathbf{r}(\mathbf{T})(x)$ - x is a theoretical description of the phenomenon that is represented by y.
$x \in \mathbf{M}(\mathbf{T})$ - x satisfies the laws of **T**.
$(x, y) \in \mathbf{AE}(\mathbf{T})$ - x is an admissible, or correct, theoretical description of the phenomenon represented by y.

Problems:

1. Abductive inferences are defeasible.
2. Satisfaction of the antecedent does not imply satisfaction of *links* to other intended applications.

Default Formulation of the Abductive Inference Rule

$$\delta(\mathbf{T}) \qquad \frac{y \in \mathbf{I}(\mathbf{T}), y = \mathbf{r}(\mathbf{T})(x), x \in \mathbf{M}(\mathbf{T}) : (x, y) \in \mathbf{AE}(\mathbf{T})}{(x, y) \in \mathbf{AE}(\mathbf{T})}$$

Now, a theoretical explanation of an empirical phenomenon can only be accepted if it is consistent with previously accepted theoretical explanations.

Problem: Which theory-elements should be applied first to empirical phenemona in the process of drawing inferences?

Priorities among Theory-Elements

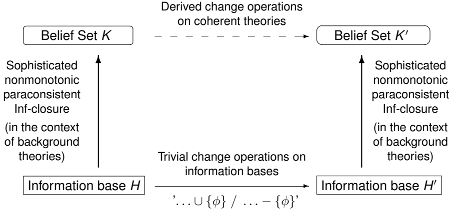
Explanation

$\delta(\mathbf{T_i}) >_R \delta(\mathbf{T}_j)$ iff $\mathbf{T}_i$ is more reliable than $\mathbf{T}_j$.

Tentatively, $\mathbf{T}_i$ is more reliable than $\mathbf{T}_j$ iff $\mathbf{T}_i$'s statistics of successful intended applications is *significantly* better than this statistics is of $\mathbf{T}_j$.

An intended application of **T** is successful iff it has **T**-theoretical conclusions that remain accepted.

Review of the Direct Mode of Base Revisions



Cf. H. Rott (2001): Change, Choice, and Inference.

A Structuralist Theory of Belief Revision

Background theory E:

1. Non-defeasible part: definitions of $\mathbf{M}(\mathbf{T})$, $\mathbf{M_p}(\mathbf{T})$, internal and external links, and postulates about $\mathbf{AE}(\mathbf{T})$
2. Defeasible part: $\delta(\mathbf{T})$ for any theory-element **T**

Belief base H:
Propositions of the form $a \in \mathbf{I}(\mathbf{T})$ for any **T**

Priority information:
Only the abductive inference rules $\delta(\mathbf{T})$ are prioritised.

Inference formalism Inf:
$A = Inf(W, D, >_R) = \{\phi \mid (W, D, <_R) \mid\!\sim_s \phi\}$,
W contains the non-defeasible part of E, and H.

Features of the System

1. Substantial simplification of the epistemic ranking that is needed to define revisions and contractions uniquely.
2. Gärdenfors's requirement that the epistemic ranking must be determinable independtly of what happens to beliefs in revisions and contractions is satisfied!
3. Minsky's request for *interpropositional* knowledge representation is satisfied since ordinary first-order propositions are represented as valuations of frames, and the frame indicates for what inference the proposition can be used.
4. Belief revision theory becomes connected with formal philosophy of science.

Summary

- Epistemic ranking is an effect of theorising!
- The epistemic ranking of a derived belief ϕ is determined by the epistemic ranking of those theory-elements (=inferential patterns) through which ϕ has been derived.
- All non-derived beliefs have equal epistemic standing.

- Outlook
  - Connections to counterfactual conditionals, DDL etc.
  - TMS for scientific knowledge

For Further Reading I

H. Rott
Change, Choice and Inference.
Oxford University Press, 2001.

G. Antoniou
Nonmonotonic Reasoning.
MIT Press, 1997.

H. Andreas.
A Structuralist Theory of Belief Revision.
Journal of Logic, Language and Information, forthcoming (conditionally accepted).

Many thanks!

Example by Gärdenfors (1988)

Oscar used to believe that he had given Victoria a gold ring at their wedding. He had bought their two rings at a jeweller's shop in Casablanca. He thought it was a bargain. The merchant had claimed that the rings were made of 24 carat gold. They certainly looked like gold, but to be on the safe side Oscar had taken the rings to the jeweller next door who has testified to their gold content. However, some time after the wedding, Oscar was repairing his boat and he noticed that the sulphuric acid he was using stained his ring. He remembered from his school chemistry that the only acid that affected gold was aqua regia. Somewhat surprised, he verified that Victoria's ring was also stained by the acid. So Oscar had to revise his beliefs because they entailed an inconsistency.

So, because he had greater confidence in what he was taught in chemistry than in his own smartness, Oscar somewhat downheartedly accepted that the rings were not made of gold after all.

(e1) $\forall x(LOOKS\_GOLDEN(x) \rightarrow MADE\_OF\_GOLD(x))$

(e2)
$\forall x \forall y(JEWELLER(x) \wedge SELLS(x, y) \wedge TESTIFIES\_GOLD(x, y) \rightarrow MADE\_OF\_GOLD(y))$

(e3) $\forall x(\neg EXPENSIVE(x) \rightarrow \neg MADE\_OF\_GOLD(x))$

(e4)
$\forall x \forall y(JEWELLER(x) \wedge \neg SELLS(x, y) \wedge TESTIFIES\_GOLD(x, y) \rightarrow MADE\_OF\_GOLD(y))$

(e5)
$\forall x \forall y \forall z(RING(x) \wedge SULPHURIC\_ACID(y) \wedge TIME\_POINT(z) \wedge EXPOSED(x, y, z) \wedge STAINED(x, z) \rightarrow \neg MADE\_OF\_GOLD(x))$