Autoepistemic logic (AEL)

Introduced by Moore [4] in 1985.

Definition 1 Let $\mathcal{L}$ be a propositional language. We define $\mathcal{L}_{\mathbf{B}}$ as the smallest set such that

- $\mathcal{L} \subseteq \mathcal{L}_{\mathbf{B}}$
- If $\varphi, \psi \in \mathcal{L}_{\mathbf{B}}$ then so are $\neg\varphi$, $(\varphi \vee \psi)$, $(\varphi \wedge \psi)$, $(\varphi \supset \psi)$
- If $\varphi \in \mathcal{L}_{\mathbf{B}}$ then $\mathbf{B}\varphi \in \mathcal{L}_{\mathbf{B}}$

Stable theories

Definition 2 An autoepistemic theory $T \subseteq \mathcal{L}_{\mathbf{B}}$ is *stable* iff

St1. $T = \{\varphi \mid T \vdash \varphi\}$

St2. If $\varphi \in T$ then $\mathbf{B}\varphi \in T$

St3. If $\varphi \notin T$ then $\neg\mathbf{B}\varphi \in T$

Note that the '$\vdash$' in (St1) is a propositional consequence operator. T is not closed under any modal system as a result of just (St1). As far as (St1) goes, $\mathbf{B}\varphi$ is an atomic formula. (St2) means that T is closed under necessitation.

Theorem 1 Let $T \subseteq \mathcal{L}_{\mathbf{B}}$ be stable. Then

1. If $\mathbf{B}\varphi \in T$ then $\varphi \in T$
2. If T is consistent and $\neg\mathbf{B}\varphi \in T$ then $\varphi \notin T$

Proof 1. Assume $\mathbf{B}\varphi \in T$ and $\varphi \notin T$. Then from (St3) $\neg\mathbf{B}\varphi \in T$. Since T is inconsistent, $\varphi \in T$, a contradiction. 2. Assume $\neg\mathbf{B}\varphi \in T$. Since T is consistent, $\mathbf{B}\varphi \notin T$. Then from (St2) $\varphi \notin T$.

Corollary 2 Let $T \subseteq \mathcal{L}_{\mathbf{B}}$ be stable and consistent. Then

1. $\varphi \in T$ iff $\mathbf{B}\varphi \in T$
2. $\varphi \notin T$ iff $\neg\mathbf{B}\varphi \in T$

Stable expansions

What beliefs should an agent have based on a set of facts $T \subseteq \mathcal{L}_{\mathbf{B}}$? They should

- include T
- allow introspection
- be grounded in T

Definition 3 E is a *stable expansion* of T iff

$$E = \{\varphi \mid T \cup \{\mathbf{B}\varphi \mid \varphi \in E\} \cup \{\neg\mathbf{B}\varphi \mid \varphi \notin E\} \vdash_{\mathrm{S}} \varphi\}$$

Here S is propositional logic (PL). If we define an expansion this way, we say that E is *weakly grounded* in T.

A stable expansion does not always exist. The theory $\{p \supset \mathbf{B}\neg p\}$ has no stable expansions. Some theories have more than one expansion. $\{\mathbf{B}p \supset p\}$ has, as we shall see two expansions.

Kernel of a stable expansion

Definition 4 The *kernel* E_0 of a stable theory $E \subseteq \mathcal{L}_{\mathbf{B}}$ is defined as the propositional subset of E.

Lemma 3 If E is a stable set, then E is an expansion of E_0.

Lemma 4 If E and F are stable sets and $E_0 = F_0$, then $E = F$.

Theorem 5 A stable set is uniquely determined by its kernel.

Proof Follows from lemmas 3 and 4.

Definition 5 $\mathrm{St}(A)$ is defined as the stable closure of a set of propositional formula A (alternatively S5 closure). This means that the kernel of $\mathrm{St}(A)$ is A, equivalently that $\mathrm{St}(A)$ is the unique stable expansion of A.

S5 consequence

Theorem 6 If $T \subseteq \mathcal{L}_{\mathbf{B}}$ is stable, then T is closed under S5 consequence.

Proof All instances of S5 axiom schemata are contained i T:

K. $\beta = \mathbf{B}(\varphi \supset \psi) \supset (\mathbf{B}\varphi \supset \mathbf{B}\psi)$
It suffices to show $\neg\mathbf{B}(\varphi \supset \psi) \vee \neg\mathbf{B}\varphi \vee \mathbf{B}\psi \in T$. Assume $\neg\mathbf{B}(\varphi \supset \psi) \notin T$ and $\neg\mathbf{B}\varphi \notin T$. Then from (St3) $\varphi \supset \psi \in T$ and $\varphi \in T$. From (St1) $\psi \in T$. From (St2) $\mathbf{B}\psi \in T$. So $\beta \in T$.

T. $\beta = \mathbf{B}\varphi \supset \varphi$
Assume $\varphi \in T$. Then $\beta \in T$. Assume $\varphi \notin T$. Then from (St3) $\neg\mathbf{B}\varphi \in T$ and $\beta \in T$.

4. $\beta = \mathbf{B}\varphi \supset \mathbf{B}\mathbf{B}\varphi$
Assume $\neg\mathbf{B}\varphi \in T$. Then $\beta \in T$. Assume $\neg\mathbf{B}\varphi \notin T$. Then from (St3) $\varphi \in T$ and from (St2) twice $\mathbf{B}\mathbf{B}\varphi \in T$. So $\beta \in T$.

5. $\beta = \neg\mathbf{B}\varphi \supset \mathbf{B}\neg\mathbf{B}\varphi$
Assume $\varphi \in T$. Then $\mathbf{B}\varphi \in T$ and $\beta \in T$. Assume $\varphi \notin T$. Then $\neg\mathbf{B}\varphi \in T$ and from (St2) $\mathbf{B}\neg\mathbf{B}\varphi \in T$. So $\beta \in T$.

S5 equivalence

Theorem 7 Let $T \subseteq \mathcal{L}_{\mathbf{B}}$. Then T is stable iff T is an S5 theory.

Proof Follows from Theorem 6 and ...

Corollary 8 The Deduction Theorem does not hold for stable sets, as it does not hold for S5.

We cannot, however use S5 reasoning to define expansions, because of axiom $\mathbf{T}$. $\{\mathbf{B}p\}$ has only one expansion $\mathrm{St}(\emptyset)$, whereas if the logic S in Definition 3 was S5, it would have another expansion $\mathrm{St}(\{p\})$.

Theorem 9 The strongest modal logic S that can be used when defining expansions is K45 (S5 without $\mathbf{T}$ or *weak S5*).

Examples

Example 1 $\{\mathrm{B}p \supset p\}$ has as we have seen two expansions $\mathrm{St}(\emptyset)$ and $\mathrm{St}(\{p\})$. Is the latter expansion reasonable? The agent believes p and accordingly $\mathbf{B}p$ but the reason it believed p in the first place was because it believed $\mathbf{B}p$.

Example 2 $\{\mathrm{B}p \supset q, \mathrm{B}q \supset p\}$ has two expansions $\mathrm{St}(\emptyset)$ and $\mathrm{St}(\{p, q\})$.

Example 3 $\{\neg\mathrm{B}p \supset q, \mathrm{B}p \supset p\}$ has two expansions $\mathrm{St}(\{p\})$ and $\mathrm{St}(\{q\})$.

Groundedness

- Weak groundedness

$$E = \{T \cup \{\mathbf{B}\varphi \mid \varphi \in E\} \cup \{\neg\mathbf{B}\varphi \mid \varphi \notin E\} \vdash \varphi\}$$

- Moderate groundedness

$$E = \{T \cup \{\neg\mathbf{B}\varphi \mid \varphi \notin E_0\} \vdash_{\mathrm{K45}} \varphi\}$$

  By eliminating $\{\mathbf{B}\varphi \mid \varphi \in E\}$ the circularity of Examples 1 and 2 disappears. Their only moderately grounded expansion is $\mathrm{St}(\emptyset)$. For expansions to remain stable we need a stronger logic, hence the K45 consequence operator.

- Strong groundedness
  $\mathrm{St}(\{p\})$ of Example 3 is not strongly grounded in T.

Algorithm

Definition 6 Let $T \subseteq \mathcal{L}_{\mathbf{B}}$. Then $\mathrm{sub}(T)$ is the union of $\mathrm{sub}(\varphi)$ for all $\varphi \in T$, where $\mathrm{sub}(\varphi)$ is defined as

$$\begin{aligned}
\mathrm{sub}(\varphi) &= \emptyset \text{ if } \varphi \in \mathcal{L} \\
\mathrm{sub}(\neg\varphi) &= \mathrm{sub}(\varphi) \\
\mathrm{sub}(\varphi \vee \psi) &= \mathrm{sub}(\varphi \wedge \psi) \\
&= \mathrm{sub}(\varphi \supset \psi) \\
&= \mathrm{sub}(\varphi) \cup \mathrm{sub}(\psi) \\
\mathrm{sub}(\mathbf{B}\varphi) &= \{\varphi\}
\end{aligned}$$

`foreach partition` E^+ `and` E^- `of` $\mathrm{sub}(T)$

  $E_0 :=$ `smallest set such that`
  $\{\varphi \mid E_0 \vdash \varphi\} =$
  $\{\psi \in \mathcal{L} \mid T \cup \{\mathbf{B}\varphi \mid \varphi \in E^+\} \cup \{\neg\mathbf{B}\varphi \mid \varphi \in E^-\} \vdash \psi\}$

  $E := \mathrm{St}(E_0)$

  `if` $E^+ \subseteq E$ `and` $E^- \cap E = \emptyset$ `then`

    `output` E_0

`endfor`

References

[1] G. Antoniou. *Nonmonotonic Reasoning*. MIT Press, Cambridge, MA, 1997.

[2] K. Konolige. Autoepistemic logic. In D. M. Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Handbook of Logic in Artificial Intelligence and Logic Programming-Nonmonotonic Reasoning and Uncertain Reasoning (Volume 3)*, pages 217–295. Clarendon Press, Oxford, 1994.

[3] V. W. Marek and M. Truszczyński. *Nonmonotonic Logic*. Springer-Verlag, Berlin, Germany, 1993.

[4] R. C. Moore. Semantical considerations on non-monotonic logic. *Artificial Intelligence*, 25:75–94, 1985.