An Abductive Reasoning Approach to the Belief-Bias Effect

Luís Moniz Pereira, Emmanuelle-Anna Dietz, Steffen Hölldobler

Syllogistic Reasoning Task by Evans, Barston & Pollard

<table>
<thead>
<tr>
<th>Premise 1</th>
<th>Premise 2</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>No millionaires are hard workers.</td>
<td>Some rich people are hard workers.</td>
<td>Therefore, some millionaires are not rich people.</td>
</tr>
</tbody>
</table>

The majority of the participants concluded $S_{\text{rich}}$ is not classical logically valid.

<table>
<thead>
<tr>
<th>Premise 1</th>
<th>Premise 2</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>No addictive things are inexpensive.</td>
<td>Some cigarettes are inexpensive.</td>
<td>Therefore, some addictive things are not cigarettes.</td>
</tr>
</tbody>
</table>

The majority of the participants concluded $S_{\text{add}}$ is classical logically valid.

This is called the belief-bias effect. It occurs when we think to judge something based on our reasoning, but are actually influenced by our beliefs.

Syntax and Semantics of Logic Programs

Logic Programs

A logic program $P$ is a finite set of clauses: $A \leftarrow L_1 \land \ldots \land L_n$, where $n \geq 0$.

- $A$ is an atom and $\text{Body } L_i$ with $1 \leq i \leq n$ are literals.
- $A \leftarrow T$ is a positive and $A \leftarrow \bot$ is a negative fact.
- $A$ is undefined in $P$ if it is not the head of some clause in $P$.
- $g P$ denotes ground $P$ and contains only the ground instances of its clauses.

Three-valued Łukasiewicz Semantics

<table>
<thead>
<tr>
<th>$I$</th>
<th>$T$</th>
<th>$U$</th>
<th>$\bot$</th>
<th>$\top$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\top$</td>
<td>$\top$</td>
<td>$\top$</td>
<td>$\bot$</td>
<td>$\top$</td>
</tr>
<tr>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\bot$</td>
<td>$\top$</td>
<td>$\bot$</td>
</tr>
<tr>
<td>$U$</td>
<td>$U$</td>
<td>$\bot$</td>
<td>$\top$</td>
<td>$U$</td>
</tr>
</tbody>
</table>

Interpretations and Models

An interpretation $I$ is a pair $\langle T^I, U^I \rangle$ of disjoint sets of ground atoms, s.t. $I^T = \{ A \in B_P \mid A \text{ is mapped to } T \}$ and $I^U = \{ A \in B_P \mid A \text{ is mapped to } U \}$, where $B_P$ is the Herbrand base with respect to a given program $P$.

Interpretation $I$ is a model for $P$ if $I$ maps each clause occurring in $P$ to $T$.

Weak Completion Semantics

Weak Completion

The following transformation for $P$ is the weak completion of $P$ (wc $P$):

1. Replace all $A \leftarrow \text{Body}_1, A \leftarrow \text{Body}_2, \ldots$ in $P$ by $A \leftarrow \text{Body}_1 \lor \text{Body}_2 \lor \ldots$
2. Replace all occurrences of $\leftarrow$ by $\lor$.

Reasoning with Respect to Least Models

Let $I$ be an interpretation in $\Phi(P) = \langle J^T, J^U \rangle$ (Stl., 2008), where $J^T = \{ A \mid \exists A \leftarrow \text{Body} \in g P \}$ with $\text{Body}(I) = T$ and $J^U = \{ A \mid A \in \text{Body} \in g P \}$ and $\forall A \leftarrow \text{Body} \in g P : \text{Body}(I) = U$.

$\text{Lfp}(\Phi(P))$ is the least model of the weak completion of $P$ (HKR, 2009).

Correspondence to Well-founded Semantics

Weak completion corresponds to well-founded semantics for tight $P^{\text{mod}}$.

- $P$ is tight, that is, it does not contain positive cycles.
- $P^{\text{mod}}$ is $P \cup \{ A \leftarrow \neg A, A, \neg A \leftarrow A \}$ for all undefined atoms $A \in P$.

Abduction

Abductive Framework

Let $(P, A, \models_{\text{lmcw}})$ be an abductive framework, literal $O$ an observation.

The set of abducibles $A$ consists of positive and negative facts for all undefined atoms in $P$. $P \models_{\text{lmcw}} F$ iff $\text{lmwc } g(P \cup \lnot F) = T$ for formula $F$. $O$ is explained by $E$ iff $E \subseteq A$. $P \cup E$ is satisfiable, and $P \cup E \models_{\text{lmcw}} O$.

Skeptical and Credulous Reasoning

- $F$ follow skeptically by abduction from $P$ and $O$ iff $O$ can be explained and for all minimal explanations $E$ such that $\text{lmwc } g(P \cup E) \models F$.
- $F$ follow credulously by abduction from $P$ and $O$ iff there exists a minimal explanation $E$ for $O$ such that $\text{lmwc } g(P \cup E) \models F$.

Contextual Abductive Reasoning

Inspection Points

Let’s introduce two reserved (meta-) predicates for atom $A$ (PP, 2011):

- inspect($A$) and inspect$_{\text{mod}}(A)$,

which are special abducibles and differ from the usual ones as they can only be abduced whenever $A$ or $\neg A$ have been abduced somewhere else already.

Syllogism $S_{\text{add}}$

We replace the $\text{ab}_{\text{add}}$-clause in $P_{\text{add}}$ where the new program, $P_{\text{add}}^{\text{insp}}$, is:

$(P_{\text{add}} \setminus \{\text{ab}_{\text{add}}(X) \leftarrow \text{cig}(X)\}) \cup \{\text{ab}_{\text{add}}(X) \leftarrow \text{inspect(cig}(X))\}$.

Suppose, $b$ is addictive, i.e. $O_{\text{add}}(b)$: $\text{cig}(b)$ cannot be abduced anymore to explain $O_{\text{add}}(b)$. The only minimal explanation for $O_{\text{add}}(b)$ is $E_{\text{cig}}(b)$.

$\text{lmwc } g(P_{\text{add}}^{\text{insp}} \cup E_{\text{cig}}(b)) = \{\text{ab}_{\text{add}}(b), \text{cig}(b), \text{inspect(cig}(b)), \text{inex}(b)\}$

Skeptically, we validate some addictive things are not cigarettes.

Conclusion

- Previous approaches have shown that weak completion semantics seems to adequately model Wason’s selection task and Byrne’s suppression task.
- We seem to adequately model the syllogistic reasoning task including the belief-bias effect under weak completion semantics with inspection points.
- This allows us to define contextual abduction for which we can specify contextual side-effects, contestable contextual side-effects, contextual relevant consequences, and jointly supported contextual relevant consequences.