An Abductive Reasoning Approach to the Belief-Bias Effect

Luís Moniz Pereira\textsuperscript{1}  Emmanuelle Dietz\textsuperscript{1,2}  Steffen Hölldobler\textsuperscript{2}

\textsuperscript{1}Centro de Inteligência Artificial (CENTRIA), Universidade Nova de Lisboa, Portugal

\textsuperscript{2}International Center for Computational Logic (ICCL) TU Dresden, Germany
Let’s consider $S_{rich}$

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

The majority of the participants concluded that $S_{rich}$ is not classical logically valid.
Let’s consider $S_{\text{rich}}$

**Premise 1**  
No millionaires are hard workers.

**Premise 2**  
Some rich people are hard workers.

**Conclusion**  
Therefore, some millionaires are not rich people.

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

Let’s consider $S_{\text{add}}$

**Premise 1**  
No addictive things are inexpensive.

**Premise 2**  
Some cigarettes are inexpensive.

**Conclusion**  
Therefore, some addictive things are not cigarettes.
Let’s consider $S_{\text{rich}}$

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

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

Let’s consider $S_{\text{add}}$

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

The majority of the participants concluded that $S_{\text{add}}$ is **classical logically valid**.
Let’s consider $S_{rich}$

**Premise 1**  
No millionaires are hard workers.

**Premise 2**  
Some rich people are hard workers.

**Conclusion**  
Therefore, some millionaires are not rich people.

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

Let’s consider $S_{add}$

**Premise 1**  
No addictive things are inexpensive.

**Premise 2**  
Some cigarettes are inexpensive.

**Conclusion**  
Therefore, some addictive things are not cigarettes.

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

Evans, Barston, and Pollard [1983] call this phenomenon the **belief-bias effect**

It occurs when we think to judge something based on our reasoning, but are actually influenced by our beliefs and our prior knowledge.
Motivation

Evans, Barston, and Pollard carried out an experiment to show that people are biased by their own beliefs but they do not propose a logic that adequately models this task.
Motivation

Evans, Barston, and Pollard carried out an experiment to show that people are biased by their own beliefs but they do not propose a logic that adequately models this task.

Classical logic cannot adequately represent this syllogistic reasoning task.

Hölldobler and Kencana Ramli [2009] propose to model human reasoning by

- logic programs
- under weak completion semantics
- based on the three-valued Łukasiewicz (1920) logic.

It seems to adequately model Byrne’s suppression task and Wason’s selection task.
Evans, Barston, and Pollard carried out an experiment to show that people are biased by their own beliefs but they do not propose a logic that adequately models this task.

Classical logic cannot adequately represent this syllogistic reasoning task.

Hölldobler and Kencana Ramli [2009] propose to model human reasoning by

- logic programs
- under weak completion semantics
- based on the three-valued Łukasiewicz (1920) logic.

It seems to adequately model Byrne’s suppression task and Wason’s selection task.

Can we also adequately model the syllogistic reasoning task under weak completion semantics?
It is commonly known that

*Cigarettes are addictive.* (1)
Background Knowledge & Abnormality Predicates

It is commonly known that

\textit{Cigarettes are addictive.} \hspace{1cm} (1)

After reading \textsc{Premise 2} participants seem to assume that

\textit{Cigarettes are inexpensive (compared to other addictive things).} \hspace{1cm} (2)

which biases the reasoning towards a representation.
Background Knowledge & Abnormality Predicates

It is commonly known that

\( \text{Cigarettes are addictive.} \) \hspace{1cm} (1)

After reading Premise 2 participants seem to assume that

\( \text{Cigarettes are inexpensive (compared to other addictive things).} \) \hspace{1cm} (2)

which biases the reasoning towards a representation.

Stenning & van Lambalgen (2008) propose to model conditionals with \( ab \) predicates. Premise 1 of \( S_{add} \) can be represented as

\( \text{If something is inexpensive and not abnormal, then it is not addictive.} \)
\( \text{Nothing (as a rule) is abnormal (wrt (3)).} \) \hspace{1cm} (3)

The belief in (1) and (2) generates an exception for cigarettes

\( \text{If something is a cigarette, then it is abnormal (wrt (3)).} \) \hspace{1cm} (4)
To explain why people validate $S_{add}$ we need to show that they reason abductively.
Abducing the Conclusion

To explain why people validate $S_{add}$ we need to show that they reason abductively.

Consider $S_{add}$ again

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

To explain why people validate $S_{\text{add}}$ we need to show that they reason abductively.

Consider $S_{\text{add}}$ again

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

1. we know that there are addictive things, let's say $b$.
   
   $b$ is addictive.
Abducting the Conclusion

To explain why people validate $S_{add}$ we need to show that they reason abductively.

Consider $S_{add}$ again

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

1. we know that there are addictive things, let’s say $b$.

   $b$ is addictive.

2. Given Premise 1 we infer that these addictive things are not inexpensive.

   $b$ is not inexpensive.
Abducing the Conclusion

To explain why people validate $S_{\text{add}}$ we need to show that they reason abductively.

Consider $S_{\text{add}}$ again

**Premise 1**  *No addictive things are inexpensive.*

**Premise 2**  *Some cigarettes are inexpensive.*

**Conclusion**  *Some addictive things are not cigarettes.*

1. We know that there are addictive things, let’s say $b$.

   \[ b \text{ is addictive.} \]

2. Given **Premise 1** we infer that these addictive things are not inexpensive.

   \[ b \text{ is not inexpensive.} \]

3. By the background knowledge generated by **Premise 2**, we abduce, because $b$ is not inexpensive, that these cannot be cigarettes.
Abducing the Conclusion

To explain why people validate $S_{\text{add}}$ we need to show that they reason abductively.

Consider $S_{\text{add}}$ again

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

1. We know that there are addictive things, let’s say $b$.
   
   $b$ is addictive.

2. Given Premise 1 we infer that these addictive things are not inexpensive.
   
   $b$ is not inexpensive.

3. By the background knowledge generated by Premise 2, we abduce, because $b$ is not inexpensive, that these cannot be cigarettes.
   
   $b$ is not a cigarette.
Modeling $S_{rich}$

If we model $S_{rich}$ analogously to $S_{add}$ we have the following background knowledge

\[ \text{Rich people are millionaires.} \]  \tag{1}
Modeling $S_{rich}$

If we model $S_{rich}$ analogously to $S_{add}$ we have the following background knowledge

$$\textit{Rich people are millionaires.} \quad (1)$$

Additionally, given $\text{Premise 2}$

$$\textit{Rich people are hard workers (compared to other millionaires).} \quad (2)$$
Modeling $S_{rich}$

If we model $S_{rich}$ analogously to $S_{add}$ we have the following background knowledge

- Rich people are millionaires. (1)
- additionally, given Premise 2

  Rich people are hard workers (compared to other millionaires). (2)

Premise 1 of $S_{rich}$ would then be represented as

- If someone is a hard worker and not abnormal, then this person is not a millionaire. (3)
  Nobody is abnormal (wrt (3)).

The belief in (1) and (2) would generate the exception for rich people

- If someone is rich, then this person is abnormal (wrt (3)). (4)
Modeling $S_{rich}$

If we model $S_{rich}$ analogously to $S_{add}$ we have the following background knowledge

Rich people are millionaires.  \hfill (1)

additionally, given **Premise 2**

Rich people are hard workers (compared to other milionaires).  \hfill (2)

**Premise 1** of $S_{rich}$ would then be represented as

If someone is a hard worker and not abnormal, then this person is not a milionaire.  \hfill (3)

Nobody is abnormal (wrt (3)).

The belief in (1) and (2) would generate the exception for rich people

If someone is rich, then this person is abnormal (wrt (3)).  \hfill (4)

Even though not tested yet, our hypothesis is, while checking $S_{rich}$, participants did not make these assumptions and thus, had not been influenced by the belief-bias effect.
Thank you very much for your attention!


We restrict ourselves to datalog programs. A logic program $\mathcal{P}$ is a finite set of clauses

\[ A \leftarrow A_1 \land \ldots \land A_n \land \neg B_1 \land \ldots \land \neg B_m, \quad (1) \]
\[ A \leftarrow \bot, \quad (2) \]

- where $A$ and $A_i$, $0 \leq i \leq n$, are atoms and $\neg B_j$, $1 \leq j \leq m$, are negated atoms.
- If $i = 0$, then we write $A \leftarrow \top$, which is called a positive fact.
- A clause of the form (2) is called a negative fact.
- $A$ is undefined if it is not the head of any clause.
- $g\mathcal{P}$ denotes ground $\mathcal{P}$, that is, it contains all ground instances of its clauses.
- $\text{undef}(\mathcal{P})$ is the set of all undefined atoms in $g\mathcal{P}$. 
Logic Programs

We restrict ourselves to datalog programs. A logic program \( \mathcal{P} \) is a finite set of clauses

\[
A \leftarrow A_1 \land \ldots \land A_n \land \neg B_1 \land \ldots \land \neg B_m, \quad (1)
\]

\[
A \leftarrow \bot, \quad (2)
\]

- where \( A \) and \( A_i, \ 0 \leq i \leq n \), are atoms and \( \neg B_j, \ 1 \leq j \leq m \), are negated atoms.
- If \( i = 0 \), then we write \( A \leftarrow \top \), which is called a positive fact.
- A clause of the form (2) is called a negative fact.
- \( A \) is undefined if it is not the head of any clause.
- \( g \mathcal{P} \) denotes ground \( \mathcal{P} \), that is, it contains all ground instances of its clauses.
- \( \text{undef}(\mathcal{P}) \) is the set of all undefined atoms in \( g \mathcal{P} \).

The following transformation is the weak completion of \( \mathcal{P} \)

1. Replace all clauses in \( g \mathcal{P} \) with the same head \( A \leftarrow \text{body}_1, \ldots, A \leftarrow \text{body}_n \) by the single expression \( A \leftarrow \text{body}_1 \lor \ldots \lor \text{body}_n \).
2. Replace all occurrences of \( \leftarrow \) by \( \leftrightarrow \).
Three-Valued Łukasiewicz Logic

<table>
<thead>
<tr>
<th>¬</th>
<th>⊤</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>⊥</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>⊥</td>
<td>T</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>∧</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>⊥</td>
<td>⊥</td>
<td>⊥</td>
<td>⊥</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>∨</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>U</td>
<td>T</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>⊥</td>
<td>T</td>
<td>U</td>
<td>⊥</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>←L</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>⊥</td>
<td>⊥</td>
<td>U</td>
<td>T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>↔L</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>T</td>
<td>U</td>
</tr>
<tr>
<td>⊥</td>
<td>⊥</td>
<td>U</td>
<td>T</td>
</tr>
</tbody>
</table>

Table:  ⊤,  ⊥,  and  U denote  true,  false,  and  unknown,  respectively.
Three-Valued Łukasiewicz Logic

<table>
<thead>
<tr>
<th></th>
<th>¬</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>⊥</td>
</tr>
<tr>
<td>⊥</td>
<td>T</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>∧</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>⊥</td>
<td>⊥</td>
<td>⊥</td>
<td>⊥</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>∨</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>U</td>
<td>T</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>⊥</td>
<td>T</td>
<td>U</td>
<td>⊥</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>←L</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>⊥</td>
<td>⊥</td>
<td>U</td>
<td>T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>↔L</th>
<th>T</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>T</td>
<td>U</td>
</tr>
<tr>
<td>⊥</td>
<td>⊥</td>
<td>U</td>
<td>T</td>
</tr>
</tbody>
</table>

**Table:** T, ⊥, and U denote *true*, *false*, and *unknown*, respectively.

An interpretation $I$ of $g \mathcal{P}$ is a mapping of the Herbrand base $B_{\mathcal{P}}$ to \{T, ⊥, U\} and is represented by an unique pair, \langle$I^\top$, $I^\bot$\rangle, where

$I^\top = \{A \in B_{\mathcal{P}} \mid A \text{ is mapped to } T\}$ and $I^\bot = \{A \in B_{\mathcal{P}} \mid A \text{ is mapped to } ⊥\}$. 

Three-Valued Łukasiewicz Logic

<table>
<thead>
<tr>
<th>¬</th>
<th>⊤</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>⊥</td>
<td>⊥</td>
<td>⊥</td>
</tr>
<tr>
<td>⊥</td>
<td>⊤</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>⊥</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>∧</th>
<th>⊤</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>⊤</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>⊥</td>
<td>⊤</td>
<td>⊥</td>
<td>⊥</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>∨</th>
<th>⊤</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>⊤</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>⊥</td>
<td>⊤</td>
<td>U</td>
<td>⊥</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>←L</th>
<th>⊤</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>⊤</td>
<td>⊤</td>
<td>⊤</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>⊥</td>
<td>⊤</td>
<td>U</td>
<td>T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>↔L</th>
<th>⊤</th>
<th>U</th>
<th>⊥</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>⊤</td>
<td>U</td>
<td>⊥</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
<td>T</td>
<td>U</td>
</tr>
<tr>
<td>⊥</td>
<td>⊤</td>
<td>U</td>
<td>T</td>
</tr>
</tbody>
</table>

Table: ⊤, ⊥, and U denote true, false, and unknown, respectively.

An interpretation \( I \) of \( \mathcal{P} \) is a mapping of the Herbrand base \( \mathcal{B}_\mathcal{P} \) to \( \{⊤, ⊥, U\} \) and is represented by an unique pair, \( \langle I^\top, I^\bot \rangle \), where

\[
I^\top = \{A \in \mathcal{B}_\mathcal{P} \mid A \text{ is mapped to } ⊤\} \text{ and } I^\bot = \{A \in \mathcal{B}_\mathcal{P} \mid A \text{ is mapped to } ⊥\}.
\]

- For every \( I \) it holds that \( I^\top \cap I^\bot = \emptyset \).
- A model of a formula \( F \) is an interpretation \( I \) such that \( F \) is true under \( I \).
- A model of \( \mathcal{P} \) is an interpretation that is a model of each clause in \( \mathcal{P} \).
Hölldobler and Kencana Ramli [2009] propose to compute the least model of the weak completion of \( P \) (\( \text{Im}_{\text{wc}} P \)) which is identical to the least fixed point of \( \Phi_P \), by an operator defined by Stenning and van Lambalgen [2008].
Hölldobler and Kencana Ramli [2009] propose to compute the least model of the weak completion of $\mathcal{P}$ ($\text{Im}_{\text{lcwc}} \mathcal{P}$) which is identical to the least fixed point of $\Phi_\mathcal{P}$, by an operator defined by Stenning and van Lambalgen [2008].

Let $I$ be an interpretation in $\Phi_\mathcal{P}(I) = \langle J^\top, J^\perp \rangle$, where

\begin{align*}
J^\top &= \{ A \mid \text{there exists } A \leftarrow \text{body} \in g\mathcal{P} \text{ with } I(\text{body}) = \top \}, \\
J^\perp &= \{ A \mid \text{there exists } A \leftarrow \text{body} \in g\mathcal{P} \text{ and } \\
    &\quad \text{for all } A \leftarrow \text{body} \in g\mathcal{P} \text{ we find } I(\text{body}) = \bot \}.
\end{align*}
Hölldobler and Kencana Ramli [2009] propose to compute the least model of the weak completion of $\mathcal{P}$ ($\text{lm}_{\text{wc}} \mathcal{P}$) which is identical to the least fixed point of $\Phi_{\mathcal{P}}$, by an operator defined by Stenning and van Lambalgen [2008].

Let $I$ be an interpretation in $\Phi_{\mathcal{P}}(I) = \langle J^\top, J^\perp \rangle$, where

\[
J^\top = \{ A \mid \text{there exists } A \leftarrow \text{body} \in g \mathcal{P} \text{ with } I(\text{body}) = \top \}, \\
J^\perp = \{ A \mid \text{there exists } A \leftarrow \text{body} \in g \mathcal{P} \text{ and} \\
\quad \text{for all } A \leftarrow \text{body} \in g \mathcal{P} \text{ we find } I(\text{body}) = \bot \}.
\]

Hölldobler and Kencana Ramli showed that the model intersection property holds for weakly completed programs. This guarantees the existence of least models for every $\mathcal{P}$. 
Hölldobler and Kencana Ramli [2009] propose to compute the least model of the weak completion of \( \mathcal{P} \) (\( \text{Im}_{\text{wc}} \mathcal{P} \)) which is identical to the least fixed point of \( \Phi_\mathcal{P} \), by an operator defined by Stenning and van Lambalgen [2008].

Let \( I \) be an interpretation in \( \Phi_\mathcal{P}(I) = \langle J^\top, J^\perp \rangle \), where

\[
J^\top = \{ A | \text{there exists } A \leftarrow \text{body} \in g \mathcal{P} \text{ with } I(\text{body}) = \top \},
\]
\[
J^\perp = \{ A | \text{there exists } A \leftarrow \text{body} \in g \mathcal{P} \text{ and for all } A \leftarrow \text{body} \in g \mathcal{P} \text{ we find } I(\text{body}) = \bot \}.
\]

Hölldobler and Kencana Ramli showed that the model intersection property holds for weakly completed programs. This guarantees the existence of least models for every \( \mathcal{P} \).

In (Dietz, Hölldobler, and Wernhard [2014]) we show that weak completion semantics corresponds to well-founded semantics for modified tight logic programs.
Abduction (Kakas, Kowalski, and Toni [1993])

Given an abductive framework $\langle \mathcal{P}, \mathcal{A}, \models_{\text{lmwc}} \rangle$ where
Abduction (Kakas, Kowalski, and Toni [1993])

Given an abductive framework \( \langle P, A, \models^{\text{lmwc}}_L \rangle \) where
- set of abducibles \( A \) contains all positive and negative facts of each \( A \in \text{undef}(P) \),
- \( E \) is an explanation and a consistent subset of \( A \),
- logical consequence relation \( \models^{\text{lmwc}}_L \), where \( P \models^{\text{lmwc}}_L F \) iff \( \text{lm}_{\text{wc}} P(F) = \top \), and
- \( O \) is an observation which is a set of (at least one) literals.
Abduction (Kakas, Kowalski, and Toni [1993])

Given an abductive framework $\langle \mathcal{P}, \mathcal{A}, \models_{\mathcal{L}}^{\text{lmwc}} \rangle$ where

- set of abducibles $\mathcal{A}$ contains all positive and negative facts of each $A \in \text{undef}(\mathcal{P})$,
- $E$ is an explanation and a consistent subset of $\mathcal{A}$,
- logical consequence relation $\models_{\mathcal{L}}^{\text{lmwc}}$, where $\mathcal{P} \models_{\mathcal{L}}^{\text{lmwc}} F$ iff $\text{lm}_{\mathcal{L}}\text{wc} \mathcal{P}(F) = \top$, and
- $O$ is an observation which is a set of (at least one) literals.

$O$ is explained by $E$ given $\mathcal{P}$ iff $\mathcal{P} \cup E \models_{\mathcal{L}}^{\text{lmwc}} O$, where $\mathcal{P} \not\models_{\mathcal{L}}^{\text{lmwc}} O$.

$O$ is explained given $\mathcal{P}$ iff there exists an $E$ such that $O$ is explained by $E$ given $\mathcal{P}$.
Abduction (Kakas, Kowalski, and Toni [1993])

Given an abductive framework $\langle \mathcal{P}, \mathcal{A}, \vdash_{\text{lmc}} \rangle$ where

- set of abducibles $\mathcal{A}$ contains all positive and negative facts of each $A \in \text{undef}(\mathcal{P})$,
- $\mathcal{E}$ is an explanation and a consistent subset of $\mathcal{A}$,
- logical consequence relation $\vdash_{\text{lmc}}$, where $\mathcal{P} \vdash_{\text{lmc}} F$ iff $\text{lmc} \mathcal{P}(F) = T$, and
- $\mathcal{O}$ is an observation which is a set of (at least one) literals.

$\mathcal{O}$ is explained by $\mathcal{E}$ given $\mathcal{P}$ iff $\mathcal{P} \cup \mathcal{E} \vdash_{\text{lmc}} \mathcal{O}$, where $\mathcal{P} \not\vdash_{\text{lmc}} \mathcal{O}$.

$\mathcal{O}$ is explained given $\mathcal{P}$ iff there exists an $\mathcal{E}$ such that $\mathcal{O}$ is explained by $\mathcal{E}$ given $\mathcal{P}$.

$F$ follows skeptically from $\mathcal{P}$, and $\mathcal{O}$ iff $\mathcal{O}$ can be explained given $\mathcal{P}$, and for all minimal explanations $\mathcal{E}$ we find that $\mathcal{P} \cup \mathcal{E} \vdash_{\text{lmc}} \mathcal{O}$.

$F$ follows credulously from $\mathcal{P}$, and $\mathcal{O}$ iff there exists a minimal explanation $\mathcal{E}$ such that $\mathcal{P} \cup \mathcal{E} \vdash_{\text{lmc}} \mathcal{O}$. 

\( F \) follows skeptically from \( \mathcal{P} \), and \( \mathcal{O} \) iff \( \mathcal{O} \) can be explained given \( \mathcal{P} \), and for all minimal explanations \( \mathcal{E} \) we find that \( \mathcal{P} \cup \mathcal{E} \vdash_{\text{lmc}} \mathcal{O} \).

\( F \) follows credulously from \( \mathcal{P} \), and \( \mathcal{O} \) iff there exists a minimal explanation \( \mathcal{E} \) such that \( \mathcal{P} \cup \mathcal{E} \vdash_{\text{lmc}} \mathcal{O} \).
Abducting the **Conclusion**

Given our background knowledge we know, there are addictive things, let’s say about $b$

$$\mathcal{O}_{add(b)} = \{add(b)\}$$

We have two minimal explanations for $\mathcal{O}_{add(b)}$

\[
\text{Im}_{Lwc}(\mathcal{P}_{add} \cup \mathcal{E}_{cig(b)}) = \langle \{add(b), cig(b), inex(b), \ldots\}, \{\ldots\}\rangle
\]

\[
\text{Im}_{Lwc}(\mathcal{P}_{add} \cup \mathcal{E}_{\neg cig(b)}) = \langle \{add(b), \ldots\}, \{cig(b), inex(b), \ldots\}\rangle
\]

Credulously, we validate some addictive things are not cigarettes.
Abducing the Conclusion

Given our background knowledge we know, there are addictive things, let’s say about $b$

$$O_{add(b)} = \{add(b)\}$$

We have two minimal explanations for $O_{add(b)}$

$$\text{Im}_Lwc(P_{add} \cup E_{cig(b)}) = \langle \{add(b), cig(b), inex(b), \ldots\}, \{\ldots\} \rangle$$

$$\text{Im}_Lwc(P_{add} \cup E_{\neg cig(b)}) = \langle \{add(b), \ldots\}, \{cig(b), inex(b), \ldots\} \rangle$$

Recall $P_{add}$. Together with $E_{cig(b)}$ it contains

$$\begin{align*}
add'(X) & \leftarrow \text{inex}(X) \land \neg\text{ab}_{add'}(X), & add(X) & \leftarrow \neg\text{add}’(X), \\
inex(X) & \leftarrow \text{cig}(X) \land \neg\text{ab}_{inex}(X), & \text{ab}_{add'}(X) & \leftarrow \text{cig}(X), \\
\text{ab}_{add'}(X) & \leftarrow \perp, & \text{ab}_{inex}(X) & \leftarrow \perp, \\
\text{cig}(b) & \leftarrow \top.
\end{align*}$$
Abducing the Conclusion

Given our background knowledge we know, there are addictive things, let’s say about $b$

$$\mathcal{O}_{add(b)} = \{add(b)\}$$

We have two minimal explanations for $\mathcal{O}_{add(b)}$

$$\text{Im}_{\text{wc}}(\mathcal{P}_{add} \cup \mathcal{E}_{cig(b)}) = \langle \{add(b), cig(b), inex(b), \ldots\}, \{\ldots\} \rangle$$

$$\text{Im}_{\text{wc}}(\mathcal{P}_{add} \cup \mathcal{E}_{\neg cig(b)}) = \langle \{add(b), \ldots\}, \{cig(b), inex(b), \ldots\} \rangle$$

Recall $\mathcal{P}_{add}$. Together with $\mathcal{E}_{\neg cig(b)}$ it contains

$$\begin{align*}
\text{add}'(X) & \iff \text{inex}(X) \land \neg \text{ab}_{add'}(X), \\
\text{add}(X) & \iff \neg \text{add}'(X), \\
\text{inex}(X) & \iff \text{cig}(X) \land \neg \text{ab}_{inex}(X), \\
\text{ab}_{add'}(X) & \iff \text{cig}(X), \\
\text{ab}_{inex}(X) & \iff \bot, \\
\text{cig}(b) & \iff \bot.
\end{align*}$$
Abducing the \textbf{Conclusion}

Given our background knowledge we know, there are addictive things, let’s say about \( b \)

\[
\mathcal{O}_{\text{add}(b)} = \{ \text{add}(b) \}
\]

We have two minimal explanations for \( \mathcal{O}_{\text{add}(b)} \)

\[
\text{Im}_{\text{Lwc}}(\mathcal{P}_{\text{add}} \cup \mathcal{E}_{\text{cig}(b)}) = \langle \{\text{add}(b), \text{cig}(b), \text{inex}(b), \ldots\}, \{\ldots\} \rangle
\]
\[
\text{Im}_{\text{Lwc}}(\mathcal{P}_{\text{add}} \cup \mathcal{E}_{\neg \text{cig}(b)}) = \langle \{\text{add}(b), \ldots\}, \{\text{cig}(b), \text{inex}(b), \ldots\} \rangle
\]

Recall \( \mathcal{P}_{\text{add}} \). Together with \( \mathcal{E}_{\neg \text{cig}(b)} \) it contains

\[
\begin{align*}
\text{add}'(X) & \leftarrow \text{inex}(X) \land \neg \text{ab}_{\text{add}'}(X), \\
\text{add}(X) & \leftarrow \neg \text{add}'(X), \\
\text{inex}(X) & \leftarrow \text{cig}(X) \land \neg \text{ab}_{\text{inex}}(X), \\
\text{ab}_{\text{add}'}(X) & \leftarrow \text{cig}(X), \\
\text{ab}_{\text{inex}}(X) & \leftarrow \bot, \\
\text{cig}(b) & \leftarrow \bot.
\end{align*}
\]

Credulously, we validate \textbf{some addictive things are not cigarettes.}
Contextual Abductive Reasoning

How to express that Premise 1 describes the usual and Premise 2 the exceptional case? Inexpensive cigarette should be the exception in the context of addictive things.
How to express that Premise 1 describes the usual and Premise 2 the exceptional case? Inexpensive cigarette should be the exception in the context of addictive things.

Introduce for every \( A \), two reserved (meta-) predicates (Pereira and Pinto [2011]).

\[
\text{inspect}(A) \quad \text{and} \quad \text{inspect}_{\text{not}}(A)
\]

These are special abducibles only to be abduced if \( A \) or \( \neg A \) are abduced somewhere else already.
How to express that Premise 1 describes the usual and Premise 2 the exceptional case? Inexpensive cigarette should be the exception in the context of addictive things.

Introduce for every $A$, two reserved (meta-) predicates (Pereira and Pinto [2011]).

$$\text{inspect}(A) \quad \text{and} \quad \text{inspect}_{\text{not}}(A)$$

These are special abducibles only to be abduced if $A$ or $\neg A$ are abduced somewhere else already. We replace the $ab_{add'}$-clause in $P_{add}$. $P_{add_{\text{insp}}}$ contains

$$\begin{align*}
    add'(X) & \leftarrow \text{inex}(X) \land \neg ab_{add'}(X), \\
    \text{inex}(X) & \leftarrow \text{cig}(X) \land \neg ab_{\text{inex}}(X), \\
    ab_{add'}(X) & \leftarrow \bot,
\end{align*}$$

$$\begin{align*}
    \text{add}(X) & \leftarrow \neg add'(X), \\
    ab_{\text{inex}}(X) & \leftarrow \bot,
\end{align*}$$

Suppose again $b$ is addictive, i.e. $O_{add}(b) = \{\text{add}(b)\}$. $E_{\text{cig}} = \{\text{cig}(b) \leftarrow \top\}$ cannot be abduced anymore to explain $O_{add}(b)$. Its only minimal explanation is $E_{\neg \text{cig}}(b)$.
Contextual Abductive Reasoning

How to express that Premise 1 describes the usual and Premise 2 the exceptional case? Inexpensive cigarette should be the exception in the context of addictive things.

Introduce for every \( A \), two reserved (meta-) predicates (Pereira and Pinto [2011]).

\[
\text{inspect}(A) \quad \text{and} \quad \text{inspect}_{\text{not}}(A)
\]

These are special abducibles only to be abduced if \( A \) or \( \neg A \) are abduced somewhere else already. We replace the \( ab_{\text{add}}' \)-clause in \( P_{\text{add}} \). \( P_{\text{add}_{\text{insp}}} \) contains

\[
\begin{align*}
\text{add}'(X) & \leftarrow \text{inex}(X) \land \neg \text{add}'(X), \\
\text{add}(X) & \leftarrow \neg \text{add}'(X), \\
\text{inex}(X) & \leftarrow \text{cig}(X) \land \neg \text{inex}_{}(X), \\
\text{ab}_{\text{add}}'(X) & \leftarrow \bot, \\
\text{ab}_{\text{inex}}(X) & \leftarrow \bot,
\end{align*}
\]

Suppose again \( b \) is addictive, i.e. \( O_{\text{add}}(b) = \{\text{add}(b)\} \). \( E_{\text{cig}} = \{\text{cig}(b) \leftarrow \top\} \) cannot be abduced anymore to explain \( O_{\text{add}}(b) \). Its only minimal explanation is \( E_{\neg \text{cig}(b)} \).
Contextual Abductive Reasoning

How to express that Premise 1 describes the usual and Premise 2 the exceptional case? Inexpensive cigarette should be the exception in the context of addictive things.

Introduce for every $A$, two reserved (meta-) predicates (Pereira and Pinto [2011]).

$$\text{inspect}(A) \quad \text{and} \quad \text{inspect}_{\text{not}}(A)$$

These are special abducibles only to be abduced if $A$ or $\neg A$ are abduced somewhere else already. We replace the $ab_{add}'$-clause in $P_{add}$. $P_{add_{\text{insp}}}$ contains

$$
\begin{align*}
\text{add}'(X) & \leftarrow \text{inex}(X) \land \neg \text{ab}_{add}'(X), \\
\text{add}(X) & \leftarrow \neg \text{add}'(X), \\
\text{inex}(X) & \leftarrow \text{cig}(X) \land \neg \text{ab}_{\text{inex}}(X), \\
\text{ab}_{\text{add}}'(X) & \leftarrow \text{inspect}(\text{cig}(X)), \\
\text{ab}_{\text{inex}}(X) & \leftarrow \bot,
\end{align*}
$$

Suppose again $b$ is addictive, i.e. $O_{\text{add}}(b) = \{\text{add}(b)\}$. $E_{\text{cig}} = \{\text{cig}(b) \leftarrow \top\}$ cannot be abduced anymore to explain $O_{\text{add}}(b)$. Its only minimal explanation is $E_{\neg \text{cig}(b)}$.

Skeptically, we validate some addictive things are not cigarettes.
Contextual Abductive Reasoning

Suppose that we observe that \( b \) is addictive and inexpensive

\[
\mathcal{O}_{add, inex} = \{\text{add}(b), \text{inex}(b)\}
\]

\[
g \overset{\text{cons}}{\underset{\text{add}_\text{insp}}{\models}} \text{where cons} = \{b\}, \text{contains}
\]

\[
inex(b) \leftarrow \text{cig}(b) \land \neg \text{ab}_{inex}(b), \quad \text{ab}_{inex}(b) \leftarrow \bot,
\]

\[
\text{ab}_{add'}(b) \leftarrow \text{inspect}(\text{cig}(b)).
\]
Contextual Abductive Reasoning

Suppose that we observe that $b$ is addictive and inexpensive

\[
\mathcal{O}_{\text{add}, \text{inex}} = \{\text{add}(b), \text{inex}(b)\}
\]

$g \mathcal{P}_{\text{cons}}^{\text{add}_\text{insp}}$ where $\text{cons} = \{b\}$, contains

\[
\begin{align*}
\text{inex}(b) & \leftarrow \text{cig}(b) \land \neg \text{ab}_{\text{inex}}(b), & \text{ab}_{\text{inex}}(b) & \leftarrow \bot, \\
\text{ab}_{\text{add}'}(b) & \leftarrow \text{inspect}(\text{cig}(b)).
\end{align*}
\]

- Because $b$ is inexpensive, $\text{cig}(b) \leftarrow \top$ can be directly abduced to explain $\text{inex}(b)$. 
Suppose that we observe that $b$ is addictive and inexpensive

$$O_{\text{add, inex}} = \{\text{add}(b), \text{inex}(b)\}$$

$g P^\text{cons}_{\text{add}_{\text{insp}}}$ where $\text{cons} = \{b\}$, contains

- $\text{inex}(b) \leftarrow \text{cig}(b) \land \neg \text{ab}_{\text{inex}}(b)$, $\text{ab}_{\text{inex}}(b) \leftarrow \bot$,
- $\text{ab}_{\text{add}'}(b) \leftarrow \text{inspect}(\text{cig}(b))$.

- Because $b$ is inexpensive, $\text{cig}(b) \leftarrow \top$ can be directly abduced to explain $\text{inex}(b)$.
- As $\text{inex}(b)$ is explained, $\text{inspect}(\text{cig}(b)) \leftarrow \top$ can be abduced to explain $\text{add}(b)$. 
Contextual Abductive Reasoning

Suppose that we observe that $b$ is addictive and inexpensive

\[ O_{\text{add}, \text{inex}} = \{ \text{add}(b), \text{inex}(b) \} \]

$g \mathcal{P}_{\text{cons}}^{\text{add}_{\text{insp}}}$ where cons $= \{b\}$, contains

\[ \text{inex}(b) \leftarrow \text{cig}(b) \land \neg \text{ab}_{\text{inex}}(b), \quad \text{ab}_{\text{inex}}(b) \leftarrow \bot, \quad \text{ab}_{\text{add}'}(b) \leftarrow \text{inspect}(\text{cig}(b)). \]

- Because $b$ is inexpensive, $\text{cig}(b) \leftarrow \top$ can be directly abduced to explain $\text{inex}(b)$.
- As $\text{inex}(b)$ is explained, $\text{inspect}(\text{cig}(b)) \leftarrow \top$ can be abduced to explain $\text{add}(b)$.

Recall the two clauses in $g \mathcal{P}_{\text{cons}}^{\text{add}_{\text{insp}}}$ representing PREMISE 1 are

\[ \text{add}'(b) \leftarrow \text{inex}(b) \land \neg \text{ab}_{\text{add}'}(b), \quad \text{add}(b) \leftarrow \neg \text{add}'(b). \]
Contextual Abductive Reasoning

Suppose that we observe that \(b\) is addictive and inexpensive

\[
\mathcal{O}_{\text{add}, \text{inex}} = \{\text{add}(b), \text{inex}(b)\}
\]

\(g \mathcal{P}^{\text{cons}}_{\text{add}, \text{insp}}\) where \(\text{cons} = \{b\}\), contains

\[
\begin{align*}
\text{inex}(b) & \leftarrow \text{cig}(b) \land \neg \text{ab}_{\text{inex}}(b), & \text{ab}_{\text{inex}}(b) & \leftarrow \bot, \\
\text{ab}_{\text{add'}}(b) & \leftarrow \text{inspect}(\text{cig}(b)).
\end{align*}
\]

- Because \(b\) is inexpensive, \(\text{cig}(b) \leftarrow \top\) can be directly abduced to explain \(\text{inex}(b)\).
- As \(\text{inex}(b)\) is explained, \(\text{inspect}(\text{cig}(b)) \leftarrow \top\) can be abduced to explain \(\text{add}(b)\).

Recall the two clauses in \(g \mathcal{P}^{\text{cons}}_{\text{add}, \text{insp}}\) representing \textsc{Premise 1} are

\[
\begin{align*}
\text{add'}(b) & \leftarrow \text{inex}(b) \land \neg \text{ab}_{\text{add'}}(b), \\
\text{add}(b) & \leftarrow \neg \text{add'}(b).
\end{align*}
\]