A Computational Logic Approach to The Belief-Bias Effect in Human Reasoning

Emmanuelle Dietz
International Center for Computational Logic
TU Dresden, Germany

Luís Moniz Pereira
Centro de Inteligência Artificial (CENTRIA)
Universidade Nova de Lisboa, Portugal
The two minds hypothesis distinguishes between:

- the reflective mind, and
- the intuitive mind

This hypothesis is supported by showing the belief-bias effect (Evans [2012]):

- The belief-bias effect is the conflict between the reflective and intuitive minds when reasoning about problems involving real-world beliefs. It is the tendency to accept or reject arguments based on own beliefs or prior knowledge rather than on the reasoning process.

How to identify the belief-bias effect?

- Do psychological studies about deductive reasoning which demonstrate possibly conflicting processes in the logical and psychological level.
Evans et al. [1983] carried out an experiment where participants were presented different syllogisms for which they should decide whether they were logically valid.

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
<th>Evans</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{dogs}$</td>
<td>valid and believable</td>
<td>No police dogs are vicious. Some highly trained dogs are vicious. Therefore, some highly trained dogs are not police dogs.</td>
</tr>
<tr>
<td>$S_{vit}$</td>
<td>valid and unbelievable</td>
<td>No nutritional things are inexpensive. Some vitamin tablets are inexpensive. Therefore, some vitamin tablets are not nutritional.</td>
</tr>
<tr>
<td>$S_{add}$</td>
<td>invalid and believable</td>
<td>No addictive things are inexpensive. Some cigarettes are inexpensive. Therefore, some addictive things are not cigarettes.</td>
</tr>
<tr>
<td>$S_{rich}$</td>
<td>invalid and unbelievable</td>
<td>No millionaires are hard workers. Some rich people are hard workers. Therefore, some millionaires are not rich people.</td>
</tr>
</tbody>
</table>

Using their reflexive minds, people read the instructions and understand that they are required to reason logically from the premises to the conclusion. However, when they look at the conclusion, their intuitive minds deliver strong tendency to say yes or no depending on whether it is believable.
How to adequately formalize human reasoning in computational logic?

Stenning and van Lambalgen (2008) propose a two step process: human reasoning should be modeled by

1. reasoning towards an appropriate representation, → conceptual adequacy
2. reasoning with respect to this representation. → inferential adequacy

The adequacy of a computational logic approach that aims a representing human reasoning should be evaluated based on how humans actually reason.
State of the Art


2. Kowalski (2011) models Wason’s (1968) Selection Task, showing that people have a matching bias, the tendency to select explicitly named values in conditionals.

3. Hölldobler and Kencana Ramli (2009a) found some technical mistakes done by Stenning and van Lambalgen and propose to model human reasoning by
   
   - logic programs
   - under weak completion semantics
   - based on the three-valued Łukasiewicz (1920) logic.

This approach seems to adequately model the Suppression and the Selection Task.

Can we adequately model the syllogisms task including the belief-bias effect under weak completion semantics?
Weak Completion Semantics
A (first-order) logic program $\mathcal{P}$ is a finite set of clauses of the form

$$p(X) \leftarrow a_1(X) \land \cdots \land a_n(X) \land \neg b_1(X) \land \cdots \land \neg b_m(X)$$

- $X$ is a variable and $p$, $a_1, \ldots, a_n$ and $b_1, \ldots, b_m$ are predicate symbols.
- $p(X)$ is an atom and head of the clause.
- $a_1(X) \land \cdots \land a_n(X) \land \neg b_1(X) \land \cdots \land \neg b_m(X)$ is a formula and body of the clause.
- $p(X) \leftarrow \top$ and $p(X) \leftarrow \bot$ denote positive and negative facts, respectively.
- Variables are written with upper case and constants with lower case ones.

- A ground formula is a formula that does not contain variables.
- A ground instance results from substituting all occurrences of each variable name by some constant of $\mathcal{P}$.
- A ground program $\text{g} \mathcal{P}$ is comprised of all ground instances of its clauses.
- The set of all atoms in $\text{g} \mathcal{P}$ is denoted by $\text{atoms}(\mathcal{P})$.
- An atom is undefined in $\text{g} \mathcal{P}$ if it is not the head of some clause in $\text{g} \mathcal{P}$. The corresponding set of these atoms is denoted by $\text{undef}(\mathcal{P})$. 
Consider the following transformation for a given $\mathcal{P}$:

1. Replace all clauses in $g\mathcal{P}$ with the same head (ground atom) $A \leftarrow body_1, A \leftarrow body_2, \ldots$ by the single expression $A \leftarrow body_1 \lor body_2, \lor \ldots$.
2. If $A \in \text{undef}(g\mathcal{P})$ then add $A \leftarrow \bot$.
3. Replace all occurrences of $\leftarrow$ by $\leftrightarrow$.

The resulting set of equivalences is called the completion of $\mathcal{P}$ (Clark [1978]).

If Step 2 is omitted, then the resulting set is called the weak completion of $\mathcal{P}$ ($\text{wc}\mathcal{P}$) (Hölldobler and Kencana Ramli [2009a,b]).
Three-Valued Łukasiewicz Logic

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

<table>
<thead>
<tr>
<th>←L</th>
<th>↔L</th>
</tr>
</thead>
<tbody>
<tr>
<td>⊤</td>
<td>⊤</td>
</tr>
<tr>
<td>U</td>
<td>U</td>
</tr>
<tr>
<td>⊥</td>
<td>U</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, $⟨I^\top, I^\bot⟩$, where

$I^\top = \{x \in \mathcal{B}_\mathcal{P} \mid x$ is mapped to $\top\}$ and $I^\bot = \{x \in \mathcal{B}_\mathcal{P} \mid x$ is mapped to $\bot\}$

- 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 $\mathcal{P}$. 
Reasoning in an Appropriate Logical Form
Premise 1 in our first case No police dogs are vicious is equivalent to

There does not exist an X such that police_dogs(X) and vicious(X).

We can write it as

For all X if police_dogs(X) and not abnormal then not vicious(X).
   By default, nothing is abnormal.

We use abnormality predicates to implement conditionals by licenses for implications (Stenning and van Lambalgen [2008]).

Problem: we only consider logic programs that allow positive heads in the clauses. We introduce \( p'(X) \) and \( p'(X) \leftarrow \neg p(X) \) for every negative conclusion \( \neg p(X) \):

\[
\begin{align*}
\text{Premise 1} & \quad \text{police_dogs}'(X) \leftarrow \text{vicious}(X) \wedge \neg ab_{1a}(X) \\
& \quad \text{police_dogs}(X) \leftarrow \neg \text{police_dogs}'(X), \\
& \quad ab_{1a}(X) \leftarrow \bot.
\end{align*}
\]

1More generally, we need an appropriate dual program like transformation when there are several positive rules.
## Commonsense Implications within the four Syllogisms

<table>
<thead>
<tr>
<th>Example</th>
<th>Commonsense Implication</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>$S_{dogs}$</strong>&lt;br&gt;No police dogs are vicious.&lt;br&gt;Some highly trained dogs are vicious.&lt;br&gt;Therefore, some highly trained dogs are not police dogs.</td>
<td>We generally assume that police dogs are highly trained.</td>
</tr>
<tr>
<td><strong>$S_{vit}$</strong>&lt;br&gt;No nutritional things are inexpensive.&lt;br&gt;Some vitamin tablets are inexpensive.&lt;br&gt;Therefore, some vitamin tablets are not nutritional.</td>
<td>The purpose of vitamin tablets is to aid nutrition.</td>
</tr>
<tr>
<td><strong>$S_{add}$</strong>&lt;br&gt;No addictive things are inexpensive.&lt;br&gt;Some cigarettes are inexpensive.&lt;br&gt;Therefore, some addictive things are not cigarettes.</td>
<td>We know that cigarettes are addictive.</td>
</tr>
<tr>
<td><strong>$S_{rich}$</strong>&lt;br&gt;No millionaires are hard workers.&lt;br&gt;Some rich people are hard workers.&lt;br&gt;Therefore, some millionaires are not rich people.</td>
<td>By definition, millionaires are rich.</td>
</tr>
</tbody>
</table>

The second premises in each case contain some background knowledge which might provide the motivation on whether to validate the syllogisms.
Modeling Syllogism $S_{dogs}$: valid argument, believable conclusion

**Premise 1**  
No police dogs are vicious.

**Premise 2**  
Some highly trained dogs are vicious.

**Conclusion**  
Therefore, some highly trained dogs are not police dogs.

**Premise 2** states facts about, let's say some $a$. The program for $S_{dogs}$ is:

\[
P_{dogs} = \{ \text{police\_dogs}'(X) \leftarrow \text{vicious}(X) \land \neg ab_{1a}(X), \]
\[
   \text{police\_dogs}(X) \leftarrow \neg \text{police\_dogs}'(X), ab_{1a}(X) \leftarrow \bot \}
\]
\[
\cup \{ \text{highly\_trained}(a) \leftarrow \top, \text{vicious}(a) \leftarrow \top \}
\]
\[
\cup \{ \text{highly\_trained}(X) \leftarrow \text{police\_dogs}(X) \land \neg ab_{1b}(X), ab_{1b}(X) \leftarrow \bot \}
\]

The corresponding weak completion is:

\[
\text{wc g } P_{dogs} = \{ \text{police\_dogs}'(a) \leftrightarrow \text{vicious}(a) \land \neg ab_{1}(a), \]
\[
   \text{police\_dogs}(a) \leftrightarrow \neg \text{police\_dogs}'(a), ab_{1}(a) \leftrightarrow \bot, \]
\[
   \text{highly\_trained}(a) \leftrightarrow \top \lor (\text{police\_dogs}(a) \land \neg ab_{1b}(a)), \]
\[
   \text{vicious}(a) \leftrightarrow \top, ab_{1b}(a) \leftrightarrow \bot \}
\]

How do we compute the intended model?
Reasoning with Respect to this Representation
Hölldobler and Kencana Ramli propose to compute the least model of the weak completion of \( P \) (\( \text{Im}_{\text{Lwc}} P \)) which is identical to the least fixed point of \( \Phi_P \), an operator defined by Stenning and van Lambalgen [2008].

Let \( I \) be an interpretation in \( \Phi_P(I) = (J^\top, J^\perp) \), where

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

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

$$\text{wc } g \mathcal{P}_{\text{dogs}} = \{ \text{police}\_\text{dogs}'(a) \leftrightarrow \text{vicious}(a) \land \neg ab_{1a}(a), \text{police}\_\text{dogs}(a) \leftrightarrow \neg \text{police}\_\text{dogs}'(a), ab_{1a}(a) \leftrightarrow \bot, \text{highly}\_\text{trained}(a) \leftrightarrow \top \lor (\text{police}\_\text{dogs}(a) \land \neg ab_{1b}(a)), \text{vicious}(a) \leftrightarrow \top, ab_{1b}(a) \leftrightarrow \bot \}$$

Let’s start with interpretation $I_0 = \langle \emptyset, \emptyset \rangle$:

$$I_1 = \Phi_g \mathcal{P}_{\text{dogs}}(I_0) = \langle \{ \text{vicious}(a), \text{highly}\_\text{trained}(a) \}, \{ ab_{1a}(a), ab_{1b}(a) \} \rangle$$

$$I_2 = \Phi_g \mathcal{P}_{\text{dogs}}(I_1) = \langle \{ \text{vicious}(a), \text{highly}\_\text{trained}(a), \text{police}\_\text{dogs}'(a) \}, \{ ab_{1a}(a), ab_{1b}(a) \} \rangle$$

$$I_3 = \Phi_g \mathcal{P}_{\text{dogs}}(I_2) = \langle \{ \text{vicious}(a), \text{highly}\_\text{trained}(a), \text{police}\_\text{dogs}'(a) \}, \{ ab_{1a}(a), ab_{1b}(a), \text{police}\_\text{dogs}(a) \} \rangle$$

$$I_4 = \Phi_g \mathcal{P}_{\text{dogs}}(I_3) = \Phi_g \mathcal{P}_{\text{dogs}}(I_2) \Leftarrow \text{lm}_{\text{wc}} (g \mathcal{P}_{\text{dogs}}(I_2))$$

This model confirms $S_{\text{dogs}}$, because $a$ is a highly trained dog and not a police dog.
Modeling Syllogism $S_{vit}$: valid argument, unbelievable conclusion

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

Premise 2 states facts about, lets say some $a$. The program for $S_{vit}$ is:

$$P_{vit} = \{ \text{nutritional'}(X) \leftarrow \text{inexp}(X) \land \neg ab_{2a}(X),$$
$$\text{nutritional}(X) \leftarrow \neg \text{nutritional'}(X), \ ab_{2a}(X) \leftarrow \bot\}$$
$$\cup \{ \text{vitamin}(a) \leftarrow \top, \ \text{inexp}(a) \leftarrow \top \} \cup \{ \text{ab}_{2a}(X) \leftarrow \text{vitamin}(X),$$
$$\text{inexp}(X) \leftarrow \text{vitamin}(X) \land \neg \text{ab}_{2b}(X), \ \text{ab}_{2b}(X) \leftarrow \bot\}$$

The least model of the weak completion of $gP_{vit}$ is

$$\text{lm}_{\text{wc}}P_{vit} = \langle \{ \text{vitamin}(a), \text{inexp}(a), \text{nutritional}(a), \text{ab}_{2a}(a) \}, \{ \text{nutritional'}(a), \text{ab}_{2b}(a) \} \rangle$$

This model does not validate $S_{vit}$ and this confirms how the participants responded.
Modeling Syllogism $S_{add}$: invalid argument, believable conclusion (1)

<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>

Premise 2 states facts about, lets say some $a$. The program for $S_{add}$ is:

$$
\mathcal{P}_{add} = \{ \text{addictive}'(X) \leftarrow \text{inexp}(X) \land \neg ab_3a(X), \\
\text{addictive}(X) \leftarrow \neg \text{addictive}'(X), \ ab_3a(X) \leftarrow \bot \}
\cup \{ \text{cigarettes}(a) \leftarrow \top, \ \text{inexp}(a) \leftarrow \top \} \cup \{ \ ab_3a(X) \leftarrow \text{cigarettes}(X), \\
\text{inexp}(X) \leftarrow \text{cigarettes}(X) \land \neg ab_3b(X), \ ab_3b(X) \leftarrow \bot \}
$$

The least model of the weak completion of $g\mathcal{P}_{add}$ is:

$$
\text{lm}_{\text{wc}} \mathcal{P}_{add} = \langle \{ \text{cigarettes}(a), \ \text{inexp}(a), \ \text{addictive}(a), \ ab_3a(a) \}, \ \{ \text{addictive}'(a), \ ab_3b(a) \} \rangle
$$

The Conclusion of $S_{add}$ states something that can obviously not be about $a$. 
Abductive Reasoning

**Definition (1)**

Let \( \langle P, A_P, \models_{L}^{\text{im wc}} \rangle \) be an abductive framework, where:

- \( P \) is a knowledge base,
- \( A_P \) a set of abducibles consisting of the (positive and negative) facts for each undefined atom in \( P \)
- \( \models_{L}^{\text{im wc}} \) a logical consequence relation where \( P \models_{L}^{\text{im wc}} F \) if and only if \( \text{Im}_{L \text{wc}} P(F) = \top \) for formula \( F \).

Let \( O \) be an observation and \( E \) be an explanation where \( O \) and \( E \) are consistent:

- \( O \) is explained by \( E \) given \( A_P \) and \( P \) iff \( P \cup E \models_{L}^{\text{im wc}} O \) where \( P \cup E \) is consistent, \( P \not\models_{L}^{\text{im wc}} O \) and \( E \subseteq A_P \).

We distinguish between two forms:

- \( F \) follows skeptically from \( P \) and \( O \) iff \( O \) can be explained, and for all minimal explanations \( E \) it holds that \( P \cup E \models_{L}^{\text{im wc}} O \),
- \( F \) follows credulously from \( P \) and \( O \) iff there exists a minimal explanation \( E \) such that \( P \cup E \models_{L}^{\text{im wc}} O \).

In the following we apply abduction with respect to **credulous reasoning**: to validate \( S_{\text{add}} \) we need one model, which contains something addictive that is not a cigarette.
We observe for something else, e.g. for $b$, that it is addictive:

$$O_{add} = \{\text{addictive}(b) \leftarrow \top\}.$$  

The set of abducibles for given $g$ ($P_{add} \cup O_{add}$), is:

$$A_g(P_{add} \cup O_{add}) = \{\text{cigarettes}(b) \leftarrow \top, \text{cigarettes}(b) \leftarrow \bot\}$$

We have the following two minimal explanations for $O_{add}$:

$$E_{\text{cig}} = \{\text{cigarettes}(b) \leftarrow \top\} \text{ and } E_{\neg\text{cig}} = \{\text{cigarettes}(b) \leftarrow \bot\}$$

The least models of the weak completion of $P_{add}$ together with each explanation are:

$$\text{lm}_{L\text{wc}} g(P_{add} \cup E_{\neg\text{cig}}) = \langle\{\text{addictive}(b)\}, \{\text{addictive}'(b), \text{cigarettes}(b), ab_{3a}(b), \text{inexpensive}(b), ab_{3b}(b)\}\rangle$$

$$\text{lm}_{L\text{wc}} g(P_{add} \cup E_{\text{cig}}) = \langle\{\text{addictive}(b), ab_{3a}(b), \text{cigarettes}(b), \text{inexpensive}(b)\}, \{\text{addictive}'(b), ab_{3b}(b)\}\rangle$$

Credulously, we conclude there exists some $b$ that is additive but not a cigarette.
Modeling Syllogism $S_{\text{rich}}$: invalid argument, unbelievable conclusion (1)

<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 program for $S_{\text{rich}}$ is:

\[
P_{\text{rich}} = \{ \text{millionaire}'(X) \leftarrow \text{hard_worker}(X) \land \neg \text{ab}_{4a}(X), \\
\text{millionaire}(X) \leftarrow \neg \text{millionaire}'(X), \text{ab}_{4a}(X) \leftarrow \bot \} \\
\cup \{ \text{rich}(a) \leftarrow \top, \text{hard_worker}(a) \leftarrow \top \} \\
\cup \{ \text{rich}(X) \leftarrow \text{millionaire}(X) \land \neg \text{ab}_{4b}(X), \text{ab}_{4b}(X) \leftarrow \bot \}\]

Again, what is talked about in PREMISE 2 cannot be the same as in the CONCLUSION.
Modeling Syllogism $S_{rich}$: invalid argument, unbelievable conclusion (2)

We observe for something else, e.g. for $b$, that is a millionaire:

$$O_{mil} = \{\text{millionaire}(b) \leftarrow \top\}$$

For $g \left( P_{rich} \cup O_{mil} \right)$ we have the following set of abducibles:

$$A_g ( P_{rich} \cup O_{mil} ) = \{ \text{hard\_worker}(b) \leftarrow \top, \text{hard\_worker}(b) \leftarrow \bot \}$$

The least model of the weak completion of $P_{rich}$ and $E_{\neg \text{hard\_worker}}$ is:

$$\text{Im}_{L \text{wc g}} ( P_{rich} \cup E_{\neg \text{hard\_worker}} ) = \langle \{\text{millionaire}(b), \text{rich}(b)\}, \{ab_4a(b), ab_4b(b), \text{millionaire}'(b), \text{hard\_worker}(b)\} \rangle$$

This model does not validate $S_{rich}$ and this confirms how the participants responded.
Different kinds of Abnormalities: the Suppression and the Cigarette Task

If she has an essay to finish, then she goes to the library.
If the library is open, then she goes to the library.
She has an essay to finish.

\[ P_{e+Add} = \{ l \leftarrow e \land \neg ab_1, ab_1 \leftarrow \neg o, l \leftarrow o \land \neg ab_2, ab_2 \leftarrow \neg e, e \leftarrow \top \} \]
\[ \text{Im}_{\text{lw}} P_{e+Add} = \langle \{ e \}, \{ ab_2 \} \rangle \]

We cannot conclude whether she goes to the library.

No addictive things are inexpensive.
Some cigarettes are inexpensive.
There are some addictive things. \( \mathcal{O}_{\text{add}} = \{ \text{addictive}(b) \leftarrow \top \} \)

We have two explanations for \( \mathcal{O}_{\text{add}} \) which have equal priority:

\[ \mathcal{E}_{\text{cig}} = \{ \text{cigarettes}(b) \leftarrow \top \} \quad \mathcal{E}_{\neg \text{cig}} = \{ \text{cigarettes}(b) \leftarrow \bot \} \]

However, instead of assuming that \( b \) is a cigarette, we would first prefer to conclude that \( b \) is not inexpensive, except if we have more information about \( b \).
Abductive Reasoning with Contextual Side-Effects
We define two forms of contextual side effects:

**Definition (2)**

Given a background knowledge \( \mathcal{P} \), two observations \( O_1 \) and \( O_2 \) and two explanations \( E_1 \) and \( E_2 \)

\( O_2 \) is a necessary contextual side-effect of explained \( O_1 \) iff for every explanation \( E_1 \) such that \( O_1 \) is explained by \( E_1 \) given \( \mathcal{P} \), there exists \( E_2 \) such that \( O_2 \) is explained by \( E_2 \) given \( \mathcal{P} \), and \( E_2 \) contains \( E_1 \).

\( O_2 \) is a strict necessary contextual side-effect of explained \( O_1 \) iff \( E_2 = E_1 \).

**Definition (3)**

Given a background knowledge \( \mathcal{P} \), two observations \( O_1 \) and \( O_2 \) and two explanations \( E_1 \) and \( E_2 \)

\( O_2 \) is a possible contextual side-effect of explained \( O_1 \) iff there exist explanations \( E_1 \) and \( E_2 \) such that \( O_1 \) is explained by \( E_1 \) given \( \mathcal{P} \), and \( O_2 \) is explained by \( E_2 \) given \( \mathcal{P} \), and \( E_2 \) contains \( E_1 \).

\( O_2 \) is a strict possible contextual side-effect of explained \( O_1 \) iff \( E_2 = E_1 \).

Definitions (2) and (3) correspond to skeptical and credulous reasoning, respectively.
Contextual Side-Effects: $S_{add}$

No addictive things are inexpensive.
Some cigarettes are inexpensive.
b is addictive. ($\mathcal{O}_{add} = \{\text{addictive}(b) \leftarrow \top\}$)

Assume, we additionally observe $\mathcal{O}_{inexp} = \{\text{inexp}(b) \leftarrow \top\}$. We can only explain it by:

$$\mathcal{E}_{cig} = \{\text{cigarettes}(b) \leftarrow \top\}$$

$\mathcal{O}_{add}$ is a strict necessary contextual side-effect of explained $\mathcal{O}_{inexp}$, given $\mathcal{P}_{add}$. But $\mathcal{O}_{inexp}$ is a strict possible contextual side-effect of explained $\mathcal{O}_{add}$ given $\mathcal{P}_{add}$ as well.

We didn’t specify which is the side-effect resulting from the primary explanation!
Pereira and Pinto [2011] distinguish between occurrences of abducibles which allow them to be produced and those which may only be consumed, if produced elsewhere. First, we introduce a reserved (meta-)predication for every abductive ground atom:

\[ U_P = \{ ud(A) \leftarrow \top \mid A \in \text{undef}(P) \} \cup \{ ud(A) \leftarrow \bot \mid A \not\in \text{undef}(P) \} \]

For every (abducible) atom \( A \) we introduce the two following clauses:

\[ \text{inspect}(A) \leftarrow A \land \neg ud(A) \quad \text{inspect}_{\neg}(A) \leftarrow \neg A \land \neg ud(A) \]

The program containing all \( \text{inspect} \) and \( \text{inspect}_{\neg} \) clauses for every such atom in \( P \) is:

\[ I_P = \{ \text{inspect}(A) \leftarrow A \land \neg ud(A) \}, \text{inspect}_{\neg}(A) \leftarrow \neg A \land \neg ud(A) \mid A \in \text{atoms}(P) \} \]

\(^2\)More generally, certain occurrences may only be used in the abductive context of some other ones.
Let us consider $S_{add}$ again: we prefer to conclude that $cigarettes(X)$ is false for any variable if not stated otherwise. We modify our original program accordingly:

$$P_{add}^{insp} = U_{P_{add}} \cup I_{P_{add}} \cup P_{add} \setminus \{\{ab_{3a}(X) \leftarrow cigarettes(X)\}\} \cup \{ab_{3a}(X) \leftarrow inspect(cigarettes(X))\}$$

If we only observe $O_{add}$, we cannot abduce $E_{cig}$ but $E_{\neg cig}$ and $b$ is not inexpensive!

However, if we additionally observe $O_{inexp} = \{inexpensive(b) \leftarrow \top\}$, $E_{cig}$ can be abduced by the following clause in $P_{add}^{insp}$:

$$\{\ldots, inexpensive(b) \leftarrow cigarettes(b) \land \neg ab_{3b}(b), \ldots \}$$

Now, $b$ is addictive, inexpensive and a cigarette!

$O_{add}$ is a strict necessary contextual side-effect of explained $O_{inexp}$ given $P_{add}^{insp}$. 
Conclusion and Future Work

Weak completion semantics
We have examined how syllogisms and the belief-bias effect can be modeled by using abnormality predicates and abduction.

Inspection Points
We propose to introduce inspect predicates to deal with various kinds of abnormalities in human reasoning which allow different treatements of abducibles within conditionals.

Future Work
We intend to investigate other applications within human reasoning studies that might confirm or refine our definitions.

We are currently working on extensions for:
1. Contextual Relevant Consequences
2. Contestable Contextual Side-effects
3. Contextual Counterfactual Side-effects
Thank you very much for your attention!
References


Let us extend \( S_{add} \) by the following conditional:

*If something is sold in the streets and not cigarettes, then it is illegal.*

Accordingly, the new program is:

\[
P_{ill} = P_{add} \cup \{ \text{illegal}(X) \leftarrow \text{streets}(X) \land \neg \text{cig}(X) \}. \]

The minimal explanation for \( O_{ill} = \{ \text{illegal}(b) \leftarrow \top \} \) is \( E_{\text{streets}, \neg \text{cig}} = \{ \text{streets}(b) \leftarrow \top, \text{cig}(b) \leftarrow \bot \}. \)

\( O_{ill} \) is a possible contextual side-effect of explained \( O_{add} \) given \( P_{ill} \).

Let us simplify \( P_{ill} \) as follows:

\[
P_{ill}^{\text{sim}} = P_{add} \cup \{ \text{illegal}(X) \leftarrow \neg \text{cig}(X) \}. \]

Explanation for \( O_{ill} \) is the same as for \( O_{add} \)!

\( \text{illegal}(b) \) is a strict possible contextual side-effect of explained \( O_{add} \), and vice versa!

We need to specify which observation can be abduced and which is a contextual side effect!
Modeling Extended Syllogism $S_{ill}$ with Inspect

We modify and extend our program accordingly:

$$P^{insp}_{ill} = \mathcal{U}_{P_{ill}} \cup \mathcal{I}_{P_{ill}} \cup P_{ill} \setminus \{\text{illegal}(X) \leftarrow \text{streets}(X) \land \neg \text{cig}(X)\} \cup \{\text{illegal}(X) \leftarrow \text{streets}(X) \land \text{inspect-}(\text{cig}(X))\}$$

where as previously defined for $\mathcal{I}_{P_{ill}}$:

$$\text{inspect-}(\text{cig}(X)) \leftarrow \neg \text{cig}(X) \land \neg \text{ud}(X)$$

We can only conclude that something is illegal when we already have abduced somewhere else that it is not a cigarette. Assume $O_{add}$ explained by $E_{\neg \text{cig}}$:

$$\text{Im}_{L \text{wc g}}(P^{insp}_{ill} \cup E_{\neg \text{cig}} \cup E_{\text{streets}}) = \langle \{\text{add}(b), \text{streets}(b), \text{illegal}(b), \text{inspect-}(\text{cig}(b))\}, \{\text{add}'(b), \text{cig}(b), ab_{1}(b), \text{inex}(b), ab_{2}(b)\}\rangle$$

Note that without $E_{\neg \text{cig}}$ for $O_{add}$ we could not have concluded $\text{illegal}(b)$. 
Contextual Relevant Consequences

Definition (4)

Given a background knowledge $\mathcal{P}$, two observations $O_1$ and $O_2$ and two explanations $\mathcal{E}_1$ and $\mathcal{E}_2$.

$O_2$ is a necessary contextual relevant consequence of explained $\mathcal{E}_1$ iff for every $\mathcal{E}_2$ such that $O_2$ is explained by $\mathcal{E}_2$ given $\mathcal{P}$, there exists an $\mathcal{E}_1$ consistent with $\mathcal{E}_2$ such that $O_1$ is explained by $\mathcal{E}_1$ given $\mathcal{P}$, and the intersection of $\mathcal{E}_1$ and $\mathcal{E}_2$ is nonempty.

Let us ensure that something is dangerous before we conclude that it is addictive:

$$\mathcal{P}_{\text{dan}} = \mathcal{P}_{\text{ill}} \setminus \{\text{add}(X) \leftarrow \neg\text{add}'(X)\} \cup \{\text{add}(X) \leftarrow \neg\text{add}'(X) \land \text{dangerous}(X)\}$$

Given $O_{\text{add}}$, we have two minimal explanations:

$$\mathcal{E}_{\text{cig},\text{dan}} = \{\text{cig}(b) \leftarrow \top, \text{dangerous}(b) \leftarrow \top\},$$
$$\mathcal{E}_{\neg\text{cig},\text{dan}} = \{\text{cig}(b) \leftarrow \bot, \text{dangerous}(b) \leftarrow \top\}$$

$O_{\text{ill}} = \{\text{illegal}(b) \leftarrow \top\}$ is a contextual relevant consequence of explained $O_{\text{add}}$.

Note again, that the original observation and the side-effect are interchangeable.
Modeling extended Syllogism, $\mathcal{P}_{\text{dan}}$ with Inspect

Again, we intend that $\mathcal{O}_{\text{ill}}$ is a contextual relevant consequence of explained $\mathcal{O}_{\text{add}}$:

\[
\mathcal{P}^{\text{insp}}_{\text{dan}} = \mathcal{U}_{\mathcal{P}_{\text{dan}}} \cup \mathcal{I}_{\mathcal{P}_{\text{dan}}} \cup \mathcal{P}_{\text{dan}} \\
\setminus \{(\text{illegal}(X) \leftarrow \text{streets}(X) \land \neg \text{cig}(X))\} \cup \\
\{\text{illegal}(X) \leftarrow \text{streets}(X) \land \text{inspect} \neg \text{cig}(X)\}\}
\]

The rule

\[
\text{inspect} \neg \text{cig}(X) \leftarrow \neg \text{cig}(X) \land \neg \text{ud}(X)
\]

stays the same and it is easy to see that the outcome is similar as just demonstrated: $\text{illegal}(b)$ can only be concluded when $\mathcal{E}_{\neg \text{cig}}$ is abduced elsewhere, e.g. for $\mathcal{O}_{\text{add}}$. 
Contestable Contextual Side-effects - Abductive Explanation Contesting

Definition (5)

Given a background knowledge $P$, two observations $O_1$ and $O_2$ and two explanations $E_1$ and $E_2$

$O_2$ is a necessarily contested contextual side-effect of explained $O_1$ iff for every $E_1$ such that $O_1$ is explained by $E_1$ given $P$, there exists $E_2$ such that $\neg O_2$ is explained by $E_2$ given $P$ and $E_1$ contains $E_2$.

$O_2$ is a possibly contested contextual side-effect of explained $O_1$ iff there exists an $E_1$ such that $O_1$ is explained by $E_1$ given $P$, there exists $E_2$ such that $\neg O_2$ is explained by $E_2$ given $P$ and $E_1$ contains $E_2$.

Let us define the program $P_{leg}$ dualistic to $P_{sim}$:

$$P_{leg} = P_{add} \cup \{leg(X) \leftarrow cig(X)\}$$

We originally observe $O_{add}$ possibly explained by $E_{\neg cig}$:

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

$\neg leg(b)$ is possible contestable contextual side-effect of explained $O_{add}$, and vice versa.
Extended Syllogism $\mathcal{P}_{\text{leg}}$ with Inspect

$\mathcal{P}_{\text{leg}}^{\text{insp}} = \mathcal{U}_{\mathcal{P}_{\text{leg}}} \cup \mathcal{I}_{\mathcal{P}_{\text{leg}}} \cup \mathcal{P}_{\text{leg}} \setminus \{\text{legal}(X) \leftarrow \text{cig}(X)\}$

$\quad \cup \{\text{legal}(X) \leftarrow \text{inspect}(\text{cig}(X))\}$

With the modified clause:

$\{\text{legal}(X) \leftarrow \text{inspect}(\text{cig}(X))\}$

We only conclude that something is legal if $\text{cig}(X)$ is used to explain something else.

Given $\mathcal{E}_{\neg \text{cig}}$, $\neg \text{leg}(b)$ is a possible contestable contextual side-effect of explained $\mathcal{O}_{\text{add}}$. 
Another variation of contestable side-effects is abductive rebuttal: the side-effect directly contradicts an observation, that is $O_2 \equiv \neg O_1$ in Definition 5.

Let us extend $P_{add}$ by a conditional obviously contradicting the one in $P_{add}$:

\[
\text{if something is a cigarette, then it is not inexpensive.}
\]

\[
P_{inex} = P_{add} \cup \{\text{inex}'(X) \leftarrow \text{cig}(X), \text{inex}(X) \leftarrow \neg \text{inex}'(X)\}
\]

We observe $O_{inex}$ and can explain it by either $E_{\neg \text{cig}}$ or by $E_{\text{cig}}$. $E_{\text{cig}}$ explains observation $O_{\text{inex}'}$ as well and obviously $O_{\text{inex}}$ and $O_{\text{inex}'}$ are dualistic.
The weak completion of $g(\mathcal{P}_{\text{inex}} \cup \mathcal{E}_{\text{cig}})$ is:

$$\text{wc } g(\mathcal{P}_{\text{inex}} \cup \mathcal{E}_{\text{cig}}) = \{\text{add}'(b) \leftrightarrow \text{inex}(b) \land \neg ab_1(b), \text{add}(b) \leftrightarrow \neg \text{add}'(b), \text{inex}(b) \leftrightarrow (\text{cig}(b) \land \neg ab_2(b)) \lor \neg \text{inex}'(b), ab_2(b) \leftrightarrow \bot, ab_1(b) \leftrightarrow \bot \lor \text{cig}(b), \text{inex}'(b) \leftrightarrow \text{cig}(b)\} \cup \{\text{cig}(b) \leftrightarrow \top\}$$

and the corresponding least model is:

$$\text{lm}_L \text{wc } g(\mathcal{P}_{\text{inex}} \cup \mathcal{E}_{\text{cig}}) = \langle\{\text{add}(b), ab_1(b), \text{inex}(b), \text{inex}'(b), \text{cig}(b)\}, \{\text{add}'(b), ab_2(b)\}\rangle$$

$\text{inex}'(b)$ is a possibly contested contextual side-effect of explained $\text{inex}(b)$. 
Extended Syllogism $\mathcal{P}_{\text{inex}}$ with Inspect

Accordingly, we can modify $\mathcal{P}_{\text{add}}^{\text{ext}}$ as follows:

$$\mathcal{P}_{\text{inex}}^{\text{insp}} = \cup \mathcal{P}_{\text{inex}} \cup \mathcal{I} \mathcal{P}_{\text{inex}} \cup \mathcal{P}_{\text{inex}} \setminus (\{\text{inex}'(X) \leftarrow \text{cig}(X)\}) \cup \{\text{inex}'(X) \leftarrow \text{inspect}(\text{cig}(X))\}$$

With the modified clause we make sure that something can only be not inexpensive if it has been explained beforehand that it is a cigarette.