TLP **14** (4–5): 633–648, 2014. © Cambridge University Press 2014
doi:10.1017/S1471068414000258

Contextual Abductive Reasoning with Side-Effects

LUÍS MONIZ PEREIRA*[1], EMMANUELLE-ANNA DIETZ†[1,2] and STEFFEN HÖLLDOBLER†[2]

[1]*Centro de Inteligência Artificial (CENTRIA), Departamento de Informática Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal*
[2]*International Center for Computational Logic TU Dresden, D-01062 Dresden, Germany*

submitted 1 January 2003; revised 1 January 2003; accepted 1 January 2003

Abstract

The belief bias effect is a phenomenon which occurs when we think that we judge an argument based on our reasoning, but are actually influenced by our beliefs and prior knowledge. Evans, Barston and Pollard carried out a psychological syllogistic reasoning task to prove this effect. Participants were asked whether they would accept or reject a given syllogism. We discuss one specific case which is commonly assumed to be believable but which is actually not logically valid. By introducing abnormalities, abduction and background knowledge, we adequately model this case under the weak completion semantics. Our formalization reveals new questions about possible extensions in abductive reasoning. For instance, observations and their explanations might include some relevant prior abductive contextual information concerning some side-effect or leading to a contestable or refutable side-effect. A weaker notion indicates the support of some relevant consequences by a prior abductive context. Yet another definition describes jointly supported relevant consequences, which captures the idea of two observations containing mutually supportive side-effects. Though motivated with and exemplified by the running psychology application, the various new general abductive context definitions are introduced here and given a declarative semantics for the first time, and have a much wider scope of application. Inspection points, a concept introduced by Pereira and Pinto, allows us to express these definitions syntactically and intertwine them into an operational semantics.

KEYWORDS: Abductive Reasoning; Contextual Reasoning; Side-effects; Human Reasoning; Belief-Bias; Inspection Points; Three-valued Łukasiewicz Logic; Weak Completion Semantics; Logic Programming

1 Introduction

In the context of abductive reasoning, whenever discovering abductive explanations for some given primary observation, one may wish to check too whether some other

* lmp@fct.unl.pt
† {dietz,sh}@iccl.tu-dresden.de

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

given additional secondary observations are true, as logical consequences of the abductive explanations found for the primary observation. In other words, whether the secondary observations are plausible in the abductive context of the primary one, is a common scientific reasoning task. Thus, for example, one may attempt to find abductive explanations for such secondary observations strictly within the context of the given abductive explanations found for the primary observation – that is, disallowing new abductions – or, nevertheless, allowing additional abductions as long as they are consistent with the primary ones. As it were, the explanations of secondary observational consequences may be consumers, but not producers, of the abductions produced in explaining the primary observation. We show this type of reasoning requires the characterization of a new abduction concept and mechanism, that of *contextual abduction*. We examine and formalize its variants, and employ these to understand and justify belief bias of human reasoning in addressing syllogisms.

Our starting point is a psychological study carried out by Evans *et al.* (1983) about deductive reasoning which demonstrated possibly conflicting processes in human reasoning. Participants were presented different syllogisms and had to decide whether they were logically valid. Consider one of them, S_{add}:

PREMISE1 *No addictive things are inexpensive.*
PREMISE2 *Some cigarettes are inexpensive.*
CONCLUSION *Therefore, some addictive things are not cigarettes.*

Even though the conclusion does not necessarily follow from the premises,[1] participants assumed the syllogism to be logically valid. They were explicitly asked to logically validate or invalidate these syllogisms, but didn't seem to have the intellectual capability to do so. Even worse, they were not at all aware about their inabilities. Repeatedly, the majority of people were proffering with certainty of confidence the wrong answer. Evans *et al.* (1983) concluded that this happened because they were being unduly influenced by their own beliefs, their belief bias. In (Evans 2010; Evans 2012), the conflict between logic and belief in human reasoning is discussed extensively.

We will show how the belief bias effect can be explained by abductive reasoning and its corresponding side-effects. This forms the basis for investigating contextual side-effects, (strict) possible side-effects, contextual contestable side-effects, and (jointly supported) contextual relevant consequences.

2 Preliminaries

We define the necessary notations, simplified for our present purposes but generalizable as usual (cf. Kowalski (2011)). We restrict ourselves to datalog programs, i.e. the set of terms consists only of constants and variables. A (first-order) logic

[1] In abstract, it could be conceived that all addictive things would be just the (expensive) cigarettes. It does not follow because the conclusion supposes an existence which is not warranted by the premises.

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

program $\mathscr{P}$ is a finite set of clauses:

$$A \leftarrow A_1 \wedge \ldots \wedge A_n \wedge \neg B_1 \wedge \ldots \wedge \neg B_m,$$

where A and A_i, $1 \leqslant i \leqslant n$, are atoms and $\neg B_j$, $1 \leqslant j \leqslant m$, are negated atoms. A is the head and $A_1 \wedge \ldots \wedge A_n \wedge \neg B_1 \wedge \ldots \wedge \neg B_m$ is the body of the clause. In the sequel we will abbreviate the body of a clause by simply writing *body*. $A \leftarrow \top$ and $A \leftarrow \bot$ are special cases of clauses denoting positive and negative facts, respectively. If an argument is written with an upper case letter, it is a variable; otherwise it is a constant. In the sequel, if not denoted otherwise, we assume $\mathscr{P}$ to be ground, containing all the ground instances of its clauses. The set of all atoms occurring in $\mathscr{P}$ is $\mathsf{atoms}(\mathscr{P})$. An atom is undefined in $\mathscr{P}$ if it is not the head of some clause in $\mathscr{P}$ and the corresponding set of these atoms is $\mathsf{undef}(\mathscr{P})$.

2.1 Three-Valued Łukasiewicz Semantics

We use the three-valued Łukasiewicz (1920) semantics. Table 1 defines the corresponding truth tables. Interpretations are represented by pairs $\langle I^\top, I^\bot \rangle$, such that

$$I^\top = \{A \in B_\mathscr{P} \mid A \text{ is mapped to } \top\} \text{ and } I^\bot = \{A \in B_\mathscr{P} \mid A \text{ is mapped to } \bot\},$$

where $\mathscr{B}_\mathscr{P}$ is the Herbrand base with respect to a given program $\mathscr{P}$. A model of $\mathscr{P}$ is an interpretation which maps each clause occurring in $\mathscr{P}$ to $\top$.

One should observe that in contrast to two-valued logic, $A \leftarrow B$ and $A \vee \neg B$ are not equivalent under three-valued Łukasiewicz Semantics. Consider, for instance, an interpretation I such that $I(A) = I(B) = \mathsf{U}$. Then, $I(A \vee \neg B) = \mathsf{U}$ whereas $I(A \leftarrow_\text{Ł} B) = \top$.

2.2 Weak Completion Semantics

Weak completion semantics has first been introduced by Hölldobler and Kencana Ramli (2009a) and seems to adequately model Byrne's (1989) suppression task (Dietz *et al.* 2012) and Wason's (1968) selection task (Dietz *et al.* 2013). Consider the following transformation for given $\mathscr{P}$:

1. Replace all clauses in $\mathscr{P}$ with the same head $A \leftarrow body_1, A \leftarrow body_2, \ldots$ by the single expression $A \leftarrow body_1 \vee body_2, \vee \ldots$.
2. if $A \in \mathsf{undef}(\mathscr{P})$ then add $A \leftarrow \bot$.
3. Replace all occurrences of $\leftarrow$ by $\leftrightarrow$.

The resulting set of equivalences is called the *completion* of $\mathscr{P}$ (Clark 1978). If Step 2 is omitted, then the resulting set is called the *weak completion* of $\mathscr{P}$ ($\mathsf{wc}\,\mathscr{P}$).

Hölldobler and Kencana Ramli (2009b) showed that the model intersection property holds for weakly completed programs. This guarantees the existence of a least model for every program. In computational logic, least models can often be computed as least fixed points of an appropriate semantic operator (Apt and van Emden 1982). Stenning and van Lambalgen (2008) devised such an operator which

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

F	$\neg F$
$\top$	$\bot$
$\bot$	$\top$
U	U

$\wedge$	$\top$	U	$\bot$
$\top$	$\top$	U	$\bot$
U	U	U	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$

$\vee$	$\top$	U	$\bot$
$\top$	$\top$	$\top$	$\top$
U	$\top$	U	U
$\bot$	$\top$	U	$\bot$

$\leftarrow_{Ł}$	$\top$	U	$\bot$
$\top$	$\top$	$\top$	$\top$
U	U	$\top$	$\top$
$\bot$	$\bot$	U	$\top$

$\leftrightarrow_{Ł}$	$\top$	U	$\bot$
$\top$	$\top$	U	$\bot$
U	U	$\top$	U
$\bot$	$\bot$	U	$\top$

Table 1. $\top$, $\bot$, *and* U *denote true, false, and unknown, respectively.*

has been generalized for first-order programs by Hölldobler and Kencana Ramli (2009a): Let I be an interpretation and $\Phi_{\mathscr{P}}(I) = \langle J^{\top}, J^{\bot} \rangle$, where

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

As shown in (Hölldobler and Kencana Ramli 2009a) the least fixed point of $\Phi_{\mathscr{P}}$ is identical to the least model of the weak completion of $\mathscr{P}$ ($\mathsf{lm}_{Ł}\mathsf{wc}\,\mathscr{P}$) under three-valued Łukasiewicz semantics. Starting with the empty interpretation $I = \langle \emptyset, \emptyset \rangle$, $\mathsf{lm}_{Ł}\mathsf{wc}\,\mathscr{P}$ can be computed by iterating $\Phi_{\mathscr{P}}$.[2]

3 Reasoning in an Appropriate Logical Form

Stenning and van Lambalgen (2005; 2008) proposed to model human reasoning by a two step process: Firstly, human reasoning should be modeled by setting up an *appropriate representation* and, secondly, the *reasoning process* should be modeled with respect to this representation. In this section we discuss the first step and show how to model syllogisms in logic programs.

3.1 Integrity Constraints

PREMISE1 of S_{add} is

No addictive things are inexpensive. (1)

and is equivalent to

If something is inexpensive, then it is not addictive. (2)

The consequence is the negation of *something is addictive*. As weak completion semantics does not allow negative heads in clauses, for every negative conclusion $\neg p(X)$ we introduce an auxiliary formula $p'(X)$, which denotes the negation of p and the clause $p(X) \leftarrow \neg p'(X)$.

We obtain the following preliminary representation of the first premise of S_{add} with regard to *addictive*: $add'(X) \leftarrow inex(X)$, and $add(X) \leftarrow \neg add'(X)$, where $add(X)$, $add'(X)$, and $inex(X)$ stand for X is addictive, not addictive, and inexpensive.

With the introduction of these auxiliary atoms, the need for integrity constraints arises. A least model of the weak completion that contains both $add(X)$ and $add'(X)$ in $I^{\top}$ should be invalidated as a model in general. This condition can be represented

[2] Weak completion semantics corresponds to well-founded semantics (Van Gelder *et al.* 1991) for tight logic programs (Dietz *et al.* 2013).

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

by a *set of integrity constraints* $\mathscr{IC}$, which contains clauses of the following form: $\mathsf{U} \leftarrow \mathscr{IC}_Body$, where the implication is understood as usual. For our example above, $\mathscr{IC}_{add}$ contains one clause: $\mathsf{U} \leftarrow add(X) \wedge add'(X)$. However, the Φ operator does not consider clauses of this form, thus such denial $\mathscr{IC}$s are not evaluated by it. We apply a two step approach to take them into consideration. First, we compute the least model of the given program and second, we verify whether it does satisfy the requirements of the $\mathscr{IC}$. Given an interpretation I and a set of integrity constraints $\mathscr{IC}$, I *satisfies* $\mathscr{IC}$ if and only if all clauses in $\mathscr{IC}$ are true under I. For the following examples, whenever there exists a $p(X)$ and its $p'(X)$ counterpart in $\mathscr{P}$, we implicitly assume that $\mathscr{IC}_p$: $\mathsf{U} \leftarrow p(X) \wedge p'(X)$.[3]

3.2 Abnormalities & Background Knowledge

A direct representation of PREMISE2 is

There exists a cigarette which is inexpensive. (3)

Additionally, it is commonly known that

Cigarettes are addictive. (4)

As discussed in (Evans *et al.* 1983), humans seem to have a background knowledge or belief, which, in this context, assuming (4), we imply that

Cigarettes are inexpensive (compared to other addictive things),

which implies (3) and biases the reasoning towards a representation. The preliminary program representing the first two premises, is the non-ground program $\mathscr{P}^{pre}_{add}$:

$$\begin{array}{lcllcl} add'(X) & \leftarrow & inex(X), & add(X) & \leftarrow & \neg add'(X), \\ inex(X) & \leftarrow & cig(X), & cig(a) & \leftarrow & \top. \end{array}$$

and its corresponding least model is: $\langle \{cig(a), inex(a), add'(a)\}, \{add(a)\} \rangle$. This model contradicts the commonly known assumption of (4). This background knowledge can be expressed in the program by applying Stenning and van Lambalgen's idea to implement conditionals by a normal default permission for implications.This can be achieved by adding an *abnormality predicate* to the antecedent of the implication and initially assuming that the abnormality predicate is false. Following this idea, the initial PREMISE1 in S_{add} is extended to:

If something is inexpensive and not abnormal, then it is not addictive.
Nothing (as a rule) is abnormal (regarding (1)).

This belief-bias together with the idea to represent conditionals by a normal default permission for implication leading to this rendering

If something is a cigarette and not abnormal, then it is inexpensive. (5)
Nothing (as a rule) is abnormal (regarding (3)).

[3] This view on $\mathscr{IC}$s corresponds to the definition applied for the well-founded semantics in (Pereira *et al.* 1991).

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

Together with (4), it leads to

If something is a cigarette, then it is abnormal (regarding (1)).

Finally, the information in the premises of S_{add} is encoded as the non-ground program $\mathscr{P}_{add}$:

$$\begin{array}{lcl@{\qquad}lcl} add'(X) & \leftarrow & inex(X) \wedge \neg ab_1(X), & add(X) & \leftarrow & \neg add'(X), \\ inex(X) & \leftarrow & cig(X) \wedge \neg ab_2(X), & ab_2(X) & \leftarrow & \bot, \\ ab_1(X) & \leftarrow & \bot, & ab_1(X) & \leftarrow & cig(X), \\ cig(a) & \leftarrow & \top. & & & \end{array}$$

$\mathscr{P}_{add}$ represents the contextual background knowledge of syllogism S_{add}.

4 Abduction and Predictability

In scientific methodology, a prediction can be made by adding hypotheses to knowledge known about the world. As specified by the classical hypothetic-deductive method (Hempel 1966), scientific inquiry is carried out in three stages: hypotheses generation, prediction, and evaluation. One or more hypotheses may be generated, by abduction, to explain observed events. A generated set of hypotheses (or assumptions) can thence be employed for predicting unseen events by means of deduction, on the implicit condition of not making further abductions (cf. Pereira and Pinto (2011)). The predicted events can then hopefully be tested against reality, in the form of such observable deduced side-effects, in order to evaluate the plausibility of the set of hypotheses.

Abduction, or inference to the best explanation (its usual designation in the philosophy of science), is a reasoning method whereby one chooses those hypotheses that would, if true, best explain observed evidence or enable to satisfy some query, whilst meeting attending constraints. Abduction has been well studied in the field of computational logic – and logic programming in particular – for a few decades now. Abduction, when added to logic programs, offers a formalism to declaratively express problems in a variety of areas and empowers many applications, e.g. in decision-making, diagnosis, planning, belief revision, and conditional reasoning. In logic programs, abductive hypotheses (or abducibles) are named given atoms of the program which have no rules, and whose truth value is not initially assumed, and hence unknown.

The approach presented in (Torasso *et al.* 1991) treats abduction with completion semantics and does not study side-effects. In the following we consider abduction with weak completion semantics and introduce the examination of side-effects in contexts afforded by abduction.

4.1 Abductive Framework

Following Kakas *et al.* (1993) we consider an *abductive framework* consisting of a program $\mathscr{P}$ as knowledge base, a collection of atoms $\mathscr{A}$ of abducibles syntactically represented by the set of the (positive and negative) facts for each undefined

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

ground atom in $\mathscr{P}$, a set of integrity constraints $\mathscr{IC}$, and the logical consequence relation $\models_{Ł}^{lmwc}$, where $\mathscr{P} \models_{Ł}^{lmwc} F$ if and only if $\mathsf{lm}_{Ł}\mathsf{wc}\,\mathscr{P}(F) = \top$ for the formula F. An *observation* is a set of (at least one) literals.

The truth value of abducibles may be independently assumed true or false, via either their positive or negated form, as the case may be, in order to produce an abductive explanation for an observation – or solution to a query –, which is a consistent set of assumed hypotheses in the form of abducibles.

An abductive solution is a consistent set of abducible instances that, when substituted by their assigned truth value in $\mathscr{P}$, affords us with a model of $\mathscr{P}$ (for the specific semantics used on $\mathscr{P}$), which satisfies the observation (or query) and any imposed integrity constraints – a so-called abductive model. In our notation this amounts to adding to $\mathscr{P}$ the corresponding positive and negative facts representing a solution's abducibles.

What we dub *observation* is analogous to a *query* whose explanation is desired, not necessarily something actually observed.

Definition 1

Let $\langle \mathscr{P}, \mathscr{A}, \mathscr{IC}, \models_{Ł}^{lmwc} \rangle$ be an abductive framework, $\mathscr{O}$ be an observation, and $\mathscr{E}$ be an explanation which is a (consistent) subset of $\mathscr{A}$, a set of integrity constraints $\mathscr{IC}$, and the consequence relation $\models_{Ł}^{lmwc}$, defined for all formulas F.

- $\mathscr{O}$ is **explained by** $\mathscr{E}$ **given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff $\mathscr{P} \cup \mathscr{E} \models_{Ł}^{lmwc} \mathscr{O}$, where $\mathscr{P} \not\models_{Ł}^{lmwc} \mathscr{O}$ and $\mathsf{lm}_{Ł}\mathsf{wc}\,(\mathscr{P} \cup \mathscr{E})$ *satisfies* $\mathscr{IC}$.
- $\mathscr{O}$ is **explained given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff there exists an $\mathscr{E}$ such that $\mathscr{O}$ is explained by $\mathscr{E}$ given $\mathscr{P}$ and $\mathscr{IC}$.

In abduction, as for its deduction counterpart, *credulous* and *skeptical reasoning* varieties are distinguished. Credulous reasoning consists in finding if there exists at least one model of the program – according to some pre-established semantics – which entails the observation to be explained. Skeptical reasoning demands that every model of the program entails the observation.

- F **follows skeptically from** $\mathscr{P}$**,** $\mathscr{IC}$ **and** $\mathscr{O}$ iff $\mathscr{O}$ can be explained given $\mathscr{P}$ and $\mathscr{IC}$, and for all minimal (or some alternative preference criterion instead) explanations $\mathscr{E}$ for $\mathscr{O}$ it holds that $\mathscr{P} \cup \mathscr{E} \models_{Ł}^{lmwc} F$.
- F **follows credulously from** $\mathscr{P}$**,** $\mathscr{IC}$ **and** $\mathscr{O}$ iff there exists a minimal (or some alternative preference criterion instead) explanation $\mathscr{E}$ for $\mathscr{O}$ and it holds that $\mathscr{P} \cup \mathscr{E} \models_{Ł}^{lmwc} F$.

Because the satisfaction of integrity constraints ($\mathscr{IC}$s) can require abductions, we must allow $\mathscr{IC}$s to be actively productive of abductions, and not just use $\mathscr{IC}$s to subsequently disallow abductive solutions that invalidate the least model with respect to them. However, we cannot actively promote that an atom is required to be false or else unknown, one of the two, in order for $\mathscr{IC}$s to be satisfied. The only possibility is to impose its falsity: because $\mathscr{IC}$s have the form of denials, we can think of $\mathsf{U} \leftarrow A$ as just an atom that we wish must never be explained by any explanation.

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

As our underlying semantics is three-valued, we may but might not make each undefined atom an abducible, so that we allow for unknown, non-abducible susceptible knowledge, which can be guaranteed by adding the clause $A \leftarrow A$ for the predicate under consideration. Furthermore, when making use of skeptical reasoning, we conclude that an abducible is unknown if it does not have the same binary truth value in all models. For instance consider the program $\mathscr{P} = \{A \leftarrow B, A \leftarrow C\}$ and observation $\mathscr{O}_A = \{A\}$, for which there are two minimal explanation $\mathscr{E}_B = \{B \leftarrow \top\}$ and $\mathscr{E}_C = \{C \leftarrow \top\}$. Under skeptical reasoning A does not follow from all minimal explanations and thus B and C stay unknown.

In previous approaches, weak completion semantics was used for cases expressed in propositional logic, and abduction under skeptical reasoning seemed adequate (Hölldobler *et al.* 2011).

4.2 Usual Contextual Abduction

One important extension of abduction pertains to the issue that, whenever discovering abductive solutions, i.e. explanations, for some given primary observation, one may wish to check too whether some other given additional secondary observations are true, being a logical consequence of the abductive explanations found for the primary observation. In other words, whether the secondary observations are plausible in the abductive context of the primary one. Indeed, often, besides needing to abductively discover which hypotheses to assume in order to satisfy some condition, we may also want to know some of the side-effects of those assumptions.

We address the issue of relaxing or loosening the implicit condition about additional abductions not being permitted whilst considering the observable side-effects explained by deductive prediction. In other words, prediction may be allowed recourse to additional assumptions, but nevertheless must make use of at least some of the initial explanations. If several such explanations exist for the observations concerned, then we might want to define alternative conditions that are less strict with regard to each set or to the collection of sets. Reuse of contextual abductions, by resorting to an implementation of tabled abduction for well-founded semantics, is reported in (Saptawijaya and Pereira 2013).

Let's consider again $\mathscr{P}_{add}$, its weak completion consists of the following equivalences:

$$\begin{array}{lcllcl} add(a) & \leftrightarrow & \neg add'(a), & add'(a) & \leftrightarrow & inex(a) \wedge \neg ab_1(a), \\ inex(a) & \leftrightarrow & cig(a) \wedge \neg ab_2(a), & ab_2(a) & \leftrightarrow & \bot, \\ ab_1(a) & \leftrightarrow & cig(a),^4 & cig(a) & \leftrightarrow & \top. \end{array}$$

Its least model, $\mathsf{lm}_{\text{Ł}}\mathsf{wc}\,\mathscr{P}_{add}$, is $\langle\{cig(a), inex(a), add(a), ab_1(a)\}, \{add'(a), ab_2(a)\}\rangle$, from which we cannot derive the CONCLUSION in S_{add}. Obviously, the CONCLUSION is something about an object which is not a. The first part of this conclusion is an observation, let's say about b: $\mathscr{O}_{add(b)} = \{add(b)\}$, which we need to explain as described in the previous subsection. The set of abducibles with respect to b is:

[4] $\bot \vee cig(a)$ is semantically equivalent to $cig(a)$ under Łukasiewicz Semantics.

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

$\mathscr{A}_{\mathscr{P}_{add}} = \{cig(b) \leftarrow \top, cig(b) \leftarrow \bot\}$. $\mathscr{O}_{add(b)}$ is true if $add'(b)$ is false which is false if $inex(b)$ is false or $ab_1(b)$ is true. Either $inex(b)$ is false but then $cig(b)$ is false or $ab_1(b)$ is true but then $cig(b)$ is true. For $\mathscr{O}_{add(b)}$ we have two minimal explanations $\mathscr{E}_{\neg cig(b)} = \{cig(b) \leftarrow \bot\}$ and $\mathscr{E}_{cig(b)} = \{cig(b) \leftarrow \top\}$. The corresponding least models of the weak completion are:

$$\mathsf{lm}_{\mathrm{L}}\mathsf{wc}\,(\mathscr{P}_{add} \cup \mathscr{E}_{\neg cig(b)}) = \langle\{\ldots, add(b), \ldots\}, \{\ldots, cig(b), inex(b), \ldots\}\rangle,$$
$$\mathsf{lm}_{\mathrm{L}}\mathsf{wc}\,(\mathscr{P}_{add} \cup \mathscr{E}_{cig(b)}) = \langle\{\ldots, add(b), cig(b), inex(b), \ldots\}, \{\ldots\}\rangle.$$

Under credulous reasoning we conclude, given explanation $\mathscr{E}_{\neg cig(b)}$, that the Conclusion of S_{add} is true, as there exists something addictive which is not a cigarette. I.e., there is a least model which abductively explains that an observed addictive b is not a cigarette.

However, we could also explain $\mathscr{O}_{add(b)}$ by $\mathscr{E}_{cig(b)}$, which is an equally justified explanation. We would prefer our formalism to reflect that the first premise describes the usual and the second premise describes the exceptional case. That is, an inexpensive cigarette is meant to be the exception not the rule in the context of things that are addictive. This exceptional case should then only be considered when more is known about b. This preference is not expressed in our framework yet. The following section discusses and proposes a solution to this issue.

5 Inspection Points

Until now, we assumed the possible observations as given, and easily identified from the corresponding context which was the exceptional and which the usual case. However, this is not explicitly encoded in our logic programs yet, and needs be syntactically indicated. For this purpose we investigate and apply inspection points, originally presented by Pereira and Pinto (2011).

In (Pereira and Pinto 2011), the authors present the concept of inspection points in abductive logic programming and show how one can employ it to investigate side-effects of interest in order to help choose among abductive solutions. In what follows we discuss this approach and show how inspection points can be modeled for our previous examples accordingly, in what concerns abducibles. Given an atom A, we introduce the following two reserved (meta-)predicates:

$$\mathsf{inspect}(A) \qquad \text{and} \qquad \mathsf{inspect}_{\neg}(A),$$

which are special cases of abducibles. They differ from the usual abducibles in the way that they can only be abduced whenever A or $\neg A$ have been abduced somewhere else already. That is, an abductive solution or explanation $\mathscr{E}$ is only valid when for each $\mathsf{inspect}(A)$ (respectively $\mathsf{inspect}_{\neg}(A)$) it contains, it also contains a corresponding A (respectively $\neg A$). That is, in a solution the consumers, here represented by the inspection points $\mathsf{inspect}(A)$ and $\mathsf{inspect}_{\neg}(A)$, respectively, must have a matching producer. The producers correspond to the usual abducibles.

One should observe that for a treatment of inspection points for all literals in a program and not just the abducible ones, we would simply need to adopt the

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

program transformation technique of Pereira and Pinto (2011), which recursively relays inspection of non-terminal literals to the base inspection of terminals.

Let us consider again the example of Section 4.2 where we stated that we would prefer to distinguish between the usual case and the exceptional case. We can now represent this in our logic program by replacing our ab_1-clause accordingly. The new non-ground program, $\mathscr{P}^{insp}_{add}$, is:

$$(\mathscr{P}_{add} \setminus \{ab_1(X) \leftarrow cig(X)\}) \cup \{ab_1(X) \leftarrow \mathsf{inspect}(cig(X))\}.$$

Suppose, b is addictive, i.e. $\mathscr{O}_{add(b)}$. As $cig(b)$ is undefined, $\mathsf{inspect}(cig(b))$ becomes false, and $ab_1(b)$ will be false rather than unknown, that is, its falsity is obtained because nothing is known about $cig(b)$. The only minimal explanation for $\mathscr{O}_{add(b)}$ is now generated by $inex(b)$ being false, which is achieved by $\mathsf{Im}_{\mathrm{L}}\mathsf{wc}(\mathscr{P}^{insp}_{add} \cup \mathscr{E}_{\neg cig(b)})$:

$$\langle\{\ldots, add(b), \ldots\}, \{\ldots, add'(b), \mathsf{inspect}(cig(b)), ab_1(b), inex(b), ab_2(b), \ldots\}\rangle.$$

Even under skeptical reasoning, there exists an addictive thing which is not a cigarette.

6 The Expressive Power of Inspection Points in Contextual Abduction

Inspection points allow us to specify various definitions, which we will provide – relative to explanation $\mathscr{E}$ of observation $\mathscr{O}$, set of abducibles $\mathscr{A}$ and background knowledge $\mathscr{P}$ –, to the effect of relaxing the impermissibility of convoking additional hypotheses to explain side-effects in context $\mathscr{E}$. Consider the following program, $\mathscr{P}_{fire}$:

$$\begin{array}{llll}
storm \leftarrow lightning \wedge \neg ab_1, & & ffire \leftarrow \mathsf{inspect}(lightning) \wedge \neg ab_3, & \\
storm \leftarrow tempest \wedge \neg ab_2, & & ffire \leftarrow barbecue \wedge \neg ab_3, & \\
ab_3 \leftarrow \neg dry_leaves, & & rained \leftarrow \mathsf{inspect}_{\neg}(dry_leaves) \wedge \neg ab_4, & \\
smoke \leftarrow fire \wedge \mathsf{inspect}(ffighters), & & sirens \leftarrow \mathsf{inspect}(fire) \wedge ffighters, & \\
ab_1 \leftarrow \bot, & ab_2 \leftarrow \bot, & ab_3 \leftarrow \bot, & ab_4 \leftarrow \bot.
\end{array}$$

where *ffire* means forest fire and *ffighters* means fire fighters. The set of abducibles, $\mathscr{A}_{\mathscr{P}_{fire}}$, is:

$$\begin{array}{lll|lll}
tempest & \leftarrow & \top, & tempest & \leftarrow & \bot, \\
barbecue & \leftarrow & \top, & barbecue & \leftarrow & \bot, \\
lightning & \leftarrow & \top, & lightning & \leftarrow & \bot, \\
\mathsf{inspect}(lightning) & \leftarrow & \top, & \mathsf{inspect}(lightning) & \leftarrow & \bot, \\
dry_leaves & \leftarrow & \top, & dry_leaves & \leftarrow & \bot, \\
\mathsf{inspect}_{\neg}(dry_leaves) & \leftarrow & \top, & \mathsf{inspect}_{\neg}(dry_leaves) & \leftarrow & \bot, \\
fire & \leftarrow & \top, & fire & \leftarrow & \bot, \\
\mathsf{inspect}(fire) & \leftarrow & \top, & \mathsf{inspect}(fire) & \leftarrow & \bot, \\
ffighters & \leftarrow & \top, & ffighters & \leftarrow & \bot, \\
\mathsf{inspect}(ffighters) & \leftarrow & \top, & \mathsf{inspect}(ffighters) & \leftarrow & \bot.
\end{array}$$

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

For simplicity, we assume that $\mathscr{IC}_{fire} = \emptyset$. In the sequel, we provide definitions and abductive examples which clarify how inspection points enrich the expressiveness of logic programs.

6.1 Contextual Side-effects

Consider the following definition which includes various notions of contextual side-effects:

Definition 2
Given background knowledge $\mathscr{P}$ and set of integrity constraints $\mathscr{IC}$, let $\mathscr{O}_1$ be an observation, $\mathscr{E}_1$ be an explanation of $\mathscr{O}_1$, $\mathscr{O}_2$ be an observation, and $\mathscr{E}_2$ be an explanation of $\mathscr{O}_2$.

$\mathscr{O}_2$ is a **necessary contextual side-effect of** $\mathscr{O}_1$ **given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff for all $\mathscr{E}_1$ there exists $\mathscr{E}_2$ such that, for any $\mathsf{inspect}(A)$, $\mathsf{inspect}_{\neg}(A) \in \mathscr{E}_2$ for which respectively $A, \neg A \notin \mathscr{E}_2$, and some such exists in $\mathscr{E}_2$, then, $A, \neg A \in \mathscr{E}_1$, accordingly.

$\mathscr{O}_2$ is a **strict necessary contextual side-effect of** $\mathscr{O}_1$ **given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff $\mathscr{O}_2$ is a necessary contextual side-effect of $\mathscr{O}_1$ given $\mathscr{P}$, and $\mathscr{E}_1 \subseteq \mathscr{E}_2$ for any $\mathscr{E}_1$.

$\mathscr{O}2$ is a **possible contextual side-effect of explained** $\mathscr{O}_1$ **given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff there exists an $\mathscr{E}_1$ and an $\mathscr{E}_2$ such that for any $\mathsf{inspect}(A)$, $\mathsf{inspect}_{\neg}(A) \in \mathscr{E}_2$ for which respectively $A, \neg A \notin \mathscr{E}_2$, and some such exists in $\mathscr{E}_2$, then $A, \neg A \in \mathscr{E}_1$, accordingly.

$\mathscr{O}_2$ is a **strict possible contextual side-effect of** $\mathscr{O}_1$ **given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff $\mathscr{O}_2$ is a necessary contextual side-effect of $\mathscr{O}_1$ given $\mathscr{P}$, and for some $\mathscr{E}_1$, $\mathscr{E}_1 \subseteq \mathscr{E}_2$.

The idea behind the definition for necessary contextual side-effects is that every explanation $\mathscr{E}_1$ for $\mathscr{O}_1$ affords us with one *complete* explanation by which any (inspection) *incomplete* explanation $\mathscr{E}_2$ for $\mathscr{O}_2$ can be necessarily completed. The idea behind the definition for possible contextual side-effects is that there is some explanation $\mathscr{E}_1$ for $\mathscr{O}_1$ which affords us with one *complete* explanation by which some (inspection) *incomplete* explanation $\mathscr{E}_2$ for $\mathscr{O}_2$ can be completed.

Consider again $\mathscr{P}_{fire}$, where there are two rules for *ffire*: one explanation is *barbecue* and another one is *lightning*. However, we assume that a lightning causing a forest fire is much more unlikely than a barbecue, and therefore, *lightning*, when not observed directly, only counts as a plausible explanation when it has been abduced by some other observation. This is expressed by $\mathsf{inspect}(lightning)$. Assume that we observe a storm and know that nothing is abnormal with respect to ab_3, thus the leaves are dry: $\mathscr{O}_{storm,dry} = \{storm, dry_leaves\}$. The minimal explanations are $\mathscr{E}_{lightning} = \{lightning \leftarrow \top\}$ and $\mathscr{E}_{tempest} = \{tempest \leftarrow \top\}$. We have only enough evidence to explain $\mathscr{O}_{ffire}$ if we abduce $\mathscr{E}_{lightning}$ for $\mathscr{O}_{storm,dry}$. By Definition 2, $\mathscr{O}_{ffire}$ is a strict possible contextual side-effect of explained $\mathscr{O}_{storm,dry}$ given $\mathscr{P}_{fire}$ and $\mathscr{IC}_{fire}$.

6.2 Contestable Contextual Side-effects

Whereas up till now we stipulated cases where known explanations gave grounds for strengthening the plausibility of contextual side-effects, we next turn our attention

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

to cases where the latter are to some extent contested and made implausible, by appealing to their negation. Analogously to Definition 2, its counterpart, contestable contextual side-effects, is defined as follows:

Definition 3
Given background knowledge $\mathscr{P}$ and set of integrity constraints $\mathscr{IC}$, let $\mathscr{O}_1$ be an observation, $\mathscr{E}_1$ be an explanation of $\mathscr{O}_1$, $\mathscr{O}_2$ be an observation, and $\mathscr{E}_2$ be an explanation of $\neg\mathscr{O}_2$.

$\mathscr{O}_2$ is a **necessarily contested contextual side-effect of** $\mathscr{O}_1$ **given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff for all explanations $\mathscr{E}_1$ there exists $\mathscr{E}_2$ such that, for any $\mathsf{inspect}(A)$, $\mathsf{inspect}_{\neg}(A) \in \mathscr{E}_2$, for which respectively A, $\neg A \notin \mathscr{E}_2$, and some such exists in $\mathscr{E}_2$, then, A, $\neg A \in \mathscr{E}_1$, accordingly.

$\mathscr{O}_2$ is a **possibly contested contextual side-effect of** $\mathscr{O}_1$ **given** $\mathscr{P}$ **and** $\mathscr{IC}$ iff there exists an $\mathscr{E}_1$ and an $\mathscr{E}_2$ such that for any $\mathsf{inspect}(A)$, $\mathsf{inspect}_{\neg}(A) \in \mathscr{E}_2$ for which respectively A, $\neg A \notin \mathscr{E}_2$, and some such exists in $\mathscr{E}_2$, then A, $\neg A \in \mathscr{E}_1$, accordingly.

The idea behind necessarily contested contextual side-effects, is, that every explanation $\mathscr{E}_1$ for $\mathscr{O}_1$ affords us with one *complete* explanation under which some *incomplete* explanation $\mathscr{E}_2$ for $\neg\mathscr{O}_2$ is necessarily completed. The idea behind the definition for the possibly contested contextual side-effects, is, that there is at least one explanation $\mathscr{E}_1$ for $\mathscr{O}_1$ which affords us with one *complete* explanation under which some *incomplete* explanation $\mathscr{E}_2$ for $\neg\mathscr{O}_2$ is necessarily completed.

Consider *rained* $\leftarrow \mathsf{inspect}_{\neg}(dry_leaves)$ in $\mathscr{P}_{fire}$: it states that if, for some other observation we explained that the leaves are not dry, then it rained. Thus, when we observe a forest fire, then, one part of the abduced explanation will be, in any case, independent of whether there was a lightning or a barbecue, that the leaves are dry. However this explanation will lead to $\mathsf{inspect}_{\neg}(dry_leaves)$ being false, which makes *rained* false as well. $\mathscr{O}_{\neg rained}$ is a consequence that follows from the explanation for $\mathscr{O}_{ffire}$. According to Definition 3, $\mathscr{O}_{\neg rained}$ is a necessarily contested contextual side-effect of $\mathscr{O}_{ff}$ given $\mathscr{P}_{fire}$ and $\mathscr{IC}_{fire}$.

Another variation of contestable side-effects is abductive rebuttal. In this case, the side-effect directly contradicts an observation. That is to say, it is the case that $\mathscr{O}_2 \equiv \neg\mathscr{O}_1$ in Definition 3. The second observation, $\mathscr{O}_2$ has to be the head of some clause and the negation of $\mathscr{O}_1$.[5]

6.3 Contextual Relevant Consequences

We identify two notions of contextual relevant consequences and define them as follows:

Definition 4
Given background knowledge $\mathscr{P}$ and set of integrity constraints $\mathscr{IC}$, let $\mathscr{O}_1$ be an observation, $\mathscr{E}_1$ be an explanation of $\mathscr{O}_1$, $\mathscr{O}_2$ be an observation, and $\mathscr{E}_2$ be an explanation of $\mathscr{O}_2$.

[5] As negative heads in our programs are not allowed, we can model these cases as described in Section 3.1.

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

$\mathcal{O}_2$ is a **necessary contextual relevant consequence of** $\mathcal{O}_1$ **given** $\mathcal{P}$ **and** $\mathcal{IC}$ iff for all $\mathcal{E}_2$ there exists $\mathcal{E}_1$ consistent with $\mathcal{E}_2$ such that there exists either $\mathsf{inspect}(A)$ or $\mathsf{inspect}_{\neg}(A) \in \mathcal{E}_2$ for which respectively A, $\neg A \notin \mathcal{E}_2$, but in $\mathcal{E}_1$.

$\mathcal{O}_2$ is a **possible contextual relevant consequence of** $\mathcal{O}_1$ **given** $\mathcal{P}$ **and** $\mathcal{IC}$ iff there exists $\mathcal{E}_2$ there exists $\mathcal{E}_1$ consistent with $\mathcal{E}_2$ such that there exists either $\mathsf{inspect}(A)$ or $\mathsf{inspect}_{\neg}(A) \in \mathcal{E}_2$ for which respectively A, $\neg A \notin \mathcal{E}_2$, but in $\mathcal{E}_1$.

Assume we only observe $\mathcal{O}_{storm}$ given $\mathcal{P}_{fire}$ and do not know whether the leaves are dry. Then $\mathcal{E}_{lightning} = \{lightning \leftarrow \top\}$ is only partly explaining $\mathcal{O}_{ffire}$, as additionally we need to abduce that the leaves are dry. However, as this explanation is not decorated with the inspect predicate, it can be abduced straightforwardly by $\mathcal{O}_{ffire}$. Accordingly, $\mathcal{O}_{ffire}$ is a possible contextual relevant consequence of $\mathcal{O}_{storm}$ given $\mathcal{P}^{fire}$ and $\mathcal{IC}_{fire}$. The case of contested relevant contextual consequence, where mere intersection is the case, could too be analogously expressed.

6.4 Jointly Supported Contextual Relevant Consequences

It might be the case that two observations contain side-effects of each other, simultaneously. That is, more generally, we can allow for each observation to need inspection of the abducibles of the other observation; that is, they are mutually plausibly explained together, but not each by itself!

Definition 5
Given background knowledge $\mathcal{P}$ and set of integrity constraints $\mathcal{IC}$, let $\mathcal{O}_1$, $\mathcal{O}_2$ be observations.

$\mathcal{O}_1$ and $\mathcal{O}_2$ are **necessarily jointly supported contextual relevant consequences given** $\mathcal{P}$ **and** $\mathcal{IC}$ iff $\mathcal{O}_1$ is a necessary contextual relevant consequence of $\mathcal{O}_2$ given $\mathcal{P}$ and $\mathcal{IC}$, and $\mathcal{O}_2$ is a necessary contextual relevant consequence of $\mathcal{O}_1$ given $\mathcal{P}$ and $\mathcal{IC}$ as defined in Definition 4.

$\mathcal{O}_1$ and $\mathcal{O}_2$ are **possibly jointly supported contextual relevant consequences given** $\mathcal{P}$ **and** $\mathcal{IC}$ iff $\mathcal{O}_1$ is a possible contextual relevant consequence of $\mathcal{O}_2$ given $\mathcal{P}$ and $\mathcal{IC}$, and $\mathcal{O}_2$ is a possible contextual relevant consequence of $\mathcal{O}_1$ given $\mathcal{P}$ and $\mathcal{IC}$, as defined in Definition 4.

Let us observe $\mathcal{O}_{smoke}$ in $\mathcal{P}_{fire}$. Then we can abduce *fire*, but *ffighters* needs to be explained by some other observation. On the other hand, by observing $\mathcal{O}_{sirens}$, we can abduce *ffighters* but not *fire*. Accordingly, $\mathcal{O}_{smoke}$ and $\mathcal{O}_{sir}$ are necessarily jointly supported contextual relevant consequences given $\mathcal{P}^{fire}$ and $\mathcal{IC}_{fire}$.

7 Conclusion and Future Work

Weak completion semantics is based on a previously proposed approach that seems to adequately model the Wason selection task and Byrne's suppression task. Yet it seems to adequately model another human reasoning task which includes the belief-bias effect. Taking our formalization as starting point, we showed by running examples the need and use for possible extensions in abductive reasoning.

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

Introducing the concept of inspection points in our current framework by applying reserved (meta-)predication for all abductive ground atoms, makes it possible to differentiate between consumed and produced abducibles. This distinction allows us to implement the concepts of contextual side-effect, (jointly supported) contextual relevant consequences and contestable contextual side-effects. Belief bias can thus be modeled using mechanisms described and formalized more abstractly, which deal with contextual abductive reasoning by means of taking side-effects under consideration, applicable to a larger scope of problems, via the notion of inspection points (references to other published examples are given). Our abstract formalism opens therefore the way to a wider use, not restricted to psychological modeling.

When examining abductive explanatory plausibility and contextual counterfactual reasoning new questions raise on whether new observations should be explained by possible or by necessary side-effects. They again might explain new and possibly unexpected side-effects. Additionally, we need to explore how to deal with inconsistency. Another aspect is about choosing the most appropriate explanations. In our examples, we follow Occam's razor with respect to competing explanations, that is, we only consider the minimal explanations. However, there are other measures of preferences which might be more appropriate (e.g. from a human reasoning point of view) and which should be further investigated.

In a future extension, some abducibles can be abduced in a three-valued way if we're trying to make the top goal unknown. This is typical of fault-finding, which concerns two separate problems: (1) finding the abducibles which, if unknown for some faulty components, would make unknown the predicted output using the correct model of the artifact, that happens to be at odds with the faulty output behavior of the artifact; (2) abducing the faulty behavior of components (using a model of the artifact comprising the modeling of faults) to actually predict the faulty behavior. This requires abducing normality of a component as a default, and going back over that normality to change it to unknown, whose technicalities are beyond the scope of this paper.
In a psychological modeling case, the artifact we might be aiming to find faults about can be some psychological model. We then may have available some real person's specific behavior that the model is not consistent with, i.e. it wrongly predicts the negation of that behavior. Consequently, we want at least render our model consistent with that behavior by forcing some abducible or other to be unknown, and thus its specific prediction is unknown, rather than the negation of the person's behavior. A second step is to improve and correct the model to make the right prediction. This is comparable to introducing a component's faulty behavior model into a correct model of the artifact, whereas now the faulty artifact is the (incorrect) model, and the missing faulty component model refers to a missing model part that would produce the correct behavior prediction.

Acknowledgments

We thank Pierangelo Dell'Acqua for his comments. This work was partly funded by DAAD's IPID program, financed by the German Federal Ministry of Education and Research (BMBF).

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

References

Apt, K. R. and van Emden, M. H. 1982. Contributions to the theory of logic programming. *Journal of the ACM 29,* 3, 841–862.

Byrne, R. M. J. 1989. Suppressing valid inferences with conditionals. *Cognition 31*, 61–83.

Clark, K. L. 1978. Negation as failure. In *Logic and Data Bases*, H. Gallaire and J. Minker, Eds. Vol. 1. Plenum Press, New York, NY, 293–322.

Dietz, E.-A., Hölldobler, S., and Ragni, M. 2012. A computational logic approach to the suppression task. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society, CogSci 2013*, N. Miyake, D. Peebles, and R. P. Cooper, Eds. Austin, TX: Cognitive Science Society, 1500–1505.

Dietz, E.-A., Hölldobler, S., and Ragni, M. 2013. A computational logic approach to the abstract and the social case of the selection task. In *Proceedings of the 11th International Symposium on Logical Formalizations of Commonsense Reasoning, COMMONSENSE 2013.* Aeya Nappa, Cyprus.

Dietz, E.-A., Hölldobler, S., and Wernhard, C. 2013. Modeling the suppression task under weak completion and well-founded semantics. *Journal of Applied Non-Classical Logics*.

Evans, J. 2010. Reasoning and imagination. In *Thinking Twice: Two Minds in One Brain.* OUP Oxford.

Evans, J. 2012. Biases in deductive reasoning. In *Cognitive Illusions: A Handbook on Fallacies and Biases in Thinking, Judgement and Memory*, R. Pohl, Ed. Psychology Press.

Evans, J., Barston, J. L., and Pollard, P. 1983. On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition 11,* 3, 295–306.

Hempel, C. G. 1966. *Philosophy of Natural Science*. Prentice Hall, Englewood Cliffs, NJ.

Hölldobler, S. and Kencana Ramli, C. D. 2009a. Logic programs under three-valued Łukasiewicz semantics. In *Logic Programming, 25th International Conference, ICLP 2009*, P. M. Hill and D. S. Warren, Eds. Lecture Notes in Computer Science, vol. 5649. Springer, Heidelberg, 464–478.

Hölldobler, S. and Kencana Ramli, C. D. 2009b. Logics and networks for human reasoning. In *International Conference on Artificial Neural Networks, ICANN 2009, Part II*, C. Alippi, M. M. Polycarpou, C. G. Panayiotou, and G. Ellinas, Eds. Lecture Notes in Computer Science, vol. 5769. Springer, Heidelberg, 85–94.

Hölldobler, S., Philipp, T., and Wernhard, C. 2011. An abductive model for human reasoning. In *Logical Formalizations of Commonsense Reasoning, Papers from the AAAI 2011 Spring Symposium.* AAAI Spring Symposium Series Technical Reports. AAAI Press, Cambridge, MA, 135–138.

Kakas, A. C., Kowalski, R. A., and Toni, F. 1993. Abductive logic programming. *Journal of Logic and Computation 2,* 6, 719–770.

Kowalski, R. 2011. *Computational Logic and Human Thinking: How to be Artificially Intelligent*. Cambridge University Press, Cambridge.

Łukasiewicz, J. 1920. O logice trójwartościowej. *Ruch Filozoficzny 5*, 169–171. English translation: On three-valued logic. In: Łukasiewicz J. and Borkowski L. (ed.). (1990). *Selected Works*, Amsterdam: North Holland, pp. 87–88.

Pereira, L. M., Aparício, J. N., and Alferes, J. J. 1991. Hypothetical reasoning with well founded semantics. In *Scandinavian Conference on Artificial Intelligence: Proc. of the SCAI'91*, B. Mayoh, Ed. IOS Press, Amsterdam, 289–300.

Pereira, L. M. and Pinto, A. M. 2011. Inspecting side-effects of abduction in logic programs. In *Logic Programming, Knowledge Representation, and Nonmonotonic Reasoning: Essays in honour of Michael Gelfond*, M. Balduccini and T. C. Son, Eds. LNAI, vol. 6565. Springer, 148–163.

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press

Saptawijaya, A. and Pereira, L. M. 2013. Tabled abduction in logic programs. *Theory and Practice of Logic Programming 13,* 4-5-Online-Supplement.

Stenning, K. and van Lambalgen, M. 2005. Semantic interpretation as computation in nonmonotonic logic: The real meaning of the suppression task. *Cognitive Science 6,* 29, 916–960.

Stenning, K. and van Lambalgen, M. 2008. *Human Reasoning and Cognitive Science*. A Bradford Book. MIT Press, Cambridge, MA.

Torasso, P., Console, L., Portinale, L., and Dupré, D. T. 1991. On the relationship between abduction and deduction. *Journal of Logic and Computation 1,* 5, 661–690.

Van Gelder, A., Ross, K. A., and Schlipf, J. S. 1991. The well-founded semantics for general logic programs. *Journal of the ACM 38,* 3, 619–649.

Wason, P. 1968. Reasoning about a rule. *Quarterly Journal of Experimental Psychology 20,* 3, 273–281.

https://doi.org/10.1017/S1471068414000258 Published online by Cambridge University Press