On Indicative Conditionals

Emmanuelle Anna Dietz[1], Steffen Hölldobler[1], and Luís Moniz Pereira[2]*

[1]International Center for Computational Logic, TU Dresden, Germany,
{dietz,sh}@iccl.tu-dresden.de
[2]NOVA Laboratory for Computer Science and Informatics, Caparica, Portugal,
lmp@fct.unl.pt

Abstract. In this paper we present a new approach to evaluate indicative conditionals with respect to some background information specified by a logic program. Because the weak completion of a logic program admits a least model under the three-valued Łukasiewicz semantics and this semantics has been successfully applied to other human reasoning tasks, conditionals are evaluated under these least Ł-models. If such a model maps the condition of a conditional to *unknown*, then abduction and revision are applied in order to satisfy the condition. Different strategies in applying abduction and revision might lead to different evaluations of a given conditional. Based on these findings we outline an experiment to better understand how humans handle those cases.

1 Indicative Conditionals

Conditionals are statements of the form *if condition then consequence*. In the literature the condition is also called *if part*, *if clause* or *protasis*, whereas the consequence is called *then part*, *then clause* or *apodosis*. Conditions as well as consequences are assumed to be finite sets (or conjunctions) of ground literals.

Indicative conditionals are conditionals whose condition may or may not be true and, consequently, whose consequence also may or may not be true; however, the consequence is asserted to be true if the condition is true. Examples for indicative conditionals are the following:

If it is raining, then he is inside. (1)

If Kennedy is dead and Oswald did not shoot him, then someone else did. (2)

If rifleman A did not shoot, then the prisoner is alive. (3)

If the prisoner is alive, then the captain did not signal. (4)

If rifleman A shot, then rifleman B shot as well. (5)

If the captain gave no signal and rifleman A decides to shoot,
then the prisoner will die and rifleman B will not shoot. (6)

Conditionals may or may not be true in a given scenario. For example, if we are told that a particular person is living in a prison cell, then most people are

* The authors are mentioned in alphabetical order.

expected to consider (1) to be true, whereas if we are told that he is living in the forest, then most people are expected to consider (1) to be false. Likewise, most people consider (2) to be true.

The question which we shall be discussing in this paper is how to automate reasoning such that conditionals are evaluated by an automated deduction system like humans do. This will be done in a context of logic programming (cf. [11]), abduction [9], Stenning and van Lambalgen's representation of conditionals as well as their semantic operator [19] and three-valued Łukasiewicz logic [12], which has been put together in [6,7,5,8,3] and has been applied to the suppression [2] and the selection task [1], as well as to model the belief-bias effect [15] and contextual abductive reasoning with side-effects [16].

The methodology of the approach presented in this paper differs significantly from methods and techniques applied in well-known approaches to evaluate (mostly subjunctive) conditionals like Ramsey's belief-retention approach [17], Lewis's maximal world-similarity one [10], Rescher's systematic reconstruction of the belief system using principles of saliency and prioritization [18], Ginsberg's possible worlds approach [4] and Pereira and Aparíco's improvements thereof by requiring relevancy [14]. Our approach is inspired by Pearl's do-calculus [13] in that it allows revisions to satisfy conditions whose truth-value is unknown and which cannot be explained by abduction, but which are amenable to hypothetical intervention instead.

2 Preliminaries

We assume the reader to be familiar with logic and logic programming. A *(logic) program* is a finite set of (program) clauses of the form $A \leftarrow B_1 \wedge \ldots \wedge B_n$ where A is an atom and B_i, $1 \leq i \leq n$, are literals or of the form $\top$ and $\bot$, denoting truth- and falsehood, respectively. A is called *head* and $B_1 \wedge \ldots \wedge B_n$ is called *body* of the clause. We restrict terms to be constants and variables only, i.e., we consider so-called data logic programs. Clauses of the form $A \leftarrow \top$ and $A \leftarrow \bot$ are called *positive* and *negative facts*, respectively.

In the this paper we assume for each program that the alphabet consists precisely of the symbols mentioned in the program. When writing sets of literals we will omit curly brackets if the set has only one element.

Let $\mathcal{P}$ be a program. $g\mathcal{P}$ denotes the set of all ground instances of clauses occurring in $\mathcal{P}$. A ground atom A is *defined* in $g\mathcal{P}$ iff $g\mathcal{P}$ contains a clause whose head is A; otherwise A is said to be *undefined*. Let $\mathcal{S}$ be a set of ground literals. $\mathit{def}(\mathcal{S}, \mathcal{P}) = \{A \leftarrow \mathit{body} \in g\mathcal{P} \mid A \in \mathcal{S} \vee \neg A \in \mathcal{S}\}$ is called *definition* of $\mathcal{S}$.

Let $\mathcal{P}$ be a program and consider the following transformation:

1. For each defined atom A, replace all clauses of the form $A \leftarrow \mathit{body}_1, \ldots,$ $A \leftarrow \mathit{body}_m$ occurring in $g\mathcal{P}$ by $A \leftarrow \mathit{body}_1 \vee \ldots \vee \mathit{body}_m$.
2. If a ground atom A is undefined in $g\mathcal{P}$, then add $A \leftarrow \bot$ to the program.
3. Replace all occurrences of $\leftarrow$ by $\leftrightarrow$.

The ground program obtained by this transformation is called *completion* of $\mathcal{P}$, whereas the ground program obtained by applying only the steps 1. and 3. is called *weak completion* of $\mathcal{P}$ or $wc\mathcal{P}$.

We consider the three-valued Łukasiewicz (or Ł-) semantics [12] and represent each interpretation I by a pair $\langle I^\top, I^\bot \rangle$, where $I^\top$ contains all atoms which are mapped to *true* by I, $I^\bot$ contains all atoms which are mapped to *false* by I, and $I^\top \cap I^\bot = \emptyset$. Atoms occurring neither in $I^\top$ not in $I^\bot$ are mapped to *unknown*. Let $\langle I^\top, I^\bot \rangle$ and $\langle J^\top, J^\bot \rangle$ be two interpretations. We define

$$\langle I^\top, I^\bot \rangle \subseteq \langle J^\top, J^\bot \rangle \text{ iff } I^\top \subseteq J^\top \text{ and } I^\bot \subseteq J^\bot.$$

Under Ł-semantics we find $F \wedge \top \equiv F \vee \bot \equiv F$ for each formula F, where $\equiv$ denotes logical equivalence. Hence, occurrences of the symbols $\top$ and $\bot$ in the bodies of clauses can be restricted to those occurring in facts.

It has been shown in [6] that logic programs as well as their weak completions admit a least model under Ł-semantics. Moreover, the least Ł-model of the weak completion of $\mathcal{P}$ can be obtained as least fixed point of the following semantic operator, which was introduced in [19]: $\Phi_{\mathcal{P}}(\langle I^\top, I^\bot \rangle) = \langle J^\top, J^\bot \rangle$, where

$$\begin{aligned} J^\top &= \{A \mid A \leftarrow \mathit{body} \in g\mathcal{P} \text{ and } \mathit{body} \text{ is } \mathit{true} \text{ under } \langle I^\top, I^\bot \rangle\}, \\ J^\bot &= \{A \mid \mathit{def}(A,\mathcal{P}) \neq \emptyset \text{ and} \\ &\qquad \mathit{body} \text{ is } \mathit{false} \text{ under } \langle I^\top, I^\bot \rangle \text{ for all } A \leftarrow \mathit{body} \in \mathit{def}(A,\mathcal{P})\}. \end{aligned}$$

We define $\mathcal{P} \models_{Ł}^{lmwc} F$ iff formula F holds in the least Ł-model of $wc\mathcal{P}$.

As shown in [2], the Ł-semantics is related to the well-founded semantics as follows: Let $\mathcal{P}$ be a program which does not contain a positive loop and let $\mathcal{P}^+ = \mathcal{P} \setminus \{A \leftarrow \bot \mid A \leftarrow \bot \in \mathcal{P}\}$. Let u be a new nullary relation symbol not occurring in $\mathcal{P}$ and B be a ground atom in

$$\mathcal{P}^* = \mathcal{P}^+ \cup \{B \leftarrow u \mid \mathit{def}(B,\mathcal{P}) = \emptyset\} \cup \{u \leftarrow \neg u\}.$$

Then, the least Ł-model of $wc\mathcal{P}$ and the well-founded model for $\mathcal{P}^*$ coincide.

An *abductive framework* consists of a logic program $\mathcal{P}$, a set of *abducibles* $\mathcal{A}_{\mathcal{P}} = \{A \leftarrow \top \mid A \text{ is undefined in } g\mathcal{P}\} \cup \{A \leftarrow \bot \mid A \text{ is undefined in } g\mathcal{P}\}$, a set of *integrity constraints* $\mathcal{IC}$, i.e., expressions of the form $\bot \leftarrow B_1 \wedge \ldots \wedge B_n$, and the entailment relation $\models_{Ł}^{lmwc}$, and is denoted by $\langle \mathcal{P}, \mathcal{A}_{\mathcal{P}}, \mathcal{IC}, \models_{Ł}^{lmwc} \rangle$.

One should observe that each finite set of positive and negative ground facts has an Ł-model. It can be obtained by mapping all heads occurring in this set to *true*. Thus, in the following definition, explanations are always satisfiable.

An *observation* $\mathcal{O}$ is a set of ground literals; it is *explainable* in the abductive framework $\langle \mathcal{P}, \mathcal{A}_{\mathcal{P}}, \mathcal{IC}, \models_{Ł}^{lmwc} \rangle$ iff there exists an $\mathcal{E} \subseteq \mathcal{A}_{\mathcal{P}}$ called *explanation* such that $\mathcal{P} \cup \mathcal{E}$ is satisfiable, the least Ł-model of the weak completion of $\mathcal{P} \cup \mathcal{E}$ satisfies $\mathcal{IC}$, and $\mathcal{P} \cup \mathcal{E} \models_{Ł}^{lmwc} L$ for each $L \in \mathcal{O}$.

3 A Reduction System for Indicative Conditionals

When parsing conditionals we assume that information concerning the mood of the conditionals has been extracted. In this paper we restrict our attention to indicative mood. In the sequel let $\mathit{cond}(\mathcal{T}, \mathcal{A})$ be a conditional with condition $\mathcal{T}$

and consequence $\mathcal{A}$, both of which are assumed to be finite sets of literals not containing a complementary pair of literals, i.e., a pair B and $\neg B$.

Conditionals are evaluated wrt background information specified as a logic program and a set of integrity constraints. More specifically, as the weak completion of each logic program always admits a least Ł-model, the conditionals are evaluated under these least Ł-models. In the reminder of this section let $\mathcal{P}$ be a program, $\mathcal{IC}$ be a finite set of integrity constraints, and $\mathcal{M}_\mathcal{P}$ be the least Ł-model of $wc\mathcal{P}$ such that $\mathcal{M}_\mathcal{P}$ satisfies $\mathcal{IC}$. A *state* is either an expression of the form $ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A})$ or *true*, *false*, *unknown*, or *vacuous*.

3.1 A Revision Operator

Let $\mathcal{S}$ be a finite set of ground literals not containing a complementary pair of literals and let B be a ground atom in

$$rev(\mathcal{P}, \mathcal{S}) = (\mathcal{P} \setminus def(\mathcal{S}, \mathcal{P})) \cup \{B \leftarrow \top \mid B \in \mathcal{S}\} \cup \{B \leftarrow \bot \mid \neg B \in \mathcal{S}\}.$$

The revision operator ensures that all literals occurring in $\mathcal{S}$ are mapped to *true* under the least Ł-model of $wc\, rev(\mathcal{P}, \mathcal{S})$.

3.2 The Abstract Reduction System

Let $cond(\mathcal{T}, \mathcal{A})$ be an indicative conditional which is to be evaluated in the context of a logic program $\mathcal{P}$ and integrity constraints $\mathcal{IC}$ such that the least Ł-model $\mathcal{M}_\mathcal{P}$ of $wc\mathcal{P}$ satisfies $\mathcal{IC}$. The initial state is $ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A})$.

If the condition of the conditional is *true*, then the conditional holds if its consequent is *true* as well; otherwise it is either *false* or *unknown*.

$$\begin{array}{lll}
ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A}) \longrightarrow_{it} true & \text{iff} & \mathcal{M}_\mathcal{P}(\mathcal{T}) = true \text{ and } \mathcal{M}_\mathcal{P}(\mathcal{A}) = true \\
ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A}) \longrightarrow_{if} false & \text{iff} & \mathcal{M}_\mathcal{P}(\mathcal{T}) = true \text{ and } \mathcal{M}_\mathcal{P}(\mathcal{A}) = false \\
ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A}) \longrightarrow_{iu} unknown & \text{iff} & \mathcal{M}_\mathcal{P}(\mathcal{T}) = true \text{ and } \mathcal{M}_\mathcal{P}(\mathcal{A}) = unknown
\end{array}$$

If the condition of the conditional is *false*, then the conditional is *true* under Ł-semantics. However, we believe that humans might make a difference between a conditional whose condition and consequence is *true* and a conditional whose condition is *false*. Hence, for the time being we consider a conditional whose condition is *false* as *vacuous*.

$$ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A}) \longrightarrow_{iv} vacuous \quad \text{iff} \quad \mathcal{M}_\mathcal{P}(\mathcal{T}) = false$$

If the condition of the conditional is *unknown*, then we could assign a truth-value to the conditional in accordance with the Ł-semantics. However, we suggest that in this case abduction and revision shall be applied in order to satisfy the condition. We start with the *abduction rule*:

$$ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A}) \longrightarrow_{ia} ic(\mathcal{P} \cup \mathcal{E}, \mathcal{IC}, \mathcal{T} \setminus \mathcal{O}, \mathcal{A})$$

iff $\mathcal{M}_\mathcal{P}(\mathcal{T}) = \mathit{unknown}$ and $\mathcal{E}$ explains $\mathcal{O} \subseteq \mathcal{T}$ in the abductive framework $\langle \mathcal{P}, \mathcal{A}_\mathcal{P}, \mathcal{IC}, \models_{Ł}^{lmwc} \rangle$ and $\mathcal{O} \neq \emptyset$. Please note that $\mathcal{T}$ may contain literals which are mapped to *true* by $\mathcal{M}_\mathcal{P}$. These literals can be removed from $\mathcal{T}$ by the rule $\longrightarrow_{ia}$ because the empty set explains them.

Now we turn to the *revision rule*:

$$ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A}) \longrightarrow_{ir} ic(\mathit{rev}(\mathcal{P}, \mathcal{S}), \mathcal{IC}, \mathcal{T} \setminus \mathcal{S}, \mathcal{A})$$

iff $\mathcal{M}_\mathcal{P}(\mathcal{T}) = \mathit{unknown}$, $\mathcal{S} \subseteq \mathcal{T}$, $\mathcal{S} \neq \emptyset$, for each $L \in \mathcal{S}$ we find $\mathcal{M}_\mathcal{P}(L) = \mathit{unknown}$, and the least Ł-model of $\mathit{wc}\,\mathit{rev}(\mathcal{P}, \mathcal{S})$ satisfies $\mathcal{IC}$.

Altogether we obtain the reduction system RIC operating on states and consisting of the rules $\{\longrightarrow_{it}, \longrightarrow_{if}, \longrightarrow_{iu}, \longrightarrow_{iv}, \longrightarrow_{ia}, \longrightarrow_{ir}\}$.

4 Examples

4.1 Al in the Jailhouse

Rainy Day Suppose we are told that Al is imprisoned in a jailhouse on a rainy day, i.e.., he is living in a cell inside the jailhouse and it is raining:

$$\mathcal{P}_1 = \{\mathit{inside}(X) \leftarrow \mathit{imprisoned}(X), \mathit{imprisoned}(\mathit{al}) \leftarrow \top, \mathit{raining} \leftarrow \top\}.$$

The least Ł-model of $\mathit{wc}\mathcal{P}_1$ is $\langle \{\mathit{imprisoned}(\mathit{al}), \mathit{inside}(\mathit{al}), \mathit{raining}\}, \emptyset \rangle$. In order to evaluate conditional (1) with respect to $\mathcal{P}_1$ we observe that this model maps *raining* and *inside* to *true*. Hence,

$$ic(\mathcal{P}_1, \emptyset, \mathit{raining}, \mathit{inside}) \longrightarrow_{it} \mathit{true}.$$

Sunny Day Let us assume that Al is still imprisoned but that it is not raining:

$$\mathcal{P}_2 = \{\mathit{inside}(X) \leftarrow \mathit{imprisoned}(X), \mathit{imprisoned}(\mathit{al}) \leftarrow \top, \mathit{raining} \leftarrow \bot\}.$$

The least Ł-model of $\mathit{wc}\mathcal{P}_2$ is $\langle \{\mathit{imprisoned}(\mathit{al}), \mathit{inside}(\mathit{al})\}, \{\mathit{raining}\} \rangle$. In order to evaluate conditional (1) wrt $\mathcal{P}_2$ we observe that this model maps *raining* to *false*. Hence,

$$ic(\mathcal{P}_2, \emptyset, \mathit{raining}, \mathit{inside}) \longrightarrow_{iv} \mathit{vacuous}.$$

No Information about the Weather Suppose we are told that Al is imprisoned in a jailhouse but we know nothing about the weather:

$$\mathcal{P}_3 = \{\mathit{inside}(X) \leftarrow \mathit{imprisoned}(X), \mathit{imprisoned}(\mathit{al}) \leftarrow \top\}.$$

The least Ł-model of $\mathit{wc}\mathcal{P}_1$ is $\langle \{\mathit{imprisoned}(\mathit{al}), \mathit{inside}(\mathit{al})\}, \emptyset \rangle$. In order to evaluate conditional (1) wrt $\mathcal{P}_3$ we observe that this model maps *raining* to *unknown*. Hence, we view *raining* as an observation which needs to be explained. The only possible explanation wrt $\langle \mathcal{P}_3, \{\mathit{raining} \leftarrow \top, \mathit{raining} \leftarrow \bot\}, \emptyset, \models_{Ł}^{lmwc} \rangle$ is $\{\mathit{raining} \leftarrow \top\}$. Altogether we obtain

$$ic(\mathcal{P}_3, \emptyset, \mathit{raining}, \mathit{inside}) \longrightarrow_{ia} ic(\mathcal{P}_1, \emptyset, \emptyset, \mathit{inside}) \longrightarrow_{it} \mathit{true}.$$

Please note that $\mathcal{P}_3 \cup \{\mathit{raining} \leftarrow \top\} = \mathcal{P}_1 = \mathit{rev}(\mathcal{P}_3, \mathit{raining})$. Hence, we could replace $\longrightarrow_{ia}$ by $\longrightarrow_{ir}$ in the previous reduction sequence.

4.2 The Shooting of Kennedy

President Kennedy was killed. There was a lengthy investigation about who actually shot the president and in the end it was determined that Oswald did it:

$$\mathcal{P}_4 = \{Kennedy\_dead \leftarrow os\_shot, Kennedy\_dead \leftarrow se\_shot, os\_shot \leftarrow \top\}.$$

The least Ł-model of $wc\mathcal{P}_4$ is $\langle\{os\_shot, Kennedy\_dead\}, \emptyset\rangle$. Evaluating the indicative conditional (2) under this model we find that its condition $\mathcal{T} = \{Kennedy\_dead, \neg os\_shot\}$ is mapped to *false*. Hence,

$$ic(\mathcal{P}_4, \emptyset, \{Kennedy\_dead, \neg os\_shot\}, se\_shot) \longrightarrow_{iv} vacuous.$$

Now consider the case that we do not know that Oswald shot the president:

$$\mathcal{P}_5 = \{Kennedy\_dead \leftarrow os\_shot, Kennedy\_dead \leftarrow se\_shot\}.$$

As least Ł-model of $wc\mathcal{P}_5$ we obtain $\langle\emptyset, \emptyset\rangle$ and find that it maps $\mathcal{T}$ to *unknown*. We may try to consider $\mathcal{T}$ as an observation and explain it wrt the abductive framework $\langle\mathcal{P}_5, \mathcal{A}_{\mathcal{P}_5}, \emptyset, \models_{L}^{lmwc}\rangle$, where $\mathcal{A}_{\mathcal{P}_5}$ consists of the positive and negative facts for *os_shot* and *se_shot*. The only possible explanation is $\mathcal{E} = \{os\_shot \leftarrow \bot, se\_shot \leftarrow \top\}$. As least Ł-model of $wc(\mathcal{P}_5 \cup \mathcal{E})$ we obtain $\langle\{Kennedy\_dead, se\_shot\}, \{os\_shot\}\rangle$. As this model maps *se_shot* to *true* we find

$$\begin{aligned} &ic(\mathcal{P}_5, \emptyset, \{Kennedy\_dead, \neg os\_shot\}, se\_shot) \\ &\longrightarrow_{ia} ic(\mathcal{P}_5 \cup \mathcal{E}, \emptyset, \emptyset, se\_shot) \longrightarrow_{it} true. \end{aligned}$$

In this example we could also apply revision. Let

$$\mathcal{P}_6 = rev(\mathcal{P}_5, \mathcal{T}) = \{Kennedy\_dead \leftarrow \top, os\_shot \leftarrow \bot\}.$$

We obtain

$$\begin{aligned} &ic(\mathcal{P}_5, \emptyset, \{Kennedy\_dead, \neg os\_shot\}, se\_shot) \\ &\longrightarrow_{ir} ic(\mathcal{P}_6, \emptyset, \emptyset, se\_shot) \longrightarrow_{iu} unknown \end{aligned}$$

because the least Ł-model of $wc\mathcal{P}_6$ is $\langle\{Kennedy\_dead\}, \{os\_shot\}\rangle$ and maps *se_shot* to *unknown*. However, as conditional (2) can be evaluated by abduction and without revising the initial program, this derivation is not preferred.

4.3 The Firing Squad

This example is presented in [13]. If the court orders an execution, then the captain will give the signal upon which riflemen A and B will shoot the prisoner. Consequently, the prisoner will be dead. We assume that the court's decision is *unknown*, that both riflemen are accurate, alert and law-abiding, and that the prisoner is unlikely to die from any other causes. Let

$$\begin{aligned} \mathcal{P}_7 = \{ \; &sig \leftarrow execution, rmA \leftarrow sig, rmB \leftarrow sig, \\ &dead \leftarrow rmA, dead \leftarrow rmB, alive \leftarrow \neg dead \; \}. \end{aligned}$$

The least Ł-model of $wc\mathcal{P}_7$ is

$$\langle\emptyset, \emptyset\rangle. \tag{7}$$

Rifleman A did not Shoot To evaluate conditional (3) wrt this model we first observe that the condition *rmA* is mapped to *unknown* by (7). Considering the abductive framework

$$\langle \mathcal{P}_7, \{execution \leftarrow \top, execution \leftarrow \bot\}, \emptyset, \models_{Ł}^{lmwc} \rangle, \tag{8}$$

$\neg rmA$ can be explained by

$$\{execution \leftarrow \bot\}. \tag{9}$$

Let $\mathcal{P}_8 = \mathcal{P}_7 \cup (9)$. The least Ł-model of $wc\mathcal{P}_8$ is

$$\langle \{alive\}, \{execution, sig, rmA, rmB, dead\} \rangle. \tag{10}$$

Because *alive* is mapped to *true* under this model, we obtain

$$ic(\mathcal{P}_7, \emptyset, \neg rmA, alive) \longrightarrow_{ia} ic(\mathcal{P}_8, \emptyset, \emptyset, alive) \longrightarrow_{it} true.$$

The Prisoner is Alive Now consider conditional (4). Because (7) maps *alive* to *unknown* we treat *alive* as an observation. Considering again the abductive framework (8) this observation can be explained by (9). Hence, we evaluate the consequence of (4) under (10) and find that the captain did not signal:

$$ic(\mathcal{P}_7, \emptyset, alive, \neg sig) \longrightarrow_{ia} ic(\mathcal{P}_8, \emptyset, \emptyset, \neg sig) \longrightarrow_{it} true.$$

Rifleman A Shot Let us turn the attention to conditional (5). Because (7) maps *rmA* to *unknown*, we treat *rmA* as an observation. Considering the abductive framework (8) this observation can be explained by

$$\{execution \leftarrow \top\}. \tag{11}$$

Let $\mathcal{P}_9 = \mathcal{P}_7 \cup (11)$. The least Ł-model of $wc\mathcal{P}_9$ is

$$\langle \{execution, sig, rmA, rmB, dead\}, \{alive\} \rangle. \tag{12}$$

Because *rmB* is mapped to *true* under this model, we obtain

$$ic(\mathcal{P}_7, \emptyset, rmA, rmB) \longrightarrow_{ia} ic(\mathcal{P}_9, \emptyset, \emptyset, rmB) \longrightarrow_{it} true.$$

The Captain Gave no Signal Let us now consider conditional (6). Its condition $\mathcal{T} = \{\neg sig, rmA\}$ is mapped to *unknown* by (7). We can only explain $\neg sig$ by (9) and *rmA* by (11), but we cannot explain $\mathcal{T}$ because

$$wc((9) \cup (11)) = \{execution \leftrightarrow \top \vee \bot\} \equiv \{execution \leftrightarrow \top\}.$$

In order to evaluate this conditional we have to consider revisions.

1. A brute force method is to revise the program wrt all conditions. Let

$$\begin{aligned}\mathcal{P}_{10} &= \mathit{rev}(\mathcal{P}_7, \{\neg \mathit{sig}, \mathit{rmA}\}) \\ &= (\mathcal{P}_7 \setminus \mathit{def}(\{\neg \mathit{sig}, \mathit{rmA}\}, \mathcal{P}_7)) \cup \{\mathit{sig} \leftarrow \bot, \mathit{rmA} \leftarrow \top\}.\end{aligned}$$

   The least Ł-model of $\mathit{wc}\mathcal{P}_{10}$ is

$$\langle \{\mathit{rmA}, \mathit{dead}\}, \{\mathit{sig}, \mathit{rmB}, \mathit{alive}\}\rangle. \tag{13}$$

   This model maps *dead* to *true* and *rmB* to *false* and we obtain

$$\begin{aligned}&\mathit{ic}(\mathcal{P}_7, \emptyset, \{\neg \mathit{sig}, \mathit{rmA}\}, \{\mathit{dead}, \neg \mathit{rmB}\}) \\ &\longrightarrow_{ir} \mathit{ic}(\mathcal{P}_{10}, \emptyset, \emptyset, \{\mathit{dead}, \neg \mathit{rmB}\}) \longrightarrow_{it} \mathit{true}.\end{aligned}$$

2. As we prefer minimal revisions let us consider

$$\mathcal{P}_{11} = \mathit{rev}(\mathcal{P}_7, \mathit{rmA}) = (\mathcal{P}_7 \setminus \mathit{def}(\mathit{rmA}, \mathcal{P}_7)) \cup \{\mathit{rmA} \leftarrow \top\}.$$

   The least Ł-model of $\mathit{wc}\mathcal{P}_{11}$ is $\langle \{\mathit{dead}, \mathit{rmA}\}, \{\mathit{alive}\}\rangle$. Unfortunately, $\neg \mathit{sig}$ is still mapped to *unknown* by this model, but it can be explained in the abductive framework $\langle \mathcal{P}_{11}, \{\mathit{execution} \leftarrow \top, \mathit{execution} \leftarrow \bot\}, \emptyset, \models_{Ł}^{lmwc}\rangle$ by (9). Let $\mathcal{P}_{12} = \mathcal{P}_{11} \cup (9)$. Because the least Ł-model of $\mathit{wc}\mathcal{P}_{12}$ is

$$\langle \{\mathit{dead}, \mathit{rmA}\}, \{\mathit{alive}, \mathit{execution}, \mathit{sig}, \mathit{rmB}\}\rangle \tag{14}$$

   we obtain

$$\begin{aligned}&\mathit{ic}(\mathcal{P}_7, \emptyset, \{\neg \mathit{sig}, \mathit{rmA}\}, \{\mathit{dead}, \neg \mathit{rmB}\}) \\ &\longrightarrow_{ir} \mathit{ic}(\mathcal{P}_{11}, \emptyset, \neg \mathit{sig}, \{\mathit{dead}, \neg \mathit{rmB}\}) \\ &\longrightarrow_{ia} \mathit{ic}(\mathcal{P}_{12}, \emptyset, \emptyset, \{\mathit{dead}, \neg \mathit{rmB}\} \qquad \longrightarrow_{it} \mathit{true}.\end{aligned}$$

   The revision leading to $\mathcal{P}_{11}$ is minimal in the sense that only the definition of *rmA* is revised and without this revision the condition of (6) cannot be explained. This is the only minimal revision as we will show in the sequel.
3. An alternative minimal revision could be the revision of $\mathcal{P}_7$ wrt to $\neg \mathit{sig}$:

$$\mathcal{P}_{13} = \mathit{rev}(\mathcal{P}_7, \neg \mathit{sig}) = (\mathcal{P}_7 \setminus \mathit{def}(\neg \mathit{sig}, \mathcal{P}_7)) \cup \{\mathit{sig} \leftarrow \bot\}.$$

   The least Ł-model of $\mathit{wc}\mathcal{P}_{13}$ is

$$\langle \{\mathit{alive}\}, \{\mathit{sig}, \mathit{dead}, \mathit{rmA}, \mathit{rmB}\}\rangle. \tag{15}$$

   Because this model maps *rmA* to *false* we obtain:

$$\begin{aligned}&\mathit{ic}(\mathcal{P}_7, \emptyset, \{\neg \mathit{sig}, \mathit{rmA}\}, \{\mathit{dead}, \neg \mathit{rmB}\}) \\ &\longrightarrow_{ir} \mathit{ic}(\mathcal{P}_{13}, \emptyset, \mathit{rmA}, \{\mathit{dead}, \neg \mathit{rmB}\}) \longrightarrow_{iv} \mathit{vacuous}.\end{aligned}$$

4. So far the first step in evaluating the conditional was a revision step. Alternatively, we could start with an abduction step. $\neg \mathit{sig}$ can be explained in the abductive framework (8) by (9) leading to the program $\mathcal{P}_8$ and the least Ł-model (10). Because this model maps *rmA* to *false* we obtain:

$$\begin{aligned}&\mathit{ic}(\mathcal{P}_7, \emptyset, \{\neg \mathit{sig}, \mathit{rmA}\}, \{\mathit{dead}, \neg \mathit{rmB}\}) \\ &\longrightarrow_{ia} \mathit{ic}(\mathcal{P}_8, \emptyset, \mathit{rmA}, \{\mathit{dead}, \neg \mathit{rmB}\}) \longrightarrow_{iv} \mathit{vacuous}.\end{aligned}$$

5. Let us now reverse the order in which the conditions are treated and start by explaining *rmA*. This has already been done before and we obtain $\mathcal{P}_9$ and the least Ł-model (12). Because this model maps $\neg sig$ to *false* we obtain:

$$\begin{array}{l} ic(\mathcal{P}_7, \emptyset, \{\neg sig, rmA\}, \{dead, \neg rmB\}) \\ \longrightarrow_{ia} ic(\mathcal{P}_9, \emptyset, \neg sig, \{dead, \neg rmB\}) \longrightarrow_{iv} vacuous. \end{array}$$

In the last example we have discussed five different approaches to handle the case that the truth value of the conditions of a conditional is *unknown* and cannot be explained: maximal (parallel) revision (MaxRev), partial (sequential) revision as well as partial (sequential) explanation, where in the sequential approaches the literals in the condition of the conditionals are treated in different orders: left-to-right and right-to-left, where we consider sets to be ordered (PRevLR,PRevRL,PExLR,PExRL). The results are summarized in Table 1, where the conditional as well as the literals are evaluated wrt the final least Ł-model computed in the different approaches.

Which approach shall be preferred? Because rifleman A causally depends on the captain's signal but not vice-versa, plus given that in this example clauses express causes, and effects come after causes, it would make sense to take the cause ordering as the preferred one for abducing the conditions. Hence, PExLR would be preferred. However, because rifleman A is an agent, the causes of his actions can be internal to him, his decisions. Hence, when autonomous agents are involved (or spontaneous phenomena like radioactivity), the ordering of abducing the conditions is independent of causal dependency.

5 Properties

In this section, let $\mathcal{P}$ be a program, $\langle I^\top, I^\bot \rangle$ the least Ł-model of $wc\mathcal{P}$, $\mathcal{IC}$ a set of integrity constraints, $\langle \mathcal{P}, \mathcal{A}_\mathcal{P}, \mathcal{IC}, \models^{lmwc}_{Ł} \rangle$ an abductive framework, and L a literal.

Proposition 1. *If $\mathcal{O}$ can be explained by $\mathcal{E} \subseteq \mathcal{A}_\mathcal{P}$ and $\langle J^\top, J^\bot \rangle$ is the least L-model of $wc(\mathcal{P} \cup \mathcal{E})$, then $\langle I^\top, I^\bot \rangle \subseteq \langle J^\top, J^\bot \rangle$.*

	MaxRev	PRevRL	PRevLR	PExLR	PExRL
final program	$\mathcal{P}_{10}$	$\mathcal{P}_{12}$	$\mathcal{P}_{13}$	$\mathcal{P}_8$	$\mathcal{P}_9$
final least Ł-model	(13)	(14)	(15)	(10)	(12)
sig	*false*	*false*	*false*	*false*	*true*
rmA	*true*	*true*	*false*	*false*	*true*
dead	*true*	*true*	*false*	*false*	*true*
rmB	*false*	*false*	*false*	*false*	*true*
alive	*false*	*false*	*true*	*true*	*false*
execution	*unknown*	*false*	*unknown*	*false*	*true*
conditional (6)	*true*	*true*	*vacuous*	*vacuous*	*vacuous*

Table 1. Different approaches to evaluate conditional (6).

Proof. The least Ł-models $\langle I^\top, I^\bot\rangle$ and $\langle J^\top, J^\bot\rangle$ are the least fixed points of the semantic operators $\Phi_\mathcal{P}$ and $\Phi_{\mathcal{P}\cup\mathcal{E}}$, respectively. Let $\langle I_n^\top, I_n^\bot\rangle$ and $\langle J_n^\top, J_n^\bot\rangle$ be the interpretations obtained after applying $\Phi_\mathcal{P}$ and $\Phi_{\mathcal{P}\cup\mathcal{E}}$ n-times to $\langle\emptyset,\emptyset\rangle$, respectively. We can show by induction on n that $\langle I_n^\top, I_n^\bot\rangle \subseteq \langle J_n^\top, J_n^\bot\rangle$. The proposition follows immediately. ∎

Proposition 1 guarantees that whenever $\longrightarrow_{ia}$ is applied, previously checked conditions of a conditional need not to be re-checked. The following Proposition 3 gives the same guarantee whenever $\longrightarrow_{ir}$ is applied.

Proposition 2. *If the least Ł-model of wc$\mathcal{P}$ maps L to unknown and $\langle J^\top, J^\bot\rangle$ is the least Ł-model of wc $rev(\mathcal{P}, L)$, then $\langle I^\top, I^\bot\rangle \subset \langle J^\top, J^\bot\rangle$.*

Proof. By induction on the number of applications of $\Phi_\mathcal{P}$ and $\Phi_{rev(\mathcal{P},L)}$. ∎

Proposition 3. RIC *is terminating.*

Proof. Each application of $\longrightarrow_{it}$, $\longrightarrow_{if}$, $\longrightarrow_{iu}$ or $\longrightarrow_{iv}$ leads to an irreducible expression. Let $cond(\mathcal{T}, \mathcal{A})$ be the conditional to which RIC is applied. Whenever $\longrightarrow_{ir}$ is applied then the definition of at least one literal L occurring in $\mathcal{T}$ is revised such that the least Ł-model of the weak completion of revised program maps L to *true*. Because $\mathcal{T}$ does not contain a complementary pair of literals this revised definition of L is never revised again. Hence, there cannot exist a rewriting sequence with infinitely many occurrences of $\longrightarrow_{ir}$. Likewise, there cannot exist a rewriting sequence with infinitely many occurrences of $\longrightarrow_{ia}$ because each application of $\longrightarrow_{ia}$ to a state $ic(\mathcal{P}, \mathcal{IC}, \mathcal{T}, \mathcal{A})$ reduces the number of literals occurring in the $\mathcal{T}$. ∎

Proposition 4. RIC *is not confluent.*

Proof. This follows immediately from the examples presented in Section 4. ∎

6 Open Questions and the Proposal of an Experiment

Open Questions The new approach gives rise to a number of questions. Which of the approaches is preferable? This may be a question of pragmatics imputable to the user. The default, because no pragmatic information has been added, is maximals revision for skepticism and minimal revisions for credulity. Do humans evaluate multiple conditions sequentially or in parallel? If multiple conditions are evaluated sequentially, are they evaluated by some preferred order? Shall explanations be computed skeptically or credulously? How can the approach be extended to handle subjunctive conditionals?

The Proposal of an Experiment Subjects are given the background information specified in the program $\mathcal{P}_9$. They are confronted with the conditionals like (6) as well as variants with different consequences (e.g., *execution* instead of $\{dead, \neg rmB\}$ or conditionals where the order of two conditions are reversed. We then ask the subjects to answer questions like: *Does the conditional hold?* or *Did the court order an execution?* Depending on the answers we may learn which approaches are preferred by humans.

Acknowledgements We thank Bob Kowalski for valuable comments on an earlier draft of the paper.

References

1. E.-A. Dietz, S. Hölldobler, and M. Ragni. A computational logic approach to the abstract and the social case of the selection task. In *Proceedings Eleventh International Symposium on Logical Formalizations of Commonsense Reasoning*, 2013.
2. E.-A. Dietz, S. Hölldobler, and M. Ragni. A computational logic approach to the suppression task. In N. Miyake, D. Peebles, and R. P. Cooper, editors, *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, pages 1500–1505. Cognitive Science Society, 2012.
3. E.-A. Dietz, S. Hölldobler, and C. Wernhard. Modelling the suppression task under weak completion and well-founded semantics. *Journal of Applied Non-Classical Logics*, 24:61–85, 2014.
4. M. L. Ginsberg. Counterfactuals. *Artificial Intelligence*, 30(1):35–79, 1986.
5. S. Hölldobler and C. D. P. Kencana Ramli. Contraction properties of a semantic operator for human reasoning. In Lei Li and K. K. Yen, editors, *Proceedings of the Fifth International Conference on Information*, pages 228–231. International Information Institute, 2009.
6. S. Hölldobler and C. D. P. Kencana Ramli. Logic programs under three-valued Łukasiewicz's semantics. In P. M. Hill and D. S. Warren, editors, *Logic Programming*, volume 5649 of *Lecture Notes in Computer Science*, pages 464–478. Springer-Verlag Berlin Heidelberg, 2009.
7. S. Hölldobler and C. D. P. Kencana Ramli. Logics and networks for human reasoning. In C. Alippi, Marios M. Polycarpou, Christos G. Panayiotou, and Georgios Ellinasetal, editors, *Artificial Neural Networks – ICANN*, volume 5769 of *Lecture Notes in Computer Science*, pages 85–94. Springer-Verlag Berlin Heidelberg, 2009.
8. S. Hölldobler, T. Philipp, and C. Wernhard. An abductive model for human reasoning. In *Proceedings Tenth International Symposium on Logical Formalizations of Commonsense Reasoning*, 2011. commonsensereasoning.org/2011/proceedings.html.
9. A. C. Kakas, R. A. Kowalski, and F. Toni. Abductive Logic Programming. *Journal of Logic and Computation*, 2(6):719–770, 1993.
10. D. Lewis. *Counterfactuals.* Blackwell Publishers, Oxford, 1973.
11. J. W. Lloyd. *Foundations of Logic Programming.* Springer, Berlin, Heidelberg, 1987.
12. J. Łukasiewicz. O logice trójwartościowej. *Ruch Filozoficzny*, 5:169–171, 1920. English translation: On Three-Valued Logic. In: *Jan Łukasiewicz Selected Works.* (L. Borkowski, ed.), North Holland, 87-88, 1990.
13. J. Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, New York, USA, 2000.
14. L. M. Pereira and J. N. Aparício. Relevant counterfactuals. In *Proceedings 4th Portuguese Conference on Artificial Intelligence (EPIA)*, volume 390 of *Lecture Notes in Computer Science*, pages 107–118. Springer, 1989.
15. L. M. Pereira, E.-A. Dietz, and S. Hölldobler. An abductive reasoning approach to the belief-bias effect. In C. Baral, G. De Giacomo, and T. Eiter, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the 14th International Conference*, pages 653–656, Cambridge, MA, 2014. AAAI Press.

16. L. M. Pereira, E.-A. Dietz, and S. Hölldobler. Contextual abductive reasoning with side-effects. In I. Niemelä, editor, *Theory and Practice of Logic Programming (TPLP)*, volume 14, pages 633–648, Cambridge, UK, 2014. Cambridge University Press.
17. F. Ramsey. *The Foundations of Mathematics and Other Logical Essays*. Harcourt, Brace and Company, 1931.
18. N. Rescher. *Conditionals*. MIT Press, Cambridge, MA, 2007.
19. K. Stenning and M. van Lambalgen. *Human Reasoning and Cognitive Science*. MIT Press, 2008.