On 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

We discuss the evaluation of conditionals. Under classical logic a conditional of the form *A implies B* is semantically equivalent to *not A or B*. However, psychological experiments have repeatedly shown that this is not how humans understand and use conditionals. We introduce an innovative abstract reduction system under the three-valued Łukasiewicz logic and the weak completion semantics, that allows us to reason abductively and by revision with respect to conditionals, in three values. We discuss the strategy of *minimal revision followed by abduction* and discuss two notions of relevance. Psychological experiments will need to ascertain if these strategies and notions, or a variant of them, correspond to how humans reason with conditionals.

1 Introduction

We present a novel approach to evaluate indicative and subjunctive (or counterfactual) conditionals. It differs from all approaches to evaluate conditionals which we are aware of like e.g. [12,13,16,19,20,24,25,27,28,30] in that it also considers conditionals whose conditions are *unknown* and evaluates them by applying revision and abduction in order to satisfy them.

Our approach is based on the ideas first expressed in [29], viz., that in a first step we reason towards a particular representation of some aspects of the world as logic programs and, in a second step, we reason with respect to the least model of this representation in a particular three-valued logic. It has been shown in [14] that the (strong) Kleene-Kleene logic proposed by [29] does not lead to the desired results, but the Łukasiewicz logic [17] does, if the weak completion of logic programs is considered. This so-called weak completion semantics (WCS) has been successfully applied to several human reasoning tasks including the *Suppression Task* [9], the *Selection Task* [10], the *Belief-Bias Effect* [21,22] and spatial reasoning [7].

In this paper we apply WCS to deal with conditionals. The background knowledge is formalized in a logic program and a set of integrity constraints following [29]. The weak completion of such a program admits a least model under Łukasiewicz logic and, hence, this model is used in the evaluation of a conditional. The evaluation is computed with respect to an abstract reduction system presented in Section 3. If the condition of the given conditional is *false*, then the conditional is considered to be a counterfactual and non-monotonic revision is applied to satisfy its conditions. If the condition of the given conditional is *unknown* or *true*, then the

---

[*]The authors are mentioned in alphabetical order.

conditional is considered to be an indicative one. If the condition is *unknown* then the conditional is evaluated by either applying monotonic revision or abduction in order to satisfy its condition. If the condition is *true* then the conditional is evaluated by inspecting the truth value of its conclusion.

The abstract reduction system is shown to be terminating but non-confluent, which leads to different evaluations of a given conditional. As we are unaware of any data from psychological experiments which would help us to identify the answers preferred by humans, we develop in Section 4 a set of questions which need to be experimentally evaluated. Following [23] we believe that humans prefer a particular model or – in our case – apply a particular strategy in evaluating conditionals, viz. minimal revision followed by abduction (see Section 5).

Finally, in Section 6 we focus on relevance: It seems to be natural and is widely assumed in the literature that humans evaluate the truth of the consequence based on whether it is supported on something in common with the support of the condition, i.e., their supports must not be completely disjoint or irrelevant to one another [18]. We discuss several concepts of relevance and show why strong relevance fits best in our system.

Our contributions are embedded into the specification of the WCS approach in Section 2 and concluding remarks in Section 7. This paper is an extension and revision of [6, 8], where we have considered indicative conditionals and have shown that this approach subsumes the proposal presented in [26].

2   Preliminaries

We assume the reader to be familiar with logic and logic programming, but recall basic notions and notations. Throughout this paper, A and B, possibly indexed, denote an atom and a literal, respectively. A *(logic) program* is a finite set of (program) clauses of the form $A \leftarrow \top$, $A \leftarrow \bot$ or $A \leftarrow B_1 \wedge \ldots \wedge B_n$, $n > 0$, where $\top$ and $\bot$ denote truth and falsehood, resp. A is called the *head* and $\top$, $\bot$ as well as $B_1 \wedge \ldots \wedge B_n$ are called the *body* of the corresponding clause. Clauses of the form $A \leftarrow \top$ and $A \leftarrow \bot$[1] are called *positive* and *negative facts*, respectively. We restrict terms to be constants and variables only, i.e., we consider *data logic programs*. Throughout this paper, $\mathcal{P}$ denotes a program. We assume for each $\mathcal{P}$ that the underlying alphabet consists precisely of the symbols occurring in $\mathcal{P}$ and that non-propositional programs contain at least one constant.

The *depends on* relation is the transitive closure of the following relation: A_1 *depends on* A_2 if $g\mathcal{P}$ contains a clause of the form $A_1 \leftarrow body$ and A_2 occurs (positively or negatively) in *body*. Let $\mathcal{S}$ be a set of ground literals; $dep(\mathcal{S}, \mathcal{P}) = \{A_2 \leftarrow body \in g\mathcal{P} \mid body \in \{\top, \bot\}, \exists B \in \mathcal{S} : B = A_1 \text{ or } B = \neg A_1, \text{ and } A_1 \text{ depends on } A_2\}$. When writing sets of literals we will omit curly brackets if a set has only one element.

$g\mathcal{P}$ 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*. $def(\mathcal{S}, \mathcal{P}) = \{A \leftarrow body \in g\mathcal{P} \mid A \in \mathcal{S} \text{ or } \neg A \in \mathcal{S}\}$ is called the *definition* of $\mathcal{S}$ in $\mathcal{P}$, where $\mathcal{S}$ is a set of ground literals. $\mathcal{S}$ is said to be *consistent* iff it does not contain a pair of complementary literals.

For a given $\mathcal{P}$, consider the following transformation: (1) For each defined atom A, replace all clauses of the form $A \leftarrow body_1, \ldots, A \leftarrow body_m$ occurring in $g\mathcal{P}$ by $A \leftarrow body_1 \vee \ldots \vee body_m$. (2) Replace all occurrences of $\leftarrow$ by $\leftrightarrow$. The obtained ground set of formulas is called *weak*

---

[1]We consider weak completion semantics and, hence, a clause of the form $A \leftarrow \bot$ is turned into $A \leftrightarrow \bot$ provided that this is the only clause in the definition of A.

completion of $\mathcal{P}$ or $wc\mathcal{P}$.[2]

We consider the three-valued Łukasiewicz (or Ł-) logic [17]. An interpretation I is a mapping from the set formulas into the truth values. It is represented by $\langle I^\top, I^\perp \rangle$, where $I^\top = \{A \mid I(A) = \top\}$, $I^\perp = \{A \mid I(A) = \perp\}$, $I^\top \cap I^\perp = \emptyset$, and each ground atom $A \notin I^\top \cup I^\perp$ is mapped to U (*unknown*). Let $\langle I^\top, I^\perp \rangle$ and $\langle J^\top, J^\perp \rangle$ be two interpretations. We define:

$$
\begin{array}{rclcl}
\langle I^\top, I^\perp \rangle & \subseteq & \langle J^\top, J^\perp \rangle & \text{iff} & I^\top \subseteq J^\top \text{ and } I^\perp \subseteq J^\perp, \\
\langle I^\top, I^\perp \rangle & \cup & \langle J^\top, J^\perp \rangle & = & \langle I^\top \cup J^\top, I^\perp \cup J^\perp \rangle.
\end{array}
$$

A *model* of $\mathcal{P}$ is an interpretation which maps each clause occurring in $\mathcal{P}$ to *true*. It has been shown in [14] that each $\mathcal{P}$ as well as its weak completion admits a least model under Ł-logic, which is the intersection of all models of $\mathcal{P}$. Moreover, the least Ł-model of $wc\mathcal{P}$ can be obtained as the least fixed point of the following semantic operator, which is due to Stenning and van Lambalgen [29]: Let I be an interpretation $\Phi_{\mathcal{P}}(I) = \langle J^\top, J^\perp \rangle$, where

$$
\begin{array}{rcl}
J^\top & = & \{A \mid A \leftarrow body \in g\mathcal{P} \text{ and } I(body) = \top\}, \\
J^\perp & = & \{A \mid def(A, \mathcal{P}) \neq \emptyset \text{ and } \forall A \leftarrow body \in def(A, \mathcal{P}): \ I(body) = \perp\}.
\end{array}
$$

Weak completion semantics (WCS) is the approach to consider weakly completed logic programs and to reason with respect to the least Ł-models of these programs.[3] We write $\mathcal{P} \models_{wcs} F$ iff formula F holds in the least Ł-model of $wc\mathcal{P}$. In the remainder of this paper, $\mathcal{M}_{\mathcal{P}}$ denotes the least Ł-model of $wc\mathcal{P}$.

An *integrity constraint* is an expression of the form $\perp \leftarrow body$. Under three-valued semantics, there are several ways to understand integrity constraints [15], two of them being the *theoremhood* and the *consistency view*. Under the theoremhood view a model satisfies an integrity constraint if *body* is *false*. Under the consistency view, a model satisfies an integrity constraint if *body* is *unknown* or *false*. In this paper, $\mathcal{P}$ *satisfies* a set $\mathcal{IC}$ of integrity constraints iff there exists a model I of $g\mathcal{P}$, and for each $\perp \leftarrow body \in IC$, we find that $I(body) \in \{\perp, U\}$.

An *abductive framework* consists of a logic program $\mathcal{P}$, a set of *abducibles* $\mathcal{A}_{\mathcal{P}} = \{A \leftarrow \top \mid def(A, \mathcal{P}) = \emptyset\} \cup \{A \leftarrow \perp \mid def(A, \mathcal{P}) = \emptyset\}$, a set of *integrity constraints* $\mathcal{IC}$, and the entailment relation $\models_{wcs}$. It is denoted by $\langle \mathcal{P}, \mathcal{A}_{\mathcal{P}}, \mathcal{IC}, \models_{wcs} \rangle$.

One should observe that each $\mathcal{P}$ and, in particular, each finite set of positive and negative ground facts has an Ł-model. For the latter, this can be obtained by mapping all heads occurring in this set to *true*. Thus, in the following definition, explanations as well as the union of a program and an explanation are satisfiable.

An *observation* $\mathcal{O}$ is a set of ground literals; it is *explainable* in the abductive framework $\langle \mathcal{P}, \mathcal{A}_{\mathcal{P}}, \mathcal{IC}, \models_{wcs} \rangle$ iff there exists a (minimal) $\mathcal{E} \subseteq \mathcal{A}_{\mathcal{P}}$ called *explanation* such that $\mathcal{M}_{\mathcal{P} \cup \mathcal{E}}$ satisfies $\mathcal{IC}$ and $\mathcal{P} \cup \mathcal{E} \models_{wcs} L$ for each $L \in \mathcal{O}$.

Let $\mathcal{S}$ be a finite and consistent set of ground literals in

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

where A denotes an atom. $rev(\mathcal{P}, \mathcal{S})$ is called the *revision of $\mathcal{P}$ with respect to $\mathcal{S}$*. Note that $\mathcal{P}$ is revised such that each element of $\mathcal{S}$ is *true* in $\mathcal{M}_{rev(\mathcal{P}, \mathcal{S})}$ (see Proposition 1).

3   ARSC–An Abstract Reduction System for Conditionals

In this paper, we consider conditionals of the form *if $\mathcal{C}$ then $\mathcal{D}$*, where $\mathcal{C}$ and $\mathcal{D}$ are finite and consistent sets of literals viewed as conjunctions of literals. Conditionals are evaluated

---

[2]Note that undefined atoms are not identified with $\perp$ as in the completion of $\mathcal{P}$ [5].
[3]WCS is related to the well-founded semantics as discussed in [11].

with respect to some background information specified as a program and a set of integrity constraints. More specifically, as the weak completion of each program admits a least L-model, conditionals are evaluated under the least L-model of a program. In the reminder, let $\mathcal{P}$ be a program, $\mathcal{IC}$ be a finite set of integrity constraints, $\mathcal{M_P}$ be the least L-model of $wc\mathcal{P}$ such that $\mathcal{M_P}$ satisfies $\mathcal{IC}$, and *if $\mathcal{C}$ then $\mathcal{D}$* be a conditional. The states of ARSC are either the truth values or quadruples containing the program, the integrity constraints and two consistent and finite sets of literals. The initial state for a given program $\mathcal{P}$, a set $\mathcal{IC}$ of integrity constraints and a conditional *if $\mathcal{C}$ then $\mathcal{D}$* is $\langle \mathcal{P}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle$. Final states are *true*, *false* and *unknown*.

The set of rules of ARSC is $\{ \longrightarrow_t, \longrightarrow_c, \longrightarrow_a, \longrightarrow_r \}$:

- $\langle \mathcal{P}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle \longrightarrow_t \mathcal{M_P}(\mathcal{D})$ iff $\mathcal{M_P}(\mathcal{C}) = true$.

- $\langle \mathcal{P}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle \longrightarrow_c \langle rev(\mathcal{P}, \mathcal{S}), \mathcal{IC}, \mathcal{C} \setminus \mathcal{S}, \mathcal{D} \rangle$
  iff $\mathcal{M_P}(\mathcal{C}) = false$, where $\mathcal{S} = \{ L \in \mathcal{C} \mid \mathcal{M_P}(L) = \bot \}$.

- $\langle \mathcal{P}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle \longrightarrow_a \langle \mathcal{P} \cup \mathcal{E}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle$
  iff $\mathcal{M_P}(\mathcal{C}) = unknown$, $\mathcal{O} \subseteq \mathcal{C}$, $\mathcal{O} \neq \emptyset$, for each $L \in \mathcal{O}$ we find $\mathcal{M_P}(L) = unknown$, and $\mathcal{E}$ explains $\mathcal{O}$ in the abductive framework $\langle \mathcal{P}, \mathcal{A_P}, \mathcal{IC}, \models_{wcs} \rangle$.

- $\langle \mathcal{P}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle \longrightarrow_r \langle rev(\mathcal{P}, \mathcal{S}), \mathcal{IC}, \mathcal{C} \setminus \mathcal{S}, \mathcal{D} \rangle$
  iff $\mathcal{M_P}(\mathcal{C}) = unknown$, $\mathcal{S} \subseteq \mathcal{C}$, $\mathcal{S} \neq \emptyset$, for each $L \in \mathcal{S}$ we find $\mathcal{M_P}(L) = unknown$, and $\mathcal{M}_{rev(\mathcal{P},\mathcal{S})}$ satisfies $\mathcal{IC}$.

ARSC extends the reduction system for indicative conditionals presented in [8] in that it can handle counterfactuals by the rule $\longrightarrow_c$; more precisely, conditionals whose condition is *false* are no longer mapped to a state *vacuous* but rather the program is revised in order to satisfy the conditions of the conditional.

ARSC modifies the reduction system presented in [8]. Explanations generated by the rule $\longrightarrow_a$ may not persist because the rule $\longrightarrow_c$ may be applied afterwards and $\longrightarrow_c$ revises the program non-monotonically (see Proposition 1). Hence, when applying $\longrightarrow_a$, observations cannot be deleted from $\mathcal{C}$ even if they are explained.

3.1 Properties

Proposition 1.   *1. rev is non-monotonic,*
    i.e., there exist $\mathcal{P}$, $\mathcal{S}$ and F such that $\mathcal{P} \models_{wcs} F$ and $rev(\mathcal{P}, \mathcal{S}) \not\models_{wcs} F$.

  2. rev is monotonic, i.e., $\mathcal{M_P} \subseteq \mathcal{M}_{rev(\mathcal{P},\mathcal{S})}$, if $\mathcal{M_P}(L) = U$ for all $L \in \mathcal{S}$.

  3. $\mathcal{M}_{rev(\mathcal{P},\mathcal{S})}(\mathcal{S}) = \top$.

For a proof see [6]. Hence, applications of $\longrightarrow_c$ are non-monotonic, whereas applications of $\longrightarrow_r$ are monotonic. Likewise, applications of $\longrightarrow_a$ are monotonic:

Proposition 2. *Let $\langle \mathcal{P}, \mathcal{A_P}, \mathcal{IC}, \models_{wcs} \rangle$ be an abductive framework. If for each $L \in \mathcal{O}$ we find $\mathcal{M_P}(L) = U$ and $\mathcal{O}$ can be explained by $\mathcal{E} \subseteq \mathcal{A_P}$, then $\mathcal{M_P} \subseteq \mathcal{M}_{\mathcal{P} \cup \mathcal{E}}$.*

Proof. By induction on $n \in \mathbb{N}$ one can show that $\Phi_\mathcal{P} \uparrow n \subseteq \Phi_{\mathcal{P} \cup \mathcal{E}} \uparrow n$, where $\Phi \uparrow 0 = \langle \emptyset, \emptyset \rangle$ and $\Phi \uparrow (n+1) = \Phi(\Phi \uparrow n)$. The result follows immediately.   □

Theorem 1. ARSC *is terminating.*

Proof. Let $\langle \mathcal{P}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle$ be a state. Applications of the rule $\longrightarrow_t$ yields a final state. Applications of the rules $\longrightarrow_c$ and $\longrightarrow_r$ reduce the number of literals occurring in $\mathcal{C}$. As $\mathcal{C}$ is finite and no rule increases $\mathcal{C}$, there cannot by an infinite chain of applications of $\longrightarrow_c$ or $\longrightarrow_r$. Finally, consider $\longrightarrow_a$: Because $\mathcal{O} \neq \emptyset$ the number of abducibles occurring in $\mathcal{A}_\mathcal{P}$ is reduced in each application. $\mathcal{A}_\mathcal{P}$ is finite for finite data logic programs $\mathcal{P}$, which prevents an infinite chain of applications of $\longrightarrow_a$.   □

Corollary 1. *Derivations in* ARSC *are of the form* $\{\longrightarrow_c, \longrightarrow_a, \longrightarrow_r\}^n \cdot \longrightarrow_t$.

Theorem 2. *Let* $\langle \mathcal{P}, \mathcal{IC}, \mathcal{C}, \mathcal{D} \rangle$ *be a state. The rules* $\{\longrightarrow_t, \longrightarrow_c, \longrightarrow_r\}$ *need to consider each element of* $\mathcal{C}$ *only once.*

Proof. If $\mathcal{M}_\mathcal{P}(\mathcal{C}) = \top$, then $\longrightarrow_t$ is applicable yielding a final state, and the result holds. If $\mathcal{M}_\mathcal{P}(\mathcal{C}) = \bot$, then $\longrightarrow_c$ is applied and the definitions for $\mathcal{S} = \{L \in \mathcal{C} \mid \mathcal{M}_\mathcal{P}(L) = \bot\}$ have been replaced by facts such that the least L-model of the revised program maps each literal occurring in $\mathcal{S}$ to *true*. As $\mathcal{C}$ is consistent, these new facts will never be revised again. Finally, if $\mathcal{M}_\mathcal{P}(\mathcal{C}) = $ U and $\longrightarrow_r$ was applied, then the definitions for $\mathcal{S} \subseteq \mathcal{C}$ have been replaced by facts such that the least L-model of the revised program maps each literal occurring in $\mathcal{S}$ to *true*, where for each $L \in \mathcal{S}$ we find $\mathcal{M}_\mathcal{P}(L) = $ U. As $\mathcal{C}$ is consistent, the new facts will never be revised again.   □

Theorem 2 does not extend to $\longrightarrow_a$: Let $\mathcal{P}_1 = \{a \leftarrow b, b \leftarrow c\}$, $\mathcal{IC} = \emptyset$, $\mathcal{C} = \{a, \neg b\}$ and $\mathcal{D} = c$. We find $\mathcal{M}_{\mathcal{P}_1} = \langle \emptyset, \emptyset \rangle$ and $\mathcal{A}_{\mathcal{P}_1} = \{c \leftarrow \top, c \leftarrow \bot\}$. As $\mathcal{M}_{\mathcal{P}_1}(\mathcal{C}) = $ U we may apply $\longrightarrow_a$ with $\mathcal{O} = a$, which can be explained by $\mathcal{E}_1 = \{c \leftarrow \top\}$:

$$\langle \mathcal{P}_1, \emptyset, \{a, \neg b\}, c \rangle \longrightarrow_a \langle \mathcal{P}_2, \emptyset, \{a, \neg b\}, c \rangle,$$

where $\mathcal{P}_2 = \mathcal{P}_1 \cup \mathcal{E}_1 = \{a \leftarrow b, b \leftarrow c, c \leftarrow \top\}$. We find $\mathcal{M}_{\mathcal{P}_2} = \langle \{a, b, c\}, \emptyset \rangle$ and $\mathcal{M}_{\mathcal{P}_2}(\{a, \neg b\}) = \bot$. Thus, we can only apply $\longrightarrow_c$ with $\neg b$ being the only literal occurring in $\mathcal{C}$, which is mapped to *false* under $\mathcal{M}_{\mathcal{P}_2}$:

$$\langle \mathcal{P}_2, \emptyset, \{a, \neg b\}, c \rangle \longrightarrow_c \langle \mathcal{P}_3, \emptyset, a, c \rangle,$$

where $\mathcal{P}_3 = rev(\mathcal{P}_2, \neg b) = \{a \leftarrow b, b \leftarrow \bot, c \leftarrow \top\}$. We find $\mathcal{M}_{\mathcal{P}_3} = \langle c, \{b, a\} \rangle$ and, hence, $a \in \mathcal{C}$ is no longer assigned to *true* and must be re-considered:

$$\langle \mathcal{P}_3, \emptyset, a, c \rangle \longrightarrow_c \langle \{a \leftarrow \top, b \leftarrow \bot, c \leftarrow \top\}, \emptyset, \emptyset, c \rangle \longrightarrow_t true.$$

Theorem 3. ARSC *is not confluent.*

Proof. Reconsider the previous example, but select $\mathcal{O} = \neg b$ in the first step. $\mathcal{O}$ can be explained by $\mathcal{E}_2 = \{c \leftarrow \bot\}$ and we obtain

$$\langle \mathcal{P}_1, \emptyset, \{a, \neg b\}, c \rangle \longrightarrow_a \langle \mathcal{P}_4, \emptyset, \{a, \neg b\}, c \rangle,$$

where $\mathcal{P}_4 = \mathcal{P}_1 \cup \mathcal{E}_2 = \{a \leftarrow b, b \leftarrow c, c \leftarrow \bot\}$. We find $\mathcal{M}_{\mathcal{P}_4} = \langle \emptyset, \{a, b, c\} \rangle$ and apply $\longrightarrow_c$ by revising the definition of a:

$$\langle \mathcal{P}_4, \emptyset, \{a, \neg b\}, c \rangle \longrightarrow_c \langle \{a \leftarrow \top, b \leftarrow c, c \leftarrow \bot\}, \emptyset, \neg b, c \rangle \longrightarrow_t false.$$

But it also possible to reduce the initial state to *unknown*:

$$\langle \mathcal{P}_1, \emptyset, \{a, \neg b\}, c \rangle \longrightarrow_r \langle \{a \leftarrow \top, b \leftarrow \bot\}, \emptyset, \emptyset, c \rangle \longrightarrow_t unknown.   □$$

3.2   Examples

The Shooting of Kennedy [1]   President Kennedy was killed. As background knowledge we consider the following implications: *If Oswald shot (os) then the president was killed (k).* *If somebody else shot (ses) then the president was killed.* Finally, it was determined that *Oswald shot.* Following the approach presented in [29], we represent implications as licenses for implications by adding abnormality predicates ab_1 and ab_2.

$$\mathcal{P}_5 = \{k \leftarrow os \wedge \neg ab_1, \ k \leftarrow ses \wedge \neg ab_2, \ os \leftarrow \top, \ ab_1 \leftarrow \bot, \ ab_2 \leftarrow \bot\}.$$

with $\mathcal{M}_{\mathcal{P}_5} = \langle \{os, k\}, \{ab_1, ab_2\}\rangle$. Consider the conditional

if Oswald had not shot, then someone else would have.

Its condition $\neg os$ is *false* under $\mathcal{M}_{\mathcal{P}_5}$ and, hence, we view it as a counterfactual.

$$\langle \mathcal{P}_5, \emptyset, \neg os, ses \rangle \longrightarrow_c \langle \mathcal{P}_6, \emptyset, \emptyset, ses \rangle \longrightarrow_t unknown,$$

where $\mathcal{P}_6 = rev(\mathcal{P}_5, \neg os) = (\mathcal{P}_5 \setminus \{os \leftarrow \top\}) \cup \{os \leftarrow \bot\}$. As *ses* is mapped to *unknown* under $\mathcal{M}_{\mathcal{P}_6} = \langle \emptyset, \{os, ab_1, ab_2\}\rangle$, the conditional is *unknown* as well. Now consider the conditional

if Kennedy was killed and Oswald did not shoot, then someone else did.

Its condition $\{k, \neg os\}$ is still *false* under $\mathcal{M}_{\mathcal{P}_5}$ and we obtain

$$\langle \mathcal{P}_5, \emptyset, \{k, \neg os\}, ses \rangle \longrightarrow_c \langle \mathcal{P}_6, \emptyset, k, ses \rangle.$$

Because $\mathcal{M}_{\mathcal{P}_6}(k) = \mathrm{U}$ we try to explain k in the abductive framework $\langle \mathcal{P}_6, \{ses \leftarrow \top, ses \leftarrow \bot\}, \emptyset, \models_{wcs}\rangle$ and find that $\mathcal{E}_3 = \{ses \leftarrow \top\}$ is the only explanation for k:

$$\langle \mathcal{P}_6, \emptyset, k, ses \rangle \longrightarrow_a \langle \mathcal{P}_6 \cup \mathcal{E}_3, \emptyset, k, ses \rangle \longrightarrow_t true,$$

where *ses* is mapped to *true* under $\mathcal{M}_{\mathcal{P}_6 \cup \mathcal{E}_3} = \langle \{ses, k\}, \{os, ab_1, ab_2\}\rangle$. Instead of abduction we could have applied revision:

$$\langle \mathcal{P}_6, \emptyset, k, ses \rangle \longrightarrow_r \langle \mathcal{P}_7, \emptyset, \emptyset, ses \rangle \longrightarrow_t unknown,$$

where $\mathcal{P}_7 = rev(\mathcal{P}_6, k) = \{k \leftarrow \top, os \leftarrow \bot, ab_1 \leftarrow \bot, ab_2 \leftarrow \bot\}$ and *ses* is mapped to *unknown* under $\mathcal{M}_{\mathcal{P}_7} = \langle k, \{os, ab_1, ab_2\}\rangle$.

Firing Squad   Pearl presents in [19] the so-called *firing squad example: If the court orders an execution (e), then the captain will give the signal (s) upon which riflemen A and B will shoot the prisoner (ra, rb). Consequently, the prisoner will be dead (d).* We assume that the court's decision is *unknown*, that the captain is law-abiding, that both riflemen are accurate, alert and law-abiding, and that the prisoner is unlikely to die from any other causes. Altogether, we obtain the program

$$\mathcal{P}_8 \ = \ \{ s \leftarrow e \wedge \neg ab_1, \ ra \leftarrow s \wedge \neg ab_2, \ rb \leftarrow s \wedge \neg ab_3,$$
$$d \leftarrow ra \wedge \neg ab_4, \ d \leftarrow rb \wedge \neg ab_5\} \cup \{ab_i \leftarrow \bot \mid 1 \le i \le 5\}$$

with $\mathcal{M}_{\mathcal{P}_8} = \langle \emptyset, Ab\rangle$, where $Ab = \{ab_i \mid 1 \le i \le 5\}$. Consider the conditional

if the captain gave no signal and rifleman A decides to shoot,
then the court did not order an execution.

Its condition $\{\neg s, ra\}$ is *unknown* under $\mathcal{M}_{\mathcal{P}_8}$ and, hence, we view it as an indicative conditional. We can revise $\mathcal{P}_8$ wrt $\{\neg s, ra\}$ to obtain

$$\langle \mathcal{P}_8, \emptyset, \{\neg s, ra\}, \neg e \rangle \longrightarrow_r \langle rev(\mathcal{P}_8, \{\neg s, ra\}), \emptyset, \emptyset, \neg e \rangle \longrightarrow_t unknown,$$

where $\mathcal{M}_{rev(\mathcal{P}_8, \{\neg s, ra\})} = \langle \{ra, d\}, \{s\} \cup Ab \rangle$ and, hence, $\neg e$ is *unknown*. Alternatively, we can revise $\mathcal{P}_8$ wrt $\neg s$ first to obtain

$$\begin{aligned}
\langle \mathcal{P}_8, \emptyset, \{\neg s, ra\}, \neg e \rangle \quad &\longrightarrow_r \quad \langle rev(\mathcal{P}_8, \neg s), \emptyset, ra, \neg e \rangle \\
&\longrightarrow_c \quad \langle rev(rev(\mathcal{P}_8, \neg s), ra), \emptyset, \emptyset, \neg e \rangle \quad \longrightarrow_t \quad unknown,
\end{aligned}$$

where $\mathcal{M}_{rev(\mathcal{P}_8, \neg s)} = \langle \emptyset, \{s, ra, rb, d\} \cup Ab \rangle$ and, hence, ra is *false*. The remaining conditional *if ra then* $\neg e$ has become a counterfactual wrt the background program $rev(\mathcal{P}_8, \neg s)$ and, consequently, the definition for ra is revised. As another alternative, we can revise $\mathcal{P}_8$ wrt ra first to obtain

$$\begin{aligned}
\langle \mathcal{P}_8, \emptyset, \{\neg s, ra\}, \neg e \rangle \quad &\longrightarrow_r \quad \langle rev(\mathcal{P}_8, ra), \emptyset, \neg s, \neg e \rangle \\
&\longrightarrow_a \quad \langle rev(\mathcal{P}_8, ra) \cup \{e \leftarrow \bot\}, \emptyset, \neg s, \neg e \rangle \quad \longrightarrow_t \quad true,
\end{aligned}$$

where $\mathcal{M}_{rev(\mathcal{P}_8, ra)} = \langle \{ra, d\}, Ab \rangle$ and, hence, $\neg s$ remains *unknown*. We could apply again revision leading to the same result as in the previous cases, but we apply abduction to explain $\neg s$ by $\{e \leftarrow \bot\}$, which yields a *true* conditional.

The condition $\{\neg s, ra\}$ cannot be explained in the abductive framework $\langle \mathcal{P}_8, \{e \leftarrow \top, e \leftarrow \bot\}, \emptyset, \models_{wcs} \rangle$. But $\{e \leftarrow \bot\}$ explains $\neg s$ and we obtain

$$\begin{aligned}
\langle \mathcal{P}_8, \emptyset, \{\neg s, ra\}, \neg e \rangle \quad &\longrightarrow_a \quad \langle \mathcal{P}_8 \cup \{e \leftarrow \bot\}, \emptyset, \{\neg s, ra\}, \neg e \rangle \\
&\longrightarrow_c \quad \langle rev(\mathcal{P}_8 \cup \{e \leftarrow \bot\}, ra), \emptyset, \neg s, \neg e \rangle \quad \longrightarrow_t \quad true,
\end{aligned}$$

where $\mathcal{M}_{\mathcal{P}_8 \cup \{e \leftarrow \bot\}} = \langle \emptyset, \{e, s, ra, rb, d\} \cup Ab \rangle$ and, hence, ra is *false*. As final alternative we observe that $\{e \leftarrow \top\}$ explains ra and we obtain

$$\begin{aligned}
\langle \mathcal{P}_8, \emptyset, \{\neg s, ra\}, \neg e \rangle \quad &\longrightarrow_a \quad \langle \mathcal{P}_8 \cup \{e \leftarrow \top\}, \emptyset, \{\neg s, ra\}, \neg e \rangle \\
&\longrightarrow_c \quad \langle rev(\mathcal{P}_8 \cup \{e \leftarrow \top\}, \neg s), \emptyset, ra, \neg e \rangle \\
&\longrightarrow_c \quad \langle rev(rev(\mathcal{P}_8 \cup \{e \leftarrow \top\}, \neg s), ra), \emptyset, \emptyset, \neg e \rangle \quad \longrightarrow_t \quad false,
\end{aligned}$$

where $\mathcal{M}_{\mathcal{P}_8 \cup \{e \leftarrow \top\}} = \langle \{e, s, ra, rb, d\}, Ab \rangle$ and, hence, $\neg s$ is *false*. After revising the program wrt $\neg s$, ra is *false* and we need to revise the program once more.

The least L-models of the weak completion of the last programs in the various reduction sequences are shown in Table 1, where we have omitted the final application of $\longrightarrow_t$ and have indexed the remaining rules by the conditions, which were revised or explained. s and ra are always *false* and *true*, respectively. d is always *true* as it depends on ra (and rb). rb is always *false* as it depends on s. But e may take any truth value depending on the sequence in which the conditions are considered and on the rules, which are applied. Thus, the conditional

> *if the captain gave no signal and rifleman A decides to shoot,*
> *then rifleman B will not shoot and the prisoner will be dead.*

will always be evaluated as *true*. The situation will change if it becomes known that *a broken firing pin leads to a malfunctioning rifle*. In this case, $\mathcal{P}_8$ is updated by replacing the definition of ab_4 with $ab_4 \leftarrow b$.[4] If rifleman A decides to shoot now, then it is *unknown* whether the prisoner will die as b is *unknown*. If $b \leftarrow \top$ is added to the program, then the prisoner will not die. Section 6 discusses an extension, which allows to abduce *unknown* consequences.

---

[4] We could update the definition of ab_5 as well, but we should then identify the firing pins of the different rifles.

	$\rightarrow_{r\{\overline{s},ra\}}$	$\rightarrow_{r\{\overline{s}\}} \rightarrow_{c\{ra\}}$	$\rightarrow_{r\{ra\}} \rightarrow_{a\{\overline{s}\}}$	$\rightarrow_{a\{\overline{s}\}} \rightarrow_{c\{ra\}}$	$\rightarrow_{a\{ra\}} \rightarrow_{c\{\overline{s}\}} \rightarrow_{c\{ra\}}$
s	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$
ra	$\top$	$\top$	$\top$	$\top$	$\top$
d	$\top$	$\top$	$\top$	$\top$	$\top$
rb	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$
e	U	U	$\bot$	$\bot$	$\top$

Table 1: The final least L-models in the firing squad example. The derivation shown in the grey box is our preferred one; it is computed by MRFA in Section 5.
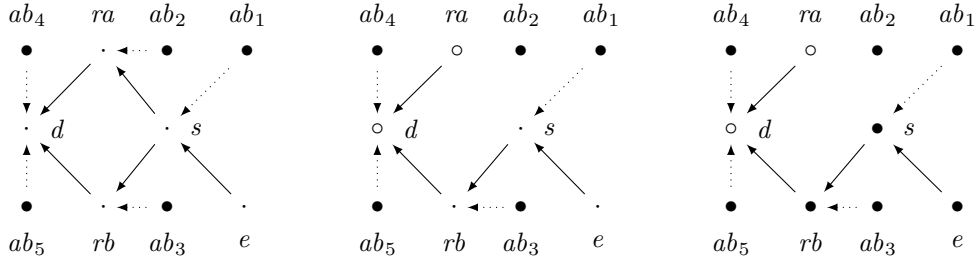

Figure 1: (Left) The dependency graph of $\mathcal{P}_8$. Positive dependencies are depicted by solid arrows, negative dependencies by dotted arrows. •, ·, and ∘ denote nodes, which are mapped to $\bot$, U and $\top$ by $\mathcal{M}_{\mathcal{P}_8}$, resp. The leaf node marked e is undefined, whereas all other nodes are defined. (Middle) The dependency graph of $rev(\mathcal{P}_8, ra)$: ra does not depend on s and ab_2 anymore and is mapped to *true*. (Right) The dependency graph of $rev(\mathcal{P}_8, ra) \cup \{e \leftarrow \bot\}$.

Fig. 1 shows the dependency graph of the program $\mathcal{P}_8$. Revision cuts the dependencies from a particular node and assigns *true* or *false* to the node. Abduction assigns *true* or *false* to the node marked e. The revision step can be understood analogously to Pearl's interventions in his *Do-Calculus* [19], where the antecedent node is isolated from its parent nodes in the network and imposed to be *true* or *false*.

4 The Need for Experimental Data

Although many papers and books have been written about conditionals we are unaware of psychological experiments, which would allow us to define adequate strategies for the application of rules in ARSC and to determine, how humans evaluate conditionals in examples like the *Shooting of Kennedy* or the *Firing Squad*.[5] From our perspective the following questions should be evaluated:

- *Do humans reason with multi-valued logics and, if they do, which multi-valued logic are they using? Can an answer 'I don't know' be qualified as a truth value assignment or is it a meta-remark?*

---

[5]Experimental data are available for examples like the *Suppression* or the *Selection Task*, but these tasks are considerably simpler than the conditionals discussed in the *Firing Squad* example.

- *What do we have to tell humans such that they fully understand the background information including, e.g., the dependency graph in the firing squad example?*

- *Do humans apply abduction and/or revision if the condition of a conditional is unknown and, if they apply both, do they prefer one over the other? Do they prefer skeptical over creduluous abduction? Do they prefer minimal revision?*

- *How important is the order in which multiple conditions of a conditional are considered?*

- *Do humans consider abduction and/or revision steps, which turn an indicative conditional into a subjunctive one like in the second, fourth and fifth reduction sequence of the firing squad example?*

We believe that humans do reason with a third truth value; we have shown that the *Suppression* and the *Selection Tasks* can be adequately modeled under WCS and, moreover, in these tasks skeptical abduction had to be applied [9, 10]. We believe that minimal revision followed by abduction are applied if the conditions of a conditional are *unknown*. Finally, we believe that humans do not consider abduction and/or revision steps, which turn an indicative conditional into a subjunctive one. Altogether, we believe that humans prefer a particular strategy in evaluating conditionals: They do not consider derivations as stated in Corollary 1, but rather search for derivations of the form $\xrightarrow{n}_c \cdot \xrightarrow{0/1}_r \cdot \xrightarrow{0/1}_a \cdot \xrightarrow{0/1}_t$, where $n \in \mathbb{N}^0$ and $\longrightarrow_r$ is only applied if needed.

In other words, the rule $\longrightarrow_c$ is only applied if the given conditional is a counterfactual, in which case the rule may be applied several times because *unknown* conditions may be turned into *false* ones by applications of $\longrightarrow_c$. But as soon as the condition of a conditional is mapped to *true* or *unknown*, $\longrightarrow_c$ will not be applied anymore. Moreover, because $rev(rev(\mathcal{P},\mathcal{S}),\mathcal{S}') = rev(\mathcal{P},\mathcal{S}\cup\mathcal{S}')$ and $(\mathcal{P}\cup\mathcal{E})\cup\mathcal{E}' = \mathcal{P}\cup(E\cup E')$, $\longrightarrow_r$ and $\longrightarrow_a$ need to be applied at most once.

If the condition of a conditional is *unknown*, then a final state can always be reached by applying the revision rule to all *unknown* conditions (first derivation in the *Firing Squad* example), but this is usually not a derivation where minimal revision is applied (third derivation). On the other hand, if $\longrightarrow_c$ cannot be applied anymore, then derivations may end in irreducible states, which are not final (see second, forth or fifth derivation of the *Firing Squad* example, which would all get stuck).

5   Minimal Revision Followed by Abduction

Our belief expressed in the last section allows us to redefine the evaluation of a conditional.

1. If $\mathcal{M}_{\mathcal{P}}(\mathcal{C}) = \top$ then *if $\mathcal{C}$ then $\mathcal{D}$* is assigned to $\mathcal{M}_{\mathcal{P}}(\mathcal{D})$.

2. If $\mathcal{M}_{\mathcal{P}}(\mathcal{C}) = \bot$, then evaluate *if $\mathcal{C}$ then $\mathcal{D}$* with respect to $\mathcal{M}_{rev(\mathcal{P},\mathcal{S})}$,
   where $\mathcal{S} = \{L \in \mathcal{C} \mid \mathcal{M}_{\mathcal{P}}(L) = \bot\}$.

3. If $\mathcal{M}_{\mathcal{P}}(\mathcal{C}) = \mathrm{U}$, then evaluate *if $\mathcal{C}$ then $\mathcal{D}$* with respect to $\mathcal{M}_{\mathcal{P}'}$, where
   - $\mathcal{P}' = rev(\mathcal{P},\mathcal{S}) \cup \mathcal{E}$,
   - $\mathcal{S}$ is a smallest subset of $\mathcal{C}$ and $\mathcal{E} \subseteq \mathcal{A}_{rev(\mathcal{P},\mathcal{S})}$ is a minimal explanation for $\mathcal{C} \setminus \mathcal{S}$ such that $\mathcal{M}_{\mathcal{P}'}(\mathcal{C}) = \top$.

If the condition $\mathcal{C}$ of a conditional is *true*, then the conditional is an indicative one and is evaluated as implication in Ł-logic. If $\mathcal{C}$ is *false*, then the conditional is a counterfactual one and non-monotonic revision is applied in order to reverse the truth value of those literals, which are mapped to *false*. If $\mathcal{C}$ is *unknown*, then we propose to split $\mathcal{C}$ into two disjoint subsets $\mathcal{S}$ and $\mathcal{C} \setminus \mathcal{S}$, where the former is treated by revision and the latter by abduction. In case $\mathcal{C}$ contains

some literals, which are *true* and some, which are *unknown* under $\mathcal{M}_\mathcal{P}$, then the former will be part of $\mathcal{C} \setminus \mathcal{S}$ because the empty explanation explains them. As we assume $\mathcal{S}$ to be minimal, this approach is called *minimal revision followed by abduction* (MRFA). Furthermore, because all revised or explained literals were assigned to *unknown*, this case is monotonic.

Reconsidering the *Firing Squad* example we find that the third derivation shown in grey in Table 1 is the only evaluation wrt MRFA.

As another example consider the *Forest Fire* scenario taken from [4]: *Lightning (ℓ) causes a forest fire (f) if nothing abnormal is taking place, lightning happened, the absence of dry leaves is an abnormality, and dry leaves (d) are present.* We obtain

$$\mathcal{P}_9 = \{f \leftarrow \ell \wedge \neg ab_1, \ \ell \leftarrow \top, \ ab_1 \leftarrow \neg d, \ d \leftarrow \top\}$$

with $\mathcal{M}_{\mathcal{P}_9} = \langle \{d, \ell, f\}, \{ab_1\}\rangle$. Now consider the conditional

if there had not been so many dry leaves on the forest floor,
then the forest fire would not have occurred.

As $\mathcal{M}_{\mathcal{P}_9}(\neg d) = \bot$, the conditional is a counterfactual and we consider $rev(\mathcal{P}_9, \neg d)$. Because $\mathcal{M}_{rev(\mathcal{P}_9, \neg d)} = \langle \{\ell, ab_1\}, \{d, f\}\rangle$ maps $\neg f$ to *true*, the conditional is *true*. Suppose we additionally learn that *arson (a) may cause a forest fire*:

$$\mathcal{P}_{10} = \mathcal{P}_9 \cup \{f \leftarrow a \wedge \neg ab_2, \ ab_2 \leftarrow \bot\}.$$

We find $\mathcal{M}_{\mathcal{P}_{10}} = \langle \{d, \ell, f\}, \{ab_1, ab_2\}\rangle$ and $\mathcal{M}_{rev(\mathcal{P}_{10}, \neg d)} = \langle \{\ell, ab_1\}, \{d, ab_2\}\rangle$. Under this model f is *unknown* and, consequently, the conditional is *unknown*.

Let us assume that arsonists would never go out when there is a storm, especially not when there is lightning: There would never be fire-raising due to arson and lightning at the same time. We represent this information by

$$\mathcal{IC} = \{\bot \leftarrow \ell \wedge a\}.$$

From $\mathcal{IC}$ and because $\mathcal{P}_{10} \models_{wcs} \ell$, we assume a certain background knowledge that can be expressed as the observation $\mathcal{O} = \{\neg a\}$. $\mathcal{P}_{10}$ together with the only explanation for $\mathcal{O}$ is

$$\mathcal{P}_{11} = \mathcal{P}_{10} \cup \{a \leftarrow \bot\}.$$

We find $\mathcal{M}_{\mathcal{P}_{11}} = \langle \{d, \ell, f\}\rangle\{ab_1, ab_2, a\}$ and $\mathcal{M}_{rev(\mathcal{P}_{11}, \neg d)} = \langle \{\ell, ab_1\}, \{d, ab_2, a, f\}\rangle$. Under this model f is *false*, and consequently the conditional is *true*.

6   Relevance

We will discuss relevance in the context of the evaluation of conditionals and present two notions of relevance, both indirectly inspired by [2,3] applied in our context, namely the interdependence of proof supports.

So far, a conditional *if $\mathcal{C}$ then $\mathcal{D}$* is mapped to *unknown* if $\mathcal{M}_P(\mathcal{C}) = \top$ and $\mathcal{M}_P(\mathcal{D}) = U$. We have seen several examples like the first derivation in the *Shooting of Kennedy* example or the last derivation discussed in the *Forest Fire* example. These conclusions are due to the fact that we are using WCS, which adopts an open-world semantics and assigns *unknown* to undefined atoms. Had we applied the well-founded or stable model semantics instead, then because they adopt a closed-world semantics, *false* would have been assigned to undefined atoms. This would have led to a positive evaluation in the last derivation of the *Forest Fire* example, as now, by the absence of dry leaves, lightning could not have caused the forest fire and by arson being

assigned to *false* by default, arson could not have caused the forest fire as well. However, in the *Shooting of Kennedy* example, by assigning *false* to *ses* by default, the conditional would be *false*, which is a rather unexpected result.

We prefer to construct a context in which the conditional is true such as:

The conditional if $\neg d$ then $\neg f$ is true in the context of a being false.

This can be achieved by allowing abduction wrt consequences if the condition of a conditional is mapped to *true*. Returning to the mentioned examples we find in the case of $rev(\mathcal{P}_{10}, \neg d)$ that $\{a \leftarrow \bot\}$ is a minimal explanation for $\neg f$ in the abductive framework $\langle rev(\mathcal{P}_{10}, \neg d), \{a \leftarrow \top, a \leftarrow \bot\}, \emptyset, \models_{wcs}\rangle$. Hence, the conditional *if there had not been so many dry leaves, then the forest fire would not have occurred* is *true* in the context of a being *false*.

In the case of $\mathcal{P}_6$ we find that $\{ses \leftarrow \top\}$ is a minimal explanation for *ses* in the abductive framework $\langle \mathcal{P}_6, \{ses \leftarrow \top, ses \leftarrow \bot\}, \emptyset, \models_{wcs}\rangle$. Hence, the conditional *if $\neg os$ then ses* is *true* in the context of *ses* being *true*.

In fact, any conditional whose conditions are *true* and whose consequences are *unknown* can be mapped to *true* in the context of its consequences being *true*. As an example consider the conditional

if Oswald had not shot, then lightning would have occurred

and suppose that it is evaluated wrt $\mathcal{P}_5$. We would revise $\mathcal{P}_5$ wrt $\neg os$ to obtain $\mathcal{P}_6$ and then explain ℓ by $\{\ell \leftarrow \top\}$. For the evaluation of conditionals this does not seem to be very helpful as it does not include any additional relevant information provided by the conditional itself. This brings us to two new aspects that need to be taken into account: Firstly, we need to restrict the set of abducibles such that the consequence cannot abduce itself and, secondly, we need to check whether the condition of a true conditional is relevant to its consequence.

6.1 Weak Relevance

We might define relevance through dependencies as follows: atom B is *relevant to* atom A iff A depends on B. Unfortunately, this definition is too weak. Let $\mathcal{P}_{12} = \{a \leftarrow b, c \leftarrow b\}$. Is c relevant to a? As a does not depend on c, the answer is *no*. Assume that we would like to evaluate *if c then a* wrt $\mathcal{P}_{12}$ using MRFA: c will be *true* by abducing the explanation $\mathcal{E}_4 = \{b \leftarrow \top\}$ and, consequently, a will be true as well. Thus, c influences a wrt $\mathcal{P}_{12} \cup \mathcal{E}_4$. But, a still does not depend on c in $\mathcal{P}_{12} \cup \mathcal{E}_4$.

Consider the following weak notion of relevance: $\mathcal{C}$ and $\mathcal{D}$ are *weakly relevant to one another wrt $\mathcal{P}$* iff

$$(\{A \leftarrow \top \mid A \in \mathcal{C}\} \cup \{A \leftarrow \bot \mid \neg A \in \mathcal{C}\} \cup dep(\mathcal{C}, \mathcal{P}))$$
$$\cap (\{A \leftarrow \top \mid A \in \mathcal{D}\} \cup \{A \leftarrow \bot \mid \neg A \in \mathcal{D}\} \cup dep(\mathcal{D}, \mathcal{P})) \neq \emptyset$$

Applied to the program $\mathcal{P}_{12} \cup \mathcal{E}_4$ and the conditional *if a then c* we find

$$(\{a \leftarrow \top\} \cup \{b \leftarrow \top\}) \cap (\{c \leftarrow \top\} \cup \{b \leftarrow \top\}) = \{b \leftarrow \top\} \neq \emptyset.$$

c and a are weakly relevant to one another.

As another example consider *if b then a* and $\mathcal{P}_{13} = \{a \leftarrow b, b \leftarrow \top, a \leftarrow \top\}$. As a depends on b, b is relevant to a. However, this seems to be too strong, because the truth of b has no influence on a. a will always be *true* in $\mathcal{M}_{\mathcal{P}_{13}}$. Our notion of weak relevance does not help here: a and b are weakly relevant to one another, because

$$(\{a \leftarrow \top\} \cup dep(a, \mathcal{P}_{12})) \cap (\{b \leftarrow \top\} \cup dep(b, \mathcal{P}_{12})) = \{b \leftarrow \top\} \neq \emptyset.$$

6.2 Strong Relevance

Consider yet another definition of relevance: $\mathcal{C}$ is *strongly relevant to $\mathcal{D}$ wrt $\mathcal{P}$* iff $\mathcal{M}_{\mathcal{P}}(\mathcal{C}) = \mathcal{M}_{\mathcal{P}}(\mathcal{D}) = \top$ and $\mathcal{M}_{\mathcal{P}'}(\mathcal{D}) \neq \top$ where $\mathcal{P}' = \mathcal{P} \setminus (def(\mathcal{C}, \mathcal{P}) \cup dep(\mathcal{C}, \mathcal{P}))$. The idea behind strong relevance is to check whether $\mathcal{D}$ looses support as soon as the support of $\mathcal{C}$ is withdrawn. Note that in contrast to weak relevance, strong relevance is not symmetrical.

Consider again $\mathcal{P}_{13}$: In order to verify whether b is strongly relevant to a wrt $\mathcal{P}_{13}$, we first need to check that both are true in $\mathcal{M}_{\mathcal{P}}$, which is indeed the case. After that, note that $\mathcal{P}'_{12} = \mathcal{P}_{12} \setminus \{b \leftarrow \top\} = \{a \leftarrow b, a \leftarrow \top\}$ where $\mathcal{M}_{\mathcal{P}'_{12}}(a) = \top$. Accordingly, b is not strongly relevant to a wrt $\mathcal{P}_{12}$.

Returning to $\mathcal{P}_{12}$ and $\mathcal{E}_4 = \{b \leftarrow \top\}$ we observe that a and b are *true* and, additionally, a is not true under $\mathcal{M}_{\mathcal{P}'_{12}}$, where $\mathcal{P}'_{12} = (\mathcal{P}_{12} \cup \mathcal{E}_4) \setminus \{def(c, \mathcal{P}_{12}) \cup dep(c, \mathcal{P}_{12})\} = \{a \leftarrow c\}$. Hence, c is strongly relevant to a.

6.3 Extending MRFA by Strong Relevance

The notion of strong relevance captures best our intention, and therefore we will assume it forthwith. Step 1 in MRFA is modified in two ways: Firstly, by checking whether for the *true* conditionals, the condition is strongly relevant to the consequence; and secondly by allowing abduction, in case the consequence is *unknown*:

1. (a) If $\mathcal{M}_{\mathcal{P}}(\mathcal{C}) = \mathcal{M}_{\mathcal{P}}(\mathcal{D}) = \top$ and $\mathcal{C}$ is strongly relevant to $\mathcal{D}$,
      then *if $\mathcal{C}$ then $\mathcal{D}$* is *true*.
   (b) If $\mathcal{M}_{\mathcal{P}}(\mathcal{C}) = \top$ and $\mathcal{M}_{\mathcal{P}}(\mathcal{D}) = \bot$, then *if $\mathcal{C}$ then $\mathcal{D}$* is *false*.
   (c) If $\mathcal{M}_{\mathcal{P}}(\mathcal{C}) = \top$ and $\mathcal{M}_{\mathcal{P}}(\mathcal{D}) = \mathrm{U}$, then:
      i. If $\mathcal{E} \subset \mathcal{A}_{\mathcal{P}} \setminus (\{A \leftarrow \top \mid A \in \mathcal{D}\} \cup \{A \leftarrow \bot \mid \neg A \in \mathcal{D}\})$ is a minimal explanation for $\mathcal{O} \subseteq \mathcal{D}$, $\mathcal{M}_{\mathcal{P} \cup \mathcal{E}}(\mathcal{D}) = \top$ and $\mathcal{C}$ is stronglyi relevant to $\mathcal{D}$ wrt $\mathcal{P} \cup \mathcal{E}$, then *if $\mathcal{C}$ then $\mathcal{D}$* is *true* in the context of $\mathcal{E}$.
      ii. Else, *if $\mathcal{C}$ then $\mathcal{D}$* is *unknown*.

If none of the cases applies because $\mathcal{C}$ is not strongly relevant to $\mathcal{D}$, then the conditional is meaningless.

Let us clarify the difference between strong and weak relevance by considering an extension of the Kennedy example from Subsection 3.2:

$$\mathcal{P}_{14} = \{k \leftarrow os \wedge \neg ab_1, \ k \leftarrow ses \wedge \neg ab_2, \ os \leftarrow \top, \ ab_1 \leftarrow \bot, \ ab_2 \leftarrow \bot, k \leftarrow \top\},$$

which additionally states that, independently of whether Oswald shot Kennedy, Kennedy is dead. Consider the conditional

If Oswald shot, then Kennedy is dead.

$\mathcal{M}_{\mathcal{P}_{14}} = \langle \{k, os\}, \{ab_1, ab_2\} \rangle$, where k and os are both *true*. os is not strongly relevant to k because k is still true in $\mathcal{M}_{\mathcal{P}'_{14}} = \langle \{k\}, \{ab_1, ab_2\} \rangle$ where

$$\mathcal{P}'_{14} = \mathcal{P}_{14} \setminus (def(os, \mathcal{P}_{14}) \cup dep(os, \mathcal{P}_{14})) = \mathcal{P}_{14} \setminus \{os \leftarrow \top\}.$$

However, os and k are weakly relevant to one another, because

$$(\{os \leftarrow \top\} \cup dep(os, \mathcal{P}_{14})) \cap (\{k \leftarrow \top\} \cup dep(k, \mathcal{P}_{14}))$$
$$= (\{os \leftarrow \top\}) \cap (\{k \leftarrow \top\} \cup \{os \leftarrow \top\}) = \{os \leftarrow \top\}.$$

In a nutshell and intuitively, in the example under consideration, os is not essential for k, and found not strongly relevant to it; but it is conceivable that os influences (the truth value of) k by revision or abduction, and so both are weakly relevant to one another.

7   Conclusion

We have presented a novel approach on conditional evaluation: ARSC is an abstract reduction system, which is flexible enough to model various evaluation steps for conditionals, possibly leading to different outcomes. Our hypothesis is that humans prefer a certain evaluation strategy, formalized in MRFA. We additionally assume that they take relevance into account and discuss the notions of weak and strong relevance. As there is not enough experimental data in the literature, we summarize the central questions that need to be investigated through psychological experiments. The results will hopefully give us insights on whether our approach or a variation thereof adequately models how humans reason about conditionals.

References

[1] E. W. Adams. Subjunctive and indicative conditionals. *Foundations of Language*, 6(1):89–94, 1970.

[2] A. R. Anderson and N.D. Belnap. *Entailment: The Logic of Relevance and Necessity, Vol. I.* Princeton University Press, NJ, 1975.

[3] A.R. Anderson, N.D. Belnap, and J.M. Dunn. *Entailment: The Logic of Relevance and Necessity, Vol. II.* Princeton University Press, NJ, 1992.

[4] R. M. J. Byrne. *The Rational Imagination: How People Create Alternatives to Reality*. MIT Press, Cambridge, MA, USA, 2007.

[5] K. L. Clark. Negation as failure. In H. Gallaire and J. Minker, editors, *Logic and Databases*, pages 293–322. Plenum, New York, 1978.

[6] E.-A. Dietz and S. Hölldobler. A new computational logic approach to reason with conditionals. In F. Calimeri, G. Ianni, and M. Truszczynski, editors, *Logic Programming and Nonmonotonic Reasoning, 13th International Conference, LPNMR*, volume 9345 of *Lecture Notes in Artificial Intelligence*. Springer, 2015.

[7] E.-A. Dietz, S. Hölldobler, and R. Höps. A computational logic approach to human spatial reasoning. Technical Report KRR-2015-02, TU Dresden, International Center for Computational Logic, 2015.

[8] E.-A. Dietz, S. Hölldobler, and L. M. Pereira. On indicative conditionals. In S. Hölldobler and Y. Liang, editors, *Proceedings of the First International Workshop on Semantic Technologies*, volume 1339 of *CEUR Workshop Proceedings*, pages 19–30. CEUR-WS.org, 2015. http://ceur-ws.org/Vol-1339/.

[9] 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.

[10] 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. commonsensereasoning.org/2013/proceedings.html.

[11] 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.

[12] J. S. B. T. Evans and D. E. Over. *If.* Oxford University Press, 2004.

[13] M. L. Ginsberg. Counterfactuals. *Artificial Intelligence*, 30(1):35–79, 1986.

[14] 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.

[15] A. C. Kakas, R. A. Kowalski, and F. Toni. Abductive Logic Programming. *Journal of Logic and Computation*, 2(6):719–770, 1993.

[16] D. Lewis. *On the Plurality of Worlds*. Blackwell Publishers, Oxford, 1986.

[17] 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.

[18] E. D. Mares. *Relevant Logic : A Philosophical Interpretation*. Cambridge University Press, 2004.

[19] J. Pearl. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York, USA, 2000.

[20] 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.

[21] 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.

[22] 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.

[23] M. Ragni and M. Knauff. A theory and a computational model of spatial reasoning. *Psychological Review*, 120:561–588, 2013.

[24] F. Ramsey. *The Foundations of Mathematics and Other Logical Essays*. Harcourt, Brace and Company, 1931.

[25] N. Rescher. *Conditionals*. MIT Press, Cambridge, MA, 2007.

[26] K. Schulz. Minimal models vs. logic programming: The case of counterfactual conditionals. *Journal of Applied Non-Classical Logics*, 24(1-2):153–168, 2014.

[27] R. C. Stalnaker. A theory of conditionals. *American Philosophical Quaterly*, pages 98–112, 1968.

[28] R. C. Stalnaker and R. H. Thomason. A semantic analysis of conditional logic. *Theoria*, 36:23–42, 1970.

[29] K. Stenning and M. van Lambalgen. *Human Reasoning and Cognitive Science*. MIT Press, 2008.

[30] J. Vennekens, M. Bruynooghe, and M. Denecker. Embracing events in causal modelling: Interventions and counterfactuals in CP-logic. In T. Janhunen and I. Niemel, editors, *Logics in Artificial Intelligence–12th European Conference, JELIA*, volume 6341 of *LNCS*, pages 313–325. Springer, 2010.