Our Themes on Abduction in Human Reasoning: A Synopsis

Emmanuelle-Anna Dietz Saldanha, Steffen Hölldobler, and Luís Moniz Pereira

Abstract Psychological experiments have shown that humans do not reason according to classical logic. Therefore, we might argue that logic-based approaches in general are not suitable for modeling human reasoning. Yet, we take a different view and are convinced that logic can help us as an underlying formalization of a cognitive theory, but claim rather that classical logic is not adequate for this purpose. In this chapter we investigate abduction and its link to human reasoning. In particular we discuss three different variations we have explored and show how they can be adequately modeled within a novel computational and integrated, cognitive theory, the Weak Completion Semantics.

1 Introduction

Originally, one of the objectives of using logic in artificial intelligence and knowledge representation and reasoning was the formalization of human and commonsense reasoning. During the past decades the original objective shifted out of focus. The problem description was not specified adequately, possibly because there was little or no communication with cognitive scientists. Non-classical logic approaches exist,

The authors are mentioned in alphabetical order.

E.-A. Dietz Saldanha (✉) · S. Hölldobler
Faculty of Computer Science, TU Dresden, 01062 Dresden, Germany
e-mail: sh43@posteo.de

S. Hölldobler
e-mail: sh@iccl.tu-dresden.de

S. Hölldobler
North-Caucasus Federal University, Stavropol, Russian Federation

L. M. Pereira
NOVA Laboratory for Computer Science and Informatics, Departamento de Informática, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829-516 Caparica, Portugal
e-mail: lmp@fct.unl.pt

J. Shook et al. (eds.), *Abduction in Cognition and Action*, Studies in Applied Philosophy, Epistemology and Rational Ethics 59,
https://doi.org/10.1007/978-3-030-61773-8_14

however most of them are purely theoretical and not applied to actual human case studies. Instead, artificial examples have been constructed, which only show that a theory works within some very specific context. But what is the value of a cognitive theory that has never been evaluated against the conclusions that humans actually draw?

Taking this observation as a starting point, several human reasoning episodes, ranging from the suppression and the selection task [8, 9], to spatial, syllogistic [5] and counterfactual reasoning [6, 7] have been successfully modeled under a new, computational and integrated, cognitive theory, the *Weak Completion Semantics*. In a nutshell, under the *Weak Completion Semantics* the following steps are taken for a given human reasoning scenario:

1. Reasoning towards a (logic) program representation.
2. Weakly completing the program.
3. Computing its least model under Łukasiewicz logic.
4. Reasoning with respect to the least model.
5. If necessary, applying abduction.

The *Weak Completion Semantics* is based on ideas first presented in [22, 23]. Unfortunately, these first ideas contained a technical bug, which was corrected in [13] by switching from three-valued Kleene logic [16] to three-valued Łukasiewicz logic [18]. As shown in [13], each weakly completed program admits a least model which can be computed as the least fixed point of an appropriate semantic operator. Unsurprisingly, it turned out that some human reasoning tasks require abduction and, hence, the *Weak Completion Semantics* has been extended with abduction.

In this chapter, we will avoid introducing formal definitions, and assume the reader to be familiar with logic and logic programming. The interested reader is referred to [12, 17]. The goal of this chapter is rather to give the intuitions behind the different facets that abduction can take within human reasoning. By means of four episodes of human reasoning, *the suppression task, counterfactual reasoning, contextual reasoning* and *reasoning with obligation and factual conditionals*, we will show that abductive reasoning is required. Interestingly, a different variation of abduction applies for each of the four cases in order to be adequately modeled under the *Weak Completion Semantics*.

2 The Suppression Task

The *suppression task* is a famous psychological experiment originally carried out by Byrne [2] showing that humans suppress previously drawn conclusion if new information becomes available.

Suppose that participants are told that

if she has an essay to finish then she will study late in the library and that

she will study late in the library.

Then, the subjects are asked whether they are willing to conclude that *she has an essay to write.* As Byrne reports, 71% of the subjects are willing to draw this conclusion.

Following [22, 23], the conditional is translated into the program

$$\{library \leftarrow essay \wedge \neg ab_1,\ ab_1 \leftarrow \bot\},$$

whose weak completion is[1]

$$\{library \leftrightarrow essay \wedge \neg ab_1,\ ab_1 \leftrightarrow \bot\},$$

where *essay* denotes that *she has an essay to finish* and *library* that *she will study late in the library*. ab_1 is an *abnormality* predicate that can deal with possible exceptions to the rule, and is introduced as needed in the program representation, depending on the conditional.

The weakly completed program admits a least model under (three-valued) Łukasiewicz logic,[2] where *library* and essay are mapped to unknown and ab_1 is mapped to false. Adding the fact $library \leftarrow \top$ to the program changes the least model in that library is now mapped to true, but essay is still unknown. Hence, we should not add it but rather consider *library* as an observation that needs to be supported by an explanation. In the given context, the minimal explanation being $\{essay \leftarrow \top\}$. If added to the program, we obtain the weakly completed program

$$\{library \leftrightarrow essay \wedge \neg ab_1,\ essay \leftrightarrow \top,\ ab_1 \leftrightarrow \bot\},$$

whose least model maps *library* and *essay* to true and ab_1 to false. Reasoning with respect to this model allows to abductively conclude *essay*, which most subjects did.

Hypothesis 1 Humans reason abductively.

In another group of the same experiment participants were told:

If she has an essay to finish, then she will study late in the library.
If she has a textbook to read, then she will study late in the library.

All other information being equal to the previous experiment, the number of subjects willing to conclude that *she has an essay to write* drops to 13% [2]. The previously drawn conclusion is suppressed as soon as the second conditional is given. Modelling this reasoning episode similarly to before, we obtain the weakly completed program

[1] One should observe that under completion as defined in [4], $essay \leftrightarrow \bot$ would be an element of this set. The *Weak Completion Semantics* however does not map undefined relations to false but rather considers them unknown.

[2] In the sequel, all models are computed with respect to three-valued Łukasiewicz logic where not explicitly mentioned otherwise.

$$\{library \leftrightarrow (essay \wedge \neg ab_1) \vee (textbook \wedge \neg ab_2),\ ab_1 \leftrightarrow \bot,\ ab_2 \leftrightarrow \bot\},$$

where *textbook* denotes that *she has a textbook to read*. Considering *library* again as an observation to be justified, there are two minimal explanations, viz. $\{essay \leftarrow \top\}$ and $\{textbook \leftarrow \top\}$. Reasoning credulously, i.e. when one explanation suffices to conclude *essay*, then one would conclude that *she has an essay to write*. However, humans do not do this. Their conclusions appear to be modelled adequately when skeptical abduction is applied.

Hypothesis 2 Humans prefer skeptical over credulous abduction.

As shown in [8, 14], all twelve cases of the suppression task can be adequately modelled under the *Weak Completion Semantics*. For six of these cases, skeptical abduction is required.

3 Counterfactual Reasoning

Consider the following scenario [1]: *President Kennedy was killed. There was a lengthy investigation about whether Oswald or somebody else shot the president. In the end, it was determined that Oswald did it.* Which of the following conditionals do humans accept easily?

If Kennedy is dead and Oswald did not shoot Kennedy then someone else did.
If Oswald had not shot Kennedy then someone else would have.

According to [1], humans accept the former, but reject the latter conditional.

In this case, the background knowledge is encoded in the program

$$\begin{array}{lll} \{\ kennedy \leftarrow oswald \wedge \neg ab_o, & ab_o \leftarrow \bot, & \\ \ \ kennedy \leftarrow someone\_else \wedge \neg ab_s, & ab_s \leftarrow \bot, & oswald \leftarrow \top\ \}, \end{array}$$

where *kennedy* denotes that *Kennedy was killed*, *oswald* denotes *Oswald shot*, *someone_else* denotes *somebody else shot*, and ab_o as well as ab_s are introduced abnormality predicates. As the weak completion of this program we obtain

$$\begin{array}{lll} \{\ kennedy \leftrightarrow (oswald \wedge \neg ab_o) \vee (someone\_else \wedge \neg ab_s), & & \\ \ \ ab_o \leftrightarrow \bot, & ab_s \leftrightarrow \bot, & oswald \leftrightarrow \top\ \ \}, \end{array}$$

whose least model maps *oswald* and *kennedy* to true and ab_o and ab_s to false.

The antecedent $\neg oswald$ of the second conditional is false in this least model. Hence, the conditional is a *counterfactual*. In order to evaluate this counterfactual we must revise the background knowledge. The *Weak Completion Semantics* does this in a straightforward and minimal way by replacing the positive fact $oswald \leftarrow \top$ with the negative assumption $oswald \leftarrow \bot$ in the program. Consequently, the least model of the weakly completed modified program maps *oswald*, ab_o and ab_s to false and *kennedy* to unknown. If we evaluate the second conditional with respect to this least model we find that the antecedent $\neg oswald$ of the conditional is true, whereas its consequence *kennedy* is unknown. Thus, the conditional itself is evaluated as unknown.

If we evaluate the first conditional with respect to the least model of the weak completion of the original program, then its antecedent $kennedy \wedge \neg oswald$ is also false as the literal $\neg oswald$ is mapped to false. If the original program is revised as before, then $\neg oswald$ is mapped to true, but now *kennedy* is unknown with respect to the least model of the weakly completed modified program. Hence, the antecedent $kennedy \wedge \neg oswald$ is unknown. However, *kennedy* can be explained by abduction. The only minimal explanation is $\{someone\_else \leftarrow \top\}$. Adding this explanation to the revised program and weakly completing it, we obtain

$$\begin{array}{l}\{\ kennedy \leftrightarrow (oswald \wedge \neg ab_o) \vee (someone\_else \wedge \neg ab_s),\\ \ \ ab_o \leftrightarrow \bot, \quad ab_s \leftrightarrow \bot, \quad oswald \leftrightarrow \bot \quad someone\_else \leftrightarrow \top\ \},\end{array}$$

whose least model maps *someone_else* and *kennedy* to true and *oswald*, ab_o and ab_s to false. Under this model the antecedent $kennedy \wedge \neg oswald$ as well as the consequent *someone_else* of the first conditional are mapped to true. Consequently, the conditional itself is evaluated to true.

Hypothesis 3 If needed, humans minimally revise their background knowledge before applying abduction.

Various counterfactual reasoning scenarios, such as Pearl's *firing squad* scenario [19] and Byrne's *forest fire* scenario [3] can also be adequately modeled by applying *minimal revision followed by abduction* under the *Weak Completion Semantics*, as detailed in [7].

4 Contextual Reasoning

How can we prefer explanations that explain the normal cases to those explanations that explain the exceptional cases? How can we express that some explanations have to be considered only if there is some evidence for considering the exception cases? We want to avoid having to consider all explanations if there is no evidence at all for considering exception cases. On the other hand, we don't want to state that all exception cases are false.

Let us consider the famous Tweety example from [20] which is as follows:

Birds usually fly. Tweety and Jerry are birds.

According to [20], if nothing else is known about Tweety and Jerry, then we can deduce that both can fly.

This scenario can be easily modeled in the *Weak Completion Semantics*. The background knowledge is encoded in the program

$$\{fly(X) \leftarrow bird(X) \wedge \neg ab(X),\ ab(X) \leftarrow \bot,\ bird(tweety) \leftarrow \top,\ bird(jerry) \leftarrow \top\},$$

where X a variable, $fly(X)$ denotes that X *can fly*, $bird(X)$ that X *is a bird* and $ab(X)$ is an abnormality predicate. The corresponding ground program is

$$\begin{array}{lll}\{\ fly(tweety) \leftarrow bird(tweety) \wedge \neg ab(tweety), & ab(tweety) \leftarrow \bot, & bird(tweety) \leftarrow \top, \\ \ \ fly(jerry) \leftarrow bird(jerry) \wedge \neg ab(jerry), & ab(jerry) \leftarrow \bot, & bird(jerry) \leftarrow \top \quad \}.\end{array}$$

Weakly completing this program we obtain

$$\begin{array}{lll}\{\ fly(tweety) \leftrightarrow bird(tweety) \wedge \neg ab(tweety), & ab(tweety) \leftrightarrow \bot, & bird(tweety) \leftrightarrow \top, \\ \ \ fly(jerry) \leftrightarrow bird(jerry) \wedge \neg ab(jerry), & ab(jerry) \leftrightarrow \bot, & bird(jerry) \leftrightarrow \top \quad \}.\end{array}$$

whose least model maps $bird(tweety)$, $bird(jerry)$, $fly(tweety)$ and $fly(jerry)$ to true and $ab(tweety)$ as well as $ab(jerry)$ to false. Reasoning with respect to this least model we conclude that *Tweety and Jerry can fly*.

Suppose we observe

Jerry does not fly.

This observation cannot be explained in the usual abductive framework as specified in [15]. In such a framework the set of abducibles is defined to be the set of undefined relations with respect to the given program, where a relation or ground atom A is said to be *undefined* given a program if (the ground instance of) the program does not contain a rule of the form $A \leftarrow Body$, where $Body$ is a conjunction of literals. In the program above all relations ($fly(jerry)$, $fly(tweety)$, $bird(jerry)$, $bird(tweety)$, $ab(jerry)$, $ab(tweety)$) are defined.

However, under the *Weak Completion Semantics*, negative assumptions like $ab(jerry) \leftarrow \bot$ can be *defeated* by positive facts like $ab(jerry) \leftarrow \top$. The weak completion of a negative assumption $A \leftarrow \bot$ and its *defeater* $A \leftarrow \top$ is $A \leftrightarrow \top \vee \bot$, which is semantically equivalent to $A \leftrightarrow \top$. We believe that humans may defeat negative assumptions. To this end we allow defeaters of negative assumption to be abducibles in an extended abductive framework. Now, the observation *Jerry does not fly* can be explained by the minimal explanation $\{ab(jerry) \leftarrow \top\}$. In other words, *Jerry does not fly because it is an abnormal bird.*

Hypothesis 4 Humans may defeat negative assumptions.

Consider this extended scenario:

Usually birds can fly, but kiwis and penguins cannot.
Tweety and Jerry are birds.

This background information can be encoded by the program

$$\begin{array}{ll}\{\ fly(X) \leftarrow bird(X) \wedge \neg ab(X), & \\ \ \ ab(X) \leftarrow kiwi(X), & ab(X) \leftarrow penguin(X), \\ \ \ bird(tweety) \leftarrow \top, & bird(jerry) \leftarrow \top \quad \}.\end{array}$$

The corresponding ground program is

$$\begin{array}{lll}\{ \mathit{fly}(\mathit{tweety}) \leftarrow \mathit{bird}(\mathit{tweety}) \wedge \neg \mathit{ab}(\mathit{tweety}), & & \\ \mathit{fly}(\mathit{jerry}) \leftarrow \mathit{bird}(\mathit{jerry}) \wedge \neg \mathit{ab}(\mathit{jerry}), & & \\ \mathit{ab}(\mathit{tweety}) \leftarrow \mathit{kiwi}(\mathit{tweety}), & \mathit{ab}(\mathit{tweety}) \leftarrow \mathit{penguin}(\mathit{tweety}), & \\ \mathit{ab}(\mathit{jerry}) \leftarrow \mathit{kiwi}(\mathit{jerry}), & \mathit{ab}(\mathit{jerry}) \leftarrow \mathit{penguin}(\mathit{jerry}), & \\ \mathit{bird}(\mathit{tweety}) \leftarrow \top, & \mathit{bird}(\mathit{jerry}) \leftarrow \top & \}.\end{array}$$

Weakly completing this program we obtain

$$\begin{array}{lll}\{ \mathit{fly}(\mathit{tweety}) \leftrightarrow \mathit{bird}(\mathit{tweety}) \wedge \neg \mathit{ab}(\mathit{tweety}), & & \\ \mathit{fly}(\mathit{jerry}) \leftrightarrow \mathit{bird}(\mathit{jerry}) \wedge \neg \mathit{ab}(\mathit{jerry}), & & \\ \mathit{ab}(\mathit{tweety}) \leftrightarrow \mathit{kiwi}(\mathit{tweety}) \vee \mathit{penguin}(\mathit{tweety}), & & \\ \mathit{ab}(\mathit{jerry}) \leftrightarrow \mathit{kiwi}(\mathit{jerry}) \vee \mathit{penguin}(\mathit{jerry}), & & \\ \mathit{bird}(\mathit{tweety}) \leftrightarrow \top, & \mathit{bird}(\mathit{jerry}) \leftrightarrow \top & \}.\end{array}$$

whose least model maps $\mathit{bird}(\mathit{jerry})$ and $\mathit{bird}(\mathit{tweety})$ to true and all other ground atoms to unknown.

Suppose we observe

Jerry flies.

This observation can be justified by the minimal explanation

$$\{\mathit{kiwi}(\mathit{jerry}) \leftarrow \bot,\ \mathit{penguin}(\mathit{jerry}) \leftarrow \bot\}.$$

One should observe that both atoms, $\mathit{kiwi}(\mathit{jerry})$ and $\mathit{penguin}(\mathit{jerry})$, must be mapped to false in order to map $\mathit{ab}(\mathit{jerry})$ to false as well, which is a prerequisite for mapping $\mathit{fly}(\mathit{jerry})$ to true. In other words, we need to assume that *Jerry is not a kiwi and not a penguin* in order to conclude that *Jerry flies*.

There are several problems with this approach. Firstly, to the best of our knowledge there are currently 41 known classes of birds which do not fly. Secondly, only specialists in biology might know these 41 classes. Thirdly, there may be classes of flightless birds which we are unaware of. Hence, it is unlikely that humans consider all known exceptions before concluding *Jerry flies*.

In order to overcome these problems, the *Weak Completion Semantics* was extended with a new truth-functional operator *ctxt* (called *context*) in [10], whose meaning is specified in Table 1. The meaning of *ctxt* can be understood as a mapping

Table 1 The truth table for *ctxt L*, where *L* denotes a literal

L	*ctxt L*
$\top$	$\top$
$\bot$	$\bot$
U	$\bot$

from three-valuedness to two-valuedness, which allows to locally capture *negation as failure* [4] under *Weak Completion Semantics*. Using the *ctxt* operator we may now specify the program for the last Tweety example as

$$\{\, fly(X) \leftarrow bird(X) \wedge \neg ab(X), \\ ab(X) \leftarrow ctxt\, kiwi(X), \quad ab(X) \leftarrow ctxt\, penguin(X), \\ bird(tweety) \leftarrow \top, \quad bird(jerry) \leftarrow \top \,\}.$$

The corresponding ground program is

$$\{\, fly(tweety) \leftarrow bird(tweety) \wedge \neg ab(tweety), \\ fly(jerry) \leftarrow bird(jerry) \wedge \neg ab(jerry), \\ ab(tweety) \leftarrow ctxt\, kiwi(tweety), \quad ab(tweety) \leftarrow ctxt\, penguin(tweety), \\ ab(jerry) \leftarrow ctxt\, kiwi(jerry), \quad ab(jerry) \leftarrow ctxt\, penguin(jerry), \\ bird(tweety) \leftarrow \top, \quad bird(jerry) \leftarrow \top \,\}.$$

Weakly completing the ground program we obtain

$$\{\, fly(tweety) \leftrightarrow bird(tweety) \wedge \neg ab(tweety), \\ fly(jerry) \leftrightarrow bird(jerry) \wedge \neg ab(jerry), \\ ab(tweety) \leftrightarrow ctxt\, kiwi(tweety) \vee ctxt\, penguin(tweety), \\ ab(jerry) \leftrightarrow ctxt\, kiwi(jerry), \vee ctxt\, penguin(jerry), \\ bird(tweety) \leftrightarrow \top, \quad bird(jerry) \leftarrow \top \,\}.$$

whose least model maps *kiwi*(*jerry*), *kiwi*(*tweety*), *penguin*(*jerry*) and *penguin*(*tweety*) to unknown, *ab*(*jerry*) and *ab*(*tweety*) to false (thanks to the *ctxt* operator), and *bird*(*jerry*) and *bird*(*tweety*) as well as *fly*(*jerry*) and *fly*(*tweety*) to true. Reasoning with respect to this least model allows to conclude that *Jerry flies* by *default* without even considering the exceptional cases of birds which do not fly.

If we learn, for example, that

Jerry is a penguin,

then the situation changes again. Our background knowledge is extended to

$$\{\, fly(X) \leftarrow bird(X) \wedge \neg ab(X), \\ ab(X) \leftarrow ctxt\, kiwi(X), \quad ab(X) \leftarrow ctxt\, penguin(X), \\ bird(tweety) \leftarrow \top, \quad bird(jerry) \leftarrow \top \\ penguin(jerry) \leftarrow \top \,\}.$$

The least model of the weak completion of this program maps *kiwi*(*jerry*), *kiwi*(*tweety*) and *penguin*(*tweety*) to unknown and *penguin*(*jerry*), *bird*(*jerry*) and *bird*(*tweety*) to true. Consequently, *ab*(*jerry*) is mapped to true, whereas *ab*(*tweety*) is still mapped to false. Hence, *fly*(*tweety*) is mapped to true, whereas *fly*(*jerry*) is mapped to false. Thus, we can conclude that *Tweety flies, but Jerry does not fly*.

Let us consider the following characterizations about kiwis and penguins:

If a bird has feathers like hair then it is likely to be a kiwi.
If a bird is black and white then it is likely to be a penguin.

Consequently, let us extend the above program with the following clauses:

$$\{\ kiwi(X) \leftarrow featherslikeHair(X) \wedge \neg ab_k(X), \qquad ab_k(X) \leftarrow \bot,$$
$$penguin(X) \leftarrow blackAndWhite(X) \wedge \neg ab_p(X), \qquad ab_p(X) \leftarrow \bot\ \}.$$

Let us assume that we observe

Tweety has feathers like hair and cannot fly.

The most plausible explanation in the context of knowing that *Tweety has feathers like hair*, is, that *Tweety is a kiwi*, and not, that *Tweety is a penguin*.

Hypothesis 5 Humans prefer certain explanations depending on the context.

In order to provide a preference among explanations, [10] provides an extension for abduction under the WCS, which refines the notion of strong dependency among literals in a program. Strong dependency among literals is given when the dependency path is not via a literal within the *ctxt* operator. Accordingly, only abducibles that strongly depend on the observation, are those that are allowed to serve as explanations. For the above observation *Tweety has feathers like hair*, can be explained by itself, which maps $kiwi(tweety)$ to true and at the same time serves as explanation for *Tweety cannot fly*.

5 Obligation Versus Factual Conditionals

This example is taken from [21] with minor modifications. Consider the background knowledge:

If it rains then the roofs are wet and she takes her umbrella.

In fact, these are two conditionals with the same antecedent which can be represented by the program

$$\{wet\_roofs \leftarrow rain \wedge \neg ab_w,\ ab_w \leftarrow \bot,\ umbrella \leftarrow rain \wedge \neg ab_u,\ ab_u \leftarrow \bot\},$$

where *wet_roofs* denotes that *the roofs are wet*, *rain* denotes that *it rains*, *umbrella* denotes that *she takes her umbrella* and ab_w and ab_u are abnormality predicates. Weakly completing this program we obtain

$$\{wet\_roofs \leftrightarrow rain \wedge \neg ab_w,\ ab_w \leftrightarrow \bot,\ umbrella \leftrightarrow rain \wedge \neg ab_u,\ ab_u \leftrightarrow \bot\},$$

whose least model maps ab_w and ab_u to false and all other atoms to unknown.

Given this background knowledge we would like to evaluate the four conditionals

If the roofs are not wet then it did not rain.
If she did not take her umbrella then it did not rain.
If the roofs are wet then it did rain.
If she took her umbrella then it did rain.

To this end we apply minimal revision followed by abduction as presented in Sect. 3. All conditionals are mapped to true. Because the antecedents of the conditionals are unknown given the least model of the weak completion of the background knowledge, abduction is applied yielding the minimal explanation $\{rain \leftarrow \bot\}$ in case of the first and second conditional and $\{rain \leftarrow \top\}$ in the third and fourth conditional. Hence, the consequences of the four conditionals will be true.

This is a bit surprising. Consider the first two conditionals, where the negative consequence of the background knowledge is abductively affirmed, i.e. the consequence is denied. It is not clear that given the information that *the roofs are not wet*, humans would conclude *it did not rain* in the same way as they would conclude *it did not rain* in case the given information was *she did not take her umbrella*. A similar difference holds regarding the third and fourth conditionals, where the consequent of the background knowledge is confirmed. Given the information that *she took her umbrella*, it is again not clear that humans would conclude *it did rain* in the same way as they would conclude *it did rain* given the information that the *roofs are wet*.

5.1 Obligation and Factual Conditionals

It appears that the two conditionals stated as background knowledge should be semantically interpreted in two different ways. Such a semantic interpretation will be developed in the remainder of this section. Consider the conditionals

if it rains then the roofs are wet

and

if it rains then she takes her umbrella.

The consequence of the first conditional is obligatory. We cannot easily imagine a case, where the antecedent is true and the consequence is not. On the other hand, we can easily imagine a situation, where the antecedent of the second conditional is true and the consequence is not. The consequence of the second conditional is not causally obligatory. We will call the first conditional an *obligation conditional* and the second one a *factual conditional*.

As explained in [3], a conditional whose consequence is denied is more likely to be evaluated to true if it is an obligation conditional. This happens because for this type of conditional there is a forbidden or unlikely possibility where antecedent and not consequent happen together and, in this case, where the consequent is known to be false, it cannot be the case that the antecedent is true as, otherwise, the forbidden possibility is violated. Thus, not antecedent is concluded. Because in the case of a factual conditional this forbidden possibility does not exist, a conditional whose consequence is denied should be evaluated as unknown.

Hypothesis 6 Humans may classify conditionals as obligation or factual conditionals.

This is an informal and pragmatic classification. It depends on the background knowledge and experience of a subject as well as on the context in which a conditional is stated.

One should observe that the conditional

Birds usually fly

considered in Sect. 4 is often classified as a factual conditional. For most humans, its consequence is not obligatory. They can imagine cases, where its antecedent *X is a bird* is true and its consequent *X can fly* is not.

5.2 Necessary and Non-necessary Antecedents

The antecedent of the first conditional in the background knowledge is *necessary*. The consequent cannot be true unless the antecedent is true. The antecedent of the second conditional in the background knowledge does not appear to be necessary. There are many different reasons for taking an umbrella like, for example, that the sun is shining. The antecedent of the second conditional is *non-necessary*.

Hypothesis 7 Humans may classify antecedents as necessary or non-necessary.

The classification is informal and pragmatic. It depends on the background knowledge and experience of a subject as well as on the context in which the condition is stated.

5.3 Reasoning with Obligation and Factual Conditionals

Obligation and factual conditionals are represented by programs as before. Thus, the program representing the background knowledge remains unchanged. However, the semantics of conditionals is taken into account by modifying the set of abducibles for a given program. As usual, undefined relations may be abduced. In addition we allow to abduce expressions of the form $A \leftarrow \top$ if A is the head of a conditional with non-necessary antecedent. The consequent A may be true even if the antecedent is not true. And we allow to abduce defeaters of abnormalities if the abnormality occurs in the body of a factual conditional. (One should observe that this is consistent with the use of defeaters in the examples discussed in Sect. 4.) In our example we obtain the set

$$\{rain \leftarrow \top,\ rain \leftarrow \bot,\ ab_u \leftarrow \top,\ umbrella \leftarrow \top\}$$

of abducibles.

Given this modified set of abducibles we may now evaluate the four conditionals using the principle *minimal revision followed by abduction*. Evaluating the conditional

if the roofs are not wet then it did not rain

yields the same result as before because the explanation $\{rain \leftarrow \bot\}$ is the only minimal explanation for the observation that the *roofs are not wet*. Hence, the conditional is true.

Evaluating the conditional

if she did not take her umbrella then it did not rain

we now find two minimal explanations for the observation that *she did not take her umbrella*, viz. $\{rain \leftarrow \bot\}$ and $\{ab_u \leftarrow \top\}$. If we add the first explanation to the background knowledge, then *rain* is mapped to false. However, if we add the second explanation to the background knowledge, then *rain* remains unknown. Reasoning skeptically the conditional is evaluated as unknown.

Evaluating the conditional

if the roofs are wet then it did rain

yields the same result as before because the explanation $\{rain \leftarrow \top\}$ is the only minimal explanation for the observation that the *roofs are wet*. Hence, the conditional is true.

Evaluating the conditional

if she did take her umbrella then it did rain

we now find two minimal explanations for the observation that *she did take her umbrella*, viz. $\{rain \leftarrow \top\}$ and $\{umbrella \leftarrow \top\}$. If we add the first explanation to the background knowledge, then *rain* is mapped to true. However, if we add the second explanation to the background knowledge, then *rain* remains unknown. Reasoning skeptically the conditional is evaluated as unknown.

According to this classification, [21] can solve the abstract as well as the social case of the selection task: In the abstract case the conditional is interpreted as a factual conditional with necessary antecedent and in the social case the conditional is interpreted as an obligation conditional with non-necessary antecedent.

6 Conclusions

We have shown that different human reasoning episodes can be adequately modeled with abductive reasoning. Interestingly, different forms of abduction seem to appear in the different cases. As exemplified by the suppression task, reasoning within conditionals requires skeptical abduction. On the other hand, when modeling reasoning within conditionals and counterfactuals, a form of minimal revision prior to abduction seems to occur. Finally, the last two cases show that the plausibility of explanations can be determined by the context and by the types of conditions and conditionals.

This leads us to two related observations: (1) From a complexity point of view, computing skeptical conclusions is quite expensive [11, 14], and in case the set of abducibles become larger, skeptical reasoning appears to be infeasible. (2) From a cognitive perspective, it does not seem adequate that humans consider all explanations as equally plausible, but rather, that they only consider a *relevant* subset *bounded* by the context, other observations and the type of conditional.

In order to investigate how these explanations have to be bounded, we will need to know more about how humans reason. In general, we need findings from other psychological experiments that can tell us whether the underlying assumptions of the *Weak Completion Semantics* are adequate. In particular, whether humans do *always* prefer skeptical over credulous abduction and how exactly do they apply *revision* when reasoning about counterfactuals.

From our perspective the following questions are desirably worth evaluating.

- *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?*
- *What do we have to tell humans such that they fully understand the background information?*
- *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 credulous abduction? Do they prefer minimal revisions?*
- *How important is the order in which multiple conditions of a conditional are considered?*
- *Do humans consider abduction and/or revision steps that turn an indicative conditional into a subjunctive one?*

We surmise that humans do reason with a third truth value; we have shown that the *Suppression* and the *Selection Tasks* can be adequately modeled under *Weak Completion Semantics* and, moreover, in these tasks skeptical abduction had to be applied [8]. However, we also believe that they only consider certain explanations as plausible. These are determined by the context, the additional observations and the type of conditional in consideration. Finally, we venture that minimal revision followed by abduction are applied if the conditions of a conditional are *unknown*.

Acknowledgements Luís Moniz Pereira acknowledges support from FCT/MEC NOVA LINCS PEst UID/CEC/04516/2019.

References

1. Adams, E.W.: Subjunctive and indicative conditionals. Found. Lang. **6**(1), 89–94 (1970)
2. Byrne, R.M.J.: Suppressing valid inferences with conditionals. Cognition **31**, 61–83 (1989)
3. Byrne, R.M.J.: The Rational Imagination: How People Create Alternatives to Reality. MIT Press, Cambridge, MA, USA (2007)
4. Clark, K.L.: Negation as failure. In: Gallaire, H., Minker, J. (eds.) Logic and Data Bases, vol. 1, pp. 293–322. Plenum Press, New York, NY (1978)
5. Costa, A., Dietz Saldanha, E.-A., Hölldobler, S., Ragni, M.: A computational logic approach to human syllogistic reasoning. In: Gunzelmann, G., Howes, A., Tenbrink, T., Davelaar, E.J. (eds.) Proceedings of the 39th Annual Conference of the Cognitive Science Society (CogSci 2017), pp. 883–888. Cognitive Science Society, Austin, TX (2017)
6. Dietz, E.-A., Hölldobler, S., Höps, R.: A computational logic approach to human spatial reasoning. In: IEEE Symposium Series on Computational Intelligence (SSCI 2015), pp. 1627–1634. IEEE (2015)
7. Dietz, E.-A., Hölldobler, S., Pereira, L.M.: On conditionals. In: Gottlob, G., Sutcliffe, G., Voronkov, A. (eds.) Global Conference on Artificial Intelligence (GCAI). Epic Series in Computing. EasyChair (2015)
8. Dietz, E.-A., Hölldobler, S., Ragni, M.: A computational logic approach to the suppression task. In: Miyake, N., Peebles, D., Cooper, R.P. (eds.) Proceedings of the 34th Annual Conference of the Cognitive Science Society, pp. 1500–1505. Cognitive Science Society (2012)
9. Dietz, E.-A., Hölldobler, S., Ragni, M.: A computational logic approach to the abstract and the social case of the selection task. In: Proceedings of the 11th International Symposium on Logical Formalizations of Commonsense Reasoning (COMMONSENSE) (2013)
10. Dietz Saldanha, E.-A., Hölldobler, S., Pereira, L.M.: Contextual reasoning: usually birds can abductively fly. In: Proceedings of 14th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR), pp. 64–77. Springer (2017)
11. Dietz Saldanha, E.-A., Hölldobler, S., Philipp, T.: Contextual abduction and its complexity issues. In: Proceedings of the 4th International Workshop on Defeasible and Ampliative Reasoning (DARe) co-located with the 14th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR). CEUR Workshop Proceedings, vol. 1872, pp. 58–70. CEUR-WS.org (2017)
12. Hölldobler, S.: Logik und Logikprogrammierung 1: Grundlagen. Kolleg Synchron, Synchron (2009)
13. Hölldobler, S., Kencana Ramli, C.D.P.: Logics and networks for human reasoning. In: Alippi, C., Polycarpou, M.M., Panayiotou, C.G., Ellinas, G. (eds.) International Conference on Artificial Neural Networks (ICANN), Part II. Lecture Notes in Computer Science, vol. 5769, pp. 85—94. Springer, Heidelberg (2009)
14. Hölldobler, S., Philipp, T., Wernhard, C.: An abductive model for human reasoning. Logical Formalizations of Commonsense Reasoning (Commonsense 2011). AAAI Spring Symposium Series Technical Reports, pp. 135–138. Cambridge, MA (2011)
15. Kakas, A., Kowalski, R., Toni, F.: Abductive logic programming. J. Logic Comput. **2**(6), 719–770 (1992)
16. Kleene, S.C.: Introduction to Metamathematics. North-Holland, Amsterdam (1952)
17. Lloyd, J.W.: Foundations of Logic Programming. Springer, New York, NY, USA (1984)
18. Łukasiewicz, J.: O logice trójwartościowej. Ruch Filozoficzny **5**, 169–171 (1920). English translation: On three-valued logic. In: Łukasiewicz, J., Borkowski, L. (eds.) Selected Works, pp. 87–88. North Holland, Amsterdam (1990)
19. Pearl, Judea: Causality: Models, Reasoning, and Inference. Cambridge University Press, New York, USA (2000)
20. Reiter, Raymond: A logic for default reasoning. Artif. Intell. **13**(1–2), 81–132 (1980)

21. Dietz Saldanha, E.-A., Hölldobler, S., Rocha, I.L.: Obligation versus factual conditionals under the weak completion semantics. In: Hölldobler, S., Malikov, A., Wernhard, C. (eds.) YSIP2—Proceedings of the Second Young Scientist's International Workshop on Trends in Information Processing, Dombai, Russian Federation, 16–20 May 2017. CEUR Workshop Proceedings, vol. 1837, pp. 55–64. CEUR-WS.org (2017)
22. Stenning, K., van Lambalgen, M.: Semantic interpretation as computation in nonmonotonic logic: the real meaning of the suppression task. Cogn Sci **6**(29), 916–960 (2005)
23. Stenning, K., van Lambalgen, M.: Human Reasoning and Cognitive Science. A Bradford Book (2008)