On Disjunctions and the Weak Completion Semantics

Islam Hamada (islamhamada222@gmail.com)
Technische Universität Dresden, Fakultät Informatik, 01062 Dresden, Germany

Steffen Hölldobler (sh43@posteo.de)
Technische Universität Dresden, Fakultät Informatik, 01062 Dresden, Germany
and North Caucasus Federal University, Stavropol, Russian Federation [1]

Abstract

The weak completion semantics is a three-valued, non-monotonic theory which has been shown to adequately model various cognitive reasoning tasks. In this paper we extend the weak completion semantics to model disjunctions and exclusive disjunctions. Such disjunctions are encoded by integrity constraints and skeptical abduction is applied to compute logical consequences. We discuss various examples and relate the approach to the elimination of disjunctions in the calculus of natural deduction.

Keywords: logic programming; human reasoning; disjunctions; weak completion semantics; natural deduction.

Introduction

Logic programs are universal in that they can compute any computable function (Sebelik & Stepanek, 1982). Hence, if high-level cognitive processes like reasoning are computable, then – from a functional point of view – they can be computed by logic programs. Although originally developed in the context of classical binary logic, modern logic programming approaches are richer: they cover non-monotonicity (Clark, 1978), they can be learned (Muggleton, 1992), they can be interpreted over multi-valued logics (Fitting, 1985), and they can be mapped onto artificial neural or connectionist networks (Hölldobler & Kalinke, 1994).

The *weak completion semantics (WCS)* is a logic programming approach to model human reasoning. Based on ideas originally developed by Stenning and van Lambalgen (2008), it is a three-valued, non-monotonic theory which is knowledge-rich, can handle inconsistent background knowledge, and has been shown to adequately model the average case in various human reasoning tasks like the suppression task (Dietz, Hölldobler, & Ragni, 2012), human syllogistic reasoning (Oliviera da Costa, Dietz Saldanha, Hölldobler, & Ragni, 2017), and human conditional reasoning (Cramer, Hölldobler, & Ragni, 2021). Thus, the WCS offers solutions for the five fundamental problems attributed to the classical binary logic approach in the psychology of reasoning by Oaksford and Chater (2020).

The WCS differs significantly from *mental logic* (Rips, 1994) and the *mental model theory* (Johnson-Laird, 1983). Mental logic is based on syntactic rules which are valid in classical binary logic. However, as pointed out by López-Astorga (2015) they are not a complete system like, for example, the calculus of natural deduction of Gentzen (1935); some problematic rules like the *introduction of disjunction* are omitted and other rules like the *introduction of implication* have certain restrictions. The WCS generates models much like the mental model theory. But, whereas the models in the mental model theory are classical binary models, the models in WCS are ternary. Moreover, the computation of WCS's models is rigorously defined by means of a fixed point construction and skeptical abduction.

The semantics of logic programs is usually defined model-theoretically, fixpoint-theoretically, and operationally. This applies to WCS as well. It has an operational semantics given by implementations in PROLOG, PYTHON, and ASP. Hölldobler and Kencana Ramli (2009) have shown that each program and its weak completion[2] admits a least model under the three-valued Łukasiewicz (1920) logic, which can be computed as the least fixed point of a semantic operator specified by Stenning and van Lambalgen (2008). In other words, in the WCS a least model is constructed and reasoning is with respect to this model. Moreover, the least model can be computed by a connectionist network (Dietz Saldanha, Hölldobler, Kencana Ramli, & Palacios Medinacelli, 2018).

As an example consider the conditional sentence *if Ella is studying, then she will be hungry.* Following Stenning and van Lambalgen (2008) this sentence is represented by the logic program $\mathcal{P} = \{h \leftarrow s \wedge \neg ab_s,\ ab_s \leftarrow \bot\}$, where h and s denote that *Ella will be hungry* and *Ella is studying*, respecively, and ab_s is an abnormality predicate. ab_s encapsulates everything that could prevent the conditional from holding and is assumed to be false initially.[3] If the conditional sentence is given as first premise, then a three-valued model is constructed, where ab_s is mapped to false and h and s are mapped to unknown. This initial model is represented by $\langle \emptyset, \{ab_s\} \rangle$. If the sentence *Ella is studying* is given as second premise, then the mapping of s is updated to true. Consequently, $s \wedge \neg ab_s$ is true and, hence, h must be true as well. This second model is represented by $\langle \{s, h\}, \{ab_s\} \rangle$. In fact, it is the least model of the weak completion of $\mathcal{P}$. Reasoning is performed with respect to this model and, hence, we conclude *Ella will be hungry*.

[1] The authors are mentioned in alphabetical order.

[2] See next section for a formal definition.

[3] In this paper the abnormalities will always be false. However, in other applications like the suppression task they are important to model exceptional cases and enabling relations (Dietz et al., 2012).

The WCS is a rigorously defined formal theory. But is it also a cognitive or psychological theory? Are the models constructed by the WCS mental models in the sense of Craik (1945) or Johnson-Laird (1983)? Is it plausible from a cognitive or psychological point of view that humans construct models in a similar way as the WCS? Is it plausible from a biological point of view that a human brain constructs models in a similar way as the connectionist encoding of WCS?

These are challenging questions. So far, we have approached them by considering human reasoning tasks and experimental data from the literature as well as conducting own experiments. But, the WCS cannot deal with disjunctions like *Ella is studying or Ella is running, or both*, or *Ella is studying or Ella is running, but not both*. It is the aim of this paper to overcome this limitation. To this end, we propose to represent disjunctions by sets of integrity constraints like $\{\bot \leftarrow \neg s \wedge \neg r\}$ or $\{\bot \leftarrow \neg s \wedge \neg r,\ \bot \leftarrow s \wedge r\}$, where the former represents the disjunction $s \vee r$ and the latter the exclusive disjunction $s \oplus r$. Furthermore, we require that models satisfy the integrity constraints. If this does not hold, then abduction (e.g. (Kakas, Kowalski, & Toni, 1992)) is applied and if several (minimal) explanations can be found, then their consequences are combined skeptically. We will demonstrate that this not only solves various human reasoning tasks involving (exclusive) disjunctions but that there is also a striking similarity to the elimination of disjunctions in the *calculus of natural deduction* as defined by Gentzen (1935). But, the WCS avoids some of the drawbacks of natural deduction like the usage of the logical law *ex falso quodlibet* (or *falsum*). The paper demonstrates that logic programming and the WCS can also model human disjunctive reasoning.

The Weak Completion Semantics

We assume the reader to be familiar with classical binary logic (e.g. (van Dalen, 1997)). Let $\top$, $\bot$, and $\cup$ be truth constants denoting *true*, *false*, and *unknown*, respectively. A *literal* is an atom or the negation of an atom. A *(logic) program* is a finite set of clauses of the form $B \leftarrow Body$, where B is an atom and $Body$ is either $\top$, or $\bot$, or a finite, non-empty conjunction of literals. Clauses of the form $B \leftarrow \top$, $B \leftarrow \bot$, and $B \leftarrow L_1 \wedge \ldots \wedge L_n$ are called *facts*, *assumptions*, and *rules*, respectively, where L_i, $1 \leq i \leq n$, are literals.

In this paper, $\mathcal{P}$ denotes a program. An atom B is *defined* in $\mathcal{P}$ if and only if $\mathcal{P}$ contains a clause of the form $B \leftarrow Body$. In the program $\mathcal{P} = \{h \leftarrow s \wedge \neg ab_s,\ ab_s \leftarrow \bot\}$ presented in the introduction the atoms h and ab_s are defined, whereas s is undefined. We restrict our attention to propositional programs although the WCS extends to first-order programs as well (see e.g. (Hölldobler, 2015) and the conclusion).

Consider the following transformation: (1) For all defined atoms B occurring in $\mathcal{P}$, replace all clauses of the form $B \leftarrow Body_1,\ B \leftarrow Body_2,\ \ldots$ by $B \leftarrow Body_1 \vee Body_2 \vee \ldots$. (2) Replace all occurrences of $\leftarrow$ by $\leftrightarrow$. The resulting set of equivalences is called the *weak completion* of $\mathcal{P}$. It differs from the completion defined by Clark (1978) in that undefined atoms are not mapped to false, but to unknown instead.

As shown by Hölldobler and Kencana Ramli (2009), each weakly completed program admits a least model under the three-valued Łukasiewicz (1920) logic (see Table 1). This model will be denoted by $\mathcal{M}_\mathcal{P}$. It can be computed as the least fixed point of a semantic operator introduced by Stenning and van Lambalgen (2008). Let $\mathcal{P}$ be a program and I a three-valued interpretation represented by the pair $\langle I^\top, I^\bot \rangle$, where $I^\top$ and $I^\bot$ are the sets of atoms mapped to true and false by I, respectively, and atoms which are not listed in either set are mapped to unknown by I. We define $\Phi_\mathcal{P} I = \langle J^\top, J^\bot \rangle$,[4] where

$$\begin{array}{rcl} J^\top & = & \{B \mid \text{there is } B \leftarrow Body \in \mathcal{P} \text{ and } I\,Body = \top\}, \\ J^\bot & = & \{B \mid \text{there is } B \leftarrow Body \in \mathcal{P} \text{ and} \\ & & \quad \text{for all } B \leftarrow Body \in \mathcal{P} \text{ we find } I\,Body = \bot\}. \end{array}$$

Following Kakas et al. (1992) we consider *abductive frameworks* $\langle \mathcal{P}, \mathcal{A}_\mathcal{P}, \mathcal{IC}, \models_{wcs} \rangle$, where $\mathcal{P}$ is a program,

$$\begin{array}{rcl} \mathcal{A}_\mathcal{P} & = & \{B \leftarrow \top \mid B \text{ is undefined in } \mathcal{P}\} \\ & \cup & \{B \leftarrow \bot \mid B \text{ is undefined in } \mathcal{P}\} \end{array}$$

is the *set of abducibles*, $\mathcal{IC}$ is a finite set of *integrity constraints* of the form $\bot \leftarrow Body$, where $Body$ is a non-empty and finite conjunction of literals, and $\mathcal{M}_\mathcal{P} \models_{wcs} L$ if and only if $\mathcal{M}_\mathcal{P}$ maps the literal L to true. Let $\mathcal{O}$ be an *observation*, i.e., a finite set of literals. $\mathcal{O}$ is *explainable* in the abductive framework $\langle \mathcal{P}, \mathcal{A}_\mathcal{P}, \mathcal{IC}, \models_{wcs} \rangle$ if and only if there exists a non-empty $\mathcal{X} \subseteq \mathcal{A}_\mathcal{P}$ called an *explanation* such that (1) $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}} \models_{wcs} L$ for all $L \in \mathcal{O}$ and (2) $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}}$ satisfies $\mathcal{IC}$. The literal L *follows credulously* from $\mathcal{P}$ and $\mathcal{O}$ if and only if there exists an explanation $\mathcal{X}$ for $\mathcal{O}$ such that $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}} \models_{wcs} L$. L *follows skeptically* from $\mathcal{P}$ and $\mathcal{O}$ if and only if $\mathcal{O}$ can be explained and for all explanations $\mathcal{X}$ for $\mathcal{O}$ we find $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}} \models_{wcs} L$. One should observe that if an observation $\mathcal{O}$ cannot be explained, then *nothing follows* credulously as well as skeptically. In case of skeptical consequences this is an application of the *Gricean implicature of an existential statement from a universal one* (Grice, 1975): humans normally do not quantify over things which do not exist.

Given premises, general knowledge, and observations, *reasoning in the WCS* is modeled in five steps:

1. Reasoning towards a program $\mathcal{P}$ following Stenning and van Lambalgen (2008).

2. Weakly completing the program.

3. Computing the least model $\mathcal{M}_\mathcal{P}$ of the weak completion of $\mathcal{P}$ under the three-valued Łukasiewicz logic.

4. Reasoning with respect to $\mathcal{M}_\mathcal{P}$.

5. If observations cannot be explained or integrity constraints are violated, then applying skeptical abduction.

[4] Whenever we apply a unary operator like $\Phi_\mathcal{P}$ to an argument like I, then we omit parenthesis and write $\Phi_\mathcal{P} I$ instead.

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

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

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

$\leftarrow$	$\top$	U	$\bot$
$\top$	$\top$	$\top$	$\top$
U	U	$\top$	$\top$
$\bot$	$\bot$	U	$\top$

$\leftrightarrow$	$\top$	U	$\bot$
$\top$	$\top$	U	$\bot$
U	U	$\top$	U
$\bot$	$\bot$	U	$\top$

Table 1: The truth tables for the Łukasiewicz logic. One should observe that $\mathsf{U} \leftarrow \mathsf{U} = \mathsf{U} \leftrightarrow \mathsf{U} = \top$ as shown in the grey cells.

Models and Integrity Constraints

Consider the following scenario: *John is playing basketball. If John is playing basketball, then his shirt is not clean anymore. What follows?* Let b and c denote that *John is playing basketball* and *John's shirt is clean*, respectively. This scenario can be represented by

$$\{b \leftarrow \top,\ \neg c \leftarrow b \wedge \neg ab_b,\ ab_b \leftarrow \bot\},$$

where ab_b is an abnormality predicate which is assumed to be false. Unfortunately, this set of formulas is not a program because $\neg c$ is not an atom. However, by introducing the antonym *dirty* and specifying that *if a shirt is not dirty, then it is clean* we can represent the scenario by the program

$$\{b \leftarrow \top,\ d \leftarrow b \wedge \neg ab_b,\ ab_b \leftarrow \bot,\ c \leftarrow \neg d \wedge \neg ab_d,\ ab_d \leftarrow \bot\},$$

where d denotes that *John's shirt is dirty* and ab_d is another abnormality predicate which is also assumed to be false. Its weak completion is

$$\{b \leftrightarrow \top,\ d \leftrightarrow b \wedge \neg ab_b,\ ab_b \leftrightarrow \bot,\ c \leftrightarrow \neg d \wedge \neg ab_d,\ ab_d \leftrightarrow \bot\}$$

and admits the least model $\langle \{b,d\}, \{c, ab_b, ab_d\} \rangle$ under Łukasiewicz logic. We conclude that *John's shirt is not clean*. But in order to complete the specification we must add the integrity constraint $\bot \leftarrow d \wedge c$ because *a shirt cannot be clean and dirty at the same time*. Luckily, the least model satisfies the integrity constraint. This idea was applied several times by Oliviera da Costa et al. (2017) in modeling human syllogistic reasoning using the WCS.

Although each weakly completed program admits a least model under Łukasiewicz logic, this model does not have to satisfy integrity constraints. As an example consider a scenario where *a judge has ordered Peter not to come near location* ℓ, which can be expressed by the integrity constraints $\bot \leftarrow \ell$. This integrity constraint is satisfied if and only if ℓ is mapped to $\bot$. Now suppose that *somebody has seen Peter at the location* ℓ, which can be expressed by the program $\mathcal{P} = \{\ell \leftarrow \top\}$. Weakly completing $\mathcal{P}$ we obtain $\{\ell \leftrightarrow \top\}$ whose least model $\mathcal{M}_\mathcal{P} = \langle \{\ell\}, \emptyset \rangle$ maps ℓ to $\top$. In this case, the conflict can not be resolved. One should observe that interpretations can be partially ordered with respect to the subset relation $\subseteq$, where for interpretations I and J we define $I = \langle I^\top, I^\bot \rangle \subseteq \langle J^\top, J^\bot \rangle = J$ if and only if $I^\top \subseteq J^\top$ and $I^\bot \subseteq J^\bot$. The partially ordered set of interpretations for the example discussed in this paragraph is shown in Figure 1.

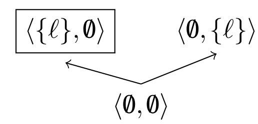

Figure 1: The partially ordered set of interpretations for the weak completion $\{\ell \leftrightarrow \top\}$ of the program $\mathcal{P} = \{\ell \leftarrow \top\}$ and the set $\mathcal{IC} = \{\bot \leftarrow \ell\}$, where ℓ is the only atom and each arrow denotes $\supseteq$. The least model of the weak completion of $\mathcal{P}$ is shown with a black border. All interpretations map the union of $\{\ell \leftrightarrow \top\}$ and $\mathcal{IC}$ to either false or unknown; this set of formulas has no model.

In classical binary logic anything follows from a set of formulas for which there is no model. For example, in the *calculus of natural deduction* there is the *falsum* rule which allows to deduce any formula from a contradiction (Gentzen, 1935). It is quite unlikely that humans do this as well and in the *mental model theory* (Johnson-Laird & Byrne, 1991) nothing follows from an empty set of models. Likewise, in the WCS, programs always have a least model and reasoning is with respect to this model. Furthermore, if skeptical abduction is applied, then there must be at least one explanation to conclude a formula from a program and an observation.

Consider a scenario where it is known that *if Linda is in Amsterdam, then she visits her most favorite club*. Ignoring abnormalities for the moment this can be encoded by the program $\mathcal{P} = \{c \leftarrow a\}$, where c and a denote that *Linda visits her most favorite club* and *Linda is in Amsterdem*, respectively. Weakly completing the program yields $\{c \leftrightarrow a\}$, whose least model $\mathcal{M}_\mathcal{P}$ is $\langle \emptyset, \emptyset \rangle$. This model does not satisfy the integrity constraint $\bot \leftarrow c$ as c is mapped to unknown. But $\bot \leftarrow c$ is satisfied by the non-least model $\langle \emptyset, \{a, c\} \rangle$ of $\{c \leftrightarrow a\}$. This model can be computed. One should observe that a is undefined in $\mathcal{P}$ and, hence, the set $\mathcal{A}_\mathcal{P}$ of abducibles is $\{a \leftarrow \top,\ a \leftarrow \bot\}$. The empty observation $O = \emptyset$ can be explained by the minimal explanation $\mathcal{X} = \{a \leftarrow \bot\}$. Adding $\mathcal{X}$ to the program $\mathcal{P}$ and weakly completing the extended program yields $\{c \leftrightarrow a,\ a \leftrightarrow \bot\}$ whose least model $\mathcal{M}_{\mathcal{P} \cup \mathcal{X}}$ is $\langle \emptyset, \{a, c\} \rangle$ (see Figure 2).

As another example consider a scenario discussed by Khemlani, Byrne, and Johnson-Laird (2018), where it is known that *Lisa is in Cambridge or Ben is in Dublin, or both*. This can be encoded by the integrity constraint $\bot \leftarrow \neg c \wedge \neg d$,

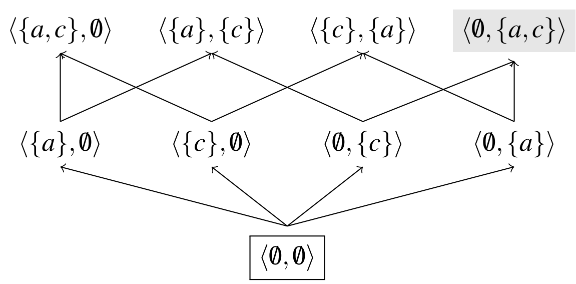

Figure 2: The partially ordered set of interpretations for the weak completion $\{c \leftrightarrow a\}$ of the program $\mathcal{P} = \{c \leftarrow a\}$ and the set $\mathcal{IC} = \{\bot \leftarrow c\}$, where a and c are the only atoms. The least model of the weak completion of $\mathcal{P}$ is shown with a black border. The model for the union of the weak completion of $\mathcal{P} \cup \{a \leftarrow \bot\}$ and $\mathcal{IC}$ is shown in grey.

where c and d denote that *Lisa is in Cambridge* and *Ben is in Dublin*, respectively. This integrity constraint is satisfied if either c is mapped to true, or d is mapped to true, or both. Now suppose that the program $\mathcal{P}$ is empty, i.e., we know nothing about Lisa and Ben except the above-mentioned disjunction. The empty interpretation $\langle\emptyset,\emptyset\rangle$ is the least model of the weak completion of the empty program and maps c and d to unknown. Thus, the least model violates the integrity constraint. In this case, we find two minimal models $\langle\{c\},\emptyset\rangle$ and $\langle\{d\},\emptyset\rangle$ of the weak completion of $\mathcal{P}$ which satisfy the integrity constraint. There are more models satisfying the integrity constraint like $\langle\{c,d\},\emptyset\rangle$ or $\langle\{c\},\{d\}\rangle$, but they are larger than at least one of the two minimal models. The minimal models can be computed by considering the empty observation $\mathcal{O} = \emptyset$. As neither c nor d are defined in the empty program $\mathcal{P}$ we find $\mathcal{A}_\mathcal{P} = \{c \leftarrow \top,\ c \leftarrow \bot,\ d \leftarrow \top,\ d \leftarrow \bot\}$. The sets $\mathcal{X}_c = \{c \leftarrow \top\}$ and $\mathcal{X}_d = \{d \leftarrow \top\}$ are the minimal explanations for $\mathcal{O}$. We obtain $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_c} = \langle\{c\},\emptyset\rangle$ and $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_d} = \langle\{d\},\emptyset\rangle$. They are the minimal models for the weak completion of the program and the integrity constraint (see Figure 3). One should observe that the non-minimal models can be computed by non-minimal explanations for the empty observation.

The models shown in Figure 3 are the three-valued models for the disjunction $c \vee d$. Moreover, if we want to compute logical consequences then it suffices to consider the minimal models and to compute their skeptical consequences.

Finally, if we modify the last example and consider the exclusive disjunction $c \oplus d$, then we obtain Figure 4. The two models $\{c, \neg d\}$ and $\{d, \neg c\}$ can be computed by explaining the empty observation. They are the three-valued models for the exclusive disjunction $c \oplus d$.

Short Summary

If integrity constraints are not satisfied by the least model of the weak completion of a program, then we try to explain the empty observation. In doing so, we may find no explanation, or a single minimal explanation, or several minimal explanations. Each explanation leads to a model of the weak completion of the program and the integrity constraints.

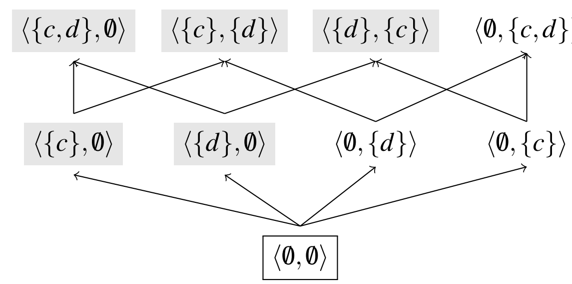

Figure 3: The partially ordered set of interpretations for the weakly completed program $\mathcal{P} = \emptyset$ and the set of integrity constraints $\mathcal{IC} = \{\bot \leftarrow \neg c \wedge \neg d\}$, where c and d are the only atoms. The least model of the weak completion of $\mathcal{P}$ is shown with a black border. Models for the union of the weak completion of $\mathcal{P}$ and $\mathcal{IC}$ are shown in grey.

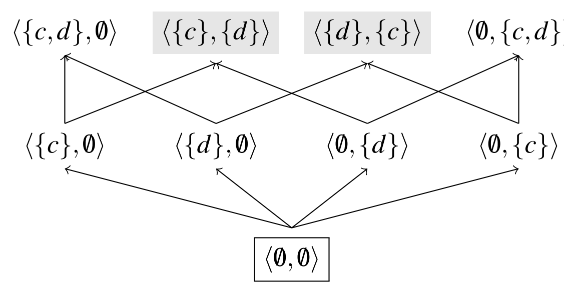

Figure 4: The partially ordered set of interpretations for the weakly completed program $\mathcal{P} = \emptyset$ and the set of integrity constraints $\mathcal{IC} = \{\bot \leftarrow \neg c \wedge \neg d,\ \bot \leftarrow c \wedge d\}$, where c and d are the only atoms. The least model of the weak completion of $\mathcal{P}$ is shown with a black border. Models for the union of the weak completion of $\mathcal{P}$ and $\mathcal{IC}$ are shown in grey.

Disjunctions

In this paper we want to extend the WCS by disjunctions. Let $\mathcal{P}$ be a program and $\mathcal{D}$ a set containing disjunctions and exclusive disjunctions of literals. If $\mathcal{D}$ contains the disjunction $L_1 \vee L_2$, where L_1 and L_2 are literals, then this is encoded by the integrity constraint $\bot \leftarrow \neg L_1 \wedge \neg L_2$, where we asssume that double negations are eliminated, i.e. $\neg\neg A$ is replaced by the semantically equivalent A for each atom A. If $\mathcal{D}$ contains the exclusive disjunction $L_1 \oplus L_2$, then this is encoded by the integrity constraints $\bot \leftarrow \neg L_1 \wedge \neg L_2$ and $\bot \leftarrow L_1 \wedge L_2$.

Example 1

Consider the following scenario discussed by Johnson-Laird, Byrne, and Schaeken (1992): *Lisa is in Cambridge or Ben is in Dublin, or both. Lisa is not in Cambridge. What follows?* Let c and d denote that *Lisa is in Cambridge* and *Ben is in*

Dublin, respectively. The disjunction is encoded by the set $\mathcal{D} = \{c \vee d\}$ and, hence, by $\mathcal{IC} = \{\bot \leftarrow \neg c \wedge \neg d\}$. The negative sentence is represented by the program $\mathcal{P} = \{c \leftarrow \bot\}$. As d is undefined in $\mathcal{P}$ we obtain the set of abducibles $\mathcal{A}_\mathcal{P} = \{d \leftarrow \top,\ d \leftarrow \bot\}$. $\mathcal{P}$ is weakly completed to $\{c \leftrightarrow \bot\}$, whose least model $\mathcal{M}_\mathcal{P}$ is $\langle \emptyset, \{c\}\rangle$. This model does not satisfy $\mathcal{IC}$ as it maps c to false, d to unknown, and $\neg c \wedge \neg d$ to unknown. But the empty observation can be explained by the minimal explanation $\mathcal{X} = \{d \leftarrow \top\}$. We obtain $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}} = \langle \{d\}, \{c\}\rangle$. We do not need to consider any other explanation as $\mathcal{X}$ is the only minimal one. $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}}$ satisfies $\mathcal{IC}$. We conclude that *Ben is in Dublin and Lisa is not in Cambridge*.

Example 2

Consider a variant of the previous example: *Lisa is in Cambridge or Ben is in Dublin, or both. Lisa is in Cambridge. What follows?* This scenario is modeled as Example 1 except that the program $\mathcal{P}$ is $\{c \leftarrow \top\}$. The program is weakly completed to $\{c \leftrightarrow \top\}$, whose least model $\mathcal{M}_\mathcal{P}$ is $\langle \{c\}, \emptyset\rangle$. This model satisfies $\mathcal{IC} = \{\bot \leftarrow \neg c \wedge \neg d\}$ as it maps c to true and $\neg c$ as well as $\neg c \wedge \neg d$ to false. $\mathcal{M}_\mathcal{P}$ satisfies $\mathcal{IC}$ and we conclude that *Lisa is in Cambridge*. But we have no idea where *Ben* is.

Example 3

Consider the following scenario which has also been discussed by Johnson-Laird et al. (1992): *Linda is in Amsterdam or Cathy is in Majorca, but not both. Cathy is not in Majorca. What follows?* Let a and m denote that *Linda is in Amsterdam* and *Cathy is in Majorca*, respectively. The exclusive disjunction is encoded by the set $\mathcal{D} = \{a \oplus m\}$ and, hence, $\mathcal{IC} = \{\bot \leftarrow \neg a \wedge \neg m,\ \bot \leftarrow a \wedge m\}$. The negated sentence is represented by the program $\mathcal{P} = \{m \leftarrow \bot\}$. As a is undefined in $\mathcal{P}$ we obtain the set of abducibles $\mathcal{A}_\mathcal{P} = \{a \leftarrow \top,\ a \leftarrow \bot\}$. Program $\mathcal{P}$ is weakly completed to $\{m \leftrightarrow \bot\}$, whose least model $\mathcal{M}_\mathcal{P}$ is $\langle \emptyset, \{m\}\rangle$. This model does not satisfy the first element of $\mathcal{IC}$ as it maps a, $\neg a$, and $\neg a \wedge \neg m$ to unknown. But the empty observation can be explained by the minimal explanation $\mathcal{X} = \{a \leftarrow \top\}$ and we obtain $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}} = \langle \{a\}, \{m\}\rangle$. $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}}$ satisfies $\mathcal{IC}$ and we conclude *Linda is in Amsterdam and Cathy is not in Majorca*.

Example 4

Consider a variant of the previous example: *Linda is in Amsterdam or Cathy is in Majorca, but not both. Cathy is in Majorca. What follows?* The scenario is modeled as Example 3 except that the program $\mathcal{P}$ is $\{m \leftarrow \top\}$. The program is weakly completed to $\{m \leftrightarrow \top\}$, whose least model $\mathcal{M}_\mathcal{P}$ is $\langle \{m\}, \emptyset\rangle$. This model satisfies the first integrity constraint but violates the second one as it maps a and $a \wedge m$ to unknown. But the empty observation can be explained by the minimal explanation $\mathcal{X} = \{a \leftarrow \bot\}$ and we obtain $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}} = \langle \{m\}, \{a\}\rangle$. $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}}$ satisfies $\mathcal{IC}$ and we conclude *Cathy is in Majorca and Linda is not in Amsterdam*.

Example 5

Consider the following scenario: *It is late in the afternoon. Ella wants to study or to go running, or both. If she is studying, she will be hungry. If she is running, she will be hungry. What follows?* Let s, r, and h denote that *Ella is studying*, *Ella is running*, and *Ella is hungry*, respectively. The disjunction is encoded by $\mathcal{D} = \{s \vee r\}$ and, hence, $\mathcal{IC} = \{\bot \leftarrow \neg s \wedge \neg r\}$. The two conditionals are represented by the program

$$\mathcal{P} = \{h \leftarrow s \wedge \neg ab_s,\ h \leftarrow r \wedge \neg ab_r,\ ab_s \leftarrow \bot,\ ab_r \leftarrow \bot\},$$

where ab_s and ab_r are two abnormality predicates which are assumed to be false. As s and r are undefined in $\mathcal{P}$ we obtain $\mathcal{A}_\mathcal{P} = \{s \leftarrow \top,\ s \leftarrow \bot,\ r \leftarrow \top,\ r \leftarrow \bot\}$. The weak completion of $\mathcal{P}$ is $\{h \leftrightarrow (s \wedge \neg ab_s) \vee (r \wedge \neg ab_r),\ ab_s \leftrightarrow \bot,\ ab_r \leftrightarrow \bot\}$. Its least model $\mathcal{M}_\mathcal{P}$ is $\langle \emptyset, \{ab_s, ab_r\}\rangle$. $\mathcal{M}_\mathcal{P}$ does not satisfy $\mathcal{IC}$ as it maps s and r to unknown. The empty observation can be explained by $\mathcal{X}_s = \{s \leftarrow \top\}$ and $\mathcal{X}_r = \{r \leftarrow \top\}$. We find $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_s} = \langle \{s,h\}, \{ab_s, ab_r\}\rangle$ and $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_r} = \langle \{r,h\}, \{ab_s, ab_r\}\rangle$. Both models satisfy $\mathcal{IC}$. We do not need to consider any other explanation as $\mathcal{X}_s$ and $\mathcal{X}_r$ are the only minimal ones. Comparing $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_s}$ and $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_r}$ we skeptically conclude that *Ella will be hungry*, but we can neither skeptically conclude that *Ella is studying* nor that *Ella is running*.

Example 6

We can modify the previous example by assuming that the disjunction is exclusive. Hence, the integrity constraints are $\{\bot \leftarrow \neg s \wedge \neg r,\ \bot \leftarrow s \wedge r\}$. This will lead to the explanations $\mathcal{X}_1 = \{s \leftarrow \top,\ r \leftarrow \bot\}$ and $\mathcal{X}_2 = \{r \leftarrow \top,\ s \leftarrow \bot\}$ and the models $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_1} = \langle \{s,h\}, \{ab_s, ab_r, r\}\rangle$ and $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_2} = \langle \{r,h\}, \{ab_s, ab_r, s\}\rangle$. Again, we skeptically conclude that *Ella will be hungry*, but we have no idea whether *she was studying* or *she was running*.

Natural Deduction

Considering Example 5, there is a striking similarity to the elimination of a disjunction in the *calculus of natural deduction*. A disjunction like $s \vee r$ can be eliminated and a formula like h can be derived if (i) h can be derived assuming s and (ii) h can be derived assuming r. (i) and (ii) hold in the given scenario: Assuming s and knowing that $\neg ab_s$ holds, we can derive $s \wedge \neg ab_s$ by the *introduction rule for conjunction*; knowing that $h \leftarrow s \wedge \neg ab_s$ holds we can derive h by the *elimination rule for implication* (see the subproof shown in light gray in Figure 5). Likewise, we can derive h assuming r by utilizing that $\neg ab_r$ and $h \leftarrow r \wedge \neg ab_r$ hold (see the subproof shown in dark gray in Figure 5). The subproof shown in light gray corresponds to the computation of $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_s}$ by iterating the $\Phi_\mathcal{P}$ operator as detailed in the introduction, whereas the subproof shown in dark grey corresponds to the computation of $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_r}$. The final application of the *elimination rule for implication* corresponds to the computation of skeptical conclusions given $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_s}$ and $\mathcal{M}_{\mathcal{P}\cup\mathcal{X}_r}$.

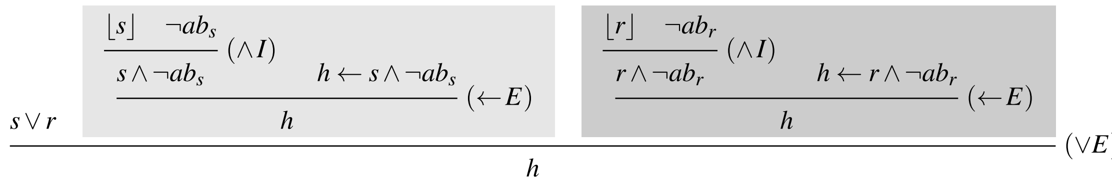

Figure 5: A natural deduction proof of $\{s \vee r,\ h \leftarrow (s \wedge \neg ab_s),\ h \leftarrow (r \wedge \neg ab_r),\ \neg ab_s,\ \neg ab_r\} \vdash h$ in the notation of Hölldobler (2009), where $(\wedge I)$, $(\leftarrow E)$, and $(\vee E)$ denote the rules *introduction of conjunction*, *elimination of implication*, and *elimination of disjunction*, respectively. $\lfloor\ \rfloor$ denotes that the enclosed hypothesis has been cancelled.

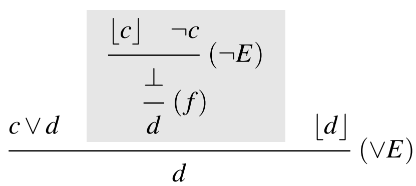

Figure 6: A natural deduction proof of $\{c \vee d, \neg c\} \vdash d$, where $(\neg E)$ and (f) denote the rules *elimination of negation* and *falsum*, respectively.

A similar observation can be made for the other examples discussed previously. E.g., Figure 6 depicts a natural deduction proof of $\{c \vee d, \neg c\} \vdash d$ corresponding to Example 1. In this proof the *falsum* rule is applied. This rule is problematic in human reasoning as well as in proof search irrespective of whether it is applied top-down or bottom-up. Such an application – and, in fact, the whole subproof shown in light grey – is avoided in the search for explanations for the empty observation in the WCS, where we simply need to consider the possible explanations $\{d \leftarrow \top\}$ and $\{d \leftarrow \bot\}$. This is because we know that c is defined in the program $\mathcal{P} = \{c \leftarrow \bot\}$ and is mapped to $\bot$ by the $\Phi_\mathcal{P}$ operator.

Conclusion

In this paper we have extended the WCS by disjunctions. Although we have only discussed examples with a single disjunction, the approach can handle finite sets of disjunctions. Each disjunction is represented by integrity constraints and skeptical abduction is applied to satisfy them. The extension is implemented in PYTHON and ASP. We are unaware of any benchmark sets for disjunctions, but have tested the discussed and related examples.

In a three-valued logic, exclusive disjunctions can also be used to assign true or false to an unknown atom. As an example consider the following puzzle which was published in the German weekly newspaper *DIE ZEIT* on July 26, 2020: *Antonia is looking at Berta while Berta is looking at Cleopatra. Antonia is wearing a red hat, Cleopatra is not wearing a hat, and it is unknown whether Berta is wearing a red hat. Is a person with a red hat looking at a person without a red hat?*

This puzzle can be solved by the techniques presented in this paper. Let a, b, and c denote *Antonia*, *Berta*, and *Cleopatra*, respectively, rX that X *is wearing a red hat*, $\ell(X,Y)$ that X *is looking at* Y and *goal* that *somebody with a red hat is looking at a person without a red hat*. The scenario can be formalized as a program $\mathcal{P}$ consisting of the following clauses:

$$\begin{array}{rcl} \ell(a,b) & \leftarrow & \top, \\ \ell(b,c) & \leftarrow & \top, \\ ra & \leftarrow & \top, \\ rc & \leftarrow & \bot, \\ goal & \leftarrow & \ell(X,Y) \wedge rX \wedge \neg rY. \end{array}$$

This is a first-order program. It can be turned into a propositional one by replacing the variables X and Y consistently by (all combinations of) the constants a, b, and c. The weak completion of this program admits the least model

$$\mathcal{M}_\mathcal{P} = \langle \{\ell(a,b), \ell(b,c), ra\}, \{rc\} \rangle.$$

The fact that it is *unknown whether Berta is wearing a red hat* can be represented by the the exclusive disjunction $rb \oplus \neg rb$ and, hence, by the set $\mathcal{IC} = \{\bot \leftarrow rb \wedge \neg rb\}$ of integrity constraints. This set is not satisfied by $\mathcal{M}_\mathcal{P}$ because $\mathcal{M}_\mathcal{P}$ maps rb to unknown. As rb is undefined in $\mathcal{P}$, the set $\mathcal{A}_\mathcal{P}$ of abducibles contains $rb \leftarrow \top$ and $rb \leftarrow \bot$. The empty observation can be explained by either $\mathcal{X}_1 = \{rb \leftarrow \top\}$ or $\mathcal{X}_2 = \{rb \leftarrow \bot\}$. We obtain

$$\begin{array}{rcl} \mathcal{M}_{\mathcal{P}\cup\mathcal{X}_1} & = & \langle \{\ell(a,b), \ell(b,c), ra, rb, goal\}, \{rc\} \rangle, \\ \mathcal{M}_{\mathcal{P}\cup\mathcal{X}_2} & = & \langle \{\ell(a,b), \ell(b,c), ra, goal\}, \{rc, rb\} \rangle. \end{array}$$

Reasoning skeptically we conclude that *goal* is true: *there is a person with a red hat looking at a person without a red hat*. If rb holds, then *Berta* is the person in question; if rb does not hold, then it is *Antonia*. One should observe that reasoning skeptically we can neither conclude that *Berta is wearing a red hat* nor that *Berta is not wearing a red hat*.

We have not yet considered nested disjunctions and disjunctive illusory inferences as discussed by Khemlani and Johnson-Laird (2009). This is one of the next goals in the development of the WCS.

Acknowledgement We like to thank Marco Ragni for many valuable comments and suggestions.

References

Clark, K. L. (1978). Negation as failure. In H. Gallaire & J. Minker (Eds.), *Logic and databases* (p. 293-322). New York: Plenum.

Craik, K. J. W. (1945). *The nature of explanation*. Cambridge: Cambridge University Press.

Cramer, M., Hölldobler, S., & Ragni, M. (2021). When are humans reasoning with modus tollens? In *Proceedings of the 43rd annual conference of the cognitive science society.* (to appear)

Dietz, E.-A., Hölldobler, S., & Ragni, M. (2012). A computational logic approach to the suppression task. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (p. 1500-1505). Cognitive Science Society.

Dietz Saldanha, E.-A., Hölldobler, S., Kencana Ramli, C. D. P., & Palacios Medinacelli, L. (2018). A Core method for the weak completion semantics with skeptical abduction. *Journal of Artificial Intelligence Research*, *63*, 51-86.

Fitting, M. (1985). A Kripke–Kleene semantics for logic programs. *Journal of Logic Programming*, *2*(4), 295–312.

Gentzen, G. (1935). Untersuchungen über das logische Schließen. *Mathematische Zeitschrift*, *39*, 176-210 und 405-431.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3, p. 41-58). Academic Press, New York.

Hölldobler, S. (2009). *Logik und Logikprogrammierung* (Vol. 1: Grundlagen). Heidelberg: Synchron Publishers GmbH.

Hölldobler, S. (2015). Weak completion semantics and its applications in human reasoning. In U. Furbach & C. Schon (Eds.), *Bridging 2015 – bridging the gap between human and automated reasoning* (Vol. 1412, p. 2-16). CEUR-WS.org. (http://ceur-ws.org/Vol-1412/)

Hölldobler, S., & Kalinke, Y. (1994). Towards a new massively parallel computational model for logic programming. In *Proceedings of the ECAI94 workshop on combining symbolic and connectionist processing* (p. 68-77).

Hölldobler, S., & Kencana Ramli, C. D. P. (2009). Logic programs under three-valued Łukasiewicz's semantics. In P. M. Hill & D. S. Warren (Eds.), *Logic programming* (Vol. 5649, p. 464-478). Springer-Verlag Berlin Heidelberg.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge: Cambridge University Press.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove and London (UK): Lawrence Erlbaum Associates.

Johnson-Laird, P. N., Byrne, R. M. J., & Schaeken, W. (1992). Propositional reasoning by model. *Psychological Review*, *99*(3), 418-439.

Kakas, A. C., Kowalski, R. A., & Toni, F. (1992). Abductive Logic Programming. *Journal of Logic and Computation*, *2(6)*, 719-770.

Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentenial reaoning. *Cognitive Science*, 1-38.

Khemlani, S. S., & Johnson-Laird, P. N. (2009). Disjunctive illusory inferences and how to eliminate them. *Memory & Cognition*, *37*, 615-623.

López-Astorga, M. (2015). The formal descipline theory and mental logic. *Praxis Filosófica*, *41*, 11-25.

Łukasiewicz, J. (1920). O logice trójwartościowej. *Ruch Filozoficzny*, *5*, 169-171. (English translation: On Three-Valued Logic. In: *Jan Łukasiewicz Selected Works*. (L. Borkowski, ed.), North Holland, 87-88, 1990.)

Muggleton, S. (1992). *Inductive logic programming*. London: Academic Press.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, *71*, 12.1-12.26.

Oliviera da Costa, A., Dietz Saldanha, E.-A., Hölldobler, S., & Ragni, M. (2017). A computational logic approach to human syllogistic reasoning. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th annual conference of the cognitive science society* (p. 883-888). Austin, TX: Cognitive Science Society.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. MIT Press.

Sebelik, J., & Stepanek, P. (1982). Horn clause programs for recursive functions. In K. Clark & S.-Å. Tärnlund (Eds.), *Logic programming* (p. 324-340). New York: Academic Press.

Stenning, K., & van Lambalgen, M. (2008). *Human reasoning and cognitive science*. MIT Press.

van Dalen, D. (1997). *Logic and structure* (3rd ed.). Berlin: Springer-Verlag.