Technische Universität Dresden
Informatik Fakultät
September 29, 2010

Human Reasoning and Abduction

Bachelor Thesis

Tobias Philipp
Autor

3413078
Matrikelnummer

15. Juni bis 29. September 2010
Bearbeitungszeit

Erklärung

Hiermit versichere ich, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe, dass alle Stellen der Arbeit, die wörtlich oder sinngem aus anderen Quellen bernommen wurden, als solche kenntlich gemacht sind und dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prfungsbehrde vorgelegt wurde.

Dresden, den 29. September 2010

Contents

Chapter 1

Introduction

1.1 Motivation

In Artificial Intelligence and Cognitive Science, it is often argued that classical logic is inadequate to model human reasoning [HR09]. By several examples, Byrne has confirmed this observation [Byr89]: She asked individuals to draw conclusions from several situations. One example is:

> *If Marian has an essay to write, she will study late in the library.*
>
> *If the library stays open, she will study late in the library.*
>
> *She has an essay to write.*

Here, 38% of subjects conclude, that Marian will study late in the library. It is clear that this behavior not consistent with respect to classical logic when using a direct translation from the natural sentences into logical formulae. Byrne argues that mental model theories explain her data.

Contrary to this view, Stenning and van Lambalgen proposed to use a Computational Logic approach [SvL08]: First, they do not use a direct translation from conditionals into logical implications. Instead, they propose to formulate conditionals as "If A and nothing abnormal occurs, then B". Then they reason with respect to the minimal models of a form of completion under the Kleene strong three valued logic with complete equivalence. They claim that humans reason consistently in this logic. Additionally, they also provide an immediate consequence operator and give a translation from logic programs to neural networks.

However, not all their claims in [SvL08] are correct, since the least fixed point of their operator is not always a model of the program, as shown in [HR09]. In the area of human reasoning, Hölldobler and Kencana Ramli suggested to reason with respect to the least model of the weakly completed program under the Łukasiewicz logic. Here, the weak completion is a form of Clark's completion [Cla78], where roughly speaking, the if-halves are replaced by "if and only if".

The difference to the completion is that no knowledge about facts is added. A three valued model I is called smaller than J, if every atom mapped to true (false) under I, is also mapped to true (false) under J. Then, a model I is called minimal, if there is no smaller one and least, if every other model J is greater than I.

Kencana Ramli has shown that the Łukasiewicz logic under the weak completion is adequate to model human reasoning [Ram09]. More precisely, she has proven that six of twelve examples by Byrne can be modeled in this logic. Consider one of the remaining examples by Byrne:

If Marian has an essay to write, she will study late in the library.

She will study late in the library.

Here, 71% of subjects conclude that she has an essay to write. Using Stenning and van Lambalgen's translation, we have the following logic program:

$$\mathcal{P} = \{l \leftarrow e \wedge \neg ab, ab \leftarrow \bot, l \leftarrow \top\}$$

The weakly completed program is as follows:

$$wc(P) = \{l \leftrightarrow (e \wedge \neg ab) \vee \top, ab \leftrightarrow \bot\}$$

Then, the least Łukasiewicz model is: $I_L = \langle \{l\}, \{ab\} \rangle$. This interpretation maps the atom l to true ($\top$), ab to false ($\bot$) and each other atom to the truth value unknown (u). However, this is incorrect with respect to Byrne's data, since $I_L \not\models e$.

This result is not surprising since we are not interested in deriving a logical consequence using deduction, but rather in inferring a reason why Marian is studying late in the library. That is abductive and not deductive reasoning [KKT98]. Stenning and van Lambalgen presented a systems on an implementation view that is adequate to model the abductive examples.

In this thesis, different abductive frameworks [KKT98] are investigated and it will be shown that the Łukasiewicz logic and the weak completion is indeed adequate to model the remaining six examples in an abductive framework. The three different logics of Kleene, Fitting and Łukasiewicz with different forms of completion are compared. In [KKT98], abductive frameworks were extended by so called integrity constraints. Since the representation of the examples do not use integrity constraints, four examples are modified in order to illustrate the use of integrity constraints. Two different semantics of integrity constraints are presented and contrasted. Moreover, the use of abductive frameworks are a natural extension that also captures the six deductive examples.

At the end, the following questions are investigated from a complexity point of view:

- Consistency: The question, whether an abductive problem has a minimal explanation,
- Relevance: The question whether a fact is part of a minimal explanation,

- Necessity: The question whether a fact is part of all minimal explanations and
- Skeptical Reasoning: The question whether a formula is a logical consequence with respect to all minimal explanations.

1.2 Thesis Structure

This thesis is structured as follows:

Chapter 2 In this chapter, logic programs and different semantics of logic programs are formally introduced. Then, some complexity classes and relationships between them are shown.

Chapter 3 Here, the general idea of abduction is shown and the formal definition is given. Then, it is shown, how one can model the remaining six examples by Byrne in an abductive framework. Finally, four examples are given that illustrate two semantics of integrity constraints.

Chapter 4 In this chapter, the relationship between three different logics and different forms of completion are shown.

Chapter 5 Chapter 5 deals with the complexity of consistency, relevance, necessity and skeptical reasoning.

Chapter 6 We summarize the work. Additionally, ideas for future work are given.

Chapter 2

Preliminaries

In this chapter, propositional logic and logic programs with negative facts are formally introduced. We present three different logics, the Łukasiewicz three valued logic [Luk70], the Kleene strong three valued logic and the Fitting logic [Fit85]. The following sections describe Clark's completion and a slightly different form of Clark's completion which is called weak completion. Then, the Stenning and van Lambalgen's immediate consequence operator is presented and the relationship between the least model of a weakly completed program and their consequence operator is given. At the end some results from complexity theory are summarized.

2.1 Propositional Language

We consider an *alphabet* consisting of an infinite set of variables, the connectives $\neg, \wedge, \vee, \leftarrow, \leftrightarrow$ and the punctuation symbols "(", "," and ")". In addition, the alphabet also contains the special symbols $\top$, $\bot$, denoting a valid and an unsatisfiable formula, respectively.

Definition 2.1.1 (Formula)**.** Let $\mathcal{R}$ be a set of propositional variables. The set of *propositional formulas* is the smallest set $\mathcal{L}(\mathcal{R})$ of strings over $\mathcal{R}$, the binary connectives $\{\wedge, \vee, \leftarrow, \leftrightarrow\}$ and the special symbols "(", ",", ")" with the following properties:

- If $p \in \mathcal{R}$, then $p \in \mathcal{L}(\mathcal{R})$.
- If $F \in \mathcal{L}(\mathcal{R})$, then $\neg F \in \mathcal{L}(\mathcal{R})$.
- If $\circ$ is a binary connective and $F, G \in \mathcal{L}(\mathcal{R})$, then $(F \circ G) \in \mathcal{L}(\mathcal{R})$.

A *literal* is either p or $\neg p$, if $p \in \mathcal{R}$. In the first case, it is called *positive* and in the second case *negaitve*.

2.2 Syntax of Logic Programs

Stenning and van Lambalgen proposed to use a specific form of formulas to represent knowledge in order to compute logical consequences efficiently.

Definition 2.2.1 (Logic Program)**.** A *logic program* $\mathcal{P}$ over the set of variables $\mathcal{R}$ is a finite set of clauses, positive and negative facts. *Clauses* are logical formulae of the form

$$H \leftarrow (\neg)B_1 \wedge (\neg)B_2 \wedge \ldots \wedge (\neg)B_n \quad n \geq 1, H, B_1, \ldots B_n \in \mathcal{R}$$

Positive facts are of the form

$$H \leftarrow \top$$

and *negative facts* of the form

$$H \leftarrow \bot$$

.

We say, H is the *head* of a clause, $B_1 \wedge \ldots \wedge B_n$ is the *body* of the clause. If H is the head of a clause in $\mathcal{P}$, we say H *is defined w.r.t.* $\mathcal{P}$. If H is not defined w.r.t. $\mathcal{P}$, it is *undefined* w.r.t. $\mathcal{P}$.

2.3 Semantics of Logic Programs

In classical logic, semantics of formulae are given by a two valued interpretation. Here, we consider three-valued logics, i.e. there exists a third, intermediate truth value u.

The semantics of logic programs or arbitrary formulae are given by a three valued interpretation.

Definition 2.3.1 (Three-Valued Interpretation)**.** A *three-valued interpretation* I is a mapping from the set of variables into the truth values $\top, \bot, u$. We represent interpretations by pairs $\langle I^\top, I^\bot \rangle$, where

- the set $I^\top$ contains all variables which are mapped to $\top$,
- the set $I^\bot$ contains all variables that are mapped to $\bot$,
- all variables, that are neither in $I^\top$ nor $I^\bot$ are mapped to u and
- we require $I^\top \cap I^\bot = \emptyset$.

The extension of I to arbitrary formulas can be obtained by Figure 2.1 in the usual way. With I_L we denote the extension obtained by using the connectives of the Łukasiewicz logic $\{\neg, \wedge, \vee, \leftarrow_L, \leftrightarrow_L\}$, I_K denotes the Kleene strong three valued logic with the connectives $\{\neg, \wedge, \vee, \leftarrow_K, \leftrightarrow_K\}$ and I_F denotes the Fitting semantics with the connectives $\{\neg, \wedge, \vee, \leftarrow_K, \leftrightarrow_C\}$.

We say, a formula F is Łukasiewicz (Kleene, Fitting) consistent iff there exists an I such that $I_L(F) = \top$ ($I_K(F) = \top, I_F(F) = \top$).

Note that a total three-valued interpretation can be seen as a two-valued interpretation.

	$\neg$
$\top$	$\bot$
$\bot$	$\top$
u	u

$*$

F	G	$\wedge$	$\vee$	$\leftarrow_L$	$\leftrightarrow_L$	$\leftarrow_K$	$\leftrightarrow_K$	$\leftrightarrow_C$
$\top$	$\top$	$\top$	$\top$	$\top$	$\top$	$\top$	$\top$	$\top$
$\top$	$\bot$	$\bot$	$\top$	$\top$	$\bot$	$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$	$\top$	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$
$\bot$	$\bot$	$\bot$	$\bot$	$\top$	$\top$	$\top$	$\top$	$\top$
$\top$	u	u	$\top$	$\top$	u	$\top$	u	$\bot$
$\bot$	u	$\bot$	u	u	u	u	u	$\bot$
u	$\top$	u	$\top$	u	u	u	u	$\bot$
u	$\bot$	$\bot$	u	$\top$	u	$\top$	u	$\bot$
u	u	u	u	$\top$	$\top$	u	u	$\top$

Figure 2.1: Truth values for Łukasiewicz logic, Kleene strong three valued logic and Fitting semantics

Definition 2.3.2 (Model)**.** Let I be an interpretation and F a formula. Then an extension I_* is called *model* for F iff $I_*(F) = \top$.

For a set of formulas $\mathcal{F}$, I_* is a model of $\mathcal{F}$, denoted by $I_* \models \mathcal{F}$, iff $I_* \models F$ for all $F \in \mathcal{F}$

Furthermore, we give a partial order on three valued interpretations.

Definition 2.3.3 (Ordering Among Interpretations)**.** Let I, J be two interpretations. If $I^\top \subseteq J^\top$ and $I^\bot \subseteq J^\bot$, we say I *is smaller than* J, denoted with $I \preceq J$.

In particular, we are interested in the least model.

Definition 2.3.4 (Least Interpretation)**.** Let $\mathcal{I}$ be a set of interpretations, then $I \in \mathcal{I}$ is called *least* in $\mathcal{I}$ iff for all interpretation $J \in \mathcal{I}$ it holds that $I \preceq J$.

We denote the least model of a set of formulae with $lm(\mathcal{F})$, if it exists. In particular, we use $lm_L(\mathcal{F})$ to denote the least Łukasiewicz model, if it exists. Then, a least model is also unique.

2.4 Program Completion

Stenning and van Lambalgen proposed to use a form of completion. Originally, completion was proposed by Clark as a semantics for negation by finite failure, stating that $\neg a$ is true, if a could not be inferred in finite time.

Definition 2.4.1 (Completion)**.** Let $\mathcal{P}$ be a logic program. Consider the following transformation:

1. Select all clauses $H \leftarrow B_1, \ldots, H \leftarrow B_n$ with the same head and replace them by the formula $H \leftarrow B_1 \vee \ldots \vee B_n$.

2. For every undefined predicate A, add $A \leftarrow \bot$.

3. Replace $\leftarrow$ by $\leftrightarrow$.

The result, obtained by this transformation, is called *completion* of $\mathcal{P}$ and is denoted with $c(\mathcal{P})$.

Note that in logic programs negative facts are allowed. For this reason, completion is unsuitable for such logic programs, because during completion these negative facts are added. If we drop the second step, we obtain the weak completion.

Definition 2.4.2 (Weak Completion)**.** Let $\mathcal{P}$ be a logic program. Consider the following transformation:

1. Select all clauses $H \leftarrow B_1, \ldots, H \leftarrow B_n$ with the same head and replace them by the formula $H \leftarrow B_1 \vee \ldots \vee B_n$.

2. Replace $\leftarrow$ by $\leftrightarrow$.

The result, obtained by this transformation, is called *weak completion* of $\mathcal{P}$ and is denoted with $wc(\mathcal{P})$.

Example 2.4.3 (Differences between Completion and Weak Completion)**.** Consider the following program:

$$\begin{aligned}
\mathcal{P} &= \{A \leftarrow \neg B_1 \wedge B_2, \\
&\quad\;\; A \leftarrow \top\} \\
c(\mathcal{P}) &= \{A \leftrightarrow (\neg B_1 \wedge B_2) \vee \top, \\
&\quad\;\; B_1 \leftrightarrow \bot, \\
&\quad\;\; B_2 \leftrightarrow \bot\} \\
wc(\mathcal{P}) &= \{A \leftrightarrow (\neg B_1 \wedge B_2) \vee \top\}
\end{aligned}$$

2.5 Consequence Operator

A *consequence operator* is a mapping from an interpretation I and a logic program $\mathcal{P}$ to an interpretation J. This operator expresses the consequences when the bodies are interpreted under I. Stenning and van Lambalgen proposed to use a slightly modified operator of Fitting.

Definition 2.5.1 (Stenning and van Lambalgen Immediate Consequence Operator)**.** Let I be an interpretation and $\mathcal{P}$ a logic program. $\Phi_{SvL,\mathcal{P}} = \langle J^\top, J^\bot \rangle$, where

$$\begin{aligned}
J^\top &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \text{ with } I(\mathit{Body}) = \top\} \\
J^\bot &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \\
&\qquad\quad \text{and for all } A \leftarrow \mathit{Body} \in \mathcal{P} \text{ we find that } I(\mathit{Body}) = \bot\}
\end{aligned}$$

.

From particular interests are least fixed points of this operator.

Definition 2.5.2 (Fixed Point)**.** Let $f : S \mapsto S$ be a function. Then, $x \in S$ is called *fixed point* iff $f(x) = x$.

Proposition 2.5.3 (Computation of the Least Fixed Point)**.** The least fixed point can be obtained by iterating the Stenning and van Lambalgen consequence operator, starting with $\langle\emptyset, \emptyset\rangle$.

Proof. See [Ram09], page 52 □

This least fixed point is of particular interest since it is the least model of the weakly completed program under the Łukasiewicz logic.

Proposition 2.5.4 (Relationship Least Fixed Points and Least Model)**.** Let $\mathcal{P}$ be a logic program. Then the following are equivalent:

(i) I is the least model of $wc(\mathcal{P})$,

(ii) I is the least fixed point of $\Phi_{SvL,\mathcal{P}}$.

Proof. See [Ram09] Corollary 6.16 □

Thus, in order to compute the least model of a weakly completed program, we can instead compute the least fixed point. Furthermore, we know that there always exists a least model of a weakly completed program under the Łukasiewicz logic.

2.6 Complexity Theory

In general, complexity theory deals with the *word-problem*, i.e. "Belongs a word w to the language L?". Here, a word is a finite string over the alphabet Σ and a language is a possibly infinite set of words over Σ. With Σ^* we denote every word over Σ. The *complement* of a language L over the alphabet Σ is $\Sigma^* \setminus L$, denoted with $\bar{L}$.

A *decision problem* is a problem, whose answer is "Yes" or "No". There is a natural correspondence between the word-problem and decision problems. Roughly speaking, the corresponding language to a decision problem contains all problems with a "Yes"-answer. Given a language L, the corresponding decision problem is, if a word belongs to L.

Decision problems that can be decided by a deterministic Turing Machine in polynomial time belong to the class P. The class of decision problems that are decidable by a non deterministic Turing Machine in polynomial time is denoted with NP. Intuitively, this class contains all decision problems, where a proof stating that the problem is a "Yes"-instance can be verified in polynomial time on a deterministic Turing Machine. More formally:

Definition 2.6.1 (Balanced Relation)**.** Let R be a binary relation on strings. R is called *balanced* if $(x, y) \in R$ implies $\mid y \mid \leq \mid x \mid^k$ for some $k \geq 1$.

Proposition 2.6.2 (Characterization of NP). Let $L \subseteq \Sigma^*$ be a language. $L \in$ NP iff there is a polynomially decidable and a polynomial balanced relation R such that $L = \{x \mid (x, y) \in R \text{ for some } y\}$.

Proof. See [Pap93] Proposition 9.1 □

The opposite of the class NP are decision problems, where a proof stating that a problem is a "No"-instance can be verified in polynomial time. More formally:

Definition 2.6.3 (coNP). The class coNP is defined as follows:

$$\text{coNP} = \{\bar{L} \mid L \in NP\}$$

Note, that coNP is not the complement of NP. In fact, they have a non empty intersection (See [Pap93] page 238).

Of particular interest are the hardest problems in a complexity class. For this purpose, a transformation from one decision problem to another one is given.

Definition 2.6.4 (Reductions). A language L is *polynomial-time reducible* to a language L', denoted by $L \leq_p L'$ if there is a polynomial-time computable function $f : \Sigma^* \mapsto \Sigma^*$ such that for every $x \in \Sigma^*$, $x \in L$ iff $f(x) \in L'$.

Here, $L \leq_p L'$ intuitively means that L is not more difficult to decide than L'. Moreover, reductions are transitive:

Proposition 2.6.5 (Transitivity of Reductions). Let L_1, L_2, L_3 be languages. If $L_1 \leq_p L_2$ and $L_2 \leq_p L_3$, then $L_1 \leq_p L_3$.

Proof. See [Pap93] Proposition 8.2 □

We say, a problem L is NP-*hard*, if every problem in NP can be polynomial reduced to L. In this sense, L is one of the hardest problems with respect to NP-problems.

Definition 2.6.6 (Hardness and Completeness). Let $\mathbf{C}$ be a complexity class. A language L is $\mathbf{C}$-*hard* if $L \leq_p L'$ for all $L' \in \mathbf{C}$. L is $\mathbf{C}$-*complete* if L is in $\mathbf{C}$ and L is $\mathbf{C}$-*hard*.

That is, an NP-*complete* language is one of the hardest problem in NP. Such languages do exist.

Theorem 2.6.7 (Cook-Levin Theorem). Let 3SAT be the language of all satisfiable formulae in conjunctive normal form where each clause has three disjuncts. Then 3SAT is NP-*complete*.

In order to show that a decision problem is NP-*complete*, one have to show that it is inside NP and is NP-*hard*. Cook and Levin have shown that every decision problem in NP can be reduced to 3SAT. By Proposition 2.6.5, it is sufficient to show that 3SAT is polynomial reducible to the problem L, to show that L is NP-*hard*. Furthermore, to show that a language L is in NP, one can use such reductions.

Proposition 2.6.8 (Closure of NP and coNP)**.** Let L_1, L_2 languages. If $L_2 \in$ NP and $L_1 \leq_p L_2$, then L_1 is in NP.

Proof. See [Pap93] Proposition 8.3 □

That is, the classes NP and coNP are *closed under polynomial-time reductions.*

If we already know that a language is NP-*complete* and it is polynomial reducible to another language and vice versa, we can conclude that the other language is also NP-*complete*.

Definition 2.6.9 (Equivalence)**.** Let L_1, L_2 be languages. If $L_1 \leq_p L_2$ and $L_2 \leq L_1$, then L_1 and L_2 are *equivalent with respect to polynomial time reductions.*

Then, we obtain the following:

Proposition 2.6.10 (Relationship Completeness and Equivalence)**.** Let L, L' be languages. If L is NP-*complete* (coNP-*complete*) and L, L' are equivalent with respect to polynomial time reductions, then L' is NP-*complete* (coNP-*complete*).

Proof. Since $L' \leq_p L$ and $L \in$ NP (coNP), we know by Proposition 2.6.8 that $L' \in$ NP (coNP). Because L is NP-*hard* (coNP-*hard*) and $L \leq_p L'$, it must follow that L' is NP-*hard* (coNP-*hard*). Thus, L' is NP-*complete* (coNP-*complete*). □

There is a symmetry between NP and coNP:

Proposition 2.6.11 (Relationship NP-*completeness* and coNP-*completeness*)**.** Let $L \subseteq \Sigma^*$ be a language. Then, L is NP-*complete* iff $\bar{L}$ is coNP-*complete*.

Proof. See [Pap93] Proposition 10.1 □

Consider the following problem: Let F and G be two formulas in conjunctive normal form. Is it true, that F is satisfiable and G not? It is clear that the first problem is in NP and the second in coNP. Such combined problems are then in the class DP. More formally:

Definition 2.6.12 (DP)**.** A language L is in the class DP iff there are two languages $L_1 \in$ NP and $L_2 \in$ coNP such that $L = L_1 \cap L_2$.

In this thesis, the complexity of several decision problems are investigated. However, we will not give a precise definition of Turing Machines. Instead, algorithms are described in a more intuitive way.

Chapter 3

Abduction

In this chapter, the general idea of abductive reasoning is described first. We continue with the approach by Stenning and van Lambalgen and the formal definition of an abductive framework, following [KKT98]. The next section presents Byrne's experiment. It is argued that the weakly completed programs under the Łukasiewicz logic is indeed adequate to model the remaining six examples by Byrne. In [KKT98] abductive frameworks extended with formulae called integrity constraints were shown . However, one can model the examples by Byrne without integrity constraints. At the end, four examples are modified to show how integrity constraints change the obtained explanations.

3.1 Idea of Abduction

The philosopher Pierce first introduced the notion of abduction. He identified three distinguished forms of reasoning [Pie]:

Deduction This is an analytic process based on the application of general rules to particular situations. Consider the following example: "All Greeks are men, all men are mortal". Then, we can infer "All Greeks are mortal".

Induction That is synthetic reasoning which infers the rule from the case and the result. Consider you observed "Socrates is a human being and is mortal". Then, one can infer "All human being are mortal".

Abduction This is another form of synthetic inference, but of the case from a rule and a result.

Let us illustrate the idea of abduction by the following example.

Example 3.1.1 (Idea of Abduction)**.** The background theory consists of the following rules:

The grass is wet, if it rained last night.

The grass is wet, if the sprinkler was on.

The shoes are wet, if the grass is wet.

If we observe that the shoes are wet, we have several explanations. One explanation is that the grass is wet. However, we can further explain, why the grass is wet. Then, we obtain the explanation "It rained last night." and the explanation "The sprinkler was on.". The last two explanations are called *basic*, since we cannot further explain them. Another explanation is "The sprinkler was on and it rained last night", which combines the last two explanations. Then, this explanation is not *minimal*. One meaningless explanation is "The shoes are wet.", i.e. the observation is explained by itself.

3.2 Approach by Stenning and van Lambalgen

Stenning and van Lambalgen describe a system that allows abductive reasoning in [SvL08]. We illustrate their system using an example:

Example 3.2.1. Consider the program

$$\mathcal{P} = \{H_1 \leftarrow A_1 \wedge A_2, H_2 \leftarrow B_1, H_2 \leftarrow B_2\}$$

If we observe H_1, we could add $H_1 \leftarrow \top$ to the program. However, if we reason with respect to the weak completion of $\mathcal{P} \cup \{H_1 \leftarrow \top\}$, we obtain a formula $H_1 \leftrightarrow (A_1 \wedge A_2) \vee \top$. From this point, we cannot conclude that A_1 and A_2 hold under the Fitting logic. This is the reason, why they proposed to use "integrity constraints". An integrity constraint is of the form "If ϕ holds, then φ must hold." . In the above program we have two such constraints:

$$\begin{aligned} &\text{If } H_1, \text{ then } A_1 \wedge A_2 \\ &\text{If } H_2, \text{ then } B_1 \vee B_2 \end{aligned}$$

If we add the fact $H_1 \leftarrow \top$ to the program, then the least fixed point of their proposed operator is $I = lfp(\Phi_{SvL,\mathcal{P}\cup\{H_1\leftarrow\top\}}) = \langle\{H_1\}, \emptyset\rangle$. Thus the first constraint is not satisfied in the sense that H_1 is mapped to true, but $A_1 \wedge A_2$ is not mapped to true under the above interpretation. In order to satisfy this constraint, we have to add further facts to the program. Their proposal is to use finitely failed sub queries to obtain such facts:

- First, they ask: "Does A_1 hold?". This is not the case in the interpretation I. Hence, they add $A_1 \leftarrow \top$ to the program.
- Then, they query "Does A_2 hold?", which is also not true in the interpretation I. Thus, they add $A_2 \leftarrow \top$ to the program.

Here, the facts $A_1 \leftarrow \top, A_2 \leftarrow \top$ must be added. The resulting program is:

$$\mathcal{P}' = \{H_1 \leftarrow A_1, A_2, H_2 \leftarrow B_1, H_2 \leftarrow B_2, H_1 \leftarrow \top, A_1 \leftarrow \top, A_2 \leftarrow \top\}$$

This program satisfies the integrity constraints. Moreover, we can conclude A_1, A_2 which is the intended explanation of H_1.

This approach seems to be curious since they use a completion based approach for deductive reasoning and they additionally use the only-if halves as integrity constraints.

In this thesis, a different formalism based on abductive frameworks is used. In contrast to the discussed scenario, we are asking for an extension $\mathcal{E}$ of the program $\mathcal{P}$ such that $\mathcal{E} \cup \mathcal{P} \models \mathcal{O}$, where $\mathcal{O}$ is the observation . One obvious difference is that the program and the observation is not joined in one program. Furthermore, the integrity constraints are not necessary anymore. Another noticeable difference is that Stenning and van Lambalgen describe a non deterministic procedure to update the program in order to satisfy the integrity constraints. Here, such precise criteria that allows a more efficient procedure, are not given.

3.3 Abductive Framework

Given a set of formulas $\mathcal{F}$ and an observation $\mathcal{O}$, abductive reasoning can be characterized in general as the problem to find an explanation $\mathcal{E}$ such that $\mathcal{O}$ can be inferred by $\mathcal{F} \cup \mathcal{E}$ by deductive reasoning.

In this thesis, $\mathcal{F}$ is a logic program and $\mathcal{O}$ is a set of literals. We always assume a fixed set of variables $\mathcal{R}$.

Definition 3.3.1 (Abductive Framework)**.** An *abductive framework* is a tuple $\langle \mathcal{P}, \mathcal{A}, \models \rangle$, where

- $\mathcal{P}$ is a logic program over $\mathcal{R}$,
- $\mathcal{A}$ is a set subset of
  $abd(\mathcal{P}) = \{A \leftarrow \top, A \leftarrow \bot \mid A \in \mathcal{R} \text{ is undefined w.r.t. } \mathcal{P}\}$, called abducibles and
- $\models \subseteq 2^{\mathcal{L}(\mathcal{R})} \times \mathcal{L}(\mathcal{R})$ is a consequence relation.

The intended meaning of $\mathcal{A}$ is that explanations are restricted to be a subset of $\mathcal{A}$. Then, every explanation is *basic*, i.e. every explanation cannot be explained by other facts. The reason is that $abd(\mathcal{P})$ only contains facts where the head is undefined.

An abductive framework serves as an environment for computing explanations. Then, an abductive problem consists of an observation and a framework.

Definition 3.3.2 (Abductive Problem)**.** Let $\langle \mathcal{P}, \mathcal{A}, \models \rangle$ be an abductive framework and $\mathcal{O}$ a set of literals called *observation*. Then, the *abductive problem* is a tuple $\langle \mathcal{P}, \mathcal{A}, \mathcal{O} \models \rangle$.

If $\mathcal{O}$ is a singleton set, we omit the the brackets. Next, the solution of an abductive problem is formally defined.

Definition 3.3.3 (Explanation, Solution)**.** Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models \rangle$ be an abductive problem. Then, $\mathcal{E}$ is an *explanation* (or *solution*) of the abductive problem iff

- $\mathcal{E}$ is a consistent subset of $\mathcal{A}$,
- $\mathcal{P} \cup \mathcal{E}$ is consistent and
- $\mathcal{P} \cup \mathcal{E} \models L$ for all $L \in \mathcal{O}$.

The first consistency requirement means that an explanation does not contain contradictory facts. For example the explanation $\{A \leftarrow \bot, B \leftarrow \top, B \leftarrow \bot\}$ is not consistent since it contains $B \leftarrow \top$ and $B \leftarrow \bot$.

Usually, the entailment relation is defined over all models of $\mathcal{P} \cup \mathcal{E}$. If there does not exist such a model, every formula is a consequence, which is unintended.

Consider the following consequence relations.

Definition 3.3.4 (Consequence Relations)**.** Let $\mathcal{P}$ be a logic program and G be a formula.

$\mathcal{P} \models_K G$	iff	for all $I : I_K \models \mathcal{P}$ implies $I_K(G) = \top$
$\mathcal{P} \models_{K,c} G$	iff	for all $I : I_K \models c(\mathcal{P})$ implies $I_K(G) = \top$
$\mathcal{P} \models_{K,wc} G$	iff	for all $I : I_K \models wc(\mathcal{P})$ implies $I_K(G) = \top$
$\mathcal{P} \models_F G$	iff	for all $I : I_F \models \mathcal{P}$ implies $I_F(G) = \top$
$\mathcal{P} \models_{F,c} G$	iff	for all $I : I_F \models c(\mathcal{P})$ implies $I_F(G) = \top$
$\mathcal{P} \models_{F,wc} G$	iff	for all $I : I_F \models wc(\mathcal{P})$ implies $I_F(G) = \top$
$\mathcal{P} \models_L G$	iff	for all $I : I_L \models \mathcal{P}$ implies $I_L(G) = \top$
$\mathcal{P} \models_{L,lm,c} G$	iff	$I = lm_L(c(\mathcal{P}))$ and $I_L(G) = \top$
$\mathcal{P} \models_{L,lm,wc} G$	iff	$I = lm_L(wc(\mathcal{P}))$ and $I_L(G) = \top$

The second consistency requirement means that there exists an interpretation I that is a model of $\mathcal{P} \cup \mathcal{E}$ w.r.t. the consequence relation. For example, if we use $\models_{L,lm,wc}$, then this states that there exists an interpretation I such that $I_L \models wc(\mathcal{P} \cup \mathcal{E})$. If one uses $\models_K$, then this states that there exists an interpretation I such that $I_K \models \mathcal{P} \cup \mathcal{E}$. In this thesis, the minimality requirement which insists that no strict subset of the explanation is also an explanation, is used.

Definition 3.3.5 (Minimal Explanation)**.** Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models, \rangle$ be an abductive problem. An explanation $\mathcal{E}$ is called minimal iff there is no $\mathcal{E}' \subset \mathcal{E}$ such that $\mathcal{E}'$ is an explanation.

Then, the explanation "The sprinkler was on and it rained last night."in Example 3.2.1 is not minimal, whereas "It rained last night." and the explanation "The sprinkler was on." are minimal.

The question is, what one can conclude with respect to explanations. Since human reasoning can be modeled by reasoning with respect to the least Łukasiewicz model of the weak completion, this logic is used to derive logical consequences of $\mathcal{P}$ and the minimal explanations. Stenning and van Lambalgen further noticed, that humans reason skeptically by explanations. Formally, that is the following:

Definition 3.3.6 (Skeptical Reasoning)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models \rangle$ be an abductive problem and F a formula.

Then F follows *skeptically* by AP, denoted by $AP \models_s F$, iff

- there exists a solution of AP and
- F is a *universal consequence* of AP, i.e. for all minimal explanations $\mathcal{E}$ of AP we find that $I = lm_L(wc(\mathcal{P} \cup \mathcal{E}))$ and $I_L(F) = \top$.

This definition is more powerful than the system by Stenning and van Lambalgen, since here F is an arbitary formula and we reason with respect to all explanations. The system by Stenning and van Lambalgen only reason with respect to one explanation, if it exists. If there are multiple explanations, nothing can be concluded. Here, we can conclude that one of the several explanations must hold. Moreover, we can simulate deductive reasoning in this framework:

Proposition 3.3.7. Let $\mathcal{P}$ be a logic program and F a formula. Then, $\mathcal{P} \models_{L,lm,wc} F$ iff $AP = \langle \mathcal{P}, \emptyset, \emptyset, \models \rangle \models_s F$

Proof.

"$\rightarrow$" Then, $I = lm_L(wc(\mathcal{P}))$ and $I(F) = \top$. Since in the above abductive problem, the observation is empty, it follows that the empty set is the only minimal explanation. Then, it immediately follows $AP \models_s F$.

"$\leftarrow$" Since the observation is the empty set, if follows that the empty explanation $\mathcal{E}$ is the only minimal explanation. Because $AP \models_s F$, it follows that $\mathcal{P} \cup \mathcal{E} = \mathcal{P} \models_{L,lm,wc} F$.

□

This shows that the definition of skeptical reasoning is a natural extension of deductive reasoning. That is, one can use this definition in order to compute the first six deductive examples by Byrne.

3.4 Byrne's Experiment

In this section, we will investigate the question, which abductive framework is adequate to model human reasoning. For this purpose, we will contrast the logics by Łukasiewicz, Fitting and Kleene with completed programs, weakly completed programs and with respect to the original program.

In [Byr89], Byrne confronted individuals with sentences like "If she has an essay to write, she will study late in the library. She will not study late in the library.". Then, the individuals were asked to draw conclusions. Here, 92% of subjects conclude that she does not have an essay to write. In Figure 3.1, the situations, observations and experimental results by Byrne are given. Moreover, the translation by Stenning and van Lambalgen from natural sentences into logic programs is given.

In Figure 3.2 minimal explanations for Byrne's experiment are summarized for different logics. One can observe the following:

Observation 3.4.1. Reasoning with respect to the original program is not adequate to model human reasoning.

Statement	If she has an essay to write she will study late in the library.
Program	$\mathcal{P}_1 = \{l \leftarrow e \wedge \neg ab, ab \leftarrow \bot\}$
Observation	Example 1: She will study late in the library.
Explanation	She has an essay to write (71%).
Observation	Example 2: She will not study late in the library.
Explanation	She does not have an essay to write (92%).
Statement	If she has an essay to write she will study late in the library. If she has a textbook to read, she will study late in the library.
Program	$\mathcal{P}_2 = \{l \leftarrow e \wedge \neg ab_1, l \leftarrow t \wedge \neg ab_2, ab_1 \leftarrow \bot, ab_2 \leftarrow \bot\}$
Observation	Example 3: She will study late in the library.
Explanation	She has an essay to write (13%).
Observation	Example 4: She will not study late in the library.
Explanation	She does not have an essay to write (96%).
Statement	If she has an essay to write she will study late in the library. If the library stays open, she will study late in the library.
Program	$\mathcal{P}_3 = \{l \leftarrow e \wedge \neg ab_1, l \leftarrow o \wedge \neg ab_2, ab_1 \leftarrow \neg o, ab_2 \leftarrow \neg e\}$
Observation	Example 5: She will study late in the library.
Explanation	She has an essay to write (54%).
Observation	Example 6: She will not study late in the library.
Explanation	She does not have an essay to write (33%).

Figure 3.1: Byrne's experiment and the translation into logic programs by Stenning and van Lambalgen

Consider Example 1 in Figure 3.2

If Marian has an essay to write, she will study late in the library.

She will study late in the library.

Here, 71% of subjects conclude that she has an essay to write. Consider the abductive problems $AP_L = \langle \mathcal{P}_1, abd(\mathcal{P}_1), l, \models_L \rangle$, $AP_F = \langle \mathcal{P}_1, abd(\mathcal{P}_1), l, \models_F \rangle$, $AP_K = \langle \mathcal{P}_1, abd(\mathcal{P}_1), l, \models_K \rangle$.

Then, l cannot be explained, if we reason with respect to the original program. Consider the interpretation $I = \langle \{e, ab\}, \{l\} \rangle$.

- $\mathcal{E}_1 = \emptyset$ is not an explanation. We obtain $I_L \models \mathcal{P}_1$, $I_F \models \mathcal{P}_1$ and $I_K \models \mathcal{P}_1$. Hence, $\mathcal{P}_1 \not\models_L l$, $\mathcal{P}_1 \not\models_F l$ and $\mathcal{P}_1 \not\models_K l$.

- $\mathcal{E}_2 = \{e \leftarrow \bot\}$ is not an explanation. We obtain $I_L \models \mathcal{P}_1 \cup \mathcal{E}_2$, $I_F \models \mathcal{P}_1 \cup \mathcal{E}_2$ and $I_K \models \mathcal{P}_1 \cup \mathcal{E}_2$. Hence, $\mathcal{P}_1 \cup \mathcal{E}_2 \not\models_L l$, $\mathcal{P}_1 \cup \mathcal{E}_2 \not\models_F l$ and $\mathcal{P}_1 \cup \mathcal{E}_2 \not\models_K l$.

- $\mathcal{E}_3 = \{e \leftarrow \top\}$ is not an explanation. We obtain $I_L \models \mathcal{P}_1 \cup \mathcal{E}_3$, $I_F \models \mathcal{P}_1 \cup \mathcal{E}_3$ and $I_K \models \mathcal{P}_1 \cup \mathcal{E}_3$. Hence, $\mathcal{P}_1 \cup \mathcal{E}_3 \not\models_L l$, $\mathcal{P}_1 \cup \mathcal{E}_3 \not\models_F l$ and $\mathcal{P}_1 \cup \mathcal{E}_3 \not\models_K l$.

Hence, there is no explanation. Then, $AP_L \not\models_s e$, $AP_F \not\models e$, $AP_K \not\models e$, which is incorrect with respect to Byrne's data.

Consequence relation	Logic Program	Observation	Minimal explanations
$\models_L, \models_F, \models_K$	$\mathcal{P}_1$	l	$\times$
		$\neg l$	$\times$
	$\mathcal{P}_2$	l	$\times$
		$\neg l$	$\times$
	$\mathcal{P}_3$	l	$\times$
		$\neg l$	$\times$
$\models_{L,lm,c}, \models_{F,c}, \models_{K,c}$	$\mathcal{P}_1$	l	$\{e \leftarrow \top\}$
		$\neg l$	$\emptyset$
	$\mathcal{P}_2$	l	$\{e \leftarrow \top\}, \{t \leftarrow \top\}$
		$\neg l$	$\emptyset$
	$\mathcal{P}_3$	l	$\{e \leftarrow \top, o \leftarrow \top\}$
		$\neg l$	$\emptyset$
$\models_{L,lm,wc}, \models_{F,wc}, \models_{K,wc}$	$\mathcal{P}_1$	l	$\{e \leftarrow \top\}$
		$\neg l$	$\{e \leftarrow \bot\}$
	$\mathcal{P}_2$	l	$\{e \leftarrow \top\}, \{t \leftarrow \top\}$
		$\neg l$	$\{e \leftarrow \bot, t \leftarrow \bot\}$
	$\mathcal{P}_3$	l	$\{e \leftarrow \top, o \leftarrow \top\}$
		$\neg l$	$\{e \leftarrow \bot\}, \{o \leftarrow \bot\}$

Figure 3.2: Minimal explanations of Byrne's experiment under different abductive frameworks. The set of abducibles is not restricted, i.e. $\mathcal{A} = abd(\mathcal{P})$. The $\times$-symbol denotes that there does not exist a minimal explanation.

Observation 3.4.2. Reasoning with respect to the completion of a program is not adequate to model abductive human reasoning.

Consider the following experiment:

If Marian has an essay to write, she will study late in the library.

She will not study late in the library.

Here, 92% of individuals conclude that she does not have an essay to write. Consider the abductive problems

$$AP_L = \langle \mathcal{P}_1, abd(\mathcal{P}_1), l, \models_{L,lm,c} \rangle$$
$$AP_F = \langle \mathcal{P}_1, abd(\mathcal{P}_1), l, \models_{F,c} \rangle$$
$$AP_K = \langle \mathcal{P}_1, abd(\mathcal{P}_1), l, \models_{K,c} \rangle$$

Next, it is shown, that the empty set is a solution of these abductive problems.

Consider the completion of $\mathcal{P}_1$:

$$c(\mathcal{P}_1) = \{l \leftrightarrow e \wedge \neg ab, e \leftrightarrow \bot, ab \leftrightarrow \bot\}$$

The least Łukasiewicz model is $I_L = \langle \emptyset, \{e, ab, l\} \rangle$. Hence, the empty set is a minimal explanation. Every Kleene and Fitting model of $c(\mathcal{P}_1)$ must map e and ab to $\bot$ and thus l as well.

Finally, we obtain that the empty set is the only minimal explanation. It is clear that $AP_L \not\models_s \neg e$, $AP_F \not\models_s \neg e$, $AP_K \not\models_s \neg e$, which is incorrect to Byrne's data.

Observation 3.4.3. The weak completion under the Łukasiewicz logic as a justification is adequate to model human reasoning.

In the following, this important observation is explained in more detail. Each example in Figure 3.2 and the associated abductive problem will be presented. For each subset of the abducibles, the least model of the resulting program will be given in order to decide if this subset is an explanation. Then, one can immediately read off the minimal explanations. Finally, we reason skeptically with these explanations according to Definition 3.3.6.

Example 1

If Marian has an essay to write, she will study late in the library.

She will study late in the library.

Here, 71% of subjects conclude that she has an essay to write. Consider the abductive problem $AP_{B,1} = \langle \mathcal{P}_1, abd(\mathcal{P}_1), l, \models_{L,lm,wc} \rangle$.

$$\begin{aligned} lm(wc(\mathcal{P}_1)) &= \langle \emptyset, \{ab\} \rangle && \not\models l \\ lm(wc(\mathcal{P}_1 \cup \{e \leftarrow \top\})) &= \langle \{e, l\}, \{ab\} \rangle && \models l \\ lm(wc(\mathcal{P}_1 \cup \{e \leftarrow \bot\})) &= \langle \{\}, \{ab, e, l\} \rangle && \not\models l \end{aligned}$$

Hence, $\{e \leftarrow \top\}$ is the only minimal explanation and thus $AP_{B,1} \models_s e$, which is correct w.r.t. Byrne's data.

Example 2

If Marian has an essay to write, she will study late in the library.

She will not study late in the library.

Here, 92% of individuals conclude that she does not have an essay to write. We obtain the abductive problem $AP_{B,2} = \langle \mathcal{P}_1, abd(\mathcal{P}_1), \neg l, \models_{L,lm,wc} \rangle$.

$$\begin{aligned} lm(wc(\mathcal{P}_1)) &= \langle \emptyset, \{ab\} \rangle && \not\models \neg l \\ lm(wc(\mathcal{P}_1 \cup \{e \leftarrow \top\})) &= \langle \{e, l\}, \{ab\} \rangle && \not\models \neg l \\ lm(wc(\mathcal{P}_1 \cup \{e \leftarrow \bot\})) &= \langle \emptyset, \{ab, e, l\} \rangle && \models \neg l \end{aligned}$$

Hence, $\{e \leftarrow \bot\}$ is the only minimal explanation and $AP_{B,2} \models_s \neg e$, which is correct w.r.t. Byrne's data.

Example 3

> *If Marian has an essay to write, she will study late in the library.*
>
> *If she has a textbook to read, she will study late in the library.*
>
> *She will study late in the library.*

Here, 13% of individuals conclude that she has an essay to write. We obtain the abductive problem $AP_{B,3} = \langle \mathcal{P}_2, abd(\mathcal{P}_2), l, \models_{L,lm,wc} \rangle$.

$$\begin{aligned}
lm(wc(\mathcal{P}_2)) &= \langle \emptyset, \{ab_1, ab_2\} \rangle && \not\models l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \top\})) &= \langle \{e, l\}, \{ab_1, ab_2\} \rangle && \models l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \bot\})) &= \langle \emptyset, \{ab_1, ab_2, e\} \rangle && \not\models l \\
lm(wc(\mathcal{P}_2 \cup \{t \leftarrow \top\})) &= \langle \{t, l\}, \{ab_1, ab_2\} \rangle && \models l \\
lm(wc(\mathcal{P}_2 \cup \{t \leftarrow \bot\})) &= \langle \emptyset, \{ab_1, ab_2, t\} \rangle && \not\models l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \bot, t \leftarrow \bot\})) &= \langle \emptyset, \{ab_1, ab_2, t, e, l\} \rangle && \not\models l
\end{aligned}$$

Hence $\{e \leftarrow \top\}$ and $\{t \leftarrow \top\}$ are minimal explanations and $AP_{B,3} \not\models_s e$, which is correct w.r.t. Byrne's data.

Example 4

> *If Marian has an essay to write, she will study late in the library.*
>
> *If she has a textbook to read, she will study late in the library.*
>
> *She will not study late in the library.*

Here, 96% of individuals conclude that she does not have an essay to write. We obtain the abductive problem $AP_{B,4} = \langle \mathcal{P}_2, abd(\mathcal{P}_2), \neg l, \models_{L,lm,wc} \rangle$.

$$\begin{aligned}
lm(wc(\mathcal{P}_2)) &= \langle \emptyset, \{ab_1, ab_2\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \top\})) &= \langle \{e, l\}, \{ab_1, ab_2\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \bot\})) &= \langle \emptyset, \{ab_1, ab_2, e\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{t \leftarrow \top\})) &= \langle \{t, l\}, \{ab_1, ab_2\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{t \leftarrow \bot\})) &= \langle \emptyset, \{ab_1, ab_2, t\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \top, t \leftarrow \top\})) &= \langle \{e, t, l\}, \{ab_1, ab_2\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \top, t \leftarrow \bot\})) &= \langle \{e, l\}, \{ab_1, ab_2, t\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \bot, t \leftarrow \top\})) &= \langle \{t, l\}, \{ab_1, ab_2, e\} \rangle && \not\models \neg l \\
lm(wc(\mathcal{P}_2 \cup \{e \leftarrow \bot, t \leftarrow \bot\})) &= \langle \emptyset, \{ab_1, ab_2, e, t, l\} \rangle && \models \neg l
\end{aligned}$$

Hence $\{e \leftarrow \bot, t \leftarrow \bot\}$ is the only minimal explanation and $AP_{B,4} \models_s \neg e$, which is correct w.r.t. Byrne's data.

Example 5

> *If Marian has an essay to write, she will study late in the library.*
>
> *If the library stays open, she will study late in the library.*
>
> *She will study late in the library.*

Here, 54% of individuals conclude that she has an essay to write. We obtain the abductive problem $AP_{B,5} = \langle \mathcal{P}_3, abd(\mathcal{P}_3), l, \models_{L,lm,wc} \rangle$.

$$
\begin{array}{rll}
lm(wc(\mathcal{P}_3)) = \langle \emptyset, \emptyset \rangle & & \not\models l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \top\})) = \langle \{e\}, \{ab_2\} \rangle & & \not\models l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \bot\})) = \langle \{ab_2\}, \{e, l\} \rangle & & \not\models l \\
lm(wc(\mathcal{P}_3 \cup \{o \leftarrow \top\})) = \langle \{o\}, \{ab_1\} \rangle & & \not\models l \\
lm(wc(\mathcal{P}_3 \cup \{o \leftarrow \bot\})) = \langle \{ab_1\}, \{o, l\} \rangle & & \not\models l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \top, o \leftarrow \top\})) = \langle \{e, o, l\}, \{ab_1, ab_2\} \rangle & & \models l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \top, o \leftarrow \bot\})) = \langle \{e, ab_1\}, \{o, ab_2, l\} \rangle & & \not\models l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \bot, o \leftarrow \top\})) = \langle \{o, ab_2\}, \{ab_1, e, l\} \rangle & & \not\models l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \bot, o \leftarrow \bot\})) = \langle \{ab_1, ab_2\}, \{e, t, l\} \rangle & & \not\models l
\end{array}
$$

Hence, $\{e \leftarrow \top, o \leftarrow \top\}$ is the only minimal explanation and $AP_{B,5} \models_s e$, which is correct w.r.t. Byrne's data.

Example 6

> *If Marian has an essay to write, she will study late in the library.*
>
> *If the library stays open, she will study late in the library.*
>
> *She will not study late in the library.*

Here, 33% of individuals conclude that she does not have an essay to write. We obtain the abductive problem $AP_{B,6} = \langle \mathcal{P}_3, abd(\mathcal{P}_3), \neg l, \models_{L,lm,wc} \rangle$.

$$
\begin{array}{rll}
lm(wc(\mathcal{P}_3)) = \langle \emptyset, \emptyset \rangle & & \not\models \neg l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \top\})) = \langle \{e\}, \{ab_2\} \rangle & & \not\models \neg l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \bot\})) = \langle \{ab_2\}, \{e, l\} \rangle & & \models \neg l \\
lm(wc(\mathcal{P}_3 \cup \{o \leftarrow \top\})) = \langle \{o\}, \{ab_1\} \rangle & & \not\models \neg l \\
lm(wc(\mathcal{P}_3 \cup \{o \leftarrow \bot\})) = \langle \{ab_1\}, \{o, l\} \rangle & & \models \neg l \\
lm(wc(\mathcal{P}_3 \cup \{e \leftarrow \top, o \leftarrow \top\})) = \langle \{e, o, l\}, \{ab_1, ab_2\} \rangle & & \not\models \neg l
\end{array}
$$

Hence, $\{e \leftarrow \bot\}$ and $\{o \leftarrow \bot\}$ are minimal explanations and $AP_{B,6} \not\models_s \neg e$, which is correct w.r.t. Byrne's data.

All in all, it is shown that skeptical reasoning under weakly completed programs and the Łukasiewicz logic is indeed adequate to model the six abductive examples by Byrne.

3.5 Integrity Constraints

Explanations can also be restricted by so called *integrity constraints*. Here, an integrity constraint is a formula. In [KKT98] two different semantics for integrity constraints are presented, the theorem-hood view and the satisfiability view.

Definition 3.5.1 (Semantics of Integrity Constraints)**.** Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem and IC a formula, called *integrity constraint*. Then, $\mathcal{E}$ is a solution iff $\mathcal{E}$ is a solution of the abductive problem and $\mathcal{E}$ *satisfies* F iff

- $\mathcal{E} \cup \mathcal{P} \models_{L,lm,wc} IC$ in *theorem-hood view* or
- there exists an interpretation I such that $I_L \models wc(\mathcal{E} \cup \mathcal{P}) \cup \{IC\}$ in the *satisfiability view*

In this section, we investigate four Byrne's examples added with integrity constraints. For each situation, the different semantics of integrity constraints are contrasted.

Example 3* Consider the following situation:

> *If she has an essay to write she will study late in the library.*
>
> *If she has a textbook to read, she will study late in the library.*
>
> *She will not read a textbook in holidays.*
>
> *There are holidays.*
>
> *She will study late in the library.*

Here, we translate "She will not read a textbook in holidays." as a constraint, i.e. we have the following:

$$\mathcal{P}_4 = \{l \leftarrow e \wedge \neg ab_1, l \leftarrow t \wedge \neg ab_2, ab_1 \leftarrow \bot, ab_2 \leftarrow \bot, h \leftarrow \top\}$$
$$IC = \bot \leftarrow t \wedge h$$

Note that "She will not read a textbook in holidays." cannot be seen as a case for ab_2, since here we mean that she will not read a textbook in holidays. If instead $ab_2 \leftarrow h$ would be added, it could be possible that she has a textbook to read in holidays.

The abductive problem is $AP_{B,7} = \langle \mathcal{P}_4, abd(\mathcal{P}_4), l, \models_{L,lm,wc} \rangle$.

$$
\begin{aligned}
lm(wc(\mathcal{P}_4)) &= \langle\{h\}, \{ab_1, ab_2\}\rangle && \not\models l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \top\})) &= \langle\{e, l, h\}, \{ab_1, ab_2\}\rangle && \models l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \bot\})) &= \langle\{h\}, \{ab_1, ab_2, e\}\rangle && \not\models l \\
lm(wc(\mathcal{P}_4 \cup \{t \leftarrow \top\})) &= \langle\{t, l, h\}, \{ab_1, ab_2\}\rangle && \models l \\
lm(wc(\mathcal{P}_4 \cup \{t \leftarrow \bot\})) &= \langle\{h\}, \{ab_1, ab_2, t\}\rangle && \not\models l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \top, t \leftarrow \top\})) &= \langle\{e, l, t, h\}, \{ab_1, ab_2\}\rangle && \models l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \top, t \leftarrow \bot\})) &= \langle\{e, l, h\}, \{ab_1, ab_2, t\}\rangle && \models l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \bot, t \leftarrow \top\})) &= \langle\{t, l, h\}, \{ab_1, ab_2, e\}\rangle && \models l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \bot, t \leftarrow \bot\})) &= \langle\{h\}, \{ab_1, ab_2, e, t, l\}\rangle && \not\models l
\end{aligned}
$$

We obtain the explanations: $\{e \leftarrow \top\}$, $\{t \leftarrow \top\}$, $\{e \leftarrow \top, t \leftarrow \top\}$, $\{e \leftarrow \top, t \leftarrow \bot\}$ and $\{e \leftarrow \bot, t \leftarrow \top\}$.

Theorem-hood View The only explanation that satisfies IC is $\{e \leftarrow \top, t \leftarrow \bot\}$. Hence, $AP_{B,7} \models_s e$ and $AP_{B,7} \models_s \neg t$

Satisfiability View Here, the minimal explanation $\{e \leftarrow \top\}$ satisfies IC with the model $\langle\{e, l, h\}, \{ab_1, ab_2, t\}\rangle$. Hence, $AP_{B,7} \models_s e$ and $AP_{B,7} \not\models_s \neg t$

Example 4* Consider the following situation:

If she has an essay to write she will study late in the library.

If she has a textbook to read, she will study late in the library.

She will not read a textbook in holidays.

There are holidays.

She will not study late in the library.

$$
\begin{aligned}
&\mathcal{P}_4 = \{l \leftarrow e \wedge \neg ab_1, l \leftarrow t \wedge \neg ab_2, ab_1 \leftarrow \bot, ab_2 \leftarrow \bot, h \leftarrow \top\} \\
&IC = \bot \leftarrow t \wedge h
\end{aligned}
$$

The abductive problem is $AP_{B,8} = \langle \mathcal{P}_4, abd(\mathcal{P}_4), \neg l, \models_{L,lm,wc}\rangle$.

$$
\begin{array}{rlr}
lm(wc(\mathcal{P}_4)) = & \langle \{h\}, \{ab_1, ab_2\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \top\})) = & \langle \{e, l, h\}, \{ab_1, ab_2\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \bot\})) = & \langle \{h\}, \{ab_1, ab_2, e\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{t \leftarrow \top\})) = & \langle \{t, l, h\}, \{ab_1, ab_2\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{t \leftarrow \bot\})) = & \langle \{h\}, \{ab_1, ab_2, t\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \top, t \leftarrow \top\})) = & \langle \{e, t, l, h\}, \{ab_1, ab_2\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \top, t \leftarrow \bot\})) = & \langle \{e, l, h\}, \{ab_1, ab_2, t\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \bot, t \leftarrow \top\})) = & \langle \{t, l, h\}, \{ab_1, ab_2, e\}\rangle & \not\models \neg l \\
lm(wc(\mathcal{P}_4 \cup \{e \leftarrow \bot, t \leftarrow \bot\})) = & \langle \{h\}, \{ab_1, ab_2, e, t, l\}\rangle & \models \neg l
\end{array}
$$

We obtain the only explanation $\{e \leftarrow \bot, t \leftarrow \bot\}$.

Theorem-hood View This explanation satisfies IC. Hence, $AP_{B,8} \models_s e$.

Satisfiability View This explanation also satisfies IC in the satisfiability view, since $\langle \{h\}, \{ab_1, ab_2, e, t, l\}\rangle$ is a model of $\mathcal{P}_4 \cup \{IC\} \cup \{e \leftarrow \bot, t \leftarrow \bot\}$. Hence, $AP_{B,8} \models_s e$.

Example 5* Consider the following situation:

If Marian has an essay to write, she will study late in the library.

If the library stays open, she will study late in the library.

The library is not open in holidays.

There are holidays.

She will study late in the library.

$$
\begin{aligned}
&\mathcal{P}_5 = \{l \leftarrow e \wedge \neg ab_1, l \leftarrow o \wedge \neg ab_2, ab_1 \leftarrow \neg o, , ab_2 \leftarrow \neg e, h \leftarrow \top\} \\
&IC = \bot \leftarrow o \wedge h
\end{aligned}
$$

The abductive problem is $AP_{B,9} = \langle \mathcal{P}_5, abd(\mathcal{P}_5), l, \models_{L,lm,wc}\rangle$.

$$\begin{aligned}
lm(wc(\mathcal{P}_5)) &= \langle\{h\},\emptyset\rangle && \not\models l\\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \top\})) &= \langle\{e,h\},\{ab_2\}\rangle && \not\models l\\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \bot\})) &= \langle\{ab_2,h\},\{e,l\}\rangle && \not\models l\\
lm(wc(\mathcal{P}_5 \cup \{o \leftarrow \top\})) &= \langle\{o,h\},\{ab_1\}\rangle && \not\models l\\
lm(wc(\mathcal{P}_5 \cup \{o \leftarrow \bot\})) &= \langle\{ab_1,h\},\{o,l\}\rangle && \not\models l\\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \top, o \leftarrow \top\})) &= \langle\{e,o,l,h\},\{ab_1,ab_2\}\rangle && \models l\\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \top, o \leftarrow \bot\})) &= \langle\{e,ab_1,h\},\{o,ab_2,l\}\rangle && \not\models l\\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \bot, o \leftarrow \top\})) &= \langle\{o,ab_2,h\},\{ab_1,e,l\}\rangle && \not\models l\\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \bot, o \leftarrow \bot\})) &= \langle\{ab_1,ab_2,h\},\{e,o,l\}\rangle && \not\models l
\end{aligned}$$

Hence, $\{e \leftarrow \top, o \leftarrow \top\}$ is the only explanation.

Theorem-hood View This explanation does not satisfy IC. Hence, $AP_{B,9} \not\models_s e$.

Satisfiability View This explanation does also not satisfy IC in the satisfiability view, since $\langle\{e,o,l,h\},\{ab_1,ab_2\}\rangle$ is the least model of $\mathcal{P}_5 \cup \{e \leftarrow \top, o \leftarrow \top\}$. Moreover, it is total. Since this model is not a model of $\mathcal{P}_5 \cup \{e \leftarrow \top, o \leftarrow \top\} \cup \{IC\}$, there is no model. Hence, this explanation does not satisfy IC. Hence $AP_{B,9} \not\models_s e$.

Example 6*

If Marian has an essay to write, she will study late in the library.

If the library stays open, she will study late in the library.

The library is not open in holidays.

There are holidays.

She will not study late in the library.

$$\mathcal{P}_5 = \{l \leftarrow e \wedge \neg ab_1, l \leftarrow o \wedge \neg ab_2, ab_1 \leftarrow \neg o, , ab_2 \leftarrow \neg e, h \leftarrow \top\}$$
$$IC = \bot \leftarrow o \wedge h$$

The abductive problem is $AP_{B,10} = \langle \mathcal{P}_5, abd(\mathcal{P}_5), \neg l, \models_{L,lm,wc}\rangle$.

$$
\begin{array}{rcl}
lm(wc(\mathcal{P}_5)) = \langle\{h\}, \emptyset\rangle & & \not\models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \top\})) = \langle\{e, h\}, \{ab_2\}\rangle & & \not\models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \bot\})) = \langle\{ab_2, h\}, \{e, l\}\rangle & & \models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{o \leftarrow \top\})) = \langle\{o, h\}, \{ab_1\}\rangle & & \not\models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{o \leftarrow \bot\})) = \langle\{ab_1, h\}, \{o, l\}\rangle & & \models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \top, o \leftarrow \top\})) = \langle\{e, o, l, h\}, \{ab_1, ab_2\}\rangle & & \not\models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \top, o \leftarrow \bot\})) = \langle\{e, h, ab_1\}, \{o, ab_2, l\}\rangle & & \models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \bot, o \leftarrow \top\})) = \langle\{ab_2, o, h\}, \{ab_1, e, l\}\rangle & & \models \neg l \\
lm(wc(\mathcal{P}_5 \cup \{e \leftarrow \bot, o \leftarrow \bot\})) = \langle\{ab_1, ab_2, h\}, \{e, o, l\}\rangle & & \models \neg l
\end{array}
$$

Here, we obtain the following explanations: $\mathcal{E}_1 = \{e \leftarrow \bot\}$, , $\mathcal{E}_2 = \{o \leftarrow \bot\}, \mathcal{E}_3 = \{e \leftarrow \top, o \leftarrow \bot\}, \mathcal{E}_4 = \{e \leftarrow \bot, o \leftarrow \top\}, \mathcal{E}_5 = \{e \leftarrow \bot, o \leftarrow \bot\}$

Theorem-hood View The explanation $\mathcal{E}_1$ does not satisfy IC, but $\mathcal{E}_2$. Hence, $\mathcal{E}_2$ is the only minimal explanation. Thus, $AP_{B,10} \not\models_s e$.

Satisfiability View Here, the explanation $\mathcal{E}_1$ does satisfy IC with the model $\langle\{ab_1, ab_2, h\}, \{e, l, o\}\rangle$. The solution $\mathcal{E}_2$ does also satisfy IC with the model $\langle\{ab_1, h\}, \{o, l\}\rangle$. Thus $AP_{B,10} \not\models_s e$

3.5.1 Discussion

Consider the modified Example 6* under the theorem-hood-view: Here, $\{e \leftarrow \bot\}$ is not an explanation. This is interesting, since one reason that she is not in the library could be that she does not have an essay to write. In this sense, the theorem-hood view eliminates meaningful explanations. This is not the case in the satisfiability view.

On the other hand, the satisfiability view seems to be inconsistent with knowledge. Consider Example 3*. Here we have

- $AP_{B,7} \models_s e \wedge \neg t$ in the theorem-hood view and
- $AP_{B,1} \not\models_s e \wedge \neg t$ in the satisfiability view

That is, in the theorem-hood view one can conclude that she will not study late in the library, which is not the case in the satisfiability view. This is inconsistent in the sense, that Marian will not read textbooks in holidays and there are holidays. From this point, one have to conclude that she has no textbook to read.

Chapter 4

Relationships

In this chapter, the relationships between the three logics by Łukasiewicz, Kleene and Fitting with the different forms of completions are investigated. If no form of completion is used, one can see that there does not exist a minimal explanation. This will be explained in the first section. Then, completion and weak completion are contrasted. Afterwards, the Fitting logic will be related to Kleene and the Łukasiewicz logic. Finally, the differences between the Łukasiewicz and Kleene logic under weakly completed programs are shown.

4.1 No Form of Completion

If no form of completion is used, we see in the examples that there does not exist a minimal explanations.

However, this is not always true. Consider the following example:

Example 4.1.1 (Existence of Explanations under no form of completion)**.** Let $\mathcal{P} = \{A \leftarrow B\}$ be a logic program, $\mathcal{A} = \{B \leftarrow \top, B \leftarrow \bot\}$ the set of abducibles and A be an observation.

Consider the abductive programs $AP_1 = \langle \mathcal{P}, \mathcal{A}, A, \models_L \rangle$, $AP_2 = \langle \mathcal{P}, \mathcal{A}, A, \models_K \rangle$, $AP_3 = \langle \mathcal{P}, \mathcal{A}, A, \models_F \rangle$.

In all abductive problems, $\mathcal{E} = \{B \leftarrow \top\}$ is a solution.

Hence, sometimes there exist explanations if no form of completion is used.

The question is, why there do not exist minimal explanations in the six examples by Byrne. Consider the program $\mathcal{P}_1$ and $\mathcal{P}_2$:

$$\mathcal{P}_1 = \{l \leftarrow e \wedge \neg ab, ab \leftarrow \bot\}$$
$$\mathcal{P}_2 = \{l \leftarrow e \wedge \neg ab_1, l \leftarrow t \wedge \neg ab_2, ab_1 \leftarrow \bot, ab_2 \leftarrow \bot\}$$

It is easy to see that for every interpretation I, we have $I_L(ab \leftarrow \bot) = I_K(ab \leftarrow \bot) = I_F(ab \leftarrow \bot) = \top$. Hence, we can map ab, ab_1, ab_2 to u. But then

l can be mapped to u, if e is mapped to true. If $I(e) = \bot$, we can map l to true or false. But then, one cannot explain l or $\neg l$.

4.2 Completion and Weak Completion

At the examples in Figure 3.2, we first observe that a minimal explanation under the completion only contains positive facts. The reason for this is that negative facts are added under the completion.

Proposition 4.2.1 (Negative Facts and Completion)**.** Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models \rangle$, where $\models \in \{\models_{L,lm,c}, \models_{K,c}, \models_{F,c}\}$ be an abductive problem and $\mathcal{E}$ a minimal explanation. Then, $\mathcal{E}$ does not contain negative facts.

Proof. Let $\mathcal{E}^+$ denote the positive facts occuring in $\mathcal{E}$ and $\mathcal{E}^-$ denote the negative facts occuring in $\mathcal{E}$. Since $\mathcal{E}^-$ only contains negative facts that are undefined w.r.t. $\mathcal{P}$ and $\mathcal{E}^+$, they will be added in the second step of completion. Hence, we can omit them. Then $c(\mathcal{P} \cup \mathcal{E}) = c(\mathcal{P} \cup \mathcal{E}^+)$. Since $\mathcal{E}^+ \subseteq \mathcal{E}$, a minimal explanation can only contain positive facts. $\square$

Moreover, we observe in Figure 3.2 that a minimal explanation under the completion can be obtained by deleting negative facts of the minimal explanations under the weak completion. This is now proven under the Łukasiewicz logic.

Proposition 4.2.2. If $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$, then $\mathcal{E}^+$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,c} \rangle$.

Proof. Since $\mathcal{E}$ is a solution, we know that

$$\mathcal{E} \cup \{A \leftarrow \bot \mid A \text{ is undefined w.r.t. } \mathcal{E} \text{ and } \mathcal{P}\}$$

is also a solution by Proposition 5.1.10. Then, we know that

$$wc(\mathcal{P} \cup \mathcal{E} \cup \{A \leftarrow \bot \mid A \text{ is undefined w.r.t. } \mathcal{E} \text{ and } \mathcal{P}\}) = c(\mathcal{P} \cup \mathcal{E}) = c(\mathcal{P} \cup \mathcal{E}^+)$$

Hence, we have that $\mathcal{E}^+$ must be a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,c} \rangle$. $\square$

The other direction does not hold in general, since we only require that $\mathcal{A}$ is a subset of the undefined predicates in $\mathcal{P}$:

Example 4.2.3. Consider the logic program $\mathcal{P} = \{p \leftarrow q\}$ and the observation $\mathcal{O} = \neg p$. The empty set is then a solution of the abductive problem $\langle \mathcal{P}, \emptyset, \mathcal{O}, \models_{L,lm,c} \rangle$, since $lm(c(\{p \leftarrow q\})) = lm(\{p \leftrightarrow q, q \leftrightarrow \bot\}) = \langle \emptyset, \{p, q\} \rangle \models \neg p$. However, the abductive problem $\langle \mathcal{P}, \emptyset, \mathcal{O}, \models_{L,lm,wc} \rangle$ has no solution, since $lm(wc(\{p \leftarrow q\})) = lm(\{p \leftrightarrow q\}) = \langle \emptyset, \emptyset \rangle \not\models \neg p$.

4.3 Fitting Logic

In this section, we discuss the Fitting logic. At the examples in Figure 3.2, the minimal explanations of the Fitting logic correspond to the Łukasiewicz logic, if weak completion or completion is used. Otherwise it corresponds to the Kleene logic. Next, it is shown that this is generally the case.

Lemma 4.3.1 (Corresponding Logics)**.** Let $\mathcal{P}$ be a logic program and I be an interpretation. Then, the following holds:

(i) $I_F \models \mathcal{P}$ iff $I_K \models \mathcal{P}$,

(ii) $I_F \models wc(\mathcal{P})$ iff $I_L \models wc(\mathcal{P})$ and

(iii) $I_F \models c(\mathcal{P})$ iff $I_L \models c(\mathcal{P})$.

Proof.

(i) This is true, since in $\mathcal{P}$, the connective $\leftrightarrow$ does not occur in $\mathcal{P}$. The remaining connectives are defined in the same way.

(ii) If we consider the completion or weak completion of $\mathcal{P}$, then we obtain a set of formulae, where each formula is of the form $A \leftrightarrow Body_1 \vee \ldots \vee Body_n$ and each body is either a conjunction of literals or one of the symbols $\top$, $\bot$. Since the Łukasiewicz logic and the Fitting logic share the semantics of the symbols $\neg, \vee, \wedge, \top, \bot$, we have the following for all interpretations: $I_L(Body_i) = I_F(Body_i)$ for all $1 \leq i \leq n$ and thus $I_L(Body_1 \vee \ldots \vee Body_n) = I_F(Body_1 \vee \ldots \vee Body_n)$. The semantic of $\leftrightarrow$ is different in the Łukasiewicz and Fitting logic. However, we have that A is an atom and thus $I_L(A) = I_F(A)$. We know that $I_L(A \leftrightarrow Body_1 \vee \ldots \vee Body_n) = \top$ iff $I_L(A) = I_L(Body_1 \vee \ldots \vee Body_n)$. This corresponds to the definition of complete equivalence. Thus $I_L(A \leftrightarrow Body_1 \vee \ldots \vee Body_n) = \top$ iff $I_F(A \leftrightarrow Body_1 \vee \ldots \vee Body_n) = \top$.

(iii) This claim is analog to (ii).

□

As an immediate consequence, we have the following:

Proposition 4.3.2 (Corresponding Solutions)**.** Let $\mathcal{P}$ be a logic program, $\mathcal{O}$ an observation and $\mathcal{A}$ a set of abducibles. Then, the following holds:

(i) $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_F \rangle$ iff $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_K \rangle$,

(ii) $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{F,wc} \rangle$ iff $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,wc} \rangle$,

(iii) $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{F,c} \rangle$ iff $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,c} \rangle$,

Proof.

(i) $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_F \rangle$ iff 1) $\mathcal{E} \subseteq \mathcal{A}$ and is consistent, 2) $\mathcal{P} \cup \mathcal{E}$ is Fitting consistent and 3) $\mathcal{P} \cup \mathcal{E} \models_F L$ for each $L \in \mathcal{O}$. The first condition is clearly satisfied. $\mathcal{P} \cup \mathcal{E}$ is Fitting consistent iff there exists a Fitting model. This Fitting model is also a Kleene model by Lemma 4.3.1. The third condition is equivalent to $\mathcal{P} \cup \mathcal{E} \models_K L$ for each $L \in \mathcal{O}$ by Lemma 4.3.1. Thus, $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_F \rangle$ iff $\mathcal{E}$ is a solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_K \rangle$.

(ii) This claim is analog to (i).

(iii) This claim is analog to (i).

□

4.4 Łukasiewicz and Kleene Logic

Figure 3.2 shows that considering the least model under the Łukasiewicz semantics or all models under the Kleene semantics leads to the same explanation. In this section we will investigate the relationship between the abductive problems $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ and $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{K,wc} \rangle$.

First, we will show that, if the Łukasiewicz least model is a model of an observation, then all Kleene models of the logic program are also models of this observation. For this purpose, we use the following lemma:

Lemma 4.4.1. Let $\mathcal{P}$ be a logic program and I be an interpretation. If $I_L \not\models wc(\mathcal{P})$, then $I_K \not\models wc(\mathcal{P})$.

Proof. Note that the following equation holds for a formula F over the connectives $\{\wedge, \vee, \neg\}$: $I_L(F) = I_K(F)$.

Since $I_L \not\models wc(\mathcal{P})$, we find $(A \leftrightarrow F) \in wc(\mathcal{P})$ such that $I_L(A \leftrightarrow F) \neq \top$. We show that $I_K(A \leftrightarrow F) \neq \top$. Consider the following three cases:

- $I_L(A) = u$. Since A is an atom, we know $I_K(A) = u$. Then we know that $I_K(u \leftrightarrow F) \neq \top$, and thus we have $I_K \not\models wc(\mathcal{P})$
- $I_L(A) = \top$. Since A is an atom, we know $I_K(A) = \top$. Because $I_L(F) = I_K(F)$, we have $I_K(F) \neq \top$ and thus $I_K(A \leftrightarrow F) \neq \top$.
- $I_L(A) = \bot$. This case is analog to the second case.

□

Then, we obtain the following result.

Proposition 4.4.2 (Relationship $\models_{L,lm,wc}$ and $\models_{K,wc}$)**.** Let $\mathcal{P}$ be a logic program and $\mathcal{O}$ a set of literals. If $I = lm_L(wc(\mathcal{P}))$ and $I_L(L) = \top$ for all $L \in \mathcal{O}$, then $wc(\mathcal{P}) \models_K L$ for all $L \in \mathcal{O}$.

Proof. We have to show: For all interpretations J: If $J_K \models wc(\mathcal{P})$, then $J_K(L) = \top$ for all $L \in \mathcal{O}$. Let J be an interpretation. Consider the following cases:

- $J \not\succeq I$. Then, $J_L \not\models wc(\mathcal{P})$, since I is the least Łukasiewicz model of $wc(\mathcal{P})$. By Lemma 4.4.1, it follows that $J_K \not\models wc(\mathcal{P})$.
- $J \succeq I$. Since $I_L(L) = \top$ for all $L \in \mathcal{O}$, we know the following: If L is a positive literal, then $L \in I^\top$. If L is a negative literal, then then $L \in I^\perp$. Because $J \succeq I$, we can conclude that $L \in J^\top$ in the first case or $L \in J^\perp$ in the second case. Hence, $J_K(L) = \top$ for all $L \in \mathcal{O}$.

Hence, we have $wc(\mathcal{P}) \models_K L$. □

However, the other direction does not hold.

Example 4.4.3. Consider the following program $\mathcal{P} = \{A \leftarrow \neg A \wedge B\}$. Then $wc(\mathcal{P}) = \{A \leftrightarrow \neg A \wedge B\}$ and we have $wc(\mathcal{P}) \models_K \neg B$, but the least Łukasiewicz model $lm(wc(\mathcal{P})) = \langle \emptyset, \emptyset \rangle \not\models \neg B$.

Moreover we have to show a relation between Łukasiewicz consistency and Kleene consistency. Under the Łukasiewicz logic, $wc(\mathcal{P} \cup \mathcal{E})$ is always consistent. However, this is not true under the Kleene logic.

Example 4.4.4 (Weakly Completed Programs, Kleene Inconsistency)**.** Consider the program $\mathcal{P} = \{A \leftarrow \neg A\}$. The weakly completed program is $wc(\mathcal{P}) = \{A \leftrightarrow \neg A\}$. There is no Kleene model of $wc(\mathcal{P})$:

$$\begin{aligned} \langle \emptyset, \emptyset \rangle &\not\models \mathcal{P} \\ \langle \{A\}, \emptyset \rangle &\not\models \mathcal{P} \\ \langle \emptyset, \{A\} \rangle &\not\models \mathcal{P} \end{aligned}$$

Thus, $wc(\mathcal{P})$ is Kleene inconsistent.

Then, we obtain the following result:

Theorem 4.4.5 (Solutions of Abductive Problems under Łukasiewicz and Kleene logic)**.** Let $AP_L = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ and $AP_K = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{K,wc} \rangle$ be two abductive problems. If $\mathcal{E}$ is a solution of AP_L and $wc(\mathcal{P} \cup \mathcal{E})$ is Kleene consistent, then $\mathcal{E}$ is a solution of AP_K.

Proof. If $\mathcal{E}$ is an explanation, then we have $\mathcal{P} \cup \mathcal{E} \models_{L,lm,wc} L$ for all $L \in \mathcal{O}$. This is the case iff $I_L = lm(\mathcal{P} \cup \mathcal{E})$ and $I_L(L) = \top$ for all $L \in \mathcal{O}$. Then, it follows that $wc(\mathcal{P} \cup \mathcal{E}) \models_K L$ for all $L \in \mathcal{O}$ by Proposition 4.4.2. Since $wc(\mathcal{P} \cup \mathcal{E})$ is Kleene consistent we can conclude that, $\mathcal{E}$ is an explanation for $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{K,wc} \rangle$. □

In principle, this means that an explanation under the Łukasiewicz logic is also an explanation under the Kleene logic with the restriction that we do not violate the consistency requirement. This is surprisingly since the two logics behave very differently with respect to the $\leftrightarrow$ junctor.

The question if the Kleene logic or the Łukasiewicz logic is more appropriate to model human reasoning arises. Therefore, consider the following situation:

If the cat is not black, then the cat is black.

If it is dark outside, then the cat is black.

Is it dark or not?

If we translate this natural sentence into the logic program

$$\mathcal{P} = \{black \leftarrow \neg black \wedge dark\}$$

Then we have:

- $\mathcal{P} \not\models_{L,lm,wc} \neg dark$, $\mathcal{P} \not\models_{L,lm,wc} dark$. That means, we do not know, if it is dark or not.
- $\mathcal{P} \models_{K,wc} \neg dark$, i.e. we know that it is not dark.

The conclusion that it is not dark under the Kleene logic, is very surprising. There is actually no hint, that it is dark or not. In this sense, the Kleene logic behave unnatural, in contrast to the Łukasiewicz logic where we actually do not know, if it is dark.

Chapter 5

Complexity of Abduction

In this chapter, we will discuss the complexity of abduction under the least model of a weakly completed logic program with negative facts according to Definition 2.2.1 under the Łukasiewicz semantics. Here, we do not consider integrity constraints.

Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. Then, one might ask, if there exists a solution at all, i.e. if one can explain the observation. Another natural question is, if a statement is relevant: Consider Example 3.1.1, where we explained, why the shoes are wet. One can explain this observation by “It rained last night” and by “The sprinkler was on”. Both explanations are relevant, whereas “Tim played socker” is not relevant. In the case of multiple explanations, one might be interested not only in relevant facts, but on statements that are universally valid. In the above scenario, if one explain why the shoes are wet and answer “It rained last night”, another person can argue that this is not the case, since possibly the sprinkler was on. Hence, one can attack such explanations. The universally valid facts cannot be attacked and we call them necesssary.

These three questions can be formalized as follows:

Consistency Does there exist a minimal solution of AP?

Relevance Given a fact $f \in \mathcal{A}$, does there exists a minimal solution $\mathcal{E}$ such that $f \in \mathcal{E}$?

Necessity Given a fact $f \in \mathcal{A}$, is $f \in \mathcal{E}$ for all minimal solutions $\mathcal{E}$ of AP?

In the first section, the consistency problem is investigated. Afterwards the complexity of relevance and necessity problem are examined. At the end of this chapter, the complexity of skeptical reasoning is analyzed.

5.1 Consistency

A naive algorithm that computes a minimal explanation is shown in Algorithm 1. This algorithm is obviously sound and complete, i.e. a decision procedure for the consistency problem. Verifying that a subset of $\mathcal{A}$ is an explanation and testing whether it is minimal, needs an exponential blowup.

Algorithm 1 NAIVEEXPLANATION $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$

guess a minimal explanation $\mathcal{E}$
compute the set of all interpretations $\mathcal{I}$ for $wc(\mathcal{P} \cup \mathcal{E})$
test whether $I(L) = \top$ for all $L \in \mathcal{O}$, where I is the least interpretation w.r.t. $\mathcal{I}$
for all $\mathcal{E}' \subset \mathcal{E}$ **do**
  test whether $\mathcal{E}'$ is an explanation
end for

In this section, it is shown that consistency is NP-*complete*. Intuitively, this means that we can guess a minimal solution and then this solution can be verified in polynomial time. For this reason, it is shown that computing the least Łukasiewicz model can be computed in polynomial time. Afterwards, it is proven that minimality can be tested in polynomial time. Then, it immediately follows that consistency is in NP. Finally, a reduction from 3SAT to consistency is presented. This shows that consistency is NP-*complete*.

5.1.1 Solution Verification

Let $AP = \langle \mathcal{P}, \mathcal{A}, L, \models_{L,lm,wc} \rangle$ be an abductive problem. The question is, if a set $\mathcal{E}$ is a solution of AP. To verify that $\mathcal{E}$ is a solution, we have to check the following:

1. $\mathcal{E}$ is a consistent subset of $\mathcal{A}$,
2. $wc(\mathcal{P} \cup \mathcal{E})$ is Łukasiewicz consistent and
3. $\mathcal{P} \cup \mathcal{E} \models_{L,lm,wc} L$.

In the following, it is shown that each step can be done in polynomial time.

Lemma 5.1.1 (Subset and Consistency Requirement)**.** Checking, whether a set $\mathcal{E}$ is consistent and a subset of $\mathcal{A}$ can be done in polynomial time.

Proof.

(i) Subset problem: We iterate over all facts $f \in \mathcal{E}$ and then check, if $f \in \mathcal{A}$. This can be easily done in $|\mathcal{E}| \cdot |\mathcal{A}|$ syntactic comparisons.

(ii) Consistent subset problem: Here, we iterate over all positive and negative facts and compare the heads of positive facts with the negative facts. Then, one need $|\mathcal{E}^+| \cdot |\mathcal{E}^-|$ syntactic comparisons.

Hence, both algorithms terminate after polynomial time in the size of $\mathcal{A}$ and $\mathcal{E}$. $\square$

The second condition ($wc(\mathcal{P} \cup \mathcal{E})$ is Łukasiewicz consistent) can be omitted, since there always exists the least Łukasiewicz model of a weakly completed program. In the following, it is shown how the third condition ($\mathcal{P} \cup \mathcal{E} \models_{L,lm,wc} L$) can be verified in polynomial time. For this purpose, the Stenning and van Lambalgen operator is used. Recall Definition 2.5.1:

Let I be an interpretation and $\mathcal{P}$ a logic program. $\Phi_{SvL,\mathcal{P}} = \langle J^\top, J^\perp \rangle$, where

$$
\begin{aligned}
J^\top &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \text{ with } I(\mathit{Body}) = \top\} \\
J^\perp &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \\
&\qquad \text{and for all } A \leftarrow \mathit{Body} \in \mathcal{P} \text{ we find that } I(\mathit{Body}) = \perp\}
\end{aligned}
$$

The relationship between this operator and the least models of a weakly completed program is stated in the following proposition.

Proposition 5.1.2 (Relationship between the Least Łukasiewicz Model and the Stenning and van Lambalgen Operator)**.** Let $\mathcal{P}$ be a logic program. Then I_L is the least fixed point of $\phi_{SvL,\mathcal{P}}$ iff I_L is the least model of $wc(\mathcal{P})$.

Proof. See [Ram09] Corollary 6.16 $\square$

Thus, instead of computing all models and afterwards taking the least model w.r.t. $\preceq$, we can use this operator to obtain the least model. The question is, how many applications of the Stenning and van Lambalgen operator are required to obtain the least fixed point.

Proposition 5.1.3 (Monotonicity of the Stenning and van Lambalgen Operator)**.** Let $\mathcal{P}$ be a logic program. Then $\Phi_{SvL,\mathcal{P}}$ is a monotonic mapping w.r.t. $\preceq$.

Proof. See [Ram09] Proposition 3.2.1 $\square$

This property of the operator roughly states, that the interpretation obtained by $\Phi_{SvL,\mathcal{P}}$ is increasing. Since we consider propositional logic programs, the least fixed point of this operator must be reached in finite time. Then, we obtain the following result:

Proposition 5.1.4 (Computation of the Least Łukasiewicz Model)**.** Computing the least model of a weakly completed program $\mathcal{P}$ under the Łukasiewicz logic can be done in polynomial time.

Proof. Computing the least model of the weak completion can be done by computing the least fixed point of the Stenning and van Lambalgen Operator by Proposition 5.1.2.

We have to show that (i) the least fixed point is reached in polynomially many applications, (ii) the set $J^\top$ can be computed in polynomial time and (iii) the set $J^\perp$ can be computed in polynomial time.

(i) Since this operator is monotone w.r.t. $\preceq$ by Proposition 5.1.3 , we have to iterate the operator n times in the worst case, where n is the number of atoms that occur in $\mathcal{P}$. After this n steps, we reached a fixed point. The number of atoms is clearly polynomial in the size of $\mathcal{P}$.

(ii) To compute the set $J^\top$, we iterate over all rules in $\mathcal{P}$ and check if the body of the rule is mapped to true under I. The second step can be done obviously in polynomial time in the length of the body. Thus, this procedure is in polynomial time.

(iii) To compute the set $J^\perp$, we iterate over all rules $A \leftarrow Body \in \mathcal{P}$ and check if the body of the rule is mapped to false under I. If this is the case, we are iterating over all rules with head A and check if all their bodies are mapped to false. Hence, this procedure is quadratic in the size of $\mathcal{P}$.

Thus, we can conclude that computing the least fixed point of a logic program $\mathcal{P}$ can be done in polynomial time. And so we know that computing the least model of a weakly completed program can be done in polynomial time. □

Proposition 5.1.5 (Solution Verification)**.** Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. Deciding, if $\mathcal{E}$ is an explanation can be done in polynomial time.

Proof. Deciding the first requirement of an explanation ($\mathcal{E}$ is a consistent subset of $\mathcal{A}$) can be done in polynomial time by Lemma 5.1.1. The second requirement ($wc(\mathcal{P} \cup \mathcal{E})$ is Łukasiewicz consistent) can be dropped since there always exist a least model. Finally computing this least model I_L can be done in polynomial time by Proposition 5.1.4. Then, one have to iterate over all literals $L \in \mathcal{O}$: If L is a positive literal, it is checked whether $L \in I^\top$. If L is a negative literal, it is checked whether $L \in I^\perp$. Hence, checking if $\mathcal{E}$ is an explanation can be done in polynomial time. □

5.1.2 Minimality Verification

The remaining problem is the decision, if an explanation $\mathcal{E}$ is minimal. In Algorithm 1, we iterated over all subsets of an explanation and checked, whether this subset is also a solution. However, there are $2^{|\mathcal{E}|} - 1$ strict subsets.

In classical logic, deciding minimality can be done in polynomial time: We are iterating over all $F \in \mathcal{E}$ and test whether $\mathcal{E} \setminus \{F\}$ is an explanation or not. If there is no such explanation, then $\mathcal{E}$ is minimal. Otherwise, $\mathcal{E}$ is not minimal. One reason, why this is correct follows by the fact, that classical logic is monotone (See [HP07] Theorem 5).

Definition 5.1.6 (Monotone Logic)**.** A logic $(\mathcal{L}, \models)$ is called monotone iff $\mathcal{F} \models G$ implies $\mathcal{F} \cup \mathcal{F}' \models G$ for all sets of formulas $\mathcal{F}, \mathcal{F}'$ and a formula G.

If we consider the least model of a weakly completed program under the Łukasiewicz semantics, then we do not have a monotone logic:

Example 5.1.7 (Non-monotonicity of Łukasiewicz Logic)**.** Consider the empty program $\mathcal{P}$ and $G = A \leftrightarrow C$. Then $lm(wc(\mathcal{P})) = \langle \emptyset, \emptyset \rangle \models G$. By adding $A \leftarrow \top$ to $\mathcal{P}$, we have $lm(wc(\mathcal{P} \cup \{A \leftarrow \top\})) = \langle \{A\}, \emptyset \rangle \not\models G$.

However, in the abductive problem, we restrict $\mathcal{F}'$ to be a subset of $abd(\mathcal{P})$, that is, each fact in $\mathcal{F}'$ has an undefined head w.r.t. $\mathcal{P}$. Additionally, G is a literal. With these restrictions, we can adopt the above algorithm to decide minimality although our logic is not monotone. In order to show this, we first prove that adding undefined facts to a program only increase the least model.

Lemma 5.1.8. Let I_1, I_2 be two interpretations, $\mathcal{P}$ a logic program and $\mathcal{F}$ a set of facts where each fact $F \in \mathcal{F}$ has an undefined head w.r.t. $\mathcal{P}$. If $I_1 \preceq I_2$, then $\phi_{SvL,\mathcal{P}}(I_1) \preceq \phi_{SvL,\mathcal{P}\cup\mathcal{F}}(I_2)$

Proof. Consider the following sets:

$$
\begin{aligned}
J_1^\top &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \text{ with } I_1(\mathit{Body}) = \top\} \\
J_2^\top &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \cup \mathcal{F} \text{ with } I_2(\mathit{Body}) = \top\} \\
J_1^\perp &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \\
&\qquad\qquad \text{and for all } A \leftarrow \mathit{Body} \in \mathcal{P} \text{ we find that } I_1(\mathit{Body}) = \perp\} \\
J_2^\perp &= \{A \mid \text{there exists } A \leftarrow \mathit{Body} \in \mathcal{P} \cup \mathcal{F} \\
&\qquad\qquad \text{and for all } A \leftarrow \mathit{Body} \in \mathcal{P} \cup \mathcal{F} \text{ we find that } I_2(\mathit{Body}) = \perp\}
\end{aligned}
$$

We have to show (i) $J_1^\top \subseteq J_2^\top$ and (ii) $J_1^\perp \subseteq J_2^\perp$.

(i) $J_1^\top \subseteq J_2^\top$: Let $A \in J_1^\top$. Then there is $A \leftarrow \mathit{Body} \in \mathcal{P}$ with $I_1(\mathit{Body}) = \top$. Since $\mathit{Body} = B_1 \wedge B_2 \wedge \ldots \wedge B_n \wedge \neg C_1 \ldots C_m$ we know $B_i \in I_1^\top$ for all $1 \leq i \leq n$ and $C_j \in I_1^\perp$ for all $1 \leq j \leq m$. Because $I_1 \preceq I_2$, we know that $I_1^\top \subseteq I_2^\top$ and $I_1^\perp \subseteq J_1^\perp$. Thus, we can conclude $J_2(\mathit{Body}) = \top$. Then, $A \in J_2^\top$.

(ii) $J_1^\perp \subseteq J_2^\perp$: Let $A \in J_1^\perp$. Then there is $A \leftarrow \mathit{Body} \in \mathcal{P}$ and for all $A \leftarrow \mathit{Body} \in \mathcal{P}$ we find that $I_1(\mathit{Body}) = \perp$. Since each fact $F \in \mathcal{F}$ has an undefined head H w.r.t. $\mathcal{P}$, we have $H \neq A$, i.e. we did not add relevant rules. Since we have for all $A \leftarrow \mathit{Body} \in \mathcal{P}$ that $I_1(\mathit{Body}) = \perp$ and $\mathit{Body} = B_1 \wedge B_2 \wedge \ldots \wedge B_n \wedge \neg C_1 \ldots \neg C_m$ there is either an $1 \leq i \leq n$ with $B_i \in I_1^\perp$ or there is a $1 \leq j \leq m$ with $C_j \in I_1^\top$. Since $I_1 \preceq I_2$, we have that $B_i \in I_2^\perp$ or $C_j \in I_2^\top$. Thus, $I_2(\mathit{Body}) = \perp$ for all $A \leftarrow \mathit{Body}$ as well. Hence, $A \in J_2^\perp$.

□

The above lemma can be generalized to several applications of the Stenning and van Lambalgen operator.

Proposition 5.1.9. Let $\mathcal{P}$ be a logic program and $\mathcal{F}$ be a set of facts where each fact has an undefined head w.r.t. $\mathcal{P}$. Then, the following holds: $lfp(\phi_{SvL,\mathcal{P}}) \preceq lfp(\phi_{SvL,\mathcal{P}\cup\mathcal{F}})$

Proof. By Proposition 2.5.3, the least fixed point is computed by repeated applications of the Stenning and van Lambalgen Immediate Consequence Operator, starting with the empty interpretation. Thus, there are n, m such that $I = lfp(\phi_{SvL,\mathcal{P}}) = \phi^n_{SvL,\mathcal{P}}(\langle\emptyset,\emptyset\rangle)$ and $J = lfp(\phi_{SvL,\mathcal{P}\cup\mathcal{F}}) = \phi^m_{SvL,\mathcal{P}\cup\mathcal{F}}(\langle\emptyset,\emptyset\rangle)$. Let $k = max(n,m)$. Since I and J are fixed points, we have: $I = \phi^k_{SvL,\mathcal{P}}(\langle\emptyset,\emptyset\rangle)$ and $J = \phi^k_{SvL,\mathcal{P}\cup\mathcal{F}}(\langle\emptyset,\emptyset\rangle)$. Since the Stenning and van Lambalgen Operator is monotone by Proposition 5.1.3, we have that $I_0 \preceq I_1 \preceq \ldots \preceq I_k$ and also $J_0 \preceq J_1 \preceq \ldots \preceq J_k$ Because they both starts with the empty interpretation, we have the following:

$$\begin{aligned}
I_0 = \langle\emptyset,\emptyset\rangle &\preceq \langle\emptyset,\emptyset\rangle = J_0\\
I_1 = \Phi_{SvL,\mathcal{P}}(I_0) &\preceq \Phi_{SvL,\mathcal{P}\cup\mathcal{F}}(J_0) = J_1 \text{ by Lemma 5.1.13}\\
I_2 = \Phi_{SvL,\mathcal{P}}(I_1) &\preceq \Phi_{SvL,\mathcal{P}\cup\mathcal{F}}(J_1) = J_2 \text{ by Lemma 5.1.13}\\
&\vdots\\
I_k = \Phi_{SvL,\mathcal{P}}(I_{k-1}) &\preceq \Phi_{SvL,\mathcal{P}\cup\mathcal{F}}(J_{k-1}) = J_k \text{ by Lemma 5.1.13}
\end{aligned}$$

Hence, we have $lfp(\Phi_{SvL,\mathcal{P}}) \preceq lfp(\Phi_{SvL,\mathcal{P}\cup\mathcal{F}})$. □

An easy consequence is that explanations are *monotone*.

Theorem 5.1.10 (Monotonicity of Explanations)**.** Let $\langle\mathcal{P},\mathcal{A},\mathcal{O},\models_{L,lm,wc}\rangle$ be an abductive problem. If $\mathcal{E}$ is an explanation, then any consistent set $\mathcal{E}'$ such that $\mathcal{E} \subseteq \mathcal{E}' \subseteq \mathcal{A}$ is a solution.

Proof. Since $\mathcal{E}' \subseteq \mathcal{A}$ and $\mathcal{E}'$ is consistent, we have to show that $I = lm_L(wc(\mathcal{P} \cup \mathcal{E}')) \models L$ for all $L \in \mathcal{O}$. Note that $\mathcal{E} \cap \mathcal{E}'$ only contains facts with an undefined head with respect to $\mathcal{P}$ and $\mathcal{E}$. By Proposition 5.1.9, it follows that $J = lfp(\Phi_{SvL,\mathcal{P}\cup\mathcal{E}}) \preceq lfp(\Phi_{SvL,\mathcal{P}\cup\mathcal{E}\cup(\mathcal{E}\cap\mathcal{E}')}) = lfp(\Phi_{SvL,\mathcal{P}\cup\mathcal{E}'}) = I$. Since $J \preceq I$ and $J(L) = \top$ for all $L \in \mathcal{O}$, it immediately follows that $\mathcal{E}'$ is an explanation. □

This result is surprising: Although the Łukasiewicz logic is not monotone, explanations are monotone. This means that one can safely extend an explanation by further non contradictory facts. The reason, why we obtain this property, is that we required that explanations cannot be further explained and that explanations do not contain contradictory facts.

Moreover, with this relation between the least models of the original program $\mathcal{P}$ and the extended program, we can decide minimality.

Proposition 5.1.11 (Characterization of Minimality)**.** Let $\langle\mathcal{P},\mathcal{A},\mathcal{O},\models_{L,lm,wc}\rangle$ be an abductive problem. An explanation $\mathcal{E}$ is minimal iff there is no $f \in \mathcal{E}$ such that $\mathcal{E} \setminus \{f\}$ is an explanation.

Proof.

"→" This part is trivial since minimality is defined that no strict subset of $\mathcal{E}$ is an explanation.

"←" Suppose, $\mathcal{E}$ is not a minimal. Thus there is $\mathcal{E}' \subset \mathcal{E}$ such that $\mathcal{E}'$ is an explanation. Since, $\mathcal{E}' \subset \mathcal{E}$, there is $f \in \mathcal{E}, f \notin \mathcal{E}'$. By Theorem 5.1.10 we obtain the following: $I_L = lfp(\Phi_{SvL,\mathcal{P}\cup\mathcal{E}'}) \preceq lfp(\Phi_{SvL,\mathcal{P}\cup(\mathcal{E}\setminus\{f\})}) = J_L$. Since $I_L(L) = \top$ for all $L \in \mathcal{O}$ and $I \preceq J$, we have $J_L(L) = \top$ for all $L \in \mathcal{O}$. Hence, $\mathcal{E} \setminus \{f\}$ is an explanation, which is a contradiction. Thus, $\mathcal{E}$ is minimal.

□

Moreover, one can implement the procedure described in Proposition 5.1.11.

Proposition 5.1.12 (Deciding Minimality)**.** Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc}\rangle$ be an abductive problem and $\mathcal{E}$ an explanation. Deciding, whether $\mathcal{E}$ is minimal can be done in polynomial time.

Proof. We are iterating over all elements $F \in \mathcal{E}$ and then check whether $\mathcal{E} \setminus \{F\}$ is an explanation or not. If there is no such a F, it follows that $\mathcal{E}$ is minimal. The iteration is finished after $|\mathcal{E}|$ steps and checking whether $\mathcal{E}\setminus\{F\}$ is an explanation can be done in polynomial time by Proposition 5.1.5. Hence, testing minimality can be done in polynomial time □

5.1.3 NP-Membership and NP-Hardness

In the previous section, a decision procedure was described that decides, if $\mathcal{E}$ is a minimal solution of an abductive framework. In this section, the complexity of the the consistency problem is analyzed.

Proposition 5.1.13 (Consistency is in NP)**.** Deciding whether $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc}\rangle$ is consistent is in NP.

Proof. Consider the binary relation $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc}\rangle \sim \mathcal{E}$ if and only if $\mathcal{E}$ is a minimal solution of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc}\rangle$.

In order to show that the problem is in NP, we will show that $\sim$ is balanced and decidable in polynomial time on a deterministic Turing Machine. Afterwards, it is shown: There exists $\mathcal{E}$ such that $AP \sim \mathcal{E}$ iff AP is consistent. This means, that we guess a solution of the abductive problem and then check, whether this solution is correct in polynomial time.

(i) $AP \sim \mathcal{E}$ is balanced: This is clear, since every explanation is restricted to be a subset of $\mathcal{A}$.

(ii) $AP \sim \mathcal{E}$ is decidable in polynomial time: By Proposition 5.1.5, it can be checked in polynomial time whether $\mathcal{E}$ is a solution of AP. Furthermore, deciding, if $\mathcal{E}$ is minimal can be done in polynomial time by Proposition 5.1.12. Hence, one can decide $\sim$ in polynomial time.

(iii) AP is consistent iff there is $\mathcal{E}$ such that $AP \sim \mathcal{E}$: This is clear, since in this way, $\sim$ is defined.

By Proposition 2.6.2, we obtain that consistency is in NP. □

Moreover, the consistency-problem is one of the hardest problems in NP.

Proposition 5.1.14 (Consistency is NP-*hard*)**.** Deciding whether $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ is consistent is NP-*hard*.

Proof. First, we show that 3SAT can be polynomially reduced to consistency.

Consider the following transformation:

Let $F = C_1 \wedge \ldots \wedge C_n$ be a 3SAT instance and $X_1 \ldots X_m$ the variables occuring in F. Then, the abductive problem is obtained as follows:

$$\begin{aligned} AP &= \langle \mathcal{P}, \{X_i \leftarrow \top, X_i \leftarrow \bot \mid 1 \leq i \leq m\}, O, \models_{L,lm,wc} \rangle \\ \mathcal{P} &= \{Y_i \leftarrow L_{i,1}, Y_i \leftarrow L_{i,2}, Y_i \leftarrow L_{i,3} \mid \text{ for each clause } C_i = L_{i,1} \vee L_{i,2} \vee L_{i,3}\} \\ &\cup \{O \leftarrow Y_1 \wedge \ldots \wedge Y_n\} \end{aligned}$$

We have to show that this transformation is polynomial-time computable and, F is satisfiable iff AP is consistent.

(i) This transformation is polynomial-time computable: This is clear, since we create for each disjunction in F three clauses in P and for each variable in F two abducibles.

(ii) If AP is consistent, then F is satisfiable: Let $\mathcal{E}$ be a minimal solution of AP, i.e. we have $I_L = lm(wc(\mathcal{P} \cup \mathcal{E}))$ and $I_L(O) = \top$. Thus, $I_L(Y_i) = \top$ and $I_L(L_{i,1} \vee L_{i,2} \vee L_{i,3}) = \top$ for all $1 \leq i \leq n$. The two valued interpretation $\overline{I}$ can be obtained as $\overline{I^\top} = I^\top \cup \{A \mid A \notin I^\bot \text{ and } A \text{ occurs in } F\}$. We have to show that $\overline{I}$ is a model of F. Since $I_L(L_{i,1} \vee L_{i,2} \vee L_{i,3}) = \top$ for all $1 \leq i \leq n$, we know that there is a $j \in \{1, 2, 3\}$ such that $I_L(L_{i,j}) = \top$. Because $I \preceq \overline{I}$, we know $\overline{I}(L_{i,j}) = \top$ and thus $\overline{I} \models C_i$ for all $1 \leq i \leq n$. Then, $\overline{I}$ is a model of F and F is satisfiable.

(iii) If F is satisfiable, then AP is consistent: There is two-valued interpretation I such that $I \models F$. This interpretation can be seen as three-valued and is used to construct an explanation:

$$\mathcal{E} = \{A \leftarrow \top \mid A \in I^\top\} \cup \{A \leftarrow \bot \mid A \in I^\bot\}$$

Let J be an interpretation such that $J = lm_L(wc(\mathcal{P} \cup \mathcal{E}))$. We have to show that $J(O) = \top$. This is the case iff $J(Y_i) = \top$ for all $1 \leq i \leq n$. Assume this is not the case, i.e. we find $1 \leq j \leq n$ such that $J(Y_j) \neq \top$. Then the literals $L_{j,1}, L_{j,2}, L_{j,3}$ are not mapped to true under J. Hence, there was no fact stating that $L_{j,i}$ must be mapped to true for $i \in \{1, 2, 3\}$ since $L_{j,i}$ is undefined w.r.t. $\mathcal{P}$. We can conclude that then $I \not\models L_{i,j}$ for $i \in \{1, 2, 3\}$. Thus, I is not a model of the clause C_j and it follows that $I \not\models F$. This is a contradiction. Hence, $J(Y_j) = \top$ for all $1 \leq j \leq n$ and it immediately follows that $J(O) = \top$ since $O \leftarrow Y_1 \wedge \ldots \wedge Y_m \in \mathcal{P}$. Then, $\mathcal{E}$ is an explanation. Because $\mathcal{E}$ is an explanation, it must follow that there exists a minimal explanation. Hence, AP is consistent.

$\square$

Finally, we obtain the following result.

Theorem 5.1.15. Deciding whether an abductive problem $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ is consistent is NP-*complete*.

Proof. This is a direct consequence, since it is already shown that the consistency-problem is in NP by Proposition 5.1.13 and is NP-*hard* by Proposition 5.1.14. Thus, consistency is NP-*complete*. □

Theorem 5.1.16. Deciding whether $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ has no solution is coNP-*complete*.

Proof. Since consistency is NP-*complete* by Theorem 5.1.15, it immediately follows that inconsistency is coNP-*complete* by Proposition 2.6.11. □

5.2 Relevance

In this section, the complexity of deciding whether a fact f is relevant, is discussed. More formally: Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem, does there exist a minimal explanation $\mathcal{E}$ such that $f \in \mathcal{E}$?

Intuitively, this problem is not harder than consistency. The idea is that one have to simply guess a minimal explanation that contains f.

Theorem 5.2.1 (Relevance is in NP)**.** Deciding whether f is relevant in $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ is in NP.

Proof. Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$. Consider the binary relation $AP \sim \mathcal{E}$ if and only if $\mathcal{E}$ is a minimal solution of AP and $f \in \mathcal{E}$.

In order to show that the problem is in NP, we will show that $\sim$ is balanced and decidable in polynomial time on a deterministic Turing Machine. Afterwards, it is shown: If there exists $\mathcal{E}$ such that $AP \sim \mathcal{E}$, then f is relevant in AP. This means, that we guess a minimal solution containing f and then check, whether this solution is correct in polynomial time.

(i) $AP \sim \mathcal{E}$ is balanced: This is clear, since every explanation is restricted to be a subset of $\mathcal{A}$.

(ii) $AP \sim \mathcal{E}$ is decidable in polynomial time: By Proposition 5.1.5, it can be checked in polynomial time whether $\mathcal{E}$ is a solution of AP. Furthermore, deciding, if $\mathcal{E}$ is minimal can be done in polynomial time by Proposition 5.1.12. Checking whether $f \in \mathcal{E}$ can obviously be done in polynomial time by iterating every element in $\mathcal{E}$ and then checking if this element is equal to f.

(iii) f is relevant in AP iff there is $\mathcal{E}$ such that $AP \sim \mathcal{E}$: This is clear, since we defined $\sim$ in this way.

By Proposition 2.6.2, we obtain that relevance is in NP. □

The NP-*hardness* proof of consistency cannot be directly copied in order to show NP-*hardness* of relevance. The reason is that in the proof of Proposition 5.1.14, one explanation was created by a model of a 3SAT-formula. This explanation must not be minimal. However, it is clear that a minimal explanation exists. Here, one have to construct a minimal explanation that contains a specific fact.

5.3 Necessity

Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. Of particular interest is the necessity-problem, i.e. is a fact f part of every minimal solution? In order to decide this question, one can enumerate all minimal solutions and then verify if f is in all minimal solutions. However, there may be exponentially many minimal solutions in the size of $\mathcal{A}$:

Proposition 5.3.1 (Sperner's Theorem)**.** The maximum cardinality of a collection of subsets of a set S, none of which contains another, is $\binom{n}{\lfloor \frac{n}{2} \rfloor}$, where $n = \mid S \mid$.

Proof. See [Spe28]. $\square$

This gives an upper bound on the number of minimal explanations.

Proposition 5.3.2 (Upper Bound on the Number of Minimal Explanations)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. Then, there exists at most $\binom{n}{\lfloor \frac{n}{2} \rfloor}$ minimal explanations.

Proof. Let S denote the set of all minimal solutions of AP and $n = |\mathcal{A}|$. Suppose, there are more minimal explanations, i.e. $|S| = \binom{n}{\lfloor \frac{n}{2} \rfloor} + x$, where $x \geq 1$. By Proposition 5.3.1, one find $\mathcal{E}, \mathcal{E}' \in S$ such that $\mathcal{E} \subset \mathcal{E}'$. However, then $\mathcal{E}'$ is not minimal. Then, S contains solutions that are not minimal, which is a contradiction. Hence, there cannot be more minimal solutions than $\binom{n}{\lfloor \frac{n}{2} \rfloor}$. $\square$

In order to decide necessity, one can use a different way. It turns out, that the inconsistency problem and necessity are equivalent with respect to polynomial time reductions.

Lemma 5.3.3 (Relationship Inconsistency and Necessity)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. The problem of deciding inconsistency is polynomial reducible to necessity.

Proof. Let $AP' = \langle \mathcal{P}, \mathcal{A} \cup \{q\}, \mathcal{O}, \models_{L,lm,wc} \rangle$, where q is a fresh atom.

(i) This transformation can be computed in polynomial time: This is clear, since we only add an atom to $\mathcal{A}$.

(ii) If AP is inconsistent, then q is necessary in AP': If AP is inconsistent, then there does not exist a minimal solution. Moreover, every solution of AP is a solution of AP'. Since AP is inconsistent, there does not exist a solution of AP' and thus $q \in \mathcal{E}$ for all solutions of AP'.

(iii) If q is necessary in AP', then AP is inconsistent: Note that q is a fresh atom. Thus, q does not distribute to any solution of AP'. Then, q is necessary iff AP' is inconsistent. Since every solution of AP' is also a solution of AP, it follows that AP must be inconsistent.

□

The converse direction also holds:

Lemma 5.3.4 (Relationship Necessity and Inconsistency)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. The problem of deciding necessity is polynomial reducible to inconsistency.

Proof. Let $AP' = \langle \mathcal{P}, \mathcal{A} \setminus \{q\}, O, \models_{L,lm,wc} \rangle$

(i) This reduction can be obviously computed in polynomial time, since we only removed the fact q.

(ii) If q is necessary in AP, then AP' is inconsistent: Since $q \in \mathcal{E}$ for all minimal solutions of AP, it follows that there does not exists a minimal solution not containing q. Then, it immediately follows that AP' must be inconsistent.

(iii) If AP' is inconsistent, then q is necessary in AP: Suppose, q is not necessary in AP, i.e. there is a minimal explanation $\mathcal{E}$ of AP such that $q \notin \mathcal{E}$. However, then $\mathcal{E}$ must be a minimal explanation of AP', which is a contradiction. Then, there does not exist such explanations. Thus, $q \in \mathcal{E}$ for all minimal explanations of AP.

□

Proposition 5.3.5 (Equivalence of Inconsistency and Necessity)**.** Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. The problems inconsistency and necessity are equivalent w.r.t. polynomial-time reductions.

Proof. By Lemma 5.3.3 and Lemma 5.3.4, it immediately follows that both are equivalent w.r.t polynomial-time reductions. □

Hence, we obtain the following:

Theorem 5.3.6. Let $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive problem. Then, deciding, if $f \in \mathcal{A}$ is necessary, is coNP-*complete*.

Proof. The inconsistency-problem is coNP-*complete* by Theorem 5.1.16. Since inconsistency and necessity are equivalent w.r.t. polynomial time reductions by Proposition 5.3.5, it follows by Proposition 2.6.10 that necessity is coNP-*complete*. □

5.4 Skeptical Reasoning

In this section, complexity of skeptical reasoning is investigated. Recall the definition of skeptical reasoning:

Definition 5.4.1 (Skeptical Reasoning)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models \rangle$ be an abductive problem and F a formula.

Then F follows *skeptically* by AP, denoted by $AP \models_s F$, iff

1 AP is consistent and

2 F is a *universal consequence* of AP, i.e. for all minimal explanations $\mathcal{E}$ of AP we find that $I = lm_L(wc(\mathcal{P} \cup \mathcal{E}))$ and $I_L(F) = \top$.

Consistency is already shown to be NP-*complete*. It is clear, that the second condition is coNP-*hard* by a reduction from necessity: A fact $A \leftarrow \top$ ($A \leftarrow \bot$) is necessary iff A ($\neg A$) is a universal consequence. In order to show coNP-*completeness*, we show that the complement is in NP. Then it is proven that the complement is NP-*hard* by a reduction from consistency.

Lemma 5.4.2 (NP-Membership)**.** Deciding whether a formula F is not a universal consequence of $\langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ is in NP.

Proof. Let $AP = \langle \mathcal{P}, \mathcal{A}, L, \models_{L,lm,wc} \rangle$ be an abductive problem and F a formula. Consider the binary relation $AP \sim \mathcal{E}$ iff $\mathcal{E}$ is a minimal solution of AP such that $\mathcal{P} \cup \mathcal{E} \not\models_{L,lm,wc} F$.

(i) $AP \sim \mathcal{E}$ is balanced: This is clear, since every explanation is restricted to be a subset of $\mathcal{A}$.

(ii) $AP \sim \mathcal{E}$ is decidable in polynomial time: By Proposition 5.1.5, it can be checked in polynomial time whether $\mathcal{E}$ is a solution of AP. Then, the model $I = lm_L(wc(\mathcal{P} \cup \mathcal{E}))$ is constructed. Deciding, if $I_L(F) \neq \top$ can be done in polynomial time in the length of the formula. Furthermore, deciding, if $\mathcal{E}$ is minimal can be done in polynomial time by Proposition 5.1.12.

(iii) F is not a universal consequence of AP iff there exists $\mathcal{E}$ such that $AP \sim \mathcal{E}$: This is clear, since in this way, $\sim$ is defined.

By Proposition 2.6.2, we know: Deciding whether F is not a universal consequence of AP is in NP. □

Next, it is shown that this problem is NP-*complete*.

Lemma 5.4.3 (NP-*completeness*)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ an abductive problem and F be a formula. Deciding, if F is not a universal consequence of AP is NP-*complete*.

Proof. Since Lemma 5.4.2, we already have that this problem is in NP. NP-*hardness* is shown by a straightforward reduction from consistency. Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive framework and $F = \bot$. Then, it is easy to see that AP is consistent iff there exists a minimal solution $\mathcal{E}$ of AP, such that $\mathcal{P} \cup \mathcal{E} \not\models_{L,lm,wc} \bot$, i.e. $\bot$ is not a universal consequence of AP. We can conclude, that that this problem is NP-*complete*. $\square$

The complement of this problem, i.e. the original second condition in skeptical reasoning, is then coNP − *complete*.

Proposition 5.4.4. Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models \rangle$ be an abductive problem and F a formula. Deciding, if F is a universal consequence of AP is coNP-*complete*.

Proof. The opposite problem is shown to be NP-*complete* by Lemma 5.4.3. By Proposition 2.6.11, we can conclude that the above problem is coNP-*complete*. $\square$

Then, we know that the first problem of skeptical reasoning is NP-*complete*, whereas the second problem is coNP-*complete*. The complexity class DP contains such merged problems.

Proposition 5.4.5 (DP-Membership)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive framework and F be a formula. Deciding, whether $AP \models_s F$ holds, is in DP.

Proof. Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive framework and F a formula. Then $AP \models_s F$ iff 1) AP is consistent and 2) F is a universal consequence of AP. It is already shown that consistency is in NP by Proposition 5.1.13 and the universal consequence problem is in coNP by Proposition 5.4.4. Hence, deciding $\models_s$ is in DP. $\square$

Proposition 5.4.6 (DP-*hardness*)**.** Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive framework and F be a formula. Deciding, whether $AP \models_s F$ holds, is DP-*hard*.

Proof. Let P be a decision problem in DP. P consists of two decision problems P_1 and P_2, where $P_1 \in$ NP and $P_2 \in$ coNP by the definition of the class DP. Since consistency is NP − *complete*, we know that P_1 is polynomially reducible to consistency. Since the universal consequence-problem is coNP-*complete*, we know that P_2 is polynomially reducible to the universal consequence problem. Hence, P can be polynomially reduced to the combined problem: Is AP consistent and is F a universal consequence of AP? Hence, $\models_s$ is DP-*hard*. $\square$

Theorem 5.4.7. Let $AP = \langle \mathcal{P}, \mathcal{A}, \mathcal{O}, \models_{L,lm,wc} \rangle$ be an abductive framework and F be a formula. Deciding, whether $AP \models_s F$ holds, is DP-*complete*.

Proof. Deciding $\models_s$ is in DP by Proposition 5.4.5 and is DP-*hard* by Proposition 5.4.6. Thus, we obtain DP-*completeness*. $\square$

In [EGL98] Eiter, Gottlob and Leone analyzed the complexity of consistency, relevance and necessity. In contrast to this thesis, they used logic programs with no negative facts, no form completion and did not restrict the abducibles to be basic. Despite these differences, the proof idea of Theorem 5.3.6 is the same as in [EGL98, Proposition 6]. However, here it is proven that relevance is in NP, whereas in [EGL98] it is shown that relevance is in NP^{NP}. This means that a better upper bound is shown in this thesis. However, if one use classical propositional logic and Horn clauses, we obtain very similar results: Namely that consistency is NP-*complete*, relevance is NP-*complete*, necessity is coNP-*complete* [EEGG93]. A horn clause is a clause of the form $L_1 \vee \dots L_n$, where at most one literal is positive. If we instead allow arbitary formulas, it is shown that consistency and relevance is DP-*complete* [EEGG93].

These results are interesting, since the complexity classes of Horn clauses under classical logic coincides with the complexity classes obtained in this chapter.

Chapter 6

Conclusion

Hölldobler and Kencana Ramli promoted the Łukasiewicz logic and weak completion to model human (deductive) reasoning [HR09]. They argue that this logic is indeed adequate to model six examples by Byrne. In this thesis, it is shown that the remaining six examples by Byrne can be modeled by abductive frameworks and skeptical reasoning. In fact, weak completion must be used in order to get correct results with respect to Byrne's data. The form of abductive reasoning presented in this thesis is more powerful than the system promoted by Stenning and van Lambalgen: They noticed that their system fails to compute consequences by multiple explanations, where the formalism we use can indeed handle multiple solutions. In [DSSd00] an experiment was investigated where subjects are allowed to give compound answers like $p \vee q$. In principle, the proposed formalism is able to decide such compound answers. The remaining question is, if these experiments can be modeled in our formalism.

Moreover, the formalism is also adequate to model deductive reasoning, which shows that it is indeed a natural extension. Then, the twelve experiments by Byrne can be modeled in such abductive frameworks.

Furthermore, the relationship between the Łukasiewicz's, Kleene's and Fitting's logic under completed and weakly completed logic programs were investigated. It turned out, that explanations under the Łukasiewicz and weak completion are also valid explanations under the Kleene logic, if there exists a Kleene model of $wc(\mathcal{P} \cup \mathcal{E})$. This is interesting, since both logic are much different w.r.t. the $\leftrightarrow$ junctor. In this sense, the Łukasiewicz logic "simulates" models under the Kleene logic. It is proven that explanations under completed logic programs can only contain positive facts. Moreover, one can easily construct explanations under completed logic programs, if we already know the explanation under weakly completed logic programs in the Łukasiewicz logic. The Fitting logic is shown to be equivalent to the Łukasiewicz logic, if a form of completion is used. Otherwise it behaves like the Kleene logic.

However, it is not clear, which logic - the Łukasiewicz or the Kleene logic - is more appropriate to model abductive reasoning. Since an abductive framework relies on deductive reasoning, a natural sentence is given, where the Łukasiewicz and Kleene logic behave differently:

If the cat is not black, then the cat is black.

If it is dark outside, then the cat is black.

Is it dark or not?

If the Kleene logic is used, then one can infer that it is not dark, whereas in the Łukasiewicz logic, it is unknown whether it is dark or not. It is argued that the Łukasiewicz logic seems to better suited than the Kleene logic in order to model human reasoning. In analogy to deductive reasoning, it follows that the Łukasiewicz logic and weak completion is adequate to model human abductive reasoning.

A noticeable property of this logic is that explanations are monotone, although the logic itself is not monotone.

Moreover, four examples were modified in the way that integrity constraints were used. Then, two different semantics of integrity constraints were presented and contrasted, the theorem-hood view and the satisfiability view. Consider the situation:

If Marian has an essay to write, she will study late in the library.

If the library stays open, she will study late in the library.

The library is not open in holidays.

There are holidays.

She will not study late in the library.

In the theorem-hood view the only minimal explanation is that the library is not open, whereas "She has not an essay to write." does not satisfy the integrity constraints. This seems unnatural, since this is actually one possible reason.

At the end, the complexity of abduction with no integrity constraints under the Łukasiewicz logic and weak completion was determined. In particular, the consistency problem is shown to be NP-*complete*, inconsistency coNP-*complete*, relevance in NP and necessity is coNP-*complete*. Finally, skeptical reasoning is proven to be DP-*complete*. These classes corresponds to abduction with horn clauses under classical logic, where the set of abducibles is not restricted [EEGG93].

Future Work The exact complexity class of relevance was not given in this work. However, there is evidence that relevance is NP-*complete*.

In general, the question arises, how integrity constraints change the complexity of abduction. Usually, the form of integrity constraints are restricted to $\bot \leftarrow B_1 \wedge \ldots B_n \wedge \neg C_1 \wedge \ldots \wedge C_m$. If such a restriction is used, it is easy to see that integrity constraints under the two semantics should not change the complexity of abductive tasks. The reason is that the theoremhood-view only increases minimal explanations while the monotonicity property is preserved. In contrast to this, the satisfiability-view gives an upper bound on explanations. Although this semantic destroys the monotonicity property, the minimality criterion in Proposition 5.1.11 remains correct. A formal proof is left open.

Stenning and van Lambalgen argued in [SvL08], that classical logic is inadequate to model human reasoning. One argument is that computing logical consequences is too expensive from a computational point of view, since one have to reason about all models. However, in the definition of skeptical reasoning, we do something similar, namely we reason with respect to all explanations. In this sense, this definition is inadequate. An approximation of skeptical reasoning using a unique least explanation, if it exists, seems to be a natural choice, if one one follow the argumentation of Stenning and van Lambalgen.

In their work, it is shown how to compute efficiently least models with neural networks. At the moment, it is not clear, how one can model abductive reasoning in neural networks.

They also considered reasoning by autistic persons in their work and presented the semantics in neural networks. It is left open, if there exists an immediate consequence operator that models the semantics of such neural networks.

Moreover, a generalization of the least model semantics to stable models is interesting. The approach by Wernhard in [Wer10], where different semantics of logic programs can be "reconstructed" using scope-wise circumscription and projection, seems to be a good candidate to give stable model semantics.

Bibliography

[Byr89] Ruth M. J. Byrne. Suppressing valid inferences with conditionals. *Cognition*, 31(1):61 – 83, 1989.

[Cla78] K.L. Clark. Negation as failure. *Logic and data bases*, pages 293–322, 1978.

[DSSd00] K. Dieussaert, W. Schaeken, W. Schroyen, and G. dYdewalle. Strategies during complex conditional inferences. *Thinking and reasoning*, 6(2):125–161, 2000.

[EEGG93] Thomas Eiter, Thomas Eiter, Georg Gottlob, and Georg Gottlob. The complexity of logic-based abduction, 1993.

[EGL98] Thomas Eiter, Georg Gottlob, and Nicola Leone. Abduction from logic programs: Semantics and complexity. *Theoretical Computer Science*, 189:129–177, 1998.

[Fit85] M. Fitting. A Kripke-Kleene semantics for logic programs. *JLP*, 2(4):295–312, 1985.

[HP07] Miki Hermann and Ecole Polytechnique. Counting complexity of propositional abduction. In *In Proc. 20th International Joint Conference on Artificial Intelligence (IJCAI07)*, pages 417–422, 2007.

[HR09] Steffen Hölldobler and Carroline Dewi Kencana Ramli. Logics and networks for human reasoning. In *ICANN '09: Proceedings of the 19th International Conference on Artificial Neural Networks*, pages 85–94, Berlin, Heidelberg, 2009. Springer-Verlag.

[KKT98] A. C. Kakas, R. A. Kowalski, and F. Toni. The role of abduction in logic programming, January 18 1998.

[Luk70] Jan Lukasiewicz. *Selected works (Studies in logic and the foundations of mathematics).* North-Holland Pub. Co, 1st edition, 1970.

[Pap93] Christos H. Papadimitriou. *Computational Complexity.* Addison Wesley, 12 1993.

[Pie] C.S. Pierce. Collected papers of charles sanders pierce. *Volume II, Elements of Logic. Harvard.*

[Ram09] Carroline Dewi Puspa Kencana Ramli. Logic Programs and Three-Valued Consequence Operators. Master's thesis, International Center for Computational Logic, Technische Universitat Dresden, 2009.

[Spe28] E. Sperner. Ein Satz über Untermengen einer endlichen Menge. *M. Z.*, 27:544–548, 1928.

[SvL08] Keith Stenning and Michiel van Lambalgen. *Human Reasoning and Cognitive Science (Bradford Books)*. The MIT Press, 1 edition, 8 2008.

[Wer10] Christoph Wernhard. Circumscription and projection as primitives of logic programming. In *Technical Communications of the 26th International Conference on Logic Programming, ICLP'10*, volume 7 of *Leibniz International Proceedings in Informatics (LIPIcs)*. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2010.