# Abduction in Nonmonotonic Theories

Claudio Delrieux
Universidad Nacional del Sur, Bahia Blanca - ARGENTINA - `claudio@acm.org`

**Abstract**. In a very broad sense we can state that abduction is the inference process that goes from observations to explanations within a more general context or theoretical framework. There are good characterizations of abduction of surprising observations in monotonic theories. Also, in normal logic programs there are a tight relation among SLDNF and the abduction of negative literals. However, a problem that faces abduction is the explanation of *anomalous* observations, *i. e.*, observations that are contradictory with the current theory. For this reason, in this work we will consider the problem of characterizing abduction in nonmonotonic theories. Our inference system is based on a natural deduction presentation of the implicational segment of a relevant logic, much similar to the $R_{\rightarrow}$ system of Anderson and Belnap. Then we will discuss some issues arising the pragmatic acceptance of abductive inferences in nonmonotonic theories. Finally, we show how our system can accommodate anomalous observations and characterize all the possible outcomes that a nonmonotonic theory may face when confronted with new evidence.

**Keywords**: Knowledge Representation and Reasoning, Logic and Inference, Relevance, Defeasible Reasoning

## 1 Introduction

Abduction plays a central rôle in many applications, such as diagnosis, expert systems, and causal reasoning [11, 20]. In a very broad sense we can state that abduction is the inference process that goes from observations to explanations within a more general context or theoretical framework. That is to say, abductive inference looks for sentences (named *explanations*), which, added to the theory, enable deductions for the observations. Most of the times there are several such explanations for a given observation. For this reason, in a narrower sense, abduction is regarded as an inference to the best explanation. There are good characterizations of abduction of surprising observations in monotonic theories [11, 12]. In normal logic programs there is a tight relationship between SLDNF and the abduction of negative literals [10]. This relation can be generalized to the SLDNF inference of evidence sets [1, 6]. However, a problem that faces abduction is the explanation of *anomalous* observations, *i. e.*, observations that are contradictory with the current theory. It is perhaps impossible to perform such inferences in monotonic theories. For this reason, in this work we will consider the problem of characterizing abduction in nonmonotonic theories.

In this work we will consider *reasoning* as the process of theory construction, and we will try to characterize the kind of reasoning that arises when surprising or anomalous observations are suddenly available. Given a theory $\mathcal{T}$ and an observed evidence $e$, then $e$ is surprising if neither $\mathcal{T} \vdash e$, nor $\mathcal{T} \vdash \neg e$, and $e$ is anomalous if $\mathcal{T} \vdash \neg e$. The first situation has received considerable interest since Peirce, who coined the word *abduction*, and characterized it as the third member of the triad of syllogistic reasoning (together with deduction and induction)[1]. The second situation (abduction of anomalous observations) has received only occasional attention. It is a well established fact that monotonic theories cannot accommodate anomalous observations. For this reason, research in this direction must focus in abduction in nonmonotonic theories. For example, Boutillier and Becher [3] propose a belief revision approach based in the notion of proximity in modal frameworks. In this approach, abduction of an anomalous observation is the process of finding the least abnormal possible world given the actual world and the new observation.

Poole [17, 18] proposes a logical framework with interests and ingredients very similar to ours, but he regards theories in which contradiction arises, as erroneous or at least incomplete. In Poole, a theory with multiple extensions (incompatible conclusions or observations) can arise only if there is missing an explicit exception for at least a defeasible rule, or any other kind of blocking defeater. This is however inadequate in most contexts, where an *a priori* enumeration

[1] This shortest account of Peirce is surely unfair, since his purpose was much wider, for in his semiotic analysis of inference, abduction was central as the source of creativity and new knowledge [7].

of all the possible exceptions to the rules is impossible. Another way to reinstate consistency can arise from an explicit ranking among default rules and plausibility among evidence items [5]). This also can be very hard to achieve in real-life situations, for instance in scientific reasoning, where there can be abundance of conflicting evidence, alternative theories, inconsistent explicative hypotheses, and many other, without a clear relation of epistemic importance among all these pieces of knowledge. For this reason, in this work we will face the crude fact that abduction in nonmonotonic theories may have to accommodate anomalous observations.

## 2 A System for Abductive Nonmonotonic Reasoning

Before discussing the worries of abduction in nonmonotonic theories, first we will give a formal characterization of a nonmonotonic reasoning system, and then include an explicit rule for abduction. In a nutshell, the system regards defeasible rules $a(X) \succ\!\!- b(X)$[2] as material implications *only* for the *modus ponens* inference rule (that is, contraposition, left strengthening, right weakening, and similar uses are explicitly left out). Defeasible rules can be "fired" in MP only when their antecedent is fully instantiated, *i. e.*, there is a ground substitution for $X$ such that all the literals in $a(X)$ have been inferred. This ground instance of $a(X)$ is an *activator* for the rule. That is, neither generic nor universally quantified inferences are allowed with defeasible rules.

The reasoning system, then, will chain inferences in a way very similar to (classical) deductions, with the addition of inferences in which a fully activated defeasible rule was used. This chains of inferences are *(sub)-theories* in Brewka [4] and Poole [19], and *arguments* in Loui [13] and Vreeswijk [23]. We will adopt this later denomination. If a defeasible rule can be regarded as a *prima facie* material implication, then an argument for $A$ is a *prima facie* proof or a *prediction* for $A$. We can then extend the (classical) consequence operator $\vdash$ to the new operator $\mid\!\sim$ , where $\mathcal{T} \mid\!\sim A$ means that there is an argument for $A$ in theory $\mathcal{T}$.

Since we may reasonably expect that these inferences will eventually generate a pair of contradictory literals, and since we want to avoid (classical) trivialization, then our reasoning system must incorporate some kind of strengthening or restriction among the structural rules. For this reasoning, we adapted a presentation of the implicational segment of a relevant logic, similar to the $R_{\rightarrow}$ system of Anderson and Belnap [2]. Then, the reiteration rule is restricted to sentences that were inferred within the same subproof. If we need reiteration of a sentence $S$ of a previous step outside the subproof, then we must either introduce $S$ as a new assumption, or reproduce the inference steps that leads to the inference of $S$. To take due care of this, we establish an index schema that labels premise introduction and its ulterior discharge by means of $\rightarrow_I$. The use of a defeasible rule is regarded as a restricted *modus ponens* that also introduces a new hypothesis. The labeling schema obeys a simple set of cases (the subindices $I$ and $J$ denote sets of indices, and $i$ and $j$ denote individual indices).

- (*Premise*) An hypothetical premise $a$ is introduced with an index $i$ never used before (we will use the sequence of natural numbers).
- ($\rightarrow_I$) From a (sub-)demonstration for $b_I$ from premise $a_j$ (with $j \in I$) to infer $(a \rightarrow b)_{I-\{j\}}$.
- (*Reit.*) Reiteration of a sentence retains the indices.
- ($\rightarrow_E$) In the *modus ponens* rule, the consequent retains the indices of the major and the minor premises: from $a_I$ and $(a \rightarrow b)_J$ to infer $b_{I \cup J}$.

We now add the case for defeasible rules.

- ($\succ\!\!-_E$) From a (sub-) demonstration of $a(t)_I$ and $a(X) \succ\!\!- b(X)$ to infer $a(t) \rightarrow b(t)_{I \cup \{k\}}$, where $k$ is an index never used before.[3].

Example **2.1** *Suppose that in our knowledge base we have*
$a, a \succ\!\!- b, b \succ\!\!- c, a \succ\!\!- \neg c$

*In this situation, we may establish the following reasoning lines:*

| | | |
|---|---|---|
| 1 | $a_{\{1\}}$ | *Premise* |
| 2 | $(a \succ\!\!- b)_{\{2\}}$ | *Defeasible rule* |
| 3 | $b_{\{1,2\}}$ | $1, 2, \succ\!\!-_E$ |
| 4 | $(b \succ\!\!- c)_{\{3\}}$ | *Defeasible rule* |
| 5 | $c_{\{1,2,3\}}$ | $3, 4, \succ\!\!-_E$ |

| | | |
|---|---|---|
| 1 | $a_{\{1\}}$ | *Premise* |
| 2 | $(a \succ\!\!- \neg c)_{\{4\}}$ | *Defeasible rule* |
| 3 | $\neg c_{\{1,4\}}$ | $1, 2, \succ\!\!-_E$ |

[2]Both antecedent and consequent of defeasible rules are restricted to be *sets of literals* (interpreted as a conjunction), and $X$ is a tuple of free variables.

[3]Both $a(X)$ and $a(t)$ denote sets of literals. $X$ is a tuple of free variables, $t$ is a tuple of ground terms such that $X$ may be substituted for $t$ in $a(X)$.

*We may also use an "intuitionistic" negation introduction rule $\neg_I$ [21], according to which if the introduction of an hypothesis leads to contradiction, then the negation of the hypothesis can be inferred.*

| | | |
|---|---|---|
| 1 | $a_{\{1\}}$ | *Premise* |
| 2 | $(a \succ\!\!- b)_{\{2\}}$ | *Defeasible rule* |
| 3 | $b_{\{1,2\}}$ | $1, 2, \succ\!\!-_E$ |
| 4 | $(b \succ\!\!- c)_{\{3\}}$ | *Defeasible rule* |
| 5 | $c_{\{1,2,3\}}$ | $3, 4, \succ\!\!-_E$ |
| 6 | $(a \succ\!\!- \neg c)_{\{4\}}$ | *Defeasible rule* |
| 7 | $\neg c_{\{1,4\}}$ | $1, 6, \succ\!\!-_E$ |
| 8 | $\neg a_{\{2,3,4\}}$ | $5, 7, \neg_I$ |

*The first two demonstrations are subsumed in the third, but we show them separately to stress the fact that the logic is showing three possible conclusions, each one founded in three different sets of premises. If we were able to establish comparisons of the assertive support (i. e., the trustability) of these sets of premises, then we may choose among these conclusions. It is remarkable that b is always consequence of the theory, because it is not contradictory with any other sentence (the logic will not try to reason by contraposition with $\neg c$ and $b \succ\!\!- c$).*

Our final step is to propose a rule for abduction. This rule is based on several considerations (which cannot be discussed at length here because of space limitations). For this reason, we will give here only a short motivation. Given a nonmonotonic theory $\mathcal{T}$ (*i. e.*, a theory that may have defeasible rules), an abduction for an observation $O$ should be a hypothetical explanation $H$ that is compatible with $\mathcal{T}$, neither $\mathcal{T}$ nor $H$ should jointly (but not separately) explain $O$, and any other explanation $H'$ should also explain $H$ itself. Formally:

- (*Abd.*) From $O_k$ to infer $H_{S\cup\{k\}}$ iff
  1. $H_{S\cup\{k\}} \cup \mathcal{T} \not\vdash \bot$, *(H is consistent with $\mathcal{T}$)*
  2. $\mathcal{T} \not\mid\!\sim O_k$, *(there is no argument for O in $\mathcal{T}$)*
  3. $H_{S\cup\{k\}} \not\mid\!\sim O_k$, *(there is no argument for O in H)*
  4. $\mathcal{T} \cup H_{S\cup\{k\}} \mid\!\sim O_k$, *(there an argument for O in $\mathcal{T}$∪H)*
  5. Any other set $H'$ that satisfies the four conditions above is such that $H' \cup \mathcal{T} \mid\!\sim H_{S\cup\{k\}}$ (*i. e.*, $H$ is the most "shallow" explanation for $O$).

Example **2.2** *Suppose that in a knowledge-based system we find the rules*

| | |
|---|---|
| $w(X) \succ\!\!- i(X)$ | *If X has work, then X receives an income.*[4] |
| $w(X) \succ\!\!- t(X)$ | *If X has work, then X pays taxes.* |
| $w(X) \succ\!\!- \neg s(X)$ | *If X has work, then X does not study.* |
| $s(X) \succ\!\!- w(X)$ | *If X studies, then X has work.* |
| $c(X) \succ\!\!- s(X)$ | *If X has a scholarship, then X studies.* |
| $c(X) \succ\!\!- i(X)$ | *If X has a scholarship, then X receives an income.* |
| $c(X) \succ\!\!- \neg t(X)$ | *If X has a scholarship, then X does not pay taxes.* |

*Given this, what can we expect about Scott, of whom we only know that he pays taxes?*

| | | |
|---|---|---|
| 1 | $t(Scott)_{\{1\}}$ | *Premise* |
| 2 | $(w(X) \succ\!\!- t(X))_{\{2\}}$ | *Defeasible rule* |
| 3 | $w(Scott)_{\{1,2\}}$ | $1, 2,$ *Abduction (Explanation)* |
| 4 | $w(Scott)_{\{1,2\}}$ | $3,$ *Reit.* |
| 5 | $(w(X) \succ\!\!- i(X))_{\{3\}}$ | *Defeasible rule* |
| 6 | $i(Scott)_{\{1,2,3\}}$ | $4, 5, \rightarrow_E$ *(Prediction)* |
| 7 | $w(Scott)_{\{1,2\}}$ | $3,$ *Reit.* |
| 8 | $(w(X) \succ\!\!- \neg s(X))_{\{4\}}$ | *Defeasible rule* |
| 9 | $\neg s(Scott)_{\{1,2,4\}}$ | $7, 8, \rightarrow_E$ *(Prediction)* |

*By abduction, we can show that $t(Scott)$ because $w(Scott)$ (he pays taxes because he works), and from this inference, we can predict that he has an income, and that he does not study. It is a desirable feature here that further (iterated) abductions (for example, $c(Scott)$ because $i(Scott)$) are blocked for being inconsistent (see the next Section).*

*If we knew about another person, say $Kim$, of whom we knew only that she received an income, then we could generate two abductive explanations for her income. The first one, $i(Kim)$ because $w(Kim)$, allows further predictions ($t(Kim)$ and $\neg s(Kim)$). The second one, $i(Kim)$ because $c(Kim)$, allows other predictions ($s(Kim)$ and $\neg t(Kim)$). In this situation we have two unrelated explanations, of which we can not choose one over the other (again, see next Section). However, knowing further that, for instance, $s(Kim)$, will block the first explanation in favor of the second.*

# 3 Some Issues Regarding Abductive Explanation

In this Section we will pose some characteristics that distinguish abductive reasoning in nonmonotonic theo-

[4]The rules in this example are defeasible, so they should be read as *"Normally if X has a work..." and so on.*

ries from other kinds of nonmonotonic reasoning. Regarding our last example, find reasonable that a theory will generate different incompatible explanations, establish new predictions from previous explanations, and even block some of these new predictions given the facts. The first of these problems regards the problem of multiple extensions [4]. These had been considered an undesirable feature of nonmonotonic theories For this reason, there is a growing interest on finding adequate comparison criteria among extensions [13, 19, 23]. When there are two or more unrelated defeasible arguments with contradictory conclusions, then it is hard to accept or believe any of the conclusions.

However, in defeasible abductive reasoning, it is natural (and indeed desirable) to have multiple explanations. This may arise when more than one activator can generate a defeasible argument whose conclusion is the surprising observation. Each of these activators gives rise to a new and different extension of the context of the theory.

Example **3.1** *(A slight modification of Nixon's diamond [9]). Suppose we have the following nonmonotonic theory $\mathcal{T}$:*

$$\begin{array}{ll} \{\ q(X) \succ\!\!- p(X) & \textit{(Quakers are pacifists)}, \\ \ r(X) \succ\!\!- \neg p(X) & \textit{(Republicans are not pacifists)}, \\ \ h(X) \succ\!\!- \neg p(X)\ \} & \textit{(Hawks are not pacifists)}. \end{array}$$

- *If we have evidence $E = \{q(Dick), r(Dick)\}$, then we can generate two defeasible arguments that respectively justify, $p(Dick)$ and $\neg p(Dick)$. However, none of these conclusions is compatible with the context $\mathcal{T} \cup E$, and for this reason no tenable extension can be generated.*
- *However if we have evidence $E = \{\neg p(Dick)\}$, then there exist two defeasible abductive explanations $r(Dick)$ and $h(Dick)$. Each of them separately, their conjunction, and also their disjunction, raise an extension that is consistent with the context.*

Another remarkable difference arises among the strength of nonmonotonic explanations and nonmonotonic predictions. We may claim that –quite on the contrary of what occurs with strict implications and monotonic theories– in nonmonotonic theories, an abductive explanation is in general stronger than a prediction (implication).

Example **3.2** *Suppose we use a defeasible rule to represent the odds of winning the lottery when a ticket is bought.*

$$\mathcal{T} = \{buys(X) \succ\!\!- wins(X)\}.$$

- *If we have evidence that $Jack$ bought a ticket $E = \{buys(Jack)\}$, then we defeasible conclude (predict) that he will win the lottery $wins(Jack)$*
- *In turn, if we have evidence that $Jack$ won the lottery $E = \{wins(Jack)\}$, then our abductive explanation is that he bought a ticket $buys(Jack)$.*

*Which of the reasoning lines seems stronger?*

There is also the issue of the *accrual of reasons*, *i. e.*, the idea that two or more arguments supporting the same conclusion give more strength to the conclusion in case there are also arguments against the conclusion. In nonmonotonic reasoning, this heuristic was proposed by Pollock [15, 16] and Verheij [22], but lately they abandoned this idea, and now nobody believes that accrual of reasons is an adequate comparison mechanism. In nonmonotonic abductive reasoning, we can certainly have an *accrual of explanations*, *i. e.*, the case where an abductive explanation for a surprising fact $f$, is also an explanation for new observed but previously unpredicted facts.

Example **3.3** *Suppose we have the following nonmonotonic theory.*

$$\mathcal{T} = \{\ \begin{array}{l} a(X) \succ\!\!- \neg b(X), \\ a(X) \succ\!\!- c(X), \\ a(X) \succ\!\!- d(X), \\ e(X) \succ\!\!- d(X)\}. \end{array}$$

*If we are now confronted with the observation $d(t)$, then there exist two possible abductive explanations $a(t)$ and $e(t)$. $a(t)$ activates also the new predictions $\neg b(t)$ and $c(t)$. For this reason, in general $a(t)$ will be preferable because of its larger empirical progress.*

*In the context of scientific explanation, this progress is regarded by many as the supreme virtue of a scientific theory, notwithstanding refutation. If we had evidence $b(t)$, $c(t)$ and $d(t)$, the only consistent explanation is $e(t)$, but $a(t)$ will continue to be preferred because it explains more observable facts, in spite of the fact that it has a counterexample.*

# 4 Combining Defeaters

A final issue we wish to discuss is the defeat among various arguments in a nonmonotonic theory. This situation arises if we admit the possibility of *iterating* abductive explanations. In Sec. 2, we introduced a "shallow" abductive operator, but it can be iterated to produce "deeper" explanations.

Example **4.1** *(After [11] and [14]). Suppose we have the following theory $\mathcal{T}$.*

$\{ r(T) \succ\!\!- wr(T)$, *(if it rains, the road is wet),*
$r(T) \succ\!\!- wl(T)$, *(if it rains, the lawn is wet),*
$r(T) \succ\!\!- \neg s(T)$, *(if it rains, it's not sunny),*
$s(T) \succ\!\!- \neg r(T)$, *(if it's sunny, it does not rain),*
$so(T) \succ\!\!- wl(T)$, *(if the sprinklers are on, the lawn is wet),*
$s(T) \wedge h(T) \succ\!\!- so(T)$, *(if it's sunny and hot, the sprinklers are on),*
$wl(T) \succ\!\!- ws(T)$, *(if the lawn is wet, the shoes are wet),*
$wr(T) \succ\!\!- ws(T) \}$. *(if the road is wet, the shoes are wet).*

*In this situation, suppose we observe that our shoes are wet ($E = ws(today)$). The possible (shallow) explanations for this are that either the road is wet, or that the lawn is wet, or both. However, none of these suffices to generate a "most specific" explanation.*

To generate a more specific explanation we can iterate the abductive inference, that is, to generate a new "evidence" set $E'$ that contains $E$ plus any of the independently generated explanations, and then use this new context to try to generate a new abductive explanation. This procedure may be easy to formalize, but, as we will see, it may be that an argument is conflicting with some of these abductive hypotheses, and a criterion for combining defeat should be taken into account.

Example **4.2** *(After [8], with slight modifications). Let us consider the following theory $\mathcal{T}$.*

$\{ q(X) \succ\!\!- p(X)$, *(Quakers are pacifists),*
$q(X) \succ\!\!- rel(X)$, *(Quakers are religious),*
$r(X) \succ\!\!- b(X)$, *(Republicans are belicists),*
$p(X) \succ\!\!- \neg b(X)$, *(pacifists are not belicists),*
$b(X) \succ\!\!- \neg p(X)$, *(belicists are not pacifists),*
$b(X) \succ\!\!- sw(X)$, *(belicists support star wars),*
$b(X) \succ\!\!- pm(X)$, *(belicists are politically motivated),*
$p(X) \succ\!\!- pm(X) \}$. *(pacifists are politically motivated).*

*Suppose that our starting point is the observation that Dick is politically motivated ($pm(Dick)$), and that we believe that an explanation for this is Dick being a Quaker ($q(Dick)$).*

*However, new information leads us to accept that Dick supports star wars ($sw(Dick)$). This is not an anomalous observation ($\neg sw(Dick)$ was not conclusion of our belief state), but any attempt to find an abductive explanation for the new observation will force us to change beliefs. In particular, the only possible explanation is ($b(Dick)$), which generates a conflict with our previous beliefs. In this state of affairs, we have two competing arguments.*

$A_1 = \{$
$q(Dick)$ *(previous assumption),*
$q(Dick) \succ\!\!- rel(Dick)$ *(prediction to be confirmed),*
$q(Dick) \succ\!\!- p(Dick)$ *(prediction to be confirmed),*
$p(Dick) \succ\!\!- \neg b(Dick)$ *(prediction to be confirmed),*
$p(Dick) \succ\!\!- pm(Dick) \}$ *(confirmed prediction).*

$A_2 = \{$
$sw(Dick)$ *(new observation),*
$b(Dick) \succ\!\!- sw(Dick)$ *(abductive explanation),*
$b(Dick) \succ\!\!- \neg p(Dick)$ *(prediction to be confirmed),*
$b(Dick) \succ\!\!- pm(Dick) \}$ *(confirmed prediction).*

*The possible conclusions to which we could arrive in this competition between $A_1$ and $A_2$ can be grouped in three cases.*

I. *We keep accepting $q(Dick)$ but we reject the abductive explanation $b(Dick)$ because it is contradictory with other knowledge, and it is the consequence of a weak inference pattern. Then, the explanation for $sw(Dick)$ must come from another rule, still unknown.*

II. *Quite on the contrary, we accept $b(Dick)$ and reject $q(Dick)$ because it was in fact an assumption (this would not be the case if $q(Dick)$ was an observation in the context).*

III. *We accept the abductive explanation $b(Dick)$ and continue to believe that $q(Dick)$, but we reject that the later is a reason to reject the former (i. e., we reject the argument $\{q(Dick) \succ\!\!- p(Dick), p(Dick) \succ\!\!- \neg b(Dick)\}$, thus establishing the conjecture that some kind of exception must be the case here (Dick is a kind of belicist Quaker).*

*Suppose now that our previous assumption that explains $pm(Dick)$, was $r(Dick)$, and that the new observation is $rel(Dick)$. This observation is not anomalous (because $\neg rel(Dick)$ is not conclusion of this belief state). However, any abductive explanation will again force us to change belief, in particular we can only explain $rel(Dick)$ with ($q(Dick)$), which conflicts with our previous beliefs. We have then two competing arguments.*

$A_3 = \{$
$r(Dick)$ *(assumption),*
$r(Dick) \succ\!\!- b(Dick)$ *(prediction to be confirmed),*
$b(Dick) \succ\!\!- dm(Dick)$ *(prediction to be confirmed),*
$b(Dick) \succ\!\!- \neg p(Dick)$ *(prediction to be confirmed),*
$b(Dick) \succ\!\!- pm(Dick) \}$ *(confirmed prediction).*

$A_4$={
$rel(Dick)$ *(observation),*
$q(Dick) \succ\!\!- rel(Dick)$ *(abductive explanation),*
$q(Dick) \succ\!\!- p(Dick)$ *(prediction to be confirmed),*
$p(Dick) \succ\!\!- \neg b(Dick)$ *(prediction to be confirmed),*
$p(Dick) \succ\!\!- pm(Dick)$ } *(confirmed prediction).*

*In this competence between arguments we can arrive at the following group of conclusions.'*

I. *We keep accepting $r(Dick)$ but we reject the abductive explanation $q(Dick)$ because it is contradictory with other knowledge, and it is the consequence of a weak inference pattern. Then, the explanation for $rel(Dick)$ must come from another rule, still unknown.*

II. *Quite on the contrary, we accept $q(Dick)$ and reject $r(Dick)$ because it was in fact an assumption.*

III. *We accept the abductive explanation $b(Dick)$ and continue to believe that $r(Dick)$, but we reject that the later is a reason to reject the former or vice versa (i. e., we reject both the argument $\{q(Dick) \succ\!\!- p(Dick), p(Dick) \succ\!\!- \neg b(Dick)\}$ –Dick is a kind of belicist Quaker– and the argument $\{r(Dick) \succ\!\!- b(Dick), b(Dick) \succ\!\!- \neg p(Dick)\}$ –Dick is a kind of pacifist Republican).*

We can summarize the possible strategies to solve the conflicts between abductive inference and arguments.

I. We include only the abductive inferences that do not generate conflicting arguments with previous beliefs.

II. We consider that abductive inferences are defeaters for arguments that supported previous beliefs.

III. Conclusions of arguments and abductive explanations are on an equal footing, and if there are contradictions, then they must be attributed to an exception in one or more defeasible rules.

Definition **4.1** *Given a context $\mathcal{T} \cup E$ with an underlying knowledge $\mathcal{K}$. Then*

1. *The set of argumentative supported conclusions $A_c$ are generated from $E \cup \mathcal{T}$.*

2. *The set of abductive explanations $B_c$ are generated from $E \cup \mathcal{T}$.*

3. *If there is a pair of contradictory literals $a \in A_c$ and $b \in B_c$, then either*

   (I) *Any argument $A$ for $a$ defeats any argument generated with $b$.*

   (II) *Any argument generated with $b$ defeats any argument generated with $a$.*

   (III) *Any of the defeasible rules used in the arguments for $a$ or in the explanations for $b$ is defeated (syntactically blocked).*

4. *Firm conclusions $C$ are the members of $A_c$ and $B_c$ that were not defeated in 3.*

*If we need to iterate abduction, then the firm conclusions $C$ are added to the "evidence" $E$, and the process is repeated.*

Example **4.3** *Suppose we are in the situation of Example 4.1, and we observe that our shoes are wet ($ws(today)$), and we remember that today it was sunny ($s(today)$). Then, what can we conclude? The most general abductive explanations are $wr(today)$ and $wl(today)$. At the moment, any of these explanations is compatible with the observations and there is no defeat. If we iterate the abductive process, we find that $r(today)$ is explanation for $wr(today)$, and $r(today)$ or $so(today)$ are explanations for $wl(today)$.*

*Following strategy I, we assimilate $so(today)$ as the only tenable explanation for $ws(today)$, that is, we conjecture that the sprinkler was on, it got the lawn wet, and then our shoes got wet. If we push this further, we can also conjecture that today it was hot in addition to being sunny. Instead, if we follow strategy II, then the explanation $r(today)$ blocks our remembrance of being sunny. Then, our explanation now is that it rained, the rain got the road and the lawn wet, and then our shoes got also wet. If we use strategy III, then both previous explanations are valid and compatible, and we reject the rules that mutually exclude $r(today)$ and $s(today)$, that is, we suppose that today it may be hot and sunny at one time, and rainy at another.*

# 5 Conclusion

We presented a treatment of abduction in nonmonotonic theories. This inference context is the only way to cope with the problem of anomalous observations without changing the underlying theory. Our inference system is based on a natural deduction presentation of the implicational segment of a relevant logic, similar to the $R_{\rightarrow}$ system of Anderson and Belnap. We will discussed some issues arising from the pragmatic acceptance of abductive inferences in nonmonotonic theories, in particular, the existence of multiple explanations, the strenght of explanations (wrt predictions) and the accrual of explanations. Finally, we discussed the problem of the combination of defeat among arguments and abductions, showing how a nonmonotonic theory evolves when confronted with new evidence. A remarkable similarity can be found among the formalization of scientific research programmes and our system. We can easily regard a scientific theory as a special case of nonmonotonic theory, where the accidental generalizations and other lawlike statements are the default rules, and the conjectures are abductions that "protect" the theory from refutation. For this reason, we are working with the reconstruction of current scientific theories and their drive when confronted with new evidence.

**Acknowledgment**: We are grateful to the anonymous reviewers for their helpful comments on this work.

# References

[1] J. Alferes, L Pereira, and T. Swift. Well-Founded Abduction via Tabled Dual Programs. In *16th. Intl. Conference on Logic Programming*, pages 426–440. MIT Press, 1999.

[2] A. Anderson and N. Belnap. *Entailment: The Logic of Necesity and Relevance*. Princeton, 1975.

[3] Craig Boutilier and Verónica Becher. Abduction as Belief Revision. *Artificial Intelligence*, 77(1):143–166, 1996.

[4] Gehrard Brewka. Cumulative Default Logic: in Defense of Nonmonotonic Default Rules. *Artificial Intelligence*, 50(2):183–205, 1991.

[5] Claudio Delrieux. Nonmonotonic Reasoning Under Uncertain Evidence. In F. Giunchiglia, editor, *Artificial Intelligence: Methodology, Systems and Applications*, (Lecture Notes in Artificial Intelligence 1480), pages 195–204. 1998.

[6] M. Denecker and A. Kakas (eds.). Special Issue on Abductive Logic Programming. In *The Journal of Logic Programming*, pages **44**(1–3), 2000.

[7] C. Hartshorne *et. al. Collected Papers of Charles S. Peirce*. Harvard U. Press, Cambridge, 1958.

[8] Matthew L. Ginsberg. *(editor) Readings in Nonmonotonic Reasoning*. Morgan Kaufmann Publishers, Los Altos, California, 1987.

[9] J. Horty, R. Thomason, and D. Touretzky. A Clash of Intuitions: The Current State of Nonmonotonic Multiple Inheritance Systems. In *Proc. Tenth International Joint Conference on Artificial Intelligence*, pages 476–482, 1987. Morgan Kaufmann, Los Altos, CA.

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

[11] Kurt Konolige. Abduction versus Closure in Causal Theories. *Artificial Intelligence*, 53(2-3):255–272, 1992.

[12] C. Lobo, and J. Uzcátegui. Abductive Consequence Relations. *Art. Intelligence*, 89(1):149–171, 1997.

[13] Ronald P. Loui. Defeat Among Arguments: A System of Defeasible Inference. *Computational Intelligence*, 3(3), 1987.

[14] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, California, 1988.

[15] John Pollock. How to Reason Defeasibily. *Artificial Intelligence*, 57(1):1–42, 1992.

[16] John Pollock. Degrees of Justification. In P. Weingartner, G. Schurz, and G. Dorn, editors, *The Role of Pragmatics in Contemporary Philosophy*, pages 207–223. Verlag Hölder-Pichler-Tempsky, Wien, Österreich, 1998.

[17] David Poole. A Logical Framework for Default Reasoning. *Artificial Intelligence*, 36(1):27–47, 1988.

[18] David Poole. Explanation and Prediction: an Architecture for Default and Abductive Reasoning. *Computational Intelligence*, 5(2):97–110, 1989.

[19] David L. Poole. On the Comparison of Theories: Preferring the Most Specific Explanation. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 144–147, 1985. Morgan Kaufmann, Los Altos, CA.

[20] R. Reiter. A Theory of Diagnosis from First Principles. *Artificial Intelligence*, 32:57–95, 1987.

[21] A. S. Troelstra and D. van Dalen. *Constructivism in Mathematics*. Elsevier Science Publishers, Amsterdam, 1988.

[22] Bart Verheij. Argue!, an Implemented System for Computer-Mediated Defeasible Argumentation. In *Proc. Tentht Netherlands/Belgium Conference on Artificial Intelligence*, pages 57–66, Netherlands, 1998.

[23] G. A. W. Vreeswijk. Abstract Argumentation Systems. *Artificial Intelligence*, 90(2):225–279, 1997.