

Chapter 1

Abduction of Multiple Explanatory Hypotheses

Abstract

In abduction – the process of finding explanatory hypotheses for puzzling phenomena – one is often confronted with multiple explanatory hypotheses. In science one generally wants to test further the different hypotheses one by one. But, if we try to model this in a logic and make it possible to derive the different hypotheses apart from each other, we generally can derive their conjunction too. An elegant solution within the framework of adaptive logics is provided in Gauderis (2011). But this approach is not restricted to science. While it is true that a lot of cases in everyday reasoning require a more practical approach – in which one acts on the knowledge that all the different hypotheses might be the case – there are also a considerable amount of situations in which the more theoretical approach of the scientist is needed. In this paper we try to illustrate this by using this logic to model reasoning within detective literature.

1.1 The Problem of Multiple Explanatory Hypotheses

Abduction and Detective literature Charles Peirce thought that there were three characteristic ways of reasoning in science. In addition to the better-known ways of deduction and induction, there was a third rational way in which scientists can reason: abduction or “the process of forming an explanatory hypothesis” (Peirce, 1958–60, CP 5.171). The logical schema of forming such an explanatory hypothesis is for Peirce (1958–60, CP 5.189) the following:

The surprising fact, C is observed;
But if A were true, C would be a matter of course,
Hence, there is reason to suspect that A is true.

Obviously, this is a form of defeasible or fallible reasoning, of which Peirce (1958–60, CP 2.777) himself was perfectly conscious: “The hypothesis which it problematically concludes is frequently utterly wrong itself, and even the method needs not ever lead to the truth.” When we translate his schema to predicative logic, we get the following schema:

$$(\forall\alpha)(A(\alpha) \supset B(\alpha)), B(\beta)/A(\beta)$$

This schema is better known as the logical fallacy *Affirming the Consequent*, but this is only a fallacy if we stick within deductive logics. In cases in which we are not able to obtain any deductive results that can explain our observations, Peirce (1958–60, CP 2.777) makes a point in stating that “. . . its method is the only way in which there can be any hope of attaining a rational explanation”.

Quite often, people have acknowledged that this is essentially the same kind of reasoning as the reasoning employed in crime investigation or detective literature. As the different articles in the book ‘The Sign of Three’¹ point out, solving a murder case by tracing back the clues is essentially an abductive operation. So, while Holmes was maybe wrong when he said to Watson that it was elementary deduction, he surely was not wrong in thinking that his reasoning was logical, it was only according to the laws of a logic for abduction.

The Problem of Multiple Explanatory Hypotheses Still, in solving a murder case, our detective is quite often confronted with two or more suspects. When we try to model this with a formal logic, this can lead to a problem. Consider the following example. Suppose we are confronted with the puzzling fact or clue Pa while our background knowledge contains two possible causes: $(\forall x)(Qx \supset Px)$ and $(\forall x)(Rx \supset Px)$. There are actually now two roads that can be taken. We could construct a logic in which we can only derive the disjunction $(Qa \vee Ra)$ and not the individual hypotheses Qa and Ra . This

¹See Sebeok and Eco (1988). This book investigates the relation between the writings of Charles Peirce on the one hand and the writings of Edgar Allen Poe (Auguste Dupin) and Arthur Conan Doyle (Sherlock Holmes) on the other hand.

road, called *practical* abduction,² is suitable to model situations in which one has to *act* on the basis of the conclusions. For instance, in medical diagnoses, a physician who finds out that two possible diseases can be the cause for the examined symptoms, needs to take appropriate steps based on the fact that both diseases might be the cause.

But our detective has a more theoretical perspective and is interested in finding out which of the hypotheses is the actual cause. Therefore, it is important that he can *abduce* the individual hypotheses Qa and Ra in order to examine them further one by one. This is because, on the one hand, one has to be able to derive Qa and Ra separately, but on the other hand, one has to prevent the derivation of their conjunction ($Qa \wedge Ra$). Not only does it seem counterintuitive to take the conjunction of two possible hypotheses as an explanation. Also, if the two hypotheses are actually incompatible – a victim cannot be murdered twice – it would lead to explosion. So it is clear that for this application, we need the second road.

Adaptive Logics for Abduction Since abduction is a defeasible type of reasoning, adaptive logics are a good tool to model this type of relation.³ The main advantages with respect to this problem can be summed up as such.

Firstly, it allows for direct implementation of defeasible reasoning steps (*in casu* applications of *Affirming the Consequent*), which makes it possible to construct logical proofs that nicely integrate defeasible (ampliative in this case) and deductive inferences. This corresponds to the natural way in which humans reason.

Secondly, the formal apparatus of an adaptive logic instructs one to specify exactly which conditions would falsify the (defeasible) reasoning step. So, if this condition is derived later on in the proof, it defeats in a formal way all steps derived on the assumption that this condition is false. As these conditions are assumed – as long as one cannot derive them – to be false, they are called *abnormalities* in adaptive logic literature. This possibility to defeat previous reasoning steps mirrors nicely the dynamics that is found in actual human reasoning.

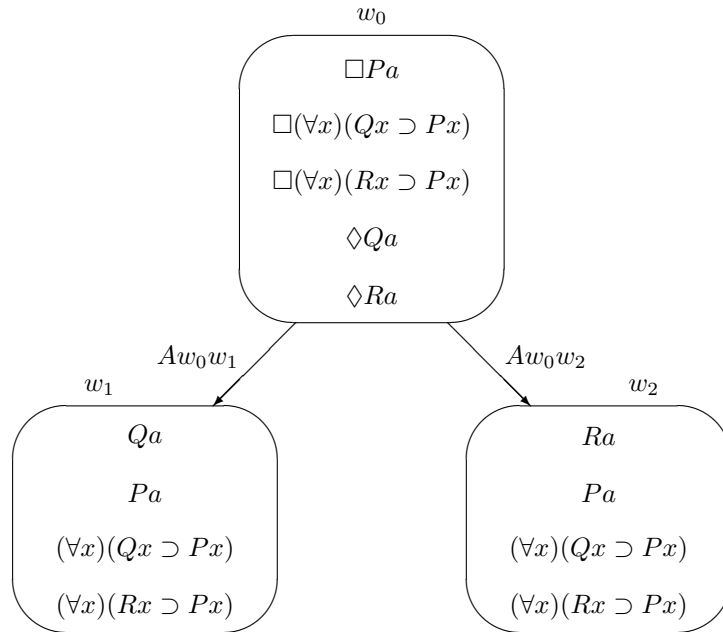
Thirdly, there are for all adaptive logics in standard format generic proofs for most of the important metatheoretical properties (including soundness and completeness).

Within the adaptive logics program, several logics have been developed that model abductive reasoning. Practical abduction – in which the disjunction of the explanatory hypotheses is derived – is, for instance, adequately modelled in

²According to the definition suggested in Meheus and Batens (2006, pp224–225) and used in Lycke (2009).

³Some features of adaptive logics will be explained in order to make the paper understandable for people not familiar with adaptive logics. But the remarks made in this paper can, due to space limitations, hardly be called a good introduction to adaptive logics. We refer readers interested in adaptive logics to Batens (2011, 2007) for a systematic overview and to Batens (2004) for a more philosophical defense of the use of adaptive logics.

the logics \mathbf{LA}^r and \mathbf{LA}_s^r .⁴ Theoretical abduction has been modelled by the logic \mathbf{AbL}^t ,⁵ but here we will concentrate on the logic \mathbf{MLA}^s , developed in Gauderis (2011) for the purpose of modelling abductive reasoning in science. This logic provides an elegant way out of this problem by adding modalities to the object language and deriving – in case of our toy example that we used to introduce the problem – the hypotheses $\diamond Qa$ and $\diamond Ra$. In this way, the scientist or detective can work further on the individual hypotheses without having to prevent the conjunction, because $(\diamond Qa \wedge \diamond Ra)$ does not imply $\diamond(Qa \wedge Ra)$ in any standard modal logic. The (Kripke-)semantics of this logic also foresees that everything that follows from investigating further one of the hypotheses is only verified within the possible world made accessible by formulating this hypothesis. The graphic representation of the semantics of our toy example below, illustrates that within this logic, different hypotheses and their conclusions are considered independently from each other.



This feature makes the logic \mathbf{MLA}^s (Modal Logic for Abduction) very apt to model, apart from its applications within science, a lot of everyday reasoning. The goal of this paper is to illustrate this with a crime investigation example. But before we present a more elaborate example, we will in the following section explain the logic \mathbf{MLA}^s .

⁴See Meheus and Batens (2006); Meheus (2007, 2010).

⁵See Lycke (2009).

1.2 Formal presentation of the logic MLA^s

As any adaptive logic in the standard format, MLA^s is defined by a triple of a lower limit logic, a set of abnormalities and a strategy. These will be introduced in the following paragraphs after we have specified the language schema.

Formal Language Schema Let \mathcal{L} be the standard predicative language of CL with logical symbols $\neg, \supset, \wedge, \vee, \equiv, \forall$ and \exists . We will further use $\mathcal{C}, \mathcal{V}, \mathcal{F}$ and \mathcal{W} to refer respectively to the sets of individual constants, individual variables, all (well-formed) formulas and the closed (well-formed) formulas of \mathcal{L} .

\mathcal{L}_M , the language of our logic, is \mathcal{L} extended with the modal operators \Box and \Diamond , where \Box is primitive and \Diamond defined in the usual way. \mathcal{W}_M , the set of closed formulas of \mathcal{L}_M is the smallest set that satisfies the following conditions:

1. if $A \in \mathcal{W}$, then $A, \Box A, \Diamond A \in \mathcal{W}_M$
2. if $A \in \mathcal{W}_M$, then $\neg A \in \mathcal{W}_M$
3. if $A, B \in \mathcal{W}_M$, then $A \wedge B, A \vee B, A \supset B, A \equiv B \in \mathcal{W}_M$

It is important to note that there are – among other things – no occurrences of modal operators within the scope of another modal operator or a quantifier. We further define the set \mathcal{W}_Γ – the subset of \mathcal{W}_M , the elements of which can act as premises in our logic – as the smallest set that satisfies the following conditions:

1. if $A \in \mathcal{W}$, then $\Box A, \Diamond A \in \mathcal{W}_\Gamma$
2. if $A, B \in \mathcal{W}_\Gamma$, then $A \wedge B \in \mathcal{W}_\Gamma$

It is easily seen that $\mathcal{W}_\Gamma \subset \mathcal{W}_M$.

Lower Limit Logic Each adaptive logic is built on the deductive frame of a Tarski-logic. This lower limit logic (**LLL**) defines the undefeasible part of our logic. Everything that follows from the premises by the **LLL** will never be defeated.

The **LLL** of MLA^s is the predicative version of **D**, restricted by the language schema. **D** is characterized by a full axiomatization of predicate **CL** together with two axioms and an inference rule:

$$\mathbf{K} \quad \Box(A \supset B) \supset (\Box A \supset \Box B)$$

$$\mathbf{D} \quad \Box A \supset \Diamond A$$

NEC if $\vdash A$, then $\vdash \Box A$

The semantics for this logic can be expressed by a standard possible world Kripke semantics where the accessibility relation A between possible worlds is *serial*, i.e. for every world w in our model, there is at least one world w' in our model such that wAw' . Soundness and completeness for **D** is – as for all normal modal logics – a well established fact.

Set of Abnormalities The defeasible part of our logic is defined by the combination of the strategy and the set of abnormalities. This is a set of **LLL**-contingent formulas characterized by a logical form (or a union of such sets) that are assumed to be false ‘as much as possible’. These assumptions allow us to derive, apart from the deductive consequences of the **LLL**, defeasible consequences that can be derived on a condition, viz. the falsehood of the abnormalities. The inference rules are reduced to three generic rules: a premise, an unconditional and a conditional inference rule. An extra element is added on each line which is set of conditions on which the formula on that line is derived.

PREM	If $A \in \Gamma$:	$\frac{\begin{array}{c} \vdots \\ \vdots \end{array}}{A \quad \emptyset}$
RU	If $A_1, \dots, A_n \vdash_{\mathbf{LLL}} B$:	$\frac{\begin{array}{c} A_1 \quad \Delta_1 \\ \vdots \\ \vdots \\ A_n \quad \Delta_n \end{array}}{B \quad \Delta_1 \cup \dots \cup \Delta_n}$
RC	If $A_1, \dots, A_n \vdash_{\mathbf{LLL}} B \vee Dab(\Theta)$	$\frac{\begin{array}{c} A_1 \quad \Delta_1 \\ \vdots \\ \vdots \\ A_n \quad \Delta_n \end{array}}{B \quad \Delta_1 \cup \dots \cup \Delta_n \cup \Theta}$

To define the set of abnormalities of \mathbf{MLA}^s ,⁶ we first need to introduce a new notation. Suppose that $A_{PCNF}(\alpha)$ is the *Prenex Conjunctive Normal Form* of $A(\alpha)$ and that for $Q_i \in \{\forall, \exists\}, \gamma_i \in \mathcal{V}, A_i(\alpha) \in \mathcal{F}$:

$$A_{PCNF}(\alpha) = (Q_1\gamma_1) \dots (Q_m\gamma_m)(A_1(\alpha) \wedge \dots \wedge A_n(\alpha))$$

Then we can define $A_i^{-1}(\alpha)$ ($1 \leq i \leq n$) as follows:

$$\begin{aligned} \text{if } n > 1 & : A_i^{-1}(\alpha) =_{df} (Q_1\gamma_1) \dots (Q_m\gamma_m)(A_1(\alpha) \wedge \dots \wedge A_{i-1}(\alpha) \wedge \\ & \quad A_{i+1}(\alpha) \wedge \dots \wedge A_n(\alpha)) \\ \text{if } n = 1 & : A_1^{-1}(\alpha) =_{df} \top \end{aligned}$$

The idea is to have a notation for the formula formed by leaving out the i^{th} conjunct. With this notation the set of abnormalities is defined as such:

⁶The set of abnormalities is, due to space limitations, presented here quite briefly, for a more elaborate explanation how this set came to be, we refer to Gauderis (2011) in which the logic was presented for the first time.

$$\Omega = \{ \Box((\forall\alpha)(A(\alpha) \supset B(\alpha)) \wedge (B(\beta) \wedge \neg A(\beta))) \mid \\ \forall\Box(\forall\alpha)B(\alpha) \vee \bigvee_{i=1}^n \Box(\forall\alpha)(A_i^{-1}(\alpha) \supset B(\alpha)) \mid \\ \text{No predicate that occurs in } B \text{ occurs in } A, \\ \alpha \in \mathcal{V}, \beta \in \mathcal{C}, A, B \in \mathcal{F} \}$$

This form might look complex, but its functioning is quite straightforward. We actually just made a disjunction of three reasons why we stop considering $A(\beta)$ as a good explanatory hypothesis for the phenomenon $B(\beta)$. These three possible reasons are (i) when $\neg A(\beta)$ is the case, (ii) when $B(\beta)$ is a tautology (and obviously, does not need an explanatory hypothesis) or (iii) when $A(\beta)$ has a redundant part and is therefore not an adequate explanatory hypothesis.

From now on, we can unambiguously shorten this logical form of the abnormalities as $!A(\beta) \triangleright B(\beta)$ which could be read as “ $A(\beta)$ is not a valid hypothesis for $B(\beta)$ ”.

Adaptive Strategy Finally, a strategy is needed to define how to interpret the idea ‘false as much as possible’ exactly. In that way, the strategy orders which defeasible reasoning steps should be marked.⁷

Definition 1 (Marking for the simple strategy). *Line i with condition Δ is marked for the simple strategy at stage s of a proof⁸, if the stage s contains a line of which an $A \in \Delta$ is the formula and \emptyset the condition.*

Definition 2. *A formula A is derived from Γ at stage s of a proof iff A is the formula of a line that is unmarked at stage s .*

Definition 3. *A formula A is finally derived from Γ at stage s of a proof iff A is derived at line i , line i is not marked at stage s and remains unmarked in every extension of the proof.⁹*

Definition 4 (Final Derivability). *For $\Gamma \subset \mathcal{W}_\Gamma$: $\Gamma \vdash_{\mathbf{MLA}^s} A$ (A is finally \mathbf{MLA}^s -derivable from Γ) iff A is finally derived in a \mathbf{MLA}^s -proof from Γ .*

⁷In adaptive logics, traditionally the mark \checkmark is used to indicate that a step is defeated.

⁸A *stage of a proof* is a sequence of lines and a proof is a chain of stages. Every proof starts off with an empty sequence (stage 0). Each time a line is added to the proof by applying one of the inference rules, the proof comes to its next stage, which is the sequence of lines written so far including the new line.

⁹This definition is slightly different from the more general definition mentioned in Batens (2011) because, using the simple strategy, it is in our case not possible that a marked line becomes unmarked at a later stage of a proof.

1.3 Application of MLA^s in Detective Literature

To illustrate the functioning of this logic, we will make up an original story. This is because we do not want to let our example grow too complex (and thus, less illustrative).¹⁰

On a certain morning, X is found murdered in mysterious circumstances. From our first investigations we are able to determine three suspects (Sx) a , b and c who could be the murderer (Mx). Confronting these suspects with the facts, only c is able to pull out a water tight alibi (Ax) for the moment of the murder, 10.30am. Further, at the crime scene we find two clues: some long blond hairs at the murder weapon and a receipt of the tailor delivered at 9.30am, both of which could not have belonged to the victim.

The whole of this data constitutes our background knowledge. Formalized we get our initial premise set Γ . The exact meaning of the predicates describing the clues is defined as follows:

B_1x “ x was in the possession of some long blond hairs at 10.30am”
 T_1x “ x was in the possession of the tailor receipt at 10.30am”
 T_2x “ x received the tailor receipt at 9.30am”

Then, we can start off our proof (which models our reasoning process):¹¹

1	$\Box(\forall x)(Mx \supset Sx)$	-;PREM	\emptyset
2	$\Box(\forall x)(Ax \supset \neg Mx)$	-;PREM	\emptyset
3	$\Box(Sa \wedge Sb \wedge Sc)$	-;PREM	\emptyset
4	$\Box(\neg Aa \wedge \neg Ab \wedge Ac)$	-;PREM	\emptyset
5	$\Box(\forall x)(Mx \equiv B_1x)$	-;PREM	\emptyset
6	$\Box(\forall x)(Mx \equiv T_1x)$	-;PREM	\emptyset

Now, the logic abduces three possible hypotheses, but the third one is - as expected - directly marked.

7	$\Diamond Ma$	1,3;RC	$\{!Ma \triangleright Sa\}$	
8	$\Diamond Mb$	1,3;RC	$\{!Mb \triangleright Sb\}$	
9	$\Diamond Mc$	1,3;RC	$\{!Mc \triangleright Sc\}$	\checkmark^{11}
10	$\Box \neg Mc$	2,4;RU	\emptyset	
11	$!Mc \triangleright Sc$	1,3,10;RU	\emptyset	

¹⁰Obviously we also don't want to spoil any plot of a classic in the genre.

¹¹The only formalization that might appear a bit odd is the first line that seems to state that the murderer is always a suspect. Still, this line is the correct formalization, since this logic models a reasoning process that leads to a murderer. Consider the following: if the murderer was never suspected, he also will never be caught. So the actual meaning of the predicate Mx is more epistemological (like the predicate Sx essentially also is): “It can be shown that x is the murderer”.

Left with two suspects, we need to investigate further the information of our clues. On the one hand, b actually has long blond hair and could be the owner of the found hairs, while a has ginger hair and could therefore not be the owner. But on the other hand, the tailor (who wasn't aware that a murder had happened) assures us that nobody with blond hair has entered his shop that morning, while several people with ginger hair did. Still, we have no reason to doubt the fact that it was the murderer himself that was that morning in the shop.¹² Puzzled by this new information, we bring the suspects a final visit to confront them with the clues. But when we ring at a 's door, the door is opened by a blond woman who says to be a 's wife.

We can now add this extra information to our proof.¹³

12 $\diamond B_1 a \wedge \square B_1 b$	-;PREM	\emptyset
13 $\diamond T_2 a \wedge \square \neg T_2 b$	-;PREM	\emptyset
14 $\square (\forall x)(T_2 x \equiv T_1 x)$	-;PREM	\emptyset

With these new data, our detective can at this point rule out one more suspect. So, only one suspect is left. It is the only hypothesis we can derive that is compatible with all the known data. Here one can see again that hypothesis formation out of a certain background knowledge is closely related with compatibility.

\vdots	\vdots	\vdots	
7 $\diamond Ma$	1,3;RC	$\{!Ma \triangleright Sa\}$	
8 $\diamond Mb$	1,3;RC	$\{!Mb \triangleright Sb\}$	\checkmark^{16}
\vdots	\vdots	\vdots	
15 $\square \neg Mb$	6,13,14;RU	\emptyset	
16 $!Mb \triangleright Sb$	1,3,15;RU	\emptyset	

Some Concluding Remarks Is it also actually possible to come with conclusive evidence? Here we have to keep in mind that we are dealing with a logic for abduction or hypothesis formation. This should not be confused with deciding whether we have conclusive evidence. In science too, the forming of hypotheses and the confirmation of theories are two different steps in the scientific process. These two different methods were actually the two things that Charles Peirce wanted to keep separate in his methodology of science by discerning *abduction* and *induction*.

¹²This assumption is made because we don't want the example grow too complex; but it actually also nicely illustrates how many 'hidden assumptions' there are in a reasoning process, assumptions that have to be spelled out fully in a formal logic.

¹³Technically speaking, we have a new premise set Γ' and need to start a new proof, but it is easily seen that we can start our new proof by copying the previous proof and continue to add the new data as premises.

At the end of the day, we illustrated in this paper that the logic **MLA**^s, developed to model abduction in science, can also be used in more everyday life situations, as, for instance, our exploration in the detective genre shows. At first sight, the reasoning process seems quite equivalent. From a formal point of view, there is actually one major difference. A detective will point his attention to which subject is the murderer; his hypotheses are, for instance, *Ma*, *Mb* and *Mc*. A scientist on the other hand, will focus more on the predicates or properties; his hypotheses for a puzzling phenomenon *Pa* will most likely be modelled with hypotheses of the form *Fa*, *Ga* and *Ha*. It is an interesting route to investigate whether this formal difference leads to more differences between the two types of reasoning.

Bibliography

- Diderik Batens. The need for adaptive logics in epistemology. In D. Gabbay, S. Rahman, J. Symons, and J.P. Van Bendegem, editors, *Logic, Epistemology and the Unity of Science*, pages 459–485. Kluwer Academic Publishers, Dordrecht, 2004.
- Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.
- Diderik Batens. *Adaptive Logics and Dynamic Proofs. A Study in the Dynamics of Reasoning*. Forthcoming, 2011.
- Tjerk Gauderis. Modelling abduction in science by means of a modal adaptive logic. *Forthcoming*, 2011.
- Hans Lycke. The adaptive logics approach to abduction. In Erik Weber, Thierry Libert, Pierre Marage, and Geert Vanpaemel, editors, *Logic, Philosophy and History of Science in Belgium. Proceedings of the Young Researchers Days 2008*, pages 35–41, Brussels, 2009. KVAB.
- Joke Meheus. Adaptive logics for abduction and the explication of explanation-seeking processes. In Olga Pombo and Alexander Gerne, editors, *Abduction and the Process of Scientific Discovery*, pages 97–119. Centro de Filosofia das Ciências da U. de Lisboa, Lisboa, 2007.
- Joke Meheus. A formal logic for the abductions of singular hypotheses. *forthcoming*, 2010.
- Joke Meheus and Diderik Batens. A formal logic for abductive reasoning. *Logic Journal of the IGPL*, 14:221–236, 2006.
- Charles S. Peirce. *Collected Papers*. Belknap Press of Harvard University Press, Cambridge, Massachusetts, 1958–60.
- Thomas A. Sebeok and Umberto Eco, editors. *The Sign of Three: Dupin, Holmes and Peirce*, Bloomington, IN, 1988. Indiana University Press.