

# Ignorance and Semantic Tableaux: Aliseda on Abduction

John WOODS

BIBLID [0495-4548 (2007) 22: 60; pp. 305-318]

ABSTRACT: This is an examination of similarities and differences between two recent models of abductive reasoning. The one is developed in Atocha Aliseda's *Abductive Reasoning: Logical Investigations into the Processes of Discovery and Evaluation* (2006). The other is advanced by Dov Gabbay and the present author in their *The Reach of Abduction: Insight and Trial* (2005). A principal difference between the two approaches is that in the Gabbay-Woods model, but not in the Aliseda model, abductive inference is ignorance-preserving. A further difference is that Aliseda reconstructs the abduction relation in a semantic tableaux environment, whereas the Woods-Gabbay model, while less systematic, is more general. Of particular note is the connection between abduction and legal reasoning.

Keywords: abduction, explanation, ignorance-problem, ignorance-preservation, legal reasoning, semantic tableaux.

## 1. Two models of abduction

A distinctive feature of Atocha Aliseda's attractive theory of abduction is its representation of it as a kind of semantic tableaux reasoning (Aliseda 2006, chapter 4). In its most general form, Aliseda sees abductive inference as instantiating what is sometimes called the *AKM* model, thus named after some of its notable supporters.<sup>1</sup> There is a good deal more to Aliseda's far-ranging investigation of abduction than the semantic tableaux-ing of it. But since this is a feature of central importance to her approach, I shall concentrate my remarks accordingly.

### *The AKM Model*

Suppose we have it that  $E$  is some true sentence, and  $K$  a knowledge-base,  $\rightarrow$  a consequence relation, and  $H$  a hypothesis. Assume that the following facts obtain:

1.  $E$
2.  $K \rightarrow E$
3.  $H \rightarrow E$

Then an abduction is the derivation of  $H$  from three further facts:

4.  $K(H)$  is consistent
5.  $K(H)$  is minimal
6.  $K(H) \rightarrow E$

Accordingly,

---

<sup>1</sup> Thus, for example, for 'A' we have Aliseda-Llera (1997), Aliseda (2006); for 'K' we have Kowalski (1979), Kuipers (1999), Kakas *et al.* (1995) and Flach and Kakas (2000); and for 'M' there is Magnani (2001) and Meheus *et al.* (2002). Needless to say, there are legions of *AKM*-proponents whose surnames are ungraced by any of these letters in first position. So the expression "*AKM*" is a loose convenience at best.



## 7. H.

A principal advantage of the Aliseda approach is that the semantic tableaux constructions are well-understood. This gives promise of a nice methodological conservatism, according to which it is always better to produce an explication of something in an analytical vocabulary that is well-understood rather than in an idiom which itself calls out for clarification.

I will briefly explain in a moment the basic mechanics of abductions as extended semantic tableaux, but for now I want to bring to the surface the issue that motivates the present note. There is not the slightest doubt, I think, that the *AKM* model is by far the dominant way of schematizing abductive inferences. It is also quite clear that there are ranges of cases in which abductive practice exhibits the features attributed to it by that schema. There is, however, a rival view of the matter proposed by Dov Gabbay and me in Gabbay and Woods (2005) and Gabbay and Woods (2006), called (unblushingly) the *GW* model. This is an approach which preserves some of the *AKM* model and is more general than it in a number of respects. This not the place to reproduce the whole apparatus of the *GW* model. It is enough for present purposes to expose its main features.

*The GW Model*

On the *GW* approach, a pivotal fact about abduction is that it is a response to an ignorance-problem. Roughly speaking, an agent's ignorance problem is constituted by his wanting to know whether this-or-that is the case or whether such-and-so is the right decision, while lacking the knowledge then and there to answer the question. When an agent is in such circumstances, we may say that he has a cognitive target *T*, and that *T* is of a kind whose attainment demands the requisite knowledge. So an agent has an ignorance-problem with respect to *T* when he lacks the knowledge to attain it.

It is easy to see that there are two quite common responses to ignorance-problems. In one of them, the agent acquires some additional knowledge. In the other, the agent abandons the target. In the first, the agent overcomes the problem with new knowledge, and in the second the agent allows the problem overcome him owing to his lack of knowledge. In the first case, the solution of *T* affords the agent a possible basis for new action. In the second case, the agent's quiescence denies him a basis for new action. What is not always recognized is that there is a third response to an ignorance-problem.

The third response is abduction. Abduction is a way of proceeding that splits the difference between the prior two. Like response one, abduction provides a basis for new action. And, like the second response, it leaves the ignorance-problem unsolved. The abductive response pivots centrally and indispensably on the role of conjecture. It was Peirce who pointed out that a conjecture is a kind of guessing (Peirce 1931-1958, 5.172). When I conjecture that *P* I attach myself to *P* in a way that reflects that I do not know it; yet it is part of this mode of attachment that I allow *P* to function, albeit defeasibly, as the basis for further action. I put *P* to work, so to speak, and wait to see what happens.

We can stitch these observations together schematically. Let  $\mathcal{A}$  be an agent and  $T_{\mathcal{A}}$  his cognitive target at a time.  $K_{\mathcal{A}}$  is his knowledge set at that time,  $K_{\mathcal{A}}^*$  his immediate-successor knowledge-set,<sup>2</sup>  $\rightarrow$  an attainment on  $T_{\mathcal{A}}$ ,  $H$  a hypothesis,  $K(H)_{\mathcal{A}}$  the revision of  $K$  by  $H$ ,  $\rightarrow^{subj}$  a subjunctive attainment relation on  $T_{\mathcal{A}}$ ,  $C(H)_{\mathcal{A}}$  the conjecture of  $H$  by  $\mathcal{A}$ , and  $H_{\mathcal{A}}^C$  the decision by  $\mathcal{A}$  to release  $H$  provisionally for premissory work in the domain of discourse in which  $\mathcal{A}$ 's ignorance-problem arose in the first place.

- |  |                                |
|--|--------------------------------|
| 1. $T_{\mathcal{A}}$                                       | $\mathcal{A}$ sets his target. |
| 2. $K_{\mathcal{A}} \rightarrow T_{\mathcal{A}}$           | fact                           |
| 3. $K_{\mathcal{A}}^* \rightarrow T_{\mathcal{A}}$         | fact                           |
| 4. $H_{\mathcal{A}} \rightarrow T_{\mathcal{A}}$           | fact                           |
| 5. $K(H)_{\mathcal{A}} \rightarrow^{subj} T_{\mathcal{A}}$ | fact                           |
| 6. $K(H)$ satisfies conditions $k_1, \dots, k_n$           | fact <sup>3</sup>              |
| 7. $C(H)_{\mathcal{A}}$                                    | 1-6, conclusion                |
| 8. $H_{\mathcal{A}}^C$                                     | 7, decision. <sup>4</sup>      |

The differences between the *AKM* and *GW* models are clear upon inspection, but it wouldn't hurt to tarry briefly with one of its most important contrasts. *AKM* lacks a subjunctive consequence relation. So *AKM* makes no express recognition of the fact that the attainment of  $T_{\mathcal{A}}$  by  $K(H)_{\mathcal{A}}$ —given that  $H$  is introduced as a hypothesis, not as a fact—is subjunctive attainment:  $H$  and  $K_{\mathcal{A}}$  are such that *were*  $H$  the case then the revision of the sentences in  $K_{\mathcal{A}}$  by addition of the hypothesis  $H$  *would* attain  $T_{\mathcal{A}}$ . That *GW* recognizes  $\rightarrow^{subj}$  can be considered an advantage. It causes it to be reflected in the abduction schema that  $K(H)_{\mathcal{A}}$  is not a solution of  $\mathcal{A}$ 's  $T_{\mathcal{A}}$ -occasioned ignorance-problem, but rather only a subjunctive solution. Accordingly, *GW* reflects, where *AKM* conceals, what is perhaps the single-most distinctive feature of abductive inference. Whereas deduction is truth-preserving and induction is likelihood-enhancing, abduction is *ignorance-preserving*.<sup>5</sup>

<sup>2</sup> Here is an informal illustration of the difference between  $K_{\mathcal{A}}$  and  $K_{\mathcal{A}}^*$ . Suppose I want to know how to spell “accommodate”. Does it have two ‘m’s or one? I can’t remember. So  $K_{john}$  does not solve my ignorance-problem. But I have a dictionary within reach, and consult it. Now I know. But this knowledge is in  $K_{john}^*$ .

<sup>3</sup> Of course, the devil is in the details. Specifying the  $k_i$  is perhaps the hardest open problem for abductive logic. In most *AKM* treatments, the  $k_i$  include the consistency and minimality constraints (lines 4 and 5 of the schema). I for one am not so sure (see, e.g. Gabbay and Woods (2006)). Still, some constraints are clearly required. I am inclined to think that likelier candidates are the relevance and plausibility of  $H$  and  $K(H)$ .

<sup>4</sup> In the experimental sciences, *GW* inferences very often stop at line 7. Instead of releasing the conjectural hypothesis for further premissory work, the hypothesis is expressly withheld from such work, and is instead to independent experimental trials. This we may think of as “partial”, as opposed to “full” abduction.

<sup>5</sup> This is not to say that successful abducers are left wholly in the dark about what they are interested in knowing, or that their abductions are blind guesses (indeed “shots in the dark”). What is required is that an agent’s target set the cognitive bar at a level that reflects his interests and his circumstances,

## 2. *Semantic tableaux abduction*

It is well worth repeating that the semantic tableaux treatment of abduction has a number of attractions. In addition to the one already mentioned, it should also be remarked that Aliseda has engineered the reconciliation of abduction and semantic tableaux with considerable technical elegance.<sup>6</sup> Still, as I say, I have reservations. They turn mainly on the ignorance-condition. I will defer consideration of criticisms that might be advanced against the ignorance-preservation claim until the section that follows. For the present, I want to expose the basic features of the tableaux interpretation of abduction, pointing out along the way the respects in which it appears not to instantiate the *GW* model. Aliseda writes that

To test if a formula  $\varphi$  follows from a set of premises  $\theta$ , a *tableau tree* for the sentences in  $\theta \cup \{\neg\varphi\}$  is constructed, denoted by  $\mathfrak{T}(\varphi \cup \{\neg\varphi\})$ . The tableau itself is a binary tree built from its initial set of sentences by using rules for each of the logical connectives that specify how the tree branches. If the tableau closes (every branch contains an atomic formula  $\psi$  and its negation), the initial set is unsatisfiable and the entailment  $\theta \vdash \varphi$  holds. Otherwise, if the resulting tableau has open branches, the formula  $\varphi$  is not a valid consequence of  $\theta$ . (Aliseda 2006, p. 98)

Assuming a language in which all formulas have either disjunctive or conjunctive normal forms, then there are two types of tableau rules to consider. Formulas  $\alpha$  that are in conjunctive normal form, the rule is to decompose the conjunction and stack its conjuncts  $\alpha_1$  and  $\alpha_2$ :

$$\text{Rule 1. } \alpha \quad \frac{\alpha_1}{\alpha_2}$$

For formulas that are in disjunctive normal form, the rule is to decompose the disjunction by branching its disjuncts  $\alpha_1$  and  $\alpha_2$ :

$$\text{Rule 2. } \alpha \quad \beta_1 \mid \beta_2.$$

These two transformations capture seven rules, according to which “[d]ouble negations are suppressed. True conjunctions add both conjuncts, negated conjunctions branch into two negated conjuncts. True disjunctions branch into two true disjuncts, while negated disjunctions add both negated disjuncts. Implications ( $\alpha \rightarrow \beta$ ) are treated as disjunctions ( $\neg\alpha \vee \beta$ )” Aliseda (2006, p. 100).

---

and, given that his response to the fact that he cannot attain that target with the cognitive resources ready to hand, his abductive response generates a hypothesis which, whatever its own cognitive merits, does not meet the standard. We have it, then, that in a not untypical case, an abducer’s *H* might well be one already independently supported to some degree by some evidence but, let us say, short of what would justify its release as provisional guide to new action. The agent would have occasion to abduce it if it were insufficient on its own to meet the agent’s tougher target and if together with his  $K_{\perp}$  it would have attained it were it true. So hypotheses can have cognitive virtues all their own, independently of their abductive involvements. What abduction adds to the picture is further reason to detach the hypothesis, notwithstanding that that reason leaves it true that the higher standard of the target is only subjunctively met.

<sup>6</sup> Of course, the semantic tableaux approach does not originate with Aliseda, as the bibliography of Aliseda (2006) makes clear. But much in the way of its development here is hers alone. For earlier work, see Mayer and Pirri (1993).

A virtue of the semantic tableau approach to validity-assessment is that whenever  $\theta \neq \varphi$ , the failure of consequence is reflected by the open branches, which may be understood as describing models for  $\theta \cup \{\neg\varphi\}$ . “This fact suggests that if these *counterexamples* were ‘corrected by amending the theory’, through adding more premises, we could perhaps make  $\varphi$  a valid consequence of some (minimally) extended theory  $\theta'$ ” (p. 107). Consider a least class of *wffs* that realize this objective. In that case,  $\varphi$  would be derivable from a minimal extension of  $\theta$ . Aliseda proposes that the finding of such *wffs* is a kind of abduction.

Aliseda is a consequentialist about abduction. In her basic treatment of it, she is also an explanationist. Accordingly, the  $\rightarrow$  of her *AKM* orientation is such that whenever we have it that  $K(H) \rightarrow E$  then  $K(H)$  is an explanation of  $E$ . This sets her up to propose (p. 108) that:

Given  $\theta$  (a set of formulae) and  $\varphi$  (a sentence),  $\alpha$  (i.e. a hypothesis) is an abductive explanation if [the following conditions hold]:

*Plain:*  $\mathfrak{I}((\theta \cup \neg\varphi) \cup \alpha) \cup \alpha$  ( $\theta, \alpha \models \varphi$ ).

*Consistent:* Plain abduction + ( $\theta \neq \varphi$ )  
 $\mathfrak{I}(\theta \cup \alpha)$  is open

*Explanatory:* Plain abduction +  
 (i)  $\mathfrak{I}(\theta \cup \neg\varphi)$  is open ( $\theta \neq \alpha$ )  
 (ii)  $\mathfrak{I}(\alpha \cup \neg\varphi)$  is open ( $\alpha \models \varphi$ ).

Further conditions are also required. Added *wffs* must be in the vocabulary of  $\theta$ , and each must be a literal, a non-repeating conjunction of literals or a non-repeating disjunction of literals. We turn now to the question of how plain abductions are computed. In the interest of space, I shall omit consideration of the computation of consistent and explanatory abduction (for which see Aliseda (2006, pp. 109-116).)

#### *Plain abduction*

*Input.* A set of *wffs* representing theory  $\theta$ . A literal  $\varphi$  representing the fact to be explained. Presentations:  $\theta \neq \varphi$ ,  $\theta \neq \neg\varphi$ .

*Output.* Generates the set of explanative abductions  $\alpha_1, \dots, \alpha_n$  such that  $\mathfrak{I}((\theta \cup \neg\varphi) \cup \alpha_i)$  is closed and the  $\alpha_i$  satisfy the previously noted lexical and syntactic conditions.

*Procedures:* Calculate  $\theta + \neg\varphi + \{\mathfrak{I}_1, \dots, \mathfrak{I}_n\}$ , where the  $\mathfrak{I}_I$  are branches of  $\mathfrak{I}$ . Select those  $\mathfrak{I}_j$  that are open.

*Atomic plain abduction.* Compute *TTC* ( $\{\mathfrak{I}_1, \dots, \mathfrak{I}_n\} = \{\gamma_1, \dots, \gamma_m\}$  = the total tableau closure = the set of literals that close all branches concurrently. Then  $\{\gamma_1, \dots,$

$\gamma_n$  is the set of literals that close all branches concurrently. Then  $\{\gamma_1, \dots, \gamma_n\}$  is the set of plain abductions.

*Conjunctive plain abduction.* For each open branch  $\mathfrak{S}_I$  construct its partial closure  $BPC(\mathfrak{S}_I)$  = the set of literals that close that branch but do not close any of the other open branches. Then determine whether all  $\mathfrak{S}_I$  have a partial closure. Otherwise there is no conjunctive solution (in which case go to END). Note that each  $BPC(\mathfrak{S}_I)$  contains those literals that partially close the tableau. To construct a conjunctive abduction, take a single literal of each  $BPC(\mathfrak{S}_I)$  and form their conjunction: For each conjunctive solution  $\beta$  delete repetitions. The solutions in conjunctive forms are given by  $\beta_1, \dots, \beta_n$ . END.

*Disjunctive plain abductions.* Construct disjunctive explanations as follows: Combine atomic abductions with atomic abductions, Conjunctive abductions with conjunctive abductions, and each and conjunctive with  $\varphi$ . Example: Construct pairs from the set of atomic abductions and form their disjunctions ( $\gamma_i, \dots, \gamma_j$ ). Example: For each atomic abduction, form a disjunction with ( $\gamma_i \vee \varphi$ ). The result of all such abductions is the set of abductions in disjunctive form. END.

Although there is considerably more detail to Aliseda's model of abduction, enough of it has been presented here to make possible a number of observations. One is that semantic tableaux abduction resembles *enthymeme resolution*. In fact, the more abstract the kind of approach to abduction the greater the similarity to enthymeme resolution. In enthymeme resolution, the task is to find a  $\varphi$  that closes a  $\vdash$ -connection between some premisses and a conclusion, where  $\vdash$  is read as deductive consequence. In other words, the task is to rehabilitate an ailing deduction. In semantic tableaux abduction, the task is to find a  $\varphi$  that closes a  $\vdash$ -connection between a theory and some (usually) empirical data, which also is the repair of a decrepit deduction. Although in Aliseda's basic theory,  $\vdash$  is deductive consequence, she expressly allows for softer interpretations in which  $\vdash$  operates as inductive or statistical consequence (p. 38). Still, the "received view" (p. 38) is that abduction is explanative and that  $\vdash$  expresses the deductive consequence relation of the deductive-nomological (*D-N*) model of scientific explanation. Treated as deductive consequence, there is nothing abductive about the closure of  $\vdash$ -connections, irrespective of whether the task is enthymematic or *D-N* explanationist. That is to say, *by the GW model*, there is nothing that is abductive about such closures. What makes this so is that  $\vdash$ -closures are completely possible without there being any ignorance-problem to which the closure is a response, much less an ignorance-preserving response. This is especially clear in the core theory, in which  $\rightarrow$  is the consequence relation of *D-N* explanation. In its most general form a *D-N* explanation is a deduction in the form:

- |                      |  |
|----------------------|--|
| 1. $C_1, \dots, C_n$ | statements of initial conditions               |
| 2. $L_1, \dots, L_m$ | general laws or lawlike sentences              |
| 3. $E$               | description of empirical data to be explained. |

This is problematic. Let  $E^*$  be some unexplained proposition. The  $D-N$  model provides that there is no set of laws  $L_i$  and no set of initial conditions  $C_i$  such that  $\{\{L_1, \dots, L_i\} \cup \{C_1, \dots, C_i\}\} \vdash E^*$ , yet for some  $\varphi$   $\{\{L_1, \dots, L_i\} \cup \{C_1, \dots, C_i\}\} \cup \{\varphi\}$  the desired entailment holds. Since it is not required for the entailment to hold that  $\varphi$  be either a law or lawlike sentence or, for that matter, an initial condition, then neither do we have an  $D-N$  explanation of  $E^*$ . However, for  $\varphi$  to work in a classic  $D-N$  explanation of  $E^*$  there is the requirement that  $\varphi$  be a proposition of which we have knowledge. (It must function as an “initial condition”). This creates two difficulties at once. The first is a difficulty for Aliseda’s core theory. The other is a difficulty by the lights of the  $GM$  model. There is nothing about the semantic tableaux model that requires abductive inferences to be ignorance-preserving. If we suppose for a moment that the  $GW$  model is sound, then it is easy to diagnose the problem with Aliseda’s account and to set out the required method of repair. The diagnosis would be that Aliseda’s model is too abstract, leaving important features out of account. Accordingly, the method of repair would be to admit those features and juggle the model to facilitate their accommodation. I daresay that it will strike the  $AKM$ -minded reader as transparently tendentious, if not outright insulting, to proffer this diagnosis and this method of repair. Doing so is tantamount to saying that what’s wrong with Aliseda’s model is that it is not Gabbay’s and Woods’s model, and that what will put the Aliseda model right are changes that transform it into the  $GW$  model. So this is not an offer that I am much inclined to make. Rather than assuming the superiority of the  $GW$  model over the Aliseda model (or any variant of the  $AKM$  model), would it not be preferable to examine whether there is any reason to assert it?

Before leaving this section, it might be useful to make some brief clarificatory remarks about the two models. These observations help us see that, in as much as the two abductive schemata are abstract constructions, inevitably they leave certain features of abduction under-represented. For example, both models assume inference to the best/a good explanation as the most common, and perhaps most intuitive, form of abductive reasoning. And so it is, properly understood. It is easy to see, however, that nothing in either schema supports the notion that the inference embedded in an explanationist abduction is an inference to the best explanation. If we look at the  $AKM$  schema, the inference it draws is performed at line 7. This is the inference to  $H$ . But it is made quite clear at line 3 that  $H$  does not explain  $E$ . So any inference that  $H$  is any explanation, never mind the best, is untenable. The same point applies to the  $GW$  schema, whose inference (also at line 7) is that  $H$  is fit to be conjectured. Yet at line 4 we have it that  $H$  does not produce any explanation that satisfies  $T_A$ . Hence the inference at line 7 is not to the best explanation, or any. If we persist in giving the phrase “inference to the best explanation” anything like its ordinary meaning, then the  $AKM$  and  $GW$  schemata are both seriously flawed. Upon reflection, however, it seems that “inference to the best explanation” is a prepositional malapropism. We should

speak instead of inference *from* the best explanation. As it happens, then, the fault here is not that the schemata underrepresent this form of abduction—at least in the aspect currently under discussion—but rather that the official name of the inference misdescribes it.

Even so, these are representational lacunae that neither model picks up on. Consider, for example, that Harry wants to know whether  $H$  explains  $E$  and doesn't, with present resources, know whether it does. Then if we consider Harry's state as triggering an  $AKM$  abduction problem for this  $H$  and  $E$ , we will have landed in the same mistake as just above; for the schema expressly says that  $H$  does not explain  $E$ . The same is true of the  $GW$  schema. A similar difficulty attends the target of wanting to know what explains  $E$ . If this generates an abductive response according to the  $AKM$  schema, the inference terminates at line 6, which says that  $K(H)$  is what explains  $E$ . Ditto for the  $GW$  schema. It terminates at line 5, with the (admittedly weaker) claim that  $K(H)_A$  subjunctively explains  $E$ . Either way, the subsequent inference to  $H$  (or in  $GW$  to  $C(H)_A$ ) is no part of the response to the present target. What these glitches get us to see is that not all ignorance-problems are triggers of abductive responses. Neither schema contradicts this fact, but neither gives it any formal recognition either.

### 3. *Is abduction ignorance-preserving?*

Part of my reservation about Aliseda's adaptation of the  $AKM$  schema is that it does not capture the ignorance-preservation property of abduction. In the interest of fairness, it must also be said that the  $GW$  schema doesn't capture it either. So here is another respect in which both models are under-representing. Since nothing in Aliseda's theory give the slightest indication of acknowledging the ignorance-condition, it must also be conceded that what Aliseda's schema leaves out is what Aliseda never thought necessary to put in. On the other hand, since Gabbay and I make so much of it, the same omission is plainly a deficiency of the  $GW$  model that warrants the judgement of under-representation. A rescue (of sorts) from this embarrassment is offered by some critics of the  $GW$  model, who would argue that, since the ignorance-preservation claim about abduction is not true, then in leaving it out the  $GW$  is not after all guilty of under-representation. Perhaps I may be forgiven for thinking that the preferred rescue is one that takes Gabbay and me out of the frying pan of schematic under-representation straight into the fire of conceptual distortion.

Aliseda is not unaware of the  $GW$  model and of its doubtful compatibility with her own approach. (See, for example, Aliseda (2005), especially p. 52, n.1.) Perhaps she thinks that such wrangles are unseemly. Perhaps she thinks that the  $GW$  model is meritless. Perhaps she favours a latitudinarian pluralism about such things. A reading of Aliseda (2005) suggests (to me) that she also may think that some of the features that Gabbay and I claim for abduction are not features that it fall within the remit of a *logic* to recognize, much less elucidate. Geoffrey Goddu makes a point to the same effect. He asks, "... why is the fundamental nature of abduction resting on the *epistemic* status of  $H$  relative to  $K$  and  $T...$ ?" (Goddu (2005, p. 293, emphasis added)), thus suggesting that Gabbay and I have strayed illicitly (irrelevantly? unwisely?) into the



precincts of epistemology. Yes we have, and have pressed the claim that in so doing we have not left the province of logic. Admittedly, we have a lot riding on the idea that a practical logic is one that must take into account the cognitive constitutions of actual reasoners, a theme we have sounded in the early chapters of Gabbay and Woods (2005) and elsewhere. Goddu himself has cordially reckoned that “Gabbay and Woods have offered a comprehensive of discussion and made an intriguing case for embedding abduction within their quite promising practical logic of cognitive systems. Indeed, I strongly suspect that the true merits of their approach will be much more evident against the backdrop of a more comprehensive theory of practical cognitive systems—but that theory must wait for, at least, the third volume in the series” (Goddu 2005, p. 294). With this, Gabbay and I heartily and gratefully concur, even where most logicians would not. So I conjecture that a reason that Aliseda doesn’t tackle the *GW* model head-on is that, since hers is expressly a work of logic, she does not see the *GW* model as a rival of it.

What, then, is the case against the ignorance-condition? Here is Goddu on the point:

... there is no argument for the ignorance-condition —(or cognitive deficit condition) on abduction—it is stipulated (repeatedly) and then used to argue that many cases of what we thought were abduction are (or perhaps might) not really be abduction at all. For example, non-subjunctive deductive-nomological explanationism and evidentially clinching inference to the best explanation are ruled non-abductive because they violate the ignorance-condition. ... Given that so much of what is taken to be examples of abductive behavior is getting thrown out, one might doubt the logical necessity of the ignorance condition on abduction. (Goddu 2005, p. 293)

How are we to respond to this objection? First a minor clarification or two. It was never Gabbay’s and my intention to stipulate the ignorance-condition into existence. Our aim was to explicate the concept of abduction, and the ignorance-condition was offered as part of that explication. In this were guided by what we think can be learned from the abductive behaviour of actual reasoners on the ground. But we also wanted to give due recognition to the modern founder of abductive logic, who emphasized the guessing character of abduction (Peirce 1931-1958, 5.172). We also wanted to acknowledge Peirce’s emphasis on its subjunctive and conjectural features.<sup>7</sup> I am mindful, of course, of Quine’s quip that one person’s explication is another’s stipulation and that there is no principled way of marking the distinction with requisite precision. So perhaps what I see as explication Goddu sees as stipulation. It doesn’t matter. If it is merely stipulation, then the explication fails. If Goddu is right, if the ignorance-condition is nothing but stipulation, then it is false that abduction is intrinsically an ignorance-preserving mode of inference. Goddu says that we offer no argument for the condition, but I would plead the efforts of chapter 4 of Gabbay and Woods (2005).

Of course, it is perfectly possible for Goddu to be quite wrong about the mere-stipulation and no-justification objections and yet to be quite right in saying that ignorance-preservation is not intrinsic to abduction. My own view is that he is indeed

---

<sup>7</sup> “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.” (Peirce 1932-1958, 5.189. Emphases added.)

wrong about the former and right about the latter. Since (*tu quoque*) Goddu presents no argument for his own objection, I shall develop one for him.

Consider the case in which the cognitive target  $T_A$  of would-be abducer  $A$  is wanting to know whether there is at least a minimally satisfactory reason to conjecture that  $H$ . Let us dub the propositional content of this want ' $M$ '. Suppose now that  $A$  performs an abduction that conforms to the  $GW$  schema. The very fact that at line 7  $A$  utters  $C(H)$  commits him to  $M$ . What is more, the very fact that  $C(H)$  is the conclusion of some correct abductive reasoning makes it true that  $M$  for this interpretation of ' $M$ '. It also constitutes a justification of  $A$ 's utterance of  $C(H)$ . In sum, then, we have it that  $M$  is true, that  $A$  is justified in accepting  $M$  and that  $A$  does indeed accept  $M$ . Accordingly, on at least one deeply entrenched model of knowledge, in performing the abduction currently under discussion,  $A$  attains his cognitive target of wanting to know whether  $M$ . So successful abduction is not always ignorance-preserving. Point to Goddu.

This is useful to know. It takes some of the pressure off my criticism of the  $AKM$  model, and it lends further emphasis to the point above that the content of an agent's cognitive target is has a bearing on the logical character of how it is responded to. What the present example tells us is that if you set your cognitive bar very low, it is not ruled out that an abductive response will close the cognitive gap between what is inferred and what the target calls for. Although the example removes the ignorance-condition from the status of logically necessary condition on abduction, it does nothing to discourage the idea that, even so, it remains a *generically* necessary condition on abduction; that is to say, that it is *typical* of abduction that whatever the level of its own cognitive accomplishment, it doesn't in general rise to the cognitive level of what the originating target calls for. If this is right, then the objection to press against the  $AKM$  model is not that it fails to capture abductions but rather that the cases it does capture are not typical of abduction.

Suppose that we now allow that abduction is typically ignorance-preserving. But why should we be in the least drawn to the idea that ignorance-preservation is distinctive of abductive reasoning? Is the same not true of any mode of reasoning under what are called "conditions of uncertainty"? Come to that, is the same not also true of every case of valid but unsound deductive reasoning? Perhaps we should concede that ignorance-preservation is not *unique* to abductive reasoning. This only intensifies the pressure to make clear the distinctiveness claim. Possibly it is a matter of degree. Virtually all the inferences we ever actually draw we draw defeasibly, that is, with a readiness to stand convicted of error. In virtually every such case, what we are able to determine falls short of what we would like to have. We quest for knowledge and settle for belief. This is the point at which I am prepared to be guided by Peirce, who insisted that belief is precisely what abduction does not secure for us (Peirce 1992, p. 178). What abduction sanctions is not belief but "a reason to suspect". Of course, it cannot be foreclosed in advance that in abducting a hypothesis  $H$  a reasoner has arrived at a proposition that he *does* believe. Peirce would say (and I too) that, even so, the belief that  $H$  is not *warranted* by the abduction. It is also clear that the belief-unwarranting conception of abduction is reinforced by Peirce's point that abductive

inference is guessing. So, then, what makes the preservation of ignorance a feature distinctive of abduction is this: Not only does abduction not secure us knowledge, it does not warrant belief.

#### 4. *Abduction in the law*

I don't know whether my typicality claim will satisfy Goddu. I don't know whether it will satisfy Aliseda. I daresay that there are plenty of people for whom what remains of this dispute with AKM theorists is a trivial quibble, a silly competition between "all" and "nearly all". Be that as it may, there are nevertheless areas of real-life application in which it matters greatly not only whether the ignorance-condition holds, but the form in which it holds. A case in point, to which I will now briefly turn, is criminal jurisprudence in the common law. In cases at the criminal bar, a central concept is that of a "theory of the case", also called a "theory of the evidence". A theory of the case is an inference to [actually, *from*] the best explanation. The prosecution's theory of the case is that the hypothesis of guilty as charged is the best explanation of the evidence heard at trial, and that on that account a guilty verdict is required. On the other hand, the defence's theory of the case (when there is one) is that the hypothesis that best explains the evidence is one that is incompatible with the hypothesis of guilt, and that, accordingly, a verdict of not guilty is required. In common law jurisdictions, these competing theories of the case are presented to the jury (or to the judge in juryless trials) after all the evidence has been heard. They are called "summations" or "closing arguments."

It is sometimes true that prosecutor's theory of the case is knock-down, that it convinces the jury of the certainty of the accused's guilt. In actual practice, it is rare for such cases to go to trial; typically counsel for the accused will negotiate a plea-bargain. Aside from prosecutions that are open-and-shut, there are three considerations that bear significantly on what we might call the "epistemic status" of convictions. One is that convictions are allowed on the basis of merely circumstantial evidence. Another is that a conviction does not exclude the existence of a reasonable case for acquittal.<sup>8</sup> The third has to do with the nature of evidence. It is all, so to speak, hearsay. Evidence is *testimony*. It is by and large testimony from complete strangers, of whose testimonial track record, honesty and reliability the jury has no independent knowledge. It is true that testimony is given under oath and that there are tough penalties for perjury. But, as anyone familiar with actual juridical practice well knows, witnesses are often unreliable, evasive or dishonest, and perjury charges are a comparative rarity. Taken together, these facts make a strong case for thinking that it

---

<sup>8</sup> I can report anecdotally that this feature of sound convictions strikes people as so shockingly permissive as to incline them to disbelieve it. Shocking or not, it should not be disbelieved. See, for example a 1978 ruling by the First Federal Appeals Court: "The prosecution may prove its case by circumstantial evidence, and it need not exclude every reasonable hypothesis of innocence . . ." (*U.S. v. Gabriner*, 571 F. 2d 48, at 50 (1st Cir., 1978)). See also its subsequent ruling: "The trier of fact is free to choose among various reasonable constructions of the evidence." (*U.S. v. Thornley*, 707 F. 2d 622, at 625 (1st Cir., 1983)).

is altogether typical that when jurors find an accused guilty of the offence with which he has been charged, they do not know whether in fact the offence was committed by him. True, the criminal proof standard requires that a juror must be satisfied beyond a reasonable doubt, but it is easy to see that satisfaction is one thing and knowledge another.

With the clarifications of the previous section at hand, we can say that on the *GW* approach abductions are typically ignorance-preserving. There isn't a juror alive (well, an honest juror) who doesn't want to know whether the accused is guilty as charged. In actual practice, the probability of his attaining this knowledge is exceedingly low. But when we examine the ratio of convictions to failures to know, we see that it is utterly common for juries to convict without knowing whether the accused is guilty. The fact is that testimony rarely closes the gap completely between what a juror desires to know and what he does know. So what a juror urgently requires *tant de mieux* is a form of reasoning that will allow in principle him to proceed reasonably to a vote to convict even in the absence of the knowledge he wishes to have. Whatever the details of such a mode of reasoning, it is clear that it is reasoning which responds to without solving an ignorance-problem. So it is, in our sense, ignorance-preserving reasoning. It is often supposed that juror's duty is to choose between the rival abductions of opposing counsel. It takes little reflection to see that occasions may present themselves in which, as finders of fact, jurors must construct their own theories of the case. Either way, whether a juror rests his decision on the theory of one or other of the counsel or on a theory of his own contrivance, his verdict to convict is based on an abduction. Putting these two characterizations together, we see that in the large majority of criminal convictions in common law verdicts are reached by ignorance-preserving abduction, which is anyhow what abduction typically is.<sup>9</sup>

It is evident that this is problematic. The standard for conviction is proof beyond a reasonable doubt. But, for all the cases that we are presently considering, convictions are actually rooted in conjectures, in inferences from testimony that are ignorance-preserving and belief-unwarranting. Since one would hardly think that a conjecture is *ever* a proof, there would appear to be a massive problem for the intellectual integrity and justice of criminal convictions.

It is worth emphasizing that if, as do I, accept a Peircean interpretation of the ignorance-preservation of abduction, then we have it that which justifies the verdict that the accused is guilty does not warrant the belief that he is, never mind that on occasion that belief may actually be present in a juror's mind. Clearly, this sharpens the string of our puzzle. How can an interpretation of the evidence which proves guilt be insufficient to justify belief?

---

<sup>9</sup> Let us also note in passing that the abductions that drive convictions are, in the sense of note 4, "full" abductions. A verdict to convict is one that releases the hypothesis of guilt as a basis for subsequent actions. In this case, the subsequent actions are sentencing and punishment. That this release is, even so, defeasible is reflected in the law's provision for the appeal of and the pardon from conviction, although in capital punishment jurisdictions this is more an honorific defeasibility than a real one.

This is not the place to determine whether this nasty-looking puzzle can satisfactorily be disarmed.<sup>10</sup> But one thing is quite plain to see. If the problem *can* be disarmed, then there must be a form of reasoning with respect to which it can be shown that it sanctions actions of deep practical import without the benefit of knowledge or warranted belief, and does so on the basis of the explanatory force of the hypothesis that generates the action. Another way of saying this is that, short of giving up on them altogether, the provisions for conviction must allow for reasoning which, while not ignorance-extinguishing or belief-warranting, is nevertheless *probative*. What this requires, in turn, is a conception of explanation according to which the explanatory force of the case for conviction is probative without being, as Goddu puts it, “evidentially clinching”. Whatever the law’s prospects for the defence of its own intellectual and moral integrity, it is evident that this is a problem for the logic of abduction, concerning which it hardly matters whether abduction is always ignorance-preserving or nearly always ignorance-preserving. Still, we should not leave it unremarked that one possible solution is that, contrary to what everyone in the mainstream of abductive logic believes, while theories of the case may be inferences from the best explanations, inference from the best explanation is not abduction; at least not intrinsically.

### 5. Conclusion

Strictly speaking, the *AKM* and *GW* models are incompatible with one another. But, when all is said and done, the principal difference between the *AKM* and the *GW* approach is generality and complexity. Even if the *GW* approach is accepted, no one has to stop mining the rich seams of the *AKM* model. From this perspective, the work done by *AKM* theorists is nothing but welcome, albeit with some tinkering here and there. Aliseda’s semantic tableaux treatment of abduction shows the *AKM* model to considerable advantage. Everyone should be grateful.<sup>11</sup>

### REFERENCES

- Aliseda Llera, A. (1997). *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. Amsterdam: Institute for Logic, Language and Computation. PhD dissertation (ILLC Dissertation Series 1997-4).
- (2005). “What is a logical system? A commentary”, in *We Will Show Them: Essays in Honour of Dov Gabbay on His 60<sup>th</sup> Birthday*, Vol. 1, pages 51-55, edited by Sergei Artimov, Howard Barringer, Artur d’Avila Garcez, Luis C. Lamb and John Woods. London: College Publications.
- (2006). *Abductive Reasoning: Logical Investigation into the Processes of Discovery and Evaluation*. Amsterdam: Springer.
- Flach, P.A., and A.C. Kakas (2000). *Editors, Abduction and Induction: Essays on Their Relation and Interpretation*. Dordrecht and Boston: Kluwer.

<sup>10</sup> For a discussion of this and related issues, see Woods (2008a), Woods (2008b) and Woods (2008c).

<sup>11</sup> For encouraging comments and for instructive demurrals I warmly thank Atocha Aliseda, Paul Bartha, Peter Bruza, Balakrishnan Chandrasekara, Erik Krabbe, Manfred Kraus, Theo Kuipers, Lorenzo Magnani, Joke Meheus, Sami Paavola, Paul Thagard and, of course, Dov Gabbay. For financial support I thank The Abductive Systems Group at UBC and the Physical Sciences and Engineering Council of the United Kingdom, and, for technical support, Carol Woods in Vancouver.

- Gabbay, D.M., and J. Woods (2005). *The Reach of Abduction: Insight and Trial*, Vol. 2 of *A Practical Logic of Cognitive Systems*. Amsterdam: North-Holland.
- Gabbay, D.M., and J. Woods (2006). "Advice on abductive logic", *Logic Journal of IGPL*, 14, 189-219.
- Goddu, G.C.(2005). "Woods and Gabbay's The Reach of Abduction: Insight and Trial", *Informal Logic* 25, 289-294.
- Kakas, A., R.A. Kowalski and F. Toni (1995). "Abductive logic programming", *Journal of Logic and Computation* 2, 719-770.
- Kowalski, R.A. (1979). *Logic for Problem Solving*. Amsterdam and New York: Elsevier.
- Magnani, L. (2001). *Abduction, Reason and Science: Processes of Discovery and Explanation*. Dordrecht: Kluwer and New York: Plenum.
- Mayer, M.C., and F. Pirri (1993). "First order abduction via tableau and sequent calculi", *Bulletin of the IGPL*, Vol. 1, 99-117.
- Meheus, J., L. Verhoeven, M. van Dyck and D. Provijn (2002). "Ampliative adaptive logics and the foundation of logic-based approaches to abduction", in L. Magnani, N.J. Nersessian and C. Pozzi (eds.), *Logical and Computational Aspects of Model-based Reasoning*. Dordrecht and Boston: Kluwer, pp. 39-77.
- Peirce, C.S. (1931-1958). *Collected Papers*. Eight volumes, edited by Charles Hartshorne, Paul Weiss and Arthur Burks. Cambridge, MA: Harvard University Press.
- (1992). *Reasoning and the Logic of Things: The Cambridge Conference Lectures of 1898*. Kenneth Laine Ketner editor, with an introduction by Kenneth Laine Kettner and Hilary Putnam. Cambridge, MA: Harvard University Press.
- Woods, J. (2008a). "The criminal abduction paradox", in Shahid Rahman (ed.), *Argumentation and Law*. Amsterdam: Springer Netherlands, to appear.
- (2008b). "Reasonable doubt and incommensurability", *Informal Logic*, to appear.
- (2008c). "Relevance in the law: A logical perspective", in D. Canale and G. Tuzet (ed.), *Inferentialism in Law and Philosophy*, to appear.

**John Woods** is Director of the Abductive Systems Group at the University of British Columbia and the Charles S. Peirce Professor of Logic at King's College London. With Dov Gabbay he is author of the multi-volume work *A Practical Logic of Cognitive Systems*, published by North-Holland, of which two volumes have appeared: *Agenda Relevance: A Study in Formal Pragmatics* (2003) and *The Reach of Abduction: Insight and Trial* (2005). Woods's other books include *The Logic of Fiction* (1974), *Fallacies: Selected Papers 1972-1982* with Douglas Walton (1989/2007), *Paradox and Paraconsistency: Conflict Resolution in the Abstract Sciences* (2003) and *The Death of Argument: Fallacies in Agent-Based Reasoning* (2004). John Woods is Fellow of the Royal Society of Canada and President Emeritus of the University of Lethbridge.

**ADDRESS:** Department of Philosophy, University of British Columbia, 1866 Main Mall, Vancouver B.C. V6T 1Z1, Canada. E-mail: jhwoods@interchange.ubc.ca. Web: www.johnwoods.ca.