Advice on Abductive Logic

January 11, 2005

File: 268/268-AAL.tex

Dov M. Gabbay
Department of Computer Science
King's College
Strand
London
WC2R 2LS
UK
e-mail: dg@dcs.kcl.ac.uk
and

John Woods
The Abductive Systems Group
Philosophy Department
University of British Columbia
Vancouver, BC
Canada
V6T 1Z1
e-mail: jhwoods@interchange.ubc.ca
and

Department of Computer Science
King's College
Strand
London
WC2R 2LS
UK
e-mail: woodsj@dcs.kcl.ac.uk

Advice on Abductive Logic

The action of thought is excited by the initiation of doubt and ceases when belief is attained; so that the production of belief is the sole function of thought.

Charles S. Peirce

In the literature there is more or less agreement about the general nature of abduction.

Theo A.F. Kuipers

Abstract

One of our purposes here is to expose something of the elementary logical structure of abductive reasoning, and to do so in a way that helps orient theorists to the various tasks that a logic of abduction should concern itself with. We are mindful of criticisms that have been levelled against the very idea of a logic of abduction; so we think it prudent to proceed with a certain diffidence. That our own account of abduction is *itself* abductive is methodological expression of this diffidence. A second objective is to test our conception of abduction's logical structure against some of the more promising going accounts of abductive reasoning.

We offer our various suggestions in a benignly advisory way.

The primary targets of our advice is ourselves, meant as guides to work we have yet to complete or, in some instances, start. It is possible that our colleagues in the abduction research communities will find our counsel to be of some interest. But we repeat that our first concern is to try to get *ourselves* straight about what a logic of abduction should encompass.

It's not usual for citations or footnotes to be included in abstracts.

1 Abduction

The term 'abduction' was introduced into logical theory by Charles Peirce in the late 19th century. The introduction was not wholly original, since 'abduction' is a passable translation of Aristotle's *apagogē*, which is also translated as 'reduction' and was given the Latin rendering *abductio* by Julius Pacius. For Aristotle, an abduction is a syllogism, from a major premiss which is certain and a minor premiss which is merely probable, to a merely probable conclusion (*Prior Analytics* 2.25 69^{a}20–36). An important modern development cites the importance of reasoning from causes to effects. An insightful discussion can be found in Laplace's *Mémoires* [Laplace, 1904].[1]

In his early attempts to characterize abduction, Peirce also takes a syllogistic approach. Later on, he saw abduction as a form of reasoning in which a new hypothesis is provisionally accepted on the grounds that it explains the available data. On this view abduction is an inference in the form

[1]Anyone interested in whether there is a "Hume" problem for for abduction might consult [Woods, 2004a]. Also, the interconnections between abductive and inductive logic are well-explored by [Flach and Kakas, 2000].

> 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, 1931–1958, 5.189,1931–1958] [2]

Peirce's explanationist leanings notwithstanding, it is necessary to take note of some significant ambiguities in the concept of abduction. In its most general sense, abduction is a process of justifying an assumption, hypothesis, or conjecture for its role in producing something in which the abducer has declared an interest. Within this category various distinctions press for recognition. The hypothesis might as Peirce says help *explain* a given set of data or some phenomenon; or, it might facilitate the generation of observationally valid *predictions*; or it might permit the *elimination of other hypotheses*, thus providing the theorist with a *simpler* and *more compact* account; or it might permit the *unification of disparate laws*. Here, then, are just four distinct reasons which an abducer might offer as a justification for making a given hypothesis or conjecture. Abductions of this sort have an unmistakably pragmatic character. They are justifications of use without being evidence of the truth of the hypotheses in question. (See below)[3]

2 Triggers

Abduction offers two faces for the investigator's scrutiny. One is abduction the process, the other is abduction the product. In a rough and ready way, abductive products are investigated by way of properties possessed by the requisite linguistic structures or of linguistic structures in relation to abstract set theoretic structures. Abductive processes are investigated by way of conditions on the success or failure of the abductive behaviour of cognitive agents in actual practice. Both product and process are important foci of the investigator's probes; but in the approach taken in our work, we try to give to considerations of process a due consideration.

In its barest form, abduction is a reaction of a certain kind to a *cognitive irritant*. As Rescher nicely observes, "The discomfort of unknowing is a natural component of human sensibility". [Rescher, 1996, p. 5]. The irritation is occasioned by the inability to hit some or other cognitive target with present epistemic resources. The cognitive target is in its turn constituted by some or other state of affairs. Putting the occasioning state of affairs as S, the set of our present cognitive resources (or knowledge-base) as K, the cognitive target occasioned by S as T, and the relation that fails to obtain between K and T as R, then (as a first, and less than adequate, pass) the basic form of an *abductive trigger*[4] is

1. S obtains
2. S occasions T
3. K does not bear R to T.

[2] A good short overview of Peirce on abduction is [Kraus, 2003].

[3] Newton, for example, accepted the action-at-a-distance theorem, but he was firm in thinking it unbelievable.

[4] We borrow this attractive metaphor from Aliseda-LLera [1997].

Targets are a kind of cognitive *agenda*. An agenda is cognitive when it admits of closure by way of an agent's cognitive states. As commonly understood, agendas can be both closed or advanced. Advancement can be likened to partial closure. This same partially also extend to abduction. An abductive inference may close an agenda or partially close it.

For ease of exposition, we here concentrate on abductions that close agendas. Agendas are described in detail in [Gabbay and Woods, 2003]. For present purposes we adopt the notion informally.

It is important to repeat that:

Proposition 2.1 (The variability of abduction) *The parameters 'S', 'T' and 'R' admit of variable instantiation.*

In one set of circumstances, S may be a newly discovered fact that cannot be explained by what is currently known. In other cases, the unmet agenda associated with an abductive trigger can be entirely non-explanationist in character. If this is right, the theories such as those of [Aliseda-LLera, 1997; Magnani, 2001], [Aliseda, forthcoming]— indeed of Peirce himself— which are explanationist accounts, cannot qualify as general theories of abduction.[5] There is much to be gained in getting straight about abduction's parts before reaching for the whole. But our advice is that sooner rather than later it would be prudent for theorists to expand their horizons.

We briefly sketch a non-explanationist example. Let the state of affairs in question be one in which a set of proof rules implies a result which is thought to be unacceptable. Suppose further that the proof in question is not taken as a *reductio*. So the fact that its conclusion is unacceptable establishes (for those who think it so) that there is something wrong with the proof. If we assume that the proof misapplies none of its proof rules, then those who find the proof defective in this way must reject one or other of the rules used. This is the situation of a "proof" "proving" the wrong thing. The target is finding the defective rule. But since the current rules encode what is currently known about these proof-structures, that agenda is not closable from K.

3 Ignorance Problems

We begin by introducing the idea of an ignorance problem IP.

Definition 3.1 (Ignorance problems) *An IP exists for a cognitive agent X iff X has a cognitive agenda T that cannot be closed from what he currently knows (or, equivalently, from K, his current knowledge base).*

IPs present cognitive agents with two options. One is to acquire new information that X will enable T to be closed. Accordingly, for an agent X,

[5]Thus Magnani: abduction is "inference to an explanatory hypothesis" [Magnani, 2001, p. xi] and Aliseda: abduction is "reasoning from an observation to its possible explanations"[Aliseda, forthcoming, p. 8]; and [Meheus *et al.*, forthcoming]: . . . logics for abductive reasoning enable one "to generate explanations for novel facts . . . as well as for anomalous facts . . . [Meheus *et al.*, forthcoming, p. 2].

IP-option (1) (X overcomes his ignorance) X extends K to some successor knowledge-base K^* such that K^* closes T.

Another option is to acknowledge that, at least *pro tem*, the pair $\{K, T\}$ constitutes for X an insolubilium. Accordingly,

IP-option (2) (X's ignorance overcomes him) Unable to succeed with option (1), X capitulates.

It is well to note the dynamic character character of this pair of options. For example, at time t_1, X might try and fail to exercise option (1). At t_2 he might acquiesce to option (2). Yet at t_3 he might recur to option (1) with good results.

It is commonly held that, when an agent is confronted with an ignorance-problem, (1) and (2) exhaust his option space. In fact, there is a third option. *It is the founding datum of abduction*. Theorists would be well-advised to give it their considered attention.

IP-option (3) (Presumptive closure) X finds an H which, if he knew it, would together with K solve his IP; and from that fact he conjectures that H.

Option (3) incorporates the element of conjecture in an essential way. This is obvious in the case of H itself, but what is often overlooked is that this does not solve the original problem. X's problem is that his T is closable only on the basis of what he now knows (K) or can readily get to know (K^*). His situation *now* is that T cannot be closed either way. If he selects an H such that the truth of K revised by H *would* hit T, then *conjecturing* H does not produce K^*. In particular, K adapted by H (hereafter $K(H)$) is not a knowledge-set for X. (It does not solve X's ignorance problem).

This highlights the second irreducible element of conjecture that option (3) embeds. $K(H)$ doesn't close T, but we may say that it closes it *presumptively*. Accordingly, option (3) offers X not a solution of his ignorance-problem, but rather closure *faute de mieux* of a lesser agenda. Instead of an agenda that admits of only *epistemic* closure, it proposes a conjectural variant of it that provides *presumptive* closure. This is deeply consequential.

Proposition 3.2 (Ignorance-preservation) *Whereas deduction is truth-preserving and induction is probability-enhancing, abduction is ignorance-preserving.*

Proposition 3.2 sets forth what we will call the *ignorance condition* on abduction.

Option (3), as we see, is not a solution of an IP; it is a *transformation* of an IP into a problem that conjecture can solve. It is a response to an IP that requires X to lower his sights with regard to T. It turns on X's willingness to satisfice rather than maximize.

Here, too, it is prudent to re-emphasize the dynamic character of IPs and the responses that they induce. A cognitive agent might try and fail with option (1), and then move to option (3). If it also failed him, option (2) might recommend itself. If option (3) succeeded, X might persist with it until, so to speak, he came to know better; in which case he might move to option (1). So we have it that, in the beginning, X might try to overcome his ignorance, and, failing that, might try to conjecture to a lesser target. If this fails, he might acknowledge that his ignorance has got the better of him. Yet even if he succeeded conjecturally, he might later chance upon the means to abandon conjecture for fact, and so solve, with new knowledge, the problem that started it all. So, we say that[6]

Proposition 3.3 (IP-relativities) *IPs arise in relation to agendas in play at a time and resources then available. Responses to IPs retain those agendas and proceed in ways permitted by subsequently available resources.*

Peirce and others have emphasized that it is a condition on the *scientific* admissibility of an abductive conjecture H that it be testable, at least in principle. By these lights, a solution to an abduction problem is also a step in a process that might eventually solve the originating ignorance problem. So, for the class of cases that Peirce has in mind,

Proposition 3.4 (Ignorance-mitigation) *Although a solution to an abduction problem preserves the ignorance that gave rise to it, it may also contribute to the solution of the originating problem by identifying candidates for the status of new knowledge.*

It is well to note that in some contexts, abductive conjectures are not scientifically testable. For example, various forms of philosophical skepticism attract inference-to-the-best-explanation abductions. It may be that the best explanation of our external world experiences is that there is an external world that produces them. But to require that the external world hypothesis be testable is to beg the question against the skeptic, which in turn, ruins the anti-skeptic's refutation. Accordingly,

Proposition 3.5 (Testability) *Testability is not intrinsic to the selection of successful abductive hypotheses.*

4 Frames

The dynamism of the IP-problematic also bears on the structure of options (1) and (2). Each turns on the availability of K^*. K^* is some future state of X's knowledge at a given time t. t is the time at which X recognizes that he has an IP, and his knowledge at that time is K. K^* is what X knows later, not anytime later, but later relative to what we might call the *frame* of his IP. It is impossible to be perfectly precise about this, save by stipulation. But intuitively the idea is sound, and clear enough to be getting on with. Consider an example. Harry wants to know whether Sarah will come to the picnic. He doesn't know. He phones her apartment; no answer. He phones her best friend; she doesn't know. There is presently no K^* for Harry that solves this problem. He has no idea, and so waits until tomorrow to see for himself. He goes to the picnic and

[6]See our companion paper in this volume "A formal model of abduction" [Gabbay and Woods, 2005c]. To construct a model of this process we need in $K + (H)$, to distinguish (via labelling) between H and the rest of the knowledge base. If later we chance upon a knowledge k_1 which obtains the target we must know that it is H which is to be deleted.

finds that Sarah isn't there. Today Harry acquiesces in option (2); but tomorrow he is able to deploy option (1). Doing so solves his ignorance-problem. But suppose instead that Harry fell ill and wasn't able to attend the picnic. Suppose that he never acquired a shred of additional information about Sarah's whereabouts on that day. Now, sixty years later, Harry is on his death-bed. Sarah appears. "Oh, Harry", she says, "how I wanted to attend that picnic all those many years ago!". Harry now knows that Sarah hadn't been there. But he hasn't resolved his IP problem. His new knowledge is outside its frame. This suggests that

Proposition 4.1 (IP-duration) *Typically an IP has a tacit "sell-by" date, after which it expires.* [7]

We now have the means to define *abduction problems* AB. With K and T set as before,

Definition 4.2 (Abduction problems) X *has an* AP *with respect to* K, T *iff he has an IP with respect to* K, T *in response to which he is disposed to exercise option (3).*

5 Generalizing IPs

APs are not natural kinds. An AP is an IP to which X responds in a particular way. X substitutes conjecture for knowledge. It is the received view that all abduction problems are transformations of ignorance problems. This is a mistake. It is easy to see that the structure of abduction problems is wholly preserved if we substitute for K *any* cognitive state in comparison with which presumption is epistemically junior (weaker belief standing in for strong belief is the obvious example). Accordingly, given that an ignorance problem represents an epistemic shortfall, a variant of it would represent a doxastic shortfall, or in some cases a plausibility shortfall. In each case, the conjecture deployed by the abducer's solution would have to meet two strong conditions.

Proposition 5.1 (Epistemic juniority) *If* H *is a solution of an* AP*,* H *has a lesser cognitive status than the cognitive standard against which the original problem arose.*

Proposition 5.2 (Effective juniority) *If* H *is a solution of an* AP*, then although there is a cognitive disparity between it and the cognitive standard against which the* AP *arose,* H*'s cognitive juniority must comport with the requirement that it produce a presumptive solution of* AP*.*

Proposition 5.1 generalizes on the ignorance-preserving character of abductive solutions to IPs. It provides that in its fully general form, abductive solutions are cognitive deficit-preserving. Proposition 5.2 offers the helpful admonition, that for all their cognitive limitations comparatively speaking, successful Hs must have the wherewithal to produce rationally adequate, though cognitively subpar, solutions of their APs. Proposition 5.1 gives us occasion to broaden the ignorance-condition. As now we see, in its

[7] The 'sell-by' date can be more precisely defined in our time-action modelling. See[Gabbay and Woods, 2003]. The IP-problem is pressingly required to enable some sequence of actions in some agenda. When the agenda is no longer relevant, we pass the 'sell-by' date.

more general form, the condition requires that abductive theories honour the *cognitive-deficit condition*. Henceforth we shall read the ignorance-condition in this more general way, in the absence of indications to the contrary.

6 Avoiding a confusion

When a reasoning agent conjectures an H that bears the presumptive attainment relation to his cognitive target T, he is operating at an *epistemic* disadvantage. If he cannot attain T on the basis K of what he now *knows*, he may conjecture a proposition H that he doesn't know but which, if it were true, would, in apposition to what he does know, attain T. Or, in a variation, if T cannot be attained on the basis of what a reasoner *strongly believes* or what he *takes to be highly probable*, his hypothesized H must be a proposition that he neither (that) strongly believes nor takes to be (that) highly probable. As we see, the epistemic juniority of H is relative to the epistemic standing of the K in relation to which the ignorance-problem arose initially. So it bears repeating that the agent's recourse to H is from a position of relative epistemic juniority, and that this aspect of juniority is expressly recognized in the fact that in selecting it, the agent is proceeding conjecturally. Note, however, that the content of the agent's conjecture of H is that H is true. This is as it should be, given that the conjecture of H turns on the fact (or what the abducer takes to be a fact) that if H were true, then H in apposition to K would attain the cognitive target T. Philosophers often characterize truth as an alethic property of propositions (or theories). Given that "alethic" derives from the Greek word for "true", the appellation has a certain redundancy about it, but not one that occasions any real harm. In fact, it is a baptism that affords us an essentially important distinction for the logic of abduction. Accordingly,

Proposition 6.1 (Epistemic v alethic factors) *While it is essential that a successfully abduced H possess the requisite epistemic juniority, it is neither necessary nor desirable that it be alethically subpar.*[8]

bf Corollary 6.1(a) If we put it that abducing a H is always a kind of guessing, it is easy to see that what the abducer hopes for is that his guess will turn out to be true. Abducers deliberately set their task as one of guessing, but they do not aspire to guess what is false.

The same lesson applies to K-parameters of strong belief or propositions held as highly probable. In conjecturing H, one's epistemic hold on it must be of a lesser grade than that of strong belief or propositions held as highly probable. But nothing precludes the abduced hypothesis hitting the alethic standard of truth. On the contrary.

[8] In classical approaches to truth, any proposition that is alethically subpar is false. In many-valued approaches, an alethically subpar proposition has a less truth-like value than the proposition to which it is subpar. In truth-approximation approaches, one proposition is alethically subpar to a second when the former is less approximately true than the latter.

7 Grounds of action

In a standard situation an ignorance-problem presents an agent with two choices. One is to acquire the knowledge that solves the problem and then to act on it in ways that may conduce to the agent's further interests. The other is (perhaps temporarily) to admit defeat and to postpone any action that would be suitably occasioned by a solution to the problem if it existed. As we have seen, there is also a third option. Perhaps its principal attraction is that it is an alternative to the passivity of giving up on one's IP. It is, of course, a qualified alternative, since it does not solve the IP but rather solves it presumptively. Notwithstanding this essential qualification, an abductive solution bears on the question of *action* in two important ways. In the one case, the abducer's embrace of H^c constitutes the *cognitive act* of releasing H for generally unfettered inferential work in the domain of enquiry within which the abducer's IP arose in the first place. In the other case, it is open to the agent to take whatever *further actions* as may comport with his other interests, on the basis of conclusions in the descendent class of inferences dependent upon H. This is far from saying that H's conjectural origins are overlooked in such cases. It means only that the actions are taken so with requisite regard to the higher risk than that that would attach to actions occasioned by what the agent does really know. Accordingly, it is a deep fact about abduction that

Proposition 7.1 (Abduction as a spring of action) *Abduced hypotheses H give agents a basis for consideration of subsequent actions involving degrees of risk concomitant with the strength of H's conjecture.*

8 The Adaptive and the Epistemically Subpar

Since it is fundamental to an abduction problem that an agenda we desire to close cannot be closed with anything that we presently know, the requisite ignorance must be invariant throughout the process of abduction. It is quite true that, once deployed, hypotheses are often made the objects of attempts to confirm them (corroborate, for those of Popperian bent). Sometimes these attempts are met with success. But this is not abductive success, but rather post-abductive. Accordingly,

Proposition 8.1 (Hypotheses and epistemic virtue) *If H is an hypothesis entertained or engaged in an abduction exercise, it is essential that H lacks some degree of epistemic virtue.*

In standard approaches to abduction, the epistemic-deficit condition is recognized only tacitly if at all. An exception is the *adaptive* orientation of Meheus and her colleagues [Meheus *et al.*, forthcoming]. In this manner of proceeding, abductive inferences are modelled as proofs in a modalized adaptive logic. In such proofs, priors carry the necessity operator $\Box$, whereas abductive conclusions carry the weaker modality of possibility, $\Diamond$. There is an apparent congruence between the $\Box$-statements and the $\Diamond$-statements of an adaptive abductive logic and the asserted priors and conjecturized conclusions of abductions according to our analysis. As might be expected, the two distinctions reflect significant conceptual dissimilarities. Even so, we have here *structural* acknowledgement of the ignorance condition.

9 Knowledge-Sets

The idea of knowledge-*sets* is something of an understatement. We may take it that the totality of what a human knower knows at a given time is a comparatively fuzzy assemblage of various modules. This modularity has something to do with the variability of our epistemic capacities and circumstances. A proposition that lacks a proof may be thought not to qualify as mathematical knowledge; a proposition that fails to negotiate the rigours of scientific method may not be thought of as scientific knowledge; a proposition that failed the requisite standard of juridical rectitude may not qualify as legal knowledge; and so on. On the other hand, propositions of this sort might well qualify as common knowledge.

There follows from this a point of considerable importance for a theory of abduction. When an abductive trigger presents itself to an agent, the failure to close his agenda is always a failure to close it in accordance with the requisite epistemic standards. If *T* is the agenda of explaining some event scientifically, then the knowledge that the abducer lacks is not all knowledge that might bear on *T*, but rather the *scientific* knowledge by which *T* could be explained. This same relativity is present in the case of hypothesis selection. In finding an *H* that would enable the abducer to produce the desired scientific explanation, the abducer must hypothesize that *H* is a serious candidate for scientific knowledge, never mind that it does not presently so qualify. It might be thought that the ignorance condition obliges the abducer to refrain from selecting *H* from anywhere in his *K*-set. This is a serious misconception. The prohibition extends only to the module that harbours (in the present case) his *scientific* knowledge, a prohibition that is trivially satisfied anyway by the very structure of an abductive trigger. The abducer is free to select a candidate from any other K-module, provided he is prepared to "bet" — apart from its requisite abductive fit — that, once discharged, it has a chance of performing in ways that would elevate it to membership in the scientific *K*-module. Subject to this key limitation, an abductive agent is free to search his own K-set for possible hypotheses. An even greater latitude is open to him with regard to his *belief*-sets and his *plausibility*-sets.

As we now see, not every abduction problem is an epistemic abduction problem. A *doxastic* abduction problem as the triggering of an agenda that can't be closed with what the abducer currently *believes* (and let us assume the modularity of belief in the same general kind of way, and to the same kind of end, as we have just done with knowledge). By the same token, it should also be possible to speak of *probability* and *plausibility* abduction problems, in which a target cannot be hit with what the abducer takes to be probable or, as the case may be, plausible. Here, too, given that probability and plausibility comes in degrees, the problem is that the target in question cannot be hit with anything he takes to be probable or plausible to a certain degree or higher. Lower probabilities and plausibilities are free to be mined with a view to their possible subsequent upgrading.

Proposition 9.1 (Multiple relativity) *Abduction problems are definable in relation to knowledge, belief, the probable and the plausible. In each case, the problem is relativized to the requisite module or modules.*

As earlier remarked, unless we indicate the contrary, our discussion of *K*-abduction will stand in for all these varieties.

The relativity of abduction in relation to the variability of its K-modules carries direct consequences for abduction's *conjectural* component. It portends a distinction between two sorts of conjecture within which the factor of variability recurs. The distinction is one between what we might call *cold-start* conjectures and *upgrade*-conjectures. Cold-start conjectures answer well to the Peircean element of abductive surprise (even more emphatically expressed by N.R. Hanson as astonishment). Cold-start conjecture is required not only when there is nothing in the reasoner's K-set that closes the requisite agenda T, but also when nothing in K-sets of lesser epistemic stripe or in belief sets, probability sets or plausibility sets appears to do the job either. *In extremis*, this requires the conjecturer to do some *originary* thinking, as Peirce calls it; that is, to think outside the box. In conditions of such austerity, prior belief must act at arm's length. There are two principal ways in which this happens. Not having any beliefs that strike him as adequate for the hitting of his abductive target, the conjecturer is free to take a novel step and reflect upon what may strike him as *possible* candidates. He may also review his prior beliefs in hopes of finding there occasion for *analogical extension*. Thus,

Proposition 9.2 (Possibility and similarity) *Two of the primary operations of cold-start conjecture are* modalization, *i.e., the recognition of possibilities, and* analogy, *i.e., the recognition of similarities* in difference.

Further,

Proposition 9.3 (Creativity and ignorance) *In solving abduction problems, the demands of originary (or creative) thinking are proportional to the depth and width of the abducer's ignorance.*

Corollary 9.3(a) *In particular, originary thinking is not intrinsic to abductive success.*

Upgrade-conjecture operates in a less arms-length way. It allows deployment even of prior beliefs the abducer holds with (up to) substantial conviction and on the basis of (up to) substantial evidence. What cannot be allowed is that these beliefs, evidenced in these ways, are such as to hit the epistemic standards of the K-set with respect to which the abducer's problem arose in the first place. It is this feature that motivates the upgrading character of conjecture. For if a candidate proposition H is already firmly believed on good evidence, there is no occasion to conjecture *that H*. Rather the conjecture can only be that H hits the epistemic standards of K, or higher. In contrast, in cold-start conjecture, there is room for the suitably modest conjecture *that H*, together with the conjecture that H *meets the abductive problem's requisite epistemic standard*.

A case in point: A crown prosecutor has come to believe that X is guilty of the crime in question. But he realises that X's confession is inadmissible for technical reasons. His belief that X is guilty fails to meet the standard of legal fact. Yet the prosecutor is free to abduce X's guilt as (part of) the best explanation of the evidence that *is* admissible. If the explanation is accepted by a jury, the proposition that X did it is upgraded to a legal fact.

10 Locating abduction on the logical map

¿From its inception, logic has served two masters, *enquiry* and *inference*. In a rough and ready way, enquiry deals with premiss-searches, and inference deals with premiss-projections. Throughout the history of logic, inference has been dominantly represented as the drawing of subsets of consequences from sets of priors. Enquiry has had a less firm grip on the evolution of logic. Aristotle makes fragmentary provision for it in *Topics* and *On Sophistical Refutations*, but in various subsequent periods enquiry (or what also could be called "discovery") was excluded from the province of logic. In the present day, discovery has not found a place in the metropolitan centre of the discipline, but it thrives in the prosperous suburbs of dialogue logic and interrogative logic; and perhaps fledgingly in the logic of abduction. From the very beginning logic has had a decidedly easier time of it with its consequentialist approach to inference. Should we expect the same of a logic of abduction?

These and other issues take on a measure of clarity when considered against the backdrop of a basic schema for abduction, a description of which we now turn.

11 Abductive schematics

Let $T!$ express that T is an agent's agenda. Let R again be the closure relation on T, R^{pres} the presumptive closure relation on T, H a hypothesis, $K(H)$, a knowledge-base revised by $H, C(H)$ a conjecture that H, and H^c a discharge of H. Then the schema for abduction begins to fall out.

1. $T!$ [declaration of T]
2. $\neg(R(K, T))$ [fact]
3. $\neg(R(K^*, T))$ [fact]
4. $R^{pres}(K(H), T)$ [fact]
5. H meets further conditions $S_1 \ldots, S_n$ [fact]
6. Therefore, $C(H)$ [conclusion]
7. Therefore, H^c [conclusion]

Remarks. $C(H)$ is read "It is justified (or reasonable) to conjecture that H". H^c, again, denotes the discharge of H. H is discharged when it is forwarded assertively and labelled in ways that reflect its conjectural origins. (Here the label is 'c' in superscript position).

We say that in the schema above the following are its *parameters*: $T!, T, K, K^*, H$, $K(H), S_1 \ldots, S_n, \therefore, C(H), H^c$.

12 Tasks for an abductive logic

Abduction is a thriving research programme in a number of disciplines, chiefly logic, philosophy of science, computer science, AI, belief dynamics, the other branches of cognitive science, including neurobiology and neurophysics, and legal reasoning. The present authors are strongly drawn to the thesis that a comprehensive logic will exploit developments from all these sources and that, in particular, a strong partnership between cognitive science and logic portends a good outcome for the theory of abduction.

If we were to take the present schema as our guide, a theory of abduction would attempt to account for all the schema's parameters. It is a fair question as to how any of these fall within the *logician's* ambit. If one takes the straight mainstream approach to logic (set theory, model theory, proof theory and recursion theory), there is not much that logic can do to elucidate these parameters. But if one takes a more laws-of-thought orientation in the manner, for example, of [Gabbay and Woods, 2001], there is substantially greater prospect for logical engagement. This is especially true for our operational test of what counts as a logician's work. By that test, a logician's proper work is what he's interested in in conjunction with what he's good at. Accordingly, the challenge posed by our abductive parameters to the logician is, in effect, "Give it your best shot and then we'll see". Here is a brief consideration of how that challenge might be met. But first we want to remind the reader of two important qualifications. One is that not even the most latitudinarian of logicians will assert exclusive domain over these issues. The other admonition that bears repeating is that working out just the logical aspects of the abductive parameters is a daunting job, made so both by its sheer size and the difficulty of some of its questions. For this reason, much of what we will have to say for ourselves here and in [Gabbay and Woods, 2005b] will be fragmentary, tentative, programmatic and promissory. As someone said at the Cognitive Science Meetings in Chicago, in August 2004, "This will give us gainful employment for at least a generation!"

$T!$ expresses the desire that some target be hit, i.e. that some agenda be closed. It is a form of expression that optative logicians have taken note of. *Optative logic* enjoyed a bit of a flurry in the 1950s, but seems not to have been an active research programme more recently. Nevertheless, the optative slack has been vigorously taken up by various kinds of *goal-directed* logics; and we may expect them to play a role in a final theory if ever it is produced. Abductive theories recognize that abduction arises from a disappointed hope, and that a successful abduction always manages to answer to some variation on that hope. In other words, T is an optative infinitive, with regard to which all known approaches to abduction attempt to specify realization conditions. This helps motivate one's choice of H and R^{pres}. both of which can be expected to meet the requisite optative-realization (or goal-directive) conditions.

K is the abducer's present knowledge-base. K^* is his knowledge base sometime later and within the frame of the problem he encountered with K. Like Peirce, we are *fallibilists* about K. K at a time is what is then taken for knowledge. Later we might come to know better, and often do. So the fallibilism at hand has a dynamic cast. Accordingly, the general setting for abductive processes can be expected to import the general structure of *dynamic logic*.

Our fallibilism complicates the structure of knowledge-succession. In particular, K^* need not contain K as a proper subset, never mind that K^* is always an adaptation K. Such problems are well-explored by epistemic logics and theories of belief-revision and belief-update. $K(H)$ also poses a belief dynamics problem.

H is a hypothesis. There exist logics of hypothetical reasoning that can be expected to play a role here. All abductive theorists recognize the necessity of subjecting H to various constraints $(S_1, \ldots, S_n)$. If, as in our approach, we do not require that $K(H)$ be consistent, then the underlying logic must be *paraconsistent* or more generally, a labelled deductive system. The base logic must also be *non-monotonic*. If, as we think, minimality is not a condition on H or on $K(H)$, some attention must nevertheless be paid to the requirement that H (and $K(H)$) contain no more *redundancy* that abets smoothness of communication. Aristotle was the first logician to impose an irredundancy condition on deductions. We see no reason to withhold the factor of redundancy from the attention of the modern logician. (See below, section 10.3.)

Other conditions are often imposed on H. One is that it be *relevant*. Another is that it be *plausible*. Logicians have produced a huge literature on relevance; and some of the basic groundwork has been done for *plausibility logics*. And if, as we believe, it is also necessary that a winning H be cognitively junior to K, this is something for the *epistemic logician* to take note of. For, again, how can it be rational to engage in a form of reasoning that is guaranteed to preserve one's original ignorance?

$\therefore$ is the abducer's conclusion operator. It faces the abductive logician with the task of specifying the inferential force of an abductive conclusion. If, as we think, its force must always be weak enough for H to honour the ignorance condition on abduction, then a logic of *plausible* (or *presumptive*, or *defeasible*) *reasoning* would have a natural place in any such specification.

$C(H)$ we have already spoken about. It is a modal sentence, with 'C' a deontic operator for permitted conjecturability. H^c denotes the release of H together with labelled recognition of its conjectural origins. So, again, $C(H)$ has a place in deontic logic, and H^c might respond well to the labelled approach to inference.

Perhaps what is most striking about abductive parameters is not their resistance to the logician's probes, but rather their collective call upon so hefty a logical pluralism. The great task of an abductive logic is to aggregate this pluralism in a systematic way.

12.1 An Alternative Schema

It will be useful to compare our own (the GW-schema) with this more standard orientation, with which it is natural to associate the names, among many others, of [Aliseda-LLera, 1997; Kowalski, 1979; Kuipers, 1999; Magnani, 2001; Kakas *et al.*, 1995]and [Meheus *et al.*, forthcoming], the AKM-schema, as we might say.

The AKM schema unfolds as follows:

1. E
2. $K(\not\looparrowright E)$
3. $H(\not\looparrowright E)$
4. $K(H)$ is consistent
5. $K(H)$ is minimal
6. $K(H) \looparrowright E$
7. Therefore, H.

Even a cursory examination of the current literature reveals a very strong inclination among researchers to take a *consequentialist* approach to abduction. On this view, it is always a condition on $K(H)$'s standing in the presumptive attainment relation to T, that there be a *consequence relation* $\rightsquigarrow$ and a *payoff proposition* E such that $K(H) \rightsquigarrow E$. As is apparent, *AKM* is structured to capture consequentialist abduction. Among philosophers of science, the standard way of approaching consequentialist abduction, is to interpret $\rightsquigarrow$ as a relation of *explanatory* consequence. Suppose, then, that there is some phenomenon E, of which the abducer wants an explanation. Notice that his target is not E; rather his target is an explanation of E. So the *AKM* model fails to specify an abduction target. This leaves the interpretation of $\rightsquigarrow$ unmotivated, in fact indeterminate. Of course, there is a reply to this. One might say that the target is to achieve an explanation of E, and accordingly that $\rightsquigarrow$ must be interpreted in ways appropriate to that target (say, as the deductive component of Deductive-Nomological-explanations) is a tacit assumption of the schema.[9] Perhaps so. But the question is whether those assumptions wouldn't better be made schematically explicit. If so, we could make a modest adjustment to the *AKM* model, as follows:

1. $T!$
2. E
3. $K \not\rightsquigarrow E$
4. $H \not\rightsquigarrow E$
5. $K(H)$ is consistent
6. $K(H)$ is minimal
7. $K(H) \rightsquigarrow E$
8. Therefore, H [10].

Of course, if we write it out in English, with T! as "I want an explanation of E", and $K(H) \rightsquigarrow E$ as "$K(H)$ explains E", and $\rightsquigarrow$ as the consequence relation of DN-explanations, it may strike us as immediate that, if true, $K(H) \rightsquigarrow E$ does indeed close T. But *schematically* this link is wholly suppressed.

In fact, however, $K(H) \rightsquigarrow E$ doesn't close T. For $K(H) \rightsquigarrow E$ to deliver the desired goods, it is necessary not only that $\rightsquigarrow$ bear the right interpretation, but also that K and H answer to the particular requirements of the deductive-nomological model. K must contain the requisite laws and H must be an initial condition. But this is a requirement that its hypothetical character precludes H from fulfilling. So, as we have it here, $K(H)$ cannot provide a DN-explanation. Repairs are readily to hand, of course. Instead of the categorical conditional

[9] A notable example of an non-DN approach to explanationist abduction is Thagard's treatment of explanatory coherence. See, e.g.,[Thagard, 1989]

[10] Consequentialism makes it necessary to point out that it is not the case that a payoff proposition is T itself. This is guaranteed by T's syntactic character. T is an infinitive construction not a statement. It expresses the agent's desire *to have* an explanation of E, *to unify* the laws of D, *to produce* a proof of Q, and so on. In the usual *AKM* notation, E is the payoff proposition — usually an explanandum. (T's presence is usually left tacit, as we have said). Accordingly, there is no requirement, and no possibility, in the *AKM*-set-up that $K(H)$ bears $\rightsquigarrow$ to T. Rather, the requirement is that $K(H)$ bears $\rightsquigarrow$ to E, where E is the payoff for T.

$$K(H) \looparrowright E$$

what's needed is the counterfactual conditional

$$K(H) > E.$$

But AKM also leaves this requirement unrepresented.

As is already apparent, a schema for abduction is open to two sorts of critical assessment. One examines whether it provides a suitably comprehensive number of parameters. The other examines whether those parameters have been adequately conceptualized. What we have been saying so far about the contrast between the GW-schema and the AKM-schema instantiates both kinds of critical approach. We are suggesting, both that the AKM-parameters are too few and that their interpretation is too narrow to afford a suitably comprehensive representation of the structure of abductive inference. Accordingly,

Proposition 12.1 (Under-representation) *The* AKM-*schema under-represents the logical structure of abduction.*

A related difficulty presents itself at (8) of the revised AKM-schema. H is detached without due regard to its intrinsically conjectural character. Against this it might be said that the "therefore" of line (8) is qualification enough, since it is obvious that it denotes a weak conclusional link, something along the lines of "it is plausible to conclude that H". But this is wrong. What (8) requires is something like "it is plausible to conclude that H is *a justified conjecture*". Again, it may very well be that this too is assumed, and left schematically implicit in the interest of clutter avoidance. Even so, how much of abduction to try to capture schematically is an important question. Omissions of abductively salient factors need to be justified.

While a dominant influence, the *AKM* model might appear *not* to be the sole model even among those who clearly are drawn to it. A case in point is what Aliseda calls *anomalous triggers* [**?**, p. 28]. Let it be the case that for some K and some true proposition $P, K \not\looparrowright P$ but $K \looparrowright \neg P$. Intuitively, this is a situation in which what one knows (up to now) is contradicted by some new fact. In its explanatory version K fails to explain P but succeeds in explaining its negation.

It is easy to see that in the first instance, what an anomaly triggers is not an abduction problem but rather a consistency-restoration problem (or, in its explanatory variation, an explanatory coherence problem) With all due recognition of what holism allows for *in principle*, this first task requires to cancel $K \looparrowright \neg P$. This in turn requires the restorer to make some deletion from K so that $K \looparrowright \neg P$ no longer holds. Since in the case before us P is not itself in K, the option of deleting P does not present itself.

Even so, this is not yet an abduction problem. To be an abduction problem, it would have to be the case that the would-be restorer has no knowledge of how to proceed. When this is so, the abductive option can be considered. In such a case, the agent's target now becomes the presumptive restoration of consistency (or explanatory coherence) by the conjecturing of an H such that $K(H) \looparrowright P$ and $K(H) \not\looparrowright \neg P$. But contrary to the appearance initially presented by the case of anomaly triggers, there is nothing in the present abduction that makes it unmodellable by the *AKM*-structure when $\looparrowright$ is read as "explanatory consequence".

12.2 The good that AKM does

At a certain level of abstraction, the *AKM*-schema does valuable work for a logic of abduction. It scores well on the following points. [11]

1. The AKM-schema acknowledges the cognitive-deficit character of abduction problems.
2. It recognizes the consequentialist character of abduction.
3. It highlights three subtasks for abductive logicians.
   a. They must give an account of $\rightsquigarrow$, where the abduction is consequentialist
   b. They must give an account of H
   c. They must give an account of the therefore-operator.

Before leaving this matter, let us attend to a slightly different example. Suppose, again, that the agent's target is to have a proof of P. Let it be that neither K nor H entails P, but that $K(H)$ does. If the abducer is satisfied with this, he is downgrading his solution in a quite crucial way. He started out questing for a proof of P, but he settles for a conditional of proof. In other words, he satisfices — a fact that is unrepresented in the AKM-model.[12]

The moral we draw from this brief discussion is that we won't get the logic of abduction right (or anyhow deeply or comprehensively right) unless we let it loose on structures that reflect all the essential peculiarities of abduction on the hoof. The GW-model is offered with this imperative in mind. It retains the programmatic virtues of the AKM-approach, but ventures beyond.

Accounts of abduction that flesh out structures such as the AKM-schema or the GW-schema are sometimes called *models of abduction*. Modelling a concept or set of concepts is a methodological commonplace for logicians. It is commonplace that counsels a considerable circumspection in attributing under-representation to a model. This is because all models, to some extent or other, are under-representations of their explicanda. How, then, can the criticism of the AKM-model embodied by our Proposition 10.1 be justified, given that the same Proposition is likewise true of the GW-model? Isn't the GW-model a standing invitation to a charge of *tu quoque* from supporters of the *AKM*-schema?

In any model of abduction (or anything else the logician turns this technique upon), some facts about real-life abduction will be suppressed or ignored. Others will be retained but also idealized. These suppressions and idealizations the modeler typically justifies sometimes on grounds of comparative unimportance or low salience, and sometimes on grounds that doing so enables the model to demonstrate interconnections or systematicities that enrich the model's clarity or explicational heft.

[11] For a more detailed discussion of the *AKM* model, see [Gabbay and Woods, 2005a].

[12] We note in passing that if an abducer is sufficiently at lease with this presumptive proof of P to detach H conjecturally, he may subsequently take "the next logical step". He may declare it an axiom! Since axioms are (save for auto-demonstration) insusceptible of proof in any system they govern, to choose an axiom is to stipulate it. But what is stipulation but conjecture with a certain swagger?

Still, it is plain that some models do better than others on the score of clarity and explicational success, and that sometimes this betterness pivots on comparative numbers of parameters and comparative scope of interpretation. So the criticism expressed by Proposition 10.1 needs to be re-phrased.

Proposition 12.2 (Under-representation again) *Considered in relation to a comprehensive logic of abduction, the* AKM-*model is under-representative to the point of significant and avoidable distortion.*

If this is right, it is advisable to see the *AKM*-model as a first (albeit important) step in exposing the logical strucutre of abduction. Undoubtedly, the *GW* alternative to *AKM* is messier to work with. But since this is messiness intrinsic to abduction, our further advice is to gird the loins and try to make it work. Even so, it remains an important virtue of the *AKM*-approach that, for those cases in which abduction takes on a consequentialist structure, *AKM* is well-suited to take note of its embedded consequence relation.

12.3 The reach of abduction

We can say that a logic of abduction will have at least two sublogics. When abduction *is* consequentialist, one gives an account of the requisite consequence relation that the abductive schema reflects (subtask (a)). The other gives an account of its conclusional operator "therefore" (subtask (c)). A contentious question in relation to subtask (b) is whether a sublogic for H exists, and, if so, how it would go. The H-factor presents the abduction theorist with at least two questions.

1. What are the conditions under which hypotheses are thought up?
2. What are the conditions under which hypotheses are deployed?

It is easy to see that part of the answer to (2) is that deployed H's should honour the abductive schema. In the *AKM*-model, H is required not to bear $\looparrowright$ to E, and not to be inconsistent with K. It is also required that $K(H)$ be minimal. In the GW model, the conditions on H are less specific. The reason for this is that we are unsure about the *AKM*-constraints. Let us take these in order.

a. $H \not\looparrowright E$ (*H's deductive independence*): The AKM-model allows for $\looparrowright$ to be a deductive consequence. There are lots of cases in which a solo P bears $\looparrowright$ to a payoff V. Why rule it out that such a P might be a candidate for H? The answer appears to be that allowing it would preclude this fact from constituting a deductive-nomological-explanation of V. So it would. But not all explanationist abduction is deductive-nomological, and not all abduction is explanationist. So we find the constraint to be over-narrow.

   Another reason for the deductive independence of H is to discourage trivial abductions in which, for example, V is abduced as an explanation of itself. But, again, independence over-determines the objective. Its more realistic accommodation is by way of abductively motivated constraints on the $\not\mapsto$ relation itself in light of the cognitive aim represented by T.

b. *H's consistency with K*[13]: There are cases in which the abducer is required to reason from data-bases that contain unresolved inconsistencies. Juries, for example, must determine the guilt or innocence of accused persons from evidence-bases that are routinely inconsistent. Verdicts are based on the acceptance or rejection of what lawyers call "theories of the evidence" or "theories of the case". A theory of the evidence is an abduction that generates a verdict on the strength of what best explains the evidence, inconsistency and all.[14] Here, too, we find the constraint excessive.[15]

c. *K(H)'s minimality*: An ambiguity lurks. Does the condition require that H be the least modification of K that delivers the intended goods? Or, does it require that H modify the least class of K that delivers the goods? Or does it mean both? What we have here, in all three cases, is a contingency elevated to the status of a logically necessary condition. It is true that abduction problems don't require for their solution everything whatever the agent may know at the time. It is also true that winning hypotheses aren't wantonly redundant. In actual practice, abductive reasoning is from subsets of K augmented by not overly redundant hypotheses. This is a fact for our schematic models to take note of. But minimization achieves this end over-aggressively.

Minimality is also a way of averting the useless proliferation of abductions by deductive closure on winning Hs. So, if H is a winning hypothesis, we don't want it to be the case in general that $H \vee Q$, for arbitrary Q, is also a winning hypothesis. But, here too, the more natural mode of discouragement is not the banishment of all redundancy, but rather constraints on the consequence relation in light of the cognitive content of T.

12.4 Simplicity

Minimality is sometimes thought to recommend itself on grounds of simplicity. There is no doubt that simplicity has its attractions. But we join with those who find, in Kuipers' words, that "simplicity may well retard empirical progress" [Kuipers, 1999; McAllister, 1996, p. 322] and [Rescher, 1996].[16]. Presumably Kuipers is drawn to

[13] See here [Boutilier and Becher, 1995].

[14] This may appear to generate a very bad problem for criminal jurisprudence. If the standard in criminal trials is *proof* beyond a reasonable doubt, how can it be envisaged than an abductive *conjecture*, however confidently made, could rise to it? This is discussed in greater detail in[Gabbay and Woods, 2005b, chapter 9] and [Gabbay and Woods, 2006].

[15] We return to the inconsistency two sections hence.

[16] Rescher: "Simplicity — is it not comitted to the idea that nature proceeds in fundamentally simple ways? By no means! We have no ground whatever for supposing the "simplicity" of nature. The so-called Principle of Simplicity is really a principle of *complexity management*"[Rescher, 1996, p. 26], emphasis added.

minimality for reasons other than simplicity. We suppose that it is an entirely natural desire to purge abductive inference from excessive redundancy.

A further reason to distrust a simplicity requirement for inference is that the most simple assumption is not always self-announcing. Here is an example drawn from [Goddu, 2002, p. 15]. Consider the argument

1. All monkeys are primates.
2. So, with certainty, all monkeys are mammals.

It may strike us, as it does Goddu, that the simplest implicit premiss that will make this argument valid is

(1^*) All primates are mammals.

If simplicity here is weakness, the present claim is false. As Hitchcock [2002, p. 158] rightly points out, the weakest missing premiss is, in fact,

(1^{**}) Either not all monkeys are primates or all monkeys are mammals.

That (1^{**}) is simpler than (1*) is shown by the fact that whereas (1^*) entails (1^{**}), it is not the case that (1^{**}) entails (1^*). Yet no one in his right mind would require that an actual reasoner conform his reasoning to the requirement that the assumption of (1^*) be rejected in favour of the assumption of (1^{**}).

13 The Cut Down Problem

Perhaps the greatest problem posed by the thinking up of hypotheses is that, on any given occasion, a candidate for selection occupies an up to arbitrarily large space of possibilities or (candidate space). Whatever the details, it appears that abductive agents manage to solve what might be called a cut down problem. One of the attractions of Atocha Aliseda's semantic tableau approach [1997] is that it reveals the structure of cutdown for certain ranges of cases. But these are rather narrow ranges, as we shall soon see, made so by the technical constraints that semantic tableaux impose. In a more general sense, it would appear that the hypotheses that an abducer actually entertains are relevant and plausible subsets of large candidate spaces. (We note in passing that the idea that the minimality condition seeks to honour is handled here non-quantitatively by relevance and plausibility filters.) It is doubtful that the full story of the dynamics of cut down can be told in any logic, no matter how capacious; but part of it, certainly, requires the logician's touch. Accordingly,

Proposition 13.1 (Relevance and plausibility) *In giving an account of H(subtask (b)), an abductive logician should deploy the resources of the appropriate logics of relevance and plausibility.*[17]

[17] For relevance, see [Gabbay and Woods, 2003]; for plausibility, see [Rescher, 1976] and [Gabbay and Woods, 2005b, Chapter 8].

It would seem that plausibility also bears in a central way on the question of hypothesis *selection*. It is implicated in a further step of the cut down process, at least for large classes of cases. It cuts down the set of *entertained* hypotheses to subsets (ideally a unit set) of the most plausible.

Abductive reasoning is shot through with considerations of plausibility and presumption. In the GW-model it is explicit that presumption plays a role. It plays it in two connected ways. If we have a successful H, then $K(H)$ will close the abducer's agenda presumptively. Correspondingly, it may plausibly be inferred that the conjecture of H is justified; that is to say, that the presumption of H is reasonable. Most of the work to date on the logic of presumption has been done by default logicians in the computer science and AI communities. As we have them now, such logics haven't adapted well to the particular requirements of abduction. There is work still to be done.

Proposition 13.2 (Presumption) *The logic of the conclusional operator "therefore" (subtask (c)) should subsume an appropriate logic of presumption.*

14 Inconsistency again

Earlier we considered and rejected the standard position concerning the consistency of K. As it happens, there are three questions about abductive consistency that need to be settled.

1. Should we require that K be consistent?
2. Should we require that $K(H)$ be consistent?
3. Should we require that H be (self-)consistent?

The position taken in that section answers question (1) in the negative but addresses neither of the remaining questions. We shall attempt to repair those omissions now.

Question (3) asks, in effect, whether we should maintain a strong anti-dialetheic stance toward winning abductive hypotheses. Dialetheism asserts that some but not all contradictions are both true and false. While we have no particular stake in dialetheism's being true, are not prepared to dismiss it out of hand. Consider a case. It is all but universally held that the proof of the Liar is a paradox. A paradox is a proof that appears to be sound, but, owing to the transparent falsity of the conclusion, cannot be taken as sound. The problem created by such proofs is that, although they demonstrate the falsity of something in or presupposed by the premiss-set, it is left wholly undetermined as to where in particular to pin the blame. Dialetheism is a different reaction to such puzzles. What best explains the appearance that the premisses are true is that they are true. What best explains the appearance that the proof is valid is that it is valid. What best explains the appearance that the conclusion is false is that it is false. What best explains why these explanations themselves are pairwise consistent is that the conclusion is also true.

While we have ourselves no stake in endorsing this solution, we fail to see how we would advance the analysis of abduction by ruling it out. In as much as the present solution is itself an abductive solution; it is perhaps appropriate that we leave official room for winning Hs to be self-inconsistent. Accordingly,

Proposition 14.1 (Self-inconsistency) *Not only may a winning H be propositionally implausible, it may also be self-inconsistent.*[18]

Question (2) asks whether we should make it a condition on H that it be consistent with K. Since we already have it that K itself might be inconsistent, the issue raised by (2), is whether to tolerate additional inconsistencies occasioned by the conversion of K to $K(H)$. If we were treating inconsistency classically, there would be trouble with the idea of "additional" inconsistencies. But since the toleration of K's negation-inconsistency requires that K be lodged in a paraconsistent or in a labelled logic, $K(H)$ can have more inconsistencies than K, since the negation-inconsistency of neither lands it in absolute inconsistency. So let us consider question (2).

Let H be a winning abductive hypothesis, and let the negation of H be a member of K. Should we prohibit this? The answer is that we should not. We should allow for the case in which, although to the best of our knowledge that P, we nevertheless conjecture that not-P. This we may do provided that any K at time t is what is taken to be known by an agent (or by the community of knowers) at t. This allows that someone might know something without realizing it. It is a natural precursor to fallibilism, the view that anything (or most things) we take for knowledge, we might be mistaken about. By these lights, conjecturing that not-P is compatible with our taking P as known. This is not to say that we have *carte blanche* in such matters. In abducing H, we decide to give it a premissory role in future reasons within or from $K(H)$. Correspondingly, we diminish $\neg H$'s premissory range, notwithstanding that $\neg H \in K$. Even though we would make paraconsistent provision for their joint use as premisses, there are lots of cases in which their joint use would contaminate the reasoning in question. So $K(H)$ should be attended by the appropriate quarantines.[19] Again, think of the historical origins of Anderson-Belnap relevance. In their attempt in 1959 to give a simple treatment of the truth functions, it happened that Disjunctive Syllogism (DS) is not a valid rule in their treatment. This was a natural occasion to conjecture the actual invalidity of DS while acknowledging that the validity of DS was currently in K. The unseemly haste with which Anderson and Belnap abandoned DS is an interesting demonstration of how quickly a conjecture can be promoted to full membership in a successor to K. But the point remains that it is perfectly tenable to conjecture the opposite of what one takes for knowledge.

What this shows is the importance of the belief/acceptance distinction to abductive logic. While a conjectured hypothesis might be something the agent believes, conjecturing it does not confer that status upon it. Conjecture is acceptance. It is acceptance for premissory work in future inferences, subject to the possibility of recall. Another way of saying this is that conjecture does not report a doxastic state. Rather it expresses a decision. From this we have it that

(1) P, but we conjecture that not-P

does not have the speech-act structure of *blindspot utterances* of the form

[18] Readers who find the latitudianarianism of this proposition excessive may wish to consult [Armour-Garb, 2004; Woods, 2004b] and [Woods, 2005, to appear].

[19] See here [Batens, 1980; Arruda, 1989; DaCosta and Bueno, 1996], [Jennings and Schotch, 1981; Priest, 2002; Detlefsen, 1986], [Jáskowski, 1948], [Gabbay and Hunter, 1993; Tanaka, 2003].

(2) P, but I believe that not-P.

The trouble with (2) is that, in the absence of additional information, it is impossible to determine what the utterer's position to P actually is. There is no such problem with (1). The utterer's position is that P is the case and nevertheless that there are adequate reasons to release not-P for properly regulated premissory work in future inferences.

In sum,

Proposition 14.2 (Paraconsistency and dialetheism) *K can be negation inconsistent but not absolutely inconsistent. The same holds of $K(H)$. This occasions the necessity to lodge abductive reasoning in a paraconsistent base logic. H itself can also be inconsistent. This makes it desirable that the underlying logic also have a dialetheic capacity.*

14.1 Abduction as Practical

In *The Reach of Abduction* [Gabbay and Woods, 2005b], we suggest that a theory of abduction would prosper if embedded in a suitably general practical logic of cognitive systems. In our conception of it, practical reasoning is reasoning transacted under press of (comparatively) scant resources (such as information, time and computational capacity.) We further propose that such constraints typically apply to an individual's reasoning, and that greater latitude is typical of institutional reasoning. So conceived of, this would seem to suggest that abduction is intrinsically, or at least dominantly, a form of reasoning attended by comparatively scant resources in quest of comparatively modest targets. We imagine that not every reader will see things in quite this way. Perhaps they would be minded to ask whether we are prepared to declare abduction off-limits for institutional reasoners. Our reply is that, like abduction itself, the practicality of abduction is a matter of degree. The ignorance-condition decrees that no matter how lofty an abducer's goal, it must be a lesser thing epistemically than that cognitive level of his knowledge-base then and there. Thus it is intrinsic to abduction that abductive reasoning is more practical than the reasoning that extends from (and in this case fails) his knowledge-set then and there. It also bears on this matter that abduction is a substitute for exploration. This is a welcome economy in as much as conjecture is often cheaper than the acquisition of relevant new knowledge. It is true that computers enable us to make exhaustive searches of enormous possibility-spaces. It is also true that some programs achieve very drastic cut downs of such spaces very quickly. To the extent that abduction involves the picking out of an H from up to arbitrarily large candidate spaces, perhaps we should say that some abduction problems are well-matched to the resources that typify institutional agency. It won't work. Abduction problems, no matter whom they are solved by, involve the timely selection of an H from sometimes huge spaces. Whether Harry performs the abduction or HAL does, H is selected in a timely way. The difference here is that it is not intrinsic to Harry's abductive success that he exhaustively search through the huge possibility spaces of which his winning H is a member. If HAL makes such a search, it is doing something not typical of abduction. So abduction does indeed retain the practical cast that we have claimed for it.

Practicality, we say, hovers at the triangulation of comparatively modest cognitive agendas and comparatively scant resources for its closure. So conceived of, practical reasoning has no non-negative natural antonym. In much of the philosophical literature, the received contrast of the practical is the theoretical. This is true of our account as well, but with a nonstandard meaning of "theoretical". On our reading, theoretical rationality aims comparatively high, and proceeds with comparatively abundant cognitive resources. This is the way in which, although it clearly has its own troubles, a theoretical agent such as NASA out-guns practical agents such as you and we. There is, however, a point at which theoretical reasoning in our sense and theoretical reasoning in the more standard philosophical sense fall into a loose confederacy. In so doing, we also capture a quite intuitive idea about practicality. This is the idea according to which the practical thing to do is that which is *timely* and *adequate*, rather than *best*. Practicality therefore deeply embeds the economist's notion of satisficing. If we now consider that in our theoretical pursuits we are routinely ready to defer to more optimal outcomes even at the cost of time spent, we see that a theorist's undertaking is, comparatively speaking, less a satisficing one than a maximizing one. Taking abduction as our example, we can put the essence of the contrast as follows. Whereas certain cognitive pursuits aim for the knowledge and will settle for nothing less, abduction aims for conjecture and can settle for nothing more. Seen this way, abduction is intrinsically a practical endeavour.

15 The Schema's Flexibility

If explanationism is the preferred version of consequentialism among philosophers of science, computer scientists and formal logicians have a marked fondness for what we might call *proof-theoretic* consequentialism. In its simplest terms, a proof-theoretic trigger is pulled when a given *wff* (or unit) Q cannot be proved from a given database K. The task this presents is that of abducing and H such that $H(K)$ does now prove Q. Proof is understood here rather generically and abstractly. Here, too, the *AKM*-model is structured to take note of these consequentialist factors. In particular $\looparrowright$ is specified as the requisite proof relation.

The schemata that we have sketched is normally associated with what are (misleadingly) called inferentialist approaches to abduction. This would be a good point at which to take brief note of different approaches that differ from (narrow) inferentialism. One such is prominent in AI, beginning with Pople's influential paper of 1973 [1973], and developed in a number of subsequent works by Pople and other investigators, especially those working in logic programming [Kakas *et al.*, 1995], knowledge assimilation [Kakas and Mancarella, 1990] and diagnostics and other forms of medical reasoning [**?**], [Peng and Reggia, 1990; Josephson and Josephson, 1994] and [Ramoni *et al.*, 1992]. Abduction is also dealt with in Bayesian networks and connectionist logics [Josephson and Josephson, 1994; Konolege, 1996; Paul, 1993; Flach and Kakas, 2000].

reference to Poole et al, 1987 missing.

Two further contexts for abduction should also be mentioned. One is linguistics [Hobbs *et al.*, 1990; Chomsky, 1972; Heim, 1983; Gervás, 1995]. The other is mathematics [Polya, 1945; Polya, 1954; Polya, 1962; Russell, 1907][Gödel, 1944;

Gödel, 1990a; Gödel, 1990b]. We take up the issue of interpretative abduction in chapter 9 of [Gabbay and Woods, 2005b]. Aspects of mathematical abduction will occupy us in chapter 5 of the same work.

Inferentialist approaches have a dominantly semantic orientation, concentrating on the specification of truth conditions on the implementation of the abductive inference schema. AI developments emphasize the role of algorithms in constructing abductions. In chapter 6 of [Gabbay and Woods, 2005b] we examine diagnostic-abduction, [Paul, 1993, pp. 109–152], and in that same chapter, as well as chapter 9, we review connectionist abduction. Here we shall give a quick sketch of the basic structure of the logic programming approach and of the rudiments of how abduction fares in theories of knowledge assimilation.

Logic programming arises from pioneering work by Robert Kowalski and Alan Colmeraner in 1974, [Kowalski, 1979; Lloyd, 1987]. One of its principal implementations is Prolog. Its underlying logic is first order. Prolog comprises a program P, queries q and a problem-solving device R, called resolution. We consider an elementary example.

Program P

lawn-wet← rain.

lawn-wet←sprinklers-on

(These are Horn-clauses in which the contained terms are literals).

Query q:lawn-wet

We say that a query *succeeds* when it is derivable from the program. In the present example q does not succeed. At this juncture, Prolog moves into an abductive mode. As can easily be seen, q would succeed if one or other of certain possibilities were added to P as hypothesis. These possibilities are: *rain*, *sprinklers-on* and *lawn-wet*. Abduction is here understood on the process evincing these possibilities. In its non-abductive mode, the failure of these to be listed in P as facts would trigger failure. In its abductive mode, the resolution mechanisms introduces these "non-facts" as hypotheses. The query now succeeds.

It is important to emphasize that the resolution device is constrained in what it can select as hypotheses. It is required to select only from sub-goal literals that fail under the backtracking operation. Thus not everything that would be reasonable to forward as a conjecture is allowed in this approach. A further limitation of the logic programming orientation — one that it shares with standard systems of diagnostics — is that hypothesis-selection must be made from a pre-determined set of abducibles [Kakas *et al.*, 1995], which in turn are required to be validated by further conditions, called integrity constraints, introduced so as to mitigate the problem of computational intractability.

Many (narrowly) inferentialist and most computer-based approaches to consequentialist abduction pivot on the fact that for some background K and a payoff V, neither $K \looparrowright V$ nor $K \looparrowright \neg V$ holds. This is a substantial constraint, excluding from consideration abductions arising from new facts that contradict what would otherwise have been expected from K. Knowledge-assimilation approaches are organized to take into account these excluded cases. They are theories of belief revision prompted by such phenomena. Typical settings for this kind of abductive trigger are diagnostics [Peng and Reggia, 1990], belief-revision in databases [Aravindan and Dung, 1994] and theory tweaking in machine learning [Ginsberg, 1988]. Ensuing from Gärdenfors theory of belief change [Gärdenfors, 1988], the principal means of incorporating new information into a database, a scientific theory or domain of common sense beliefs are the operations of *expansion*, *contraction* and *revision*, none of which is intrinsically abductive, contrary to what is rather widely supposed. However, certain of these operations is adaptable to the conjectural requirements of abduction. Expansion is simply a matter of adding a new fact P to K. Doing so enlarges K to $K(P) = K \cup \{P\}$. But this is not abduction. P is not conjectural, and $K(P) \nrightarrow P$, where $K(P)$ is itself a knowledge-set at epistemic par with the original K. Contraction is different. It requires the deletion of a subset of K, with which the new fact P is inconsistent. Except in those rare cases in which the structures of K and P permit unique determination of the candidate for deletion, there is room in reaching these judgements for conjecture. The same is true of the revision operation, which is a composite of expansion and contraction. In standard systems of knowledge assimilation, additional constraints are imposed. Two of the most prominent involve closure under belief-change operations and consistency of outcome. As these tend to be classical constraints, they alienate such systems from the give-and-take of real-life belief change in real time. The approach that we favour melds various of the features of inferential, computational and knowledge assimilation approaches. But in a general sense, all the standard approaches to abduction are inferential.

We have remarked abduction's connection to what Reichenbach called the 'context of discovery', which he contrasted with the 'context of justification' [1938]. Abductions in this sense are said to be the business of the logic of discovery. Reichenbach was not alone in thinking it possible to have a logic of scientific justification, which, as he supposed, is precisely what an inductive logic is designed to be. But a logic of discovery, or an abductive logic, Reichenbach regarded as a mistake in principle, because it confuses psychological considerations with logical considerations. Reichenbach's skepticism was shared by most logical positivists, but as early as 1958, Hanson [1958] worked up a contrast between reasons for accepting a given hypothesis and reasons that suggest that the correct hypothesis will be one of a particular kind or description. [20] A theory which analyzes reasons of this second kind Hanson called a logic of retroductive reasoning. In Hanson's account, a logic of retroduction resembles a logic of analogical reasoning. Hanson's efforts were criticized — even pilloried. This had as much to do with their novelty as with their deficiencies. Even so, the very idea of a logic of discovery trails some important questions in its wash. One deals with the extent to which the supposed contrast between contexts of justification and contexts of discovery catches a hard and fast distinction. Another question is whether the contrast between a psychological account of hypothesis formation and a logical theory of the same thing will hold up. A related issue is the extent to which a heuristics for a set of reasoning problems can be distinguished in a principled way from a logic for such

[20] More recent discussions of the logic of discovery include [Laudan, 1980; Nickles, 1980; Musgrave, 1989; Kelly, 1989; Savary, 1995; Kuipers, 2000].

problems. A further matter — also closely connected with the others — is whether a theory of hypothesis formation is able to retain a sharp distinction between descriptive adequacy and normative soundness.

16 Empirical Progress

Deduction is truth-preserving and induction is probability-enhancing, and abduction is ignorance-preserving. In what remains of this paper, we propose to test this latter view by examining the semantic tableaux approach to putative abductions in the context of determinations of empirical progress [Kuipers, 1999].

Suppose that an agent's agenda T is to produce a revision $K(H)$ of a given theory K under the condition that $K(H)$ represents a degree of empirical progress over K. Let us also take it that a necessary condition on the closure of T is that $K(H)$ exceeds the explanatory power (or reach) of K and that in so doing it is no worse off than K on the score of observational anomalies. In other words, the agent's goal is to find a H such that the explanatory improvements afforded by $K(H)$ are not offset by observational degradation. Given these facts, we may assume that our agent infers H on the grounds that it engenders the requisite $K(H)$.

The agent's reasoning is both backwards-chaining and explanationist. But it is not successfully abductive. To see why, it is necessary to observe that our agent is performing not one but two cognitive tasks and that they are linked. In the first instance, he infers H. In the second, he infers that $K(H)$ is a better theory than K, and does so only because he has the required confidence in his inference to H. If his inference of H is abductive, then H is cognitively junior to K. Accordingly, so too is $K(H)$. But if $K(H)$ is cognitively junior to K, how can $K(H)$ be better science than K? Accordingly,

Proposition 16.1 (Kuipers' Dilemma) *Either the backward chaining reflected Kuipers' revision of K to $K(H)$ is not abductive or it is not successful.*

We see, then, that

Proposition 16.2 (Backwards Chaining) *Backwards chaining reasoning is not intrinsically abductive.*

17 Semantic Tableaux

Semantic tableaux constitute a refutation method for designated formal languages [Hintikka, 1955; Beth, 1969]. An attractive exposition is that of Smullyan [1968]. To the best of our knowledge, [Mayer and Pirri, 1993] is the first to etend the methodology of semantic tableaux to abductive contexts. Another adaptation is that of [Aliseda-LLera, 1997]. Given that one of Aliseda's purposes is an elucidation of the structure of empirical progress, we shall concentrate our remarks upon her version (see also [Aliseda, forthcoming]. Aliseda writes

To test if a formula Φ follows from a set of premisses Θ, a *tableau* for the sentence $\Theta \cup \{\neg\Phi\}$ is constructed. 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, the initial set is unsatisfiable and the entailment $\Theta \models \Phi$ holds. Otherwise, if the resulting tableau has open branches, the formula Φ is not a valid consequence of Θ. A tableau closes if every branch contains an atomic formula β and its negation [1997, p. 83].

According to the tableau rules, double negations are suppressed, conjunctions add both conjuncts; negated conjunctions branch to two negated conjuncts; disjunctions branch into two disjuncts; negated disjunctions add both negated disjuncts; and conditionals reduce via negation and disjunction. Accordingly,

Negation

$\neg\neg\Phi \to \Phi$

Conjunction

$\Phi \wedge \Psi \to \frac{\Phi}{\Psi}$

$\neg(\Phi \wedge \Psi) \to \neg\Phi \mid \neg\Psi$

Disjunction

$\Phi \vee \Psi \to \Phi \mid \Psi$

$\neg(\Phi \vee \Psi) \to \frac{\neg\Phi}{\neg\Psi}$

Conditional

$\Phi \to \Psi \to \neg\Phi \mid \Psi$

$\neg(\Phi \to \Psi) \to \frac{\Phi}{\neg\Psi}$

In languages in which all wffs have either disjunctive or conjunctive normal forms, the above rules reduce to one or other of two types of rule, a conjunctive (α-type) rule, and a disjunctive (β-type rule).

Rule 1. $\alpha \to \frac{\alpha_1}{\alpha_2}$

Rule 2. $\beta \to \beta_1 \mid \beta_2$

Suppose that $T(\Theta)$ is a tableau for a theory Θ. Then it is known that

a. If $T(\Theta)$ has open branches, Θ is consistent. An open branch represents a verifying model for Θ.

b. If $T(\Theta)$ has only closed branches, Θ is inconsistent.

c. Semantic tableau constitute a sound a complete system for truth functional languages.

d. A branch of tableau is closed if it contains some formula and its negation.

e. A branch is open if it is not closed.

f. A branch B of a tableau is closed if (recall rules 1 and 2 just above) for every α occurring in B both α_1 and α_2 occur in B, and for every β occurring in B, at least one of β_1, β_2 occurs in B.

g. A tableau is completed if every branch is either closed or complete.

h. A proof of a wff Φ is a closed tableau for $\ulcorner\neg\Phi\urcorner$.

i. A proof of $\Theta \vDash \Phi$ is a closed tableau for $\Theta \cup \{\neg\Phi\}$.

Aliseda construes abductions as a kind of extended tableau. An extended tableau is the result of adding new formulas to a tableau. Consider the case in which for some theory Θ, Φ is not a valid consequence. In the tableau for $\Theta \vDash \{\Phi\}$ there are open branches. Each such branch is a counterexample to the claim that Φ is a consequence of Θ. Of course, if further wffs were added to the open branches it is possible that they might now close. Consider a least class of wffs that fulfill this function. Then Φ would be derivable from a minimal extension of Θ. Finding such wffs is a kind of abduction.

Accordingly, Aliseda proposes [1997, p. 91] the following conditions.

Plain Abduction $T((\Theta) \cup \{\neg\Phi\} \cup \{\alpha\})$ is closed. (Hence $\Theta, \alpha \vDash \Phi$.)
Consistent Abduction Plain abduction $+T(\Theta \cup \{\alpha\})$ is open. (Hence $\Theta \nvDash \Phi$.)
Explanatory Abduction Plain abduction $+T(\Theta \cup \{\neg\Phi\})$ is open (hence $\Theta \nvDash \Phi$) and $T(\{\alpha\} \cup \{\neg\Phi\})$ is open (hence $\alpha \nvDash \Phi$).

Further restrictions are required. Added wffs must be in the vocabulary of Θ. Each such wff must be either a literal, a non-repeating conjunction of literals or a non-repeating disjunction of literals.

Given a Θ and a Φ, *plain* abductive explanations are those wffs that close the open branches Γ of $T(\Theta \cup \{\neg\Phi\})$. *Consistent* abductive explanations are subsets of wffs which close some but not all open branches Γ of $T(\Theta)$.

A *total closure* of a tableau is the set of all literals that close each open branch Γ. A *partial closure* of a tableau is the set of those literals that close some but not all open branches Γ. (Such closures are definable for both branches and tableau.)

The *negation* of a literal is its ordinary negation of atomic or the embedded atom if negative. Thus $\neg \pm \Phi = \neg\Phi$ or Φ.

We consider now an algorithm for computing plain abductions [Aliseda-LLera, 1997, pp. 102–103].

Input

A set of wffs representing theory Θ. A literal Φ representing the fact to be explained.
Preconditions: $\Theta \nvDash \{\Phi\}, \Theta \nvDash \{\neg\Phi\}$.

Output

Generates the set of abductive explanations $\alpha_1, \ldots, \alpha_n$ such that $T((\Theta \cup \{\neg\Phi\}) \cup \{\alpha_i\})$ is closed and the α_i satisfy the previously mentioned lexical and syntactic restrictions.

Procedures

Calculate $\Theta + \neg\Phi = \{\Gamma_1, \ldots, \Gamma_n\}$. Select those Γ_i that are open branches

Atomic plain explanations

1. Compute $TTC(\Gamma_1, \ldots, \Gamma_n) = \{\gamma_1, \ldots, \gamma_m\}$ = the total tableau closure = the set of literals which close all branches concurrently.
2. $\{\gamma_1, \ldots, \gamma_m\}$ is the set of plain abductions.

Conjunctive plain explanations

1. For each open branch Γ_i construct its partial closure $BPC(\Gamma_i)$ = the set of literals that close that branch but do not close any of the other open branches.
2. Determine whether all Γ_i have a partial closure. Otherwise there is no conjunctive solution (hence go to END).
3. Each $BPC(\Gamma_i)$ contains those literals that partially close the tableau. To construct a conjunctive explanation take a single literal of each $BPC(\Gamma_i)$ and form their conjunction.
4. For each conjunctive solution β delete repetitions. The solutions in conjunctive form is $\beta_1, \ldots, \beta_h$.
5. END.

Disjunctive plain explanations

1. combine atomic with atomic explanations, conjunctive with conjunctive explanations, conjunctive with atomic, and each atomic and conjunctive with Φ.
2. Example. Construct pairs from the set of atomic explanations, and form their disjunctions $(\gamma_i \vee \gamma_j)$.
3. Example. For each atomic explanation, form a disjunction with Φ (viz. $(\gamma_i \vee \Phi)$).
4. The result of all such combinations is the set of explanations in disjunctive form.
5. END.

In the interest of space, we omit the specification of algorithms for constructing consistent abductive explanation. The interested reader should consult [Aliseda-LLera, 1997, pp. 103–106].

There are attractions to Aliseda's approach to abduction, not least of which is its employment of the well-understood machinery of semantic tableaux. Another is that it extends in a natural way to accommodate (most of) the structural features of Theo Kuipers' theory of empirical progress [Aliseda, forthcoming, ch. 6] and [Kuipers, 2000, p. 112]. Let us say that at a given time t a theory Θ_2 is at least as successful as a theory Θ_1 if and only if the set of failures in Θ_1; that the set of successes in Θ_1 is a subset of the set of successes in Θ_2; and at least one of these subsets is proper. Intuitively a theory Θ has a success with respect so some data E, and some background conditions K, if $K \cup \{\Theta\} \vDash E$; and that it has a failure with respect to these same parameters if $K \cup \{\Theta\} \vDash \neg E$. We may read F as confirmation. For ease of reference, we follow Kuipers in characterizing these cases as a success (failure) by E of Θ relative to C. Clearly we also have it that E may sometimes be such that it is neither a success nor a failure of Θ relative to C, but rather is a *lacuna* of H with respect to C.

Progress in science is a matter of moving from given theories to better ones. Roughly speaking, a theory Θ_2 is better than a theory Θ_1, if Θ_2 it has more successes than Θ_1 One standard way of achieving this kind of empirical progress is by revising Θ_1 in ways that produce Θ_2. Kuipers [1999] sets out the following

> *Instrumentalist abduction task*: Search for a revision Θ_2 of a theory Θ_1 such that H_2 is more successful than Θ_1 having regard to the available data.

Aliseda approaches this task by showing that the concepts of lacuna, success and failure admit of precise characterization in the language of semantic tableaux. Where $\mathcal{T}(\Theta)$ is a tableau for Θ, then if $\mathcal{T}(\Theta) + \{\neg E\}$ and $\mathcal{T}(\Theta) + \{E\}$ are both open extensions of $\mathcal{T}g(\Theta)$, then E is a lacuna of Θ. Similarly, if E is a success of Θ, then $\mathcal{T}(\Theta) = \{\neg E\}$ is a semi-closed extension and there is an initial condition K such that $\mathcal{T}(\Theta) + \{K\}$ is a semi-closed extension (Similarly for failure, putting E for $\neg E$). A standard way of achieving empirical progress is to convert a theory for which E is lacuna into a theory for which E is a success. This involves finding a set of wffs H to add to the theory such that $K \cup \Theta \cup \{H\}$ is a revision of Θ relative to K for which E is now a success. This particular transformation is achieved when $K \cup \Theta \cup \{H\}$ confirms neither E nor $\neg E$.

17.1 Assessing Semantic Tableau Abduction

Semantic tableau abduction considerably resembles enthymeme resolution, (concerning which, see chapter 9 of [Gabbay and Woods, 2005b]). The more abstractly formal its presentation, the closer the similarity is. In enthymeme resolution is to find a P that closes a $\vDash$-connection between some premisses and a conclusion, where $\vDash$ is read as deductive consequence. In semantic tableau abduction, the task is to find a P that closes a $\vDash$-connection between a theory and some empirical data. Although the general theory doesn't require it, Θ in its adaptation to the ends of Kuipers' theory of empirical progress, $\vDash$ may be read as a confirmation relation. What this shows us is the importance of the interpretation given to $\vDash$ (or, in the GW-schema, $\looparrowright$). Treated as *any* consequence relation, as in Aliseda's core theory, there is nothing abductive about the closure of $\vDash$-connections. Treated as confirmation, things aren't quite so bleak.

So, then, what *do* we make of a situation in which adding P to Θ closes the confirmation-connection to some data E? Is this not some reason not only to send P to trial, but also to make the conjecture that P is true? The answer depends in part on the size and importance of E and of the collateral costs (if any) that attach to conjecturing P's truth. But let us concede what is plain to see. In ranges of cases, this kind of confirmation is reason, albeit sometimes modest reason, to make the conjecture. But this is not abduction either. Justified conjecturability is part of what is required, and we may suppose that we have it for large classes of cases. What is also required is ignorance, and nowhere is that factor addressed in semantic tableau abduction. Of course, it is a requirement of abductions of this sort that the P that turns the trick not be derivable in the original theory Θ. But it is not ruled out that P be true or that it be known to be true, then it is eligible for consideration by semantic tableau theorists even so. For it suffices to add it to Θ as a new axiom. Suppose we do. Where now is the case that P's role in closing the $\vDash$-connection to E offers reasons for conjecturing that P is true? If P is known to be true, there is no room for any such conjecture; and anything purporting to constitute grounds for conjecture will have turned out to be epistemically defective. It hardly needs saying that needed repairs are not hard to make. One might require, for example, that any P selected in a semantic tableau exercise not have a degree of epistemic virtue equal or greater to that evinced by the original theory Θ. Still, there are two admonitions that require gentle sounding. The first we've already met with. It is

Proposition 17.1 (Arbitrary interpretations of $\vDash$) *The freer the range of interpretations of $\vDash$ in semantic tableau abductions, the less genuinely abductive they are.*

Corollary 17.1(a) *Apart from that, in semantic tableau abductions, $\vDash$ can bear only those interpretations that answer to the relevant closure conditions. In particular, $\vDash$ cannot be interpreted as* plausible consequence*, since plausibility is not closed under negation.*

17.2 Is It Abduction?

In our discussion in the section just above we have been assuming, with Aliseda, that Kuipers' theory reveals empirical progress to have an abductive character. Armed with that assumption, we have been considering Aliseda's interesting attempt to elucidate the abductive structure of empirical progress by adapting Kuipers' account to her own semantic tableau account. It is essential to the success of her project that she be right in making this assumption. Given our discussion of empirical progress in the previous section, we are unable to agree that the assumption is sound. According to what we called *Kuipers' Dilemma*, empirical progress inferences are either not abductive or not sound. Accordingly, we are skeptical of Aliseda semantic tableau accommodation of Kuipers as a contribution to the logic of abduction. Of course, this does nothing to disturb the importance of the link Aliseda has been able to forge between the methodology of semantic tableaux and the analysis of empirical proggress.

When we say that the ignorance-preservation aspect of abduction puts pressure on the presumedly abductive character of (Aliseda's approach to) empirical-progress inferences in the manner of Kuipers, we are mindful that a law of equal and opposite reaction applies here, and that the Aliseda-Kuipers approach pushes right back, putting pressure on our claim of ignorance-preservation. We don't for a moment think that we have settled this issue here, once and for all. But, unless we are mistaken, it is an issue that will require settlement sooner or later.

18 Concluding Remarks

Peirce saw abduction as a kind of guessing. In doing so, the modern discoverer of abduction took empathetic note of abduction's epistemically constrained state. By our lights, this is the defining insight of abduction and is central to any purported account of it.

In the present essay, we have concentrated on *one step* and *fully executed* abductions. These may be thought of as the core building-blocks of abductive reasoning. Clearly, however, abduction admits of degrees of complexity not captured by our present remarks. Our last bit of advice is that a logic of abduction should not overlook these complexities, including among others the following (which we state for the consequentialist case; they generalize to the *GW* model quite straightforwardly):

a. *Multiple target abduction*, in which the closure of an agenda by way of a conditional in the form $K(H) \looparrowright V$ is itself a state of affairs that closes a further agenda.

b. *Compound abduction*, in which for different submodels K_1 and K_2 of K, and different payoff propositions V_1 and V_2, we have it that $K_1(H) \looparrowright V_1$ and $k_2(H) \looparrowright V_2$.

c. *Transitive closure abduction*, in which for certain (but not all) interpretations of $\looparrowright$, we have it that $K(H) \looparrowright V_1, V_1 \looparrowright V_2$, hence that $K(H) \looparrowright V_2$. For example, this would work when $\looparrowright$ is construed as causal implication, but not for all interpretations in which $\looparrowright$ is explanatory consequence.

d. *Iterated abduction*, in which a reasoning $K(H) \looparrowright V$ sets up a further trigger $\{K(H), T\}$, with respect to which for some T', V' and H' there is a further winner $K(H)(H') \looparrowright V'$.

e. *Partial abduction*, in which either or both the consequence relation $\looparrowright$ and the closure relation R admit of degrees of satisfaction. A case in point: if $K(H)$ partly explains V, and the abducer's agenda is to have an explanation, then the truth of $K(H) \looparrowright V$ constitutes *partial closure* of that agenda.

[21]

[21]We owe a considerable debt of gratitude to our valued colleague Lorenzo Magnani for having organized two superb meetings of interest to abduction theorists – the Symposium on Abduction of the Cognitive Science Society, in Chicago, August 2004, and the Model Based Reasoning Conference, at the University of Pavia, December 2004. We were greatly assisted in achieving better grasp on the structure of abduction by a number of lectures in both places and comments from interested parties. We would especially like to thank, in addition to Professor Magnani, Atocha Aliseda, Tom Addis, Peter Bruza, Balakishnan Chandrasekaran, Theo Kuipers, Joke Meheus, Sami Paavola and Matti Sintonen. Early versions benefitted considerably from discussion at Woods" seminars on abduction during his appointment in 2001 as Vonhoff Professor at the University of Groningen. We wish to thank in particular Atocha Aliseda and Theo Kuipers (again), David Atkinson, Erik Krabbe, Jeanne Peijnenburg, menno Rol, Jan-Willem Romeyn, Victor Sánchez Valencia, Alice ter Meulen, Alexander van den Bosch and Jan Albert van Laar. Our research on abduction is supported by grants from the Engineering and Physical Sciences Research Council of the United Kingdom and the Social Sciences and Humanities Research Council of Canada. We are exceedingly grateful for this support. Special thanks for unstinting technical support is reserved by Jane Spurr in London and Carol Woods in Vancouver.

References

[Aliseda-LLera, 1997] Atocha Aliseda-LLera. *Seeking Explanations: Abduction in Logic, Philosophy of Science and Artificial Intelligence*. Amsterdam: Institute for Logic, Language and Computation, 1997. PhD dissertation (ILLC Dissertation Series 1997-4).

[Aliseda, forthcoming] Atocha Aliseda. *Abductive Reasoning: Logical Investigation into the Processes of Discovery and Evaluation*. Dordrecht: Kluwer, forthcoming.

[Aravindan and Dung, 1994] C. Aravindan and P.M. Dung. Belief dynamics, abduction and databases. In C. MacNish, D. Pearch, and L.M. Pereira, editors, *Logics in Artificial Intelligence*, pages 66–85. Berlin: Springer-Verlag, 1994.

[Armour-Garb, 2004] Brad Armour-Garb. Wrestling with (and without) dialetheism. *Australasian Journal of Philosophy*, 2004. To appear.

[Arruda, 1989] A.I. Arruda. Aspects of the historical development of paraconsistent logic. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic: Essays on the Inconsistent*, Munich: Philosophia Verlag, 1989.

[Batens, 1980] D. Batens. Paraconsistent extensional propositional logic. *Logique et Analyse*, 23:195–234, 1980.

[Beth, 1969] E.W. Beth. Semantic entailment and formal derivability. In J. Hintikka, editor, *The Philosophy of Mathematics*, pages 9–41. Oxford: Oxford University Press, 1969.

[Boutilier and Becher, 1995] C. Boutilier and V. Becher. Abduction as belief revision. *Artificial Intelligence*, 77:43–94, 1995.

[Chomsky, 1972] Noam Chomsky. *Language and Mind*. New York: Harcourt Brace Jovanovich, enlarged edition, 1972.

[DaCosta and Bueno, 1996] N. DaCosta and O. Bueno. Consistency, paraconsistency and truth. *Ideas y Valores*, 100:48–60, 1996.

[Detlefsen, 1986] M. Detlefsen. *Hilbert's Program: An Essay on Mathematical Instrumentalism*. Dordrecht and Boston: Reidel, 1986.

[Flach and Kakas, 2000] Peter A. Flach and Antonis C. Kakas, editors. *Abduction and Induction: Essays on Their Relation and Integration*. Dordrecht and Boston: Kluwer, 2000.

[Gabbay and Hunter, 1993] Dov M. Gabbay and A. Hunter. Making inconsistency respectable part 2: Meta-level handling of inconsistency. In *LNCS 747*. Berlin: Springer-Verlag, 1993.

[Gabbay and Woods, 2001] Dov M. Gabbay and John Woods. The new logic. *Logic Journal of the IGPL*, 9:157–190, 2001.

[Gabbay and Woods, 2002] Dov M. Gabbay and John Woods. Formal approaches to practical reasoning: A survey. In Dov M. Gabbay, Ralph H. Johnson, Hans Jürgen Ohlbach, and John Woods, editors, *Handbook of the Logic of Argument and Inference: The Turn Towards the Practical*, volume 1 of *Studies in Logic and Practical Reasoning*, pages 445–478. Amsterdam: North-Holland, 2002.

[Gabbay and Woods, 2003] Dov M. Gabbay and John Woods. *Agenda Relevance: A Study in Formal Pragmatics*, volume 1 of *A Practical Logic of Cognitive Systems*. Amsterdam: North-Holland, 2003.

[Gabbay and Woods, 2005a] Dov M. Gabbay and John Woods. A formal model of abduction. In L. Magnani, editor, *Model-based reasoning in science and engineering. Abduction, visualization and simulation*. 2005.

[Gabbay and Woods, 2005b] Dov M. Gabbay and John Woods. *The Reach of Abduction: Insight and Trial*, volume 2 of *A Practical Logic of Cognitive Systems*. Amsterdam: North-Holland, 2005.

[Gabbay and Woods, 2005c] Dov M. Gabbay and John Woods. *Seductions and Shortcuts: Fallacies in the Cognitive Economy*, volume 3 of *A Practical Logic of Cognitive Systems*. Amsterdam: North Holland, 2005.

[Gabbay and Woods, 2006] Dov M. Gabbay and John Woods. *The Logical Structure of English Criminal Law*. Springer, Berlin and Dordrecht, 2006. to appear.

[Gärdenfors, 1988] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Cambridge Massachusetts:Bradford Books, The MIT Press, 1988.

[Gervás, 1995] P. Gervás. *Logical Considerations in the Interpretation of Presuppositional Sentences*. PhD thesis, Department of Computing, Imperial College, London, 1995.

[Ginsberg, 1988] A. Ginsberg. Theory of revision via prior operationalization. In Jeff Shrager and Pat Langley, editors, *Computational Models of Scientific Discovery and Theory Formation*. San Mateo, CA: Morgan Kaufmann Publishers, 1988.

[Goddu, 2002] G.C. Goddu. The most important and fundamental distinction in logic. *Informal Logic*, 22:1–17, 2002.

[Gödel, 1944] Kurt Gödel. Russell's mathematical logic. In *The Philosophy of Bertrand Russell*, pages 123–153. New York: Tudor, 3rd edition, 1951.

[Gödel, 1990a] Kurt Gödel. Remarks before the Princeton bicentennial conference on problems in mathematics. In Solomon Feferman, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort, editors, *Kurt Gödel Collected Works, Volume II, Publications 1938–1974*, pages 150–153. New York and Oxford: Oxford University Press, 1990. Originally published in 1946.

[Gödel, 1990b] Kurt Gödel. What is Cantor's continuum problem? In Solomon Feferman, John W. Dawson Jr., Stephen C. Kleene, Gregory H. Moore, Robert M. Solovay, and Jean van Heijenoort, editors, *Kurt Gödel Collected Works, Volume II, Publications 1938–1974*, pages 176–187. New York and Oxford: Oxford University Press, 1990. Originally published in 1947.

[Hanson, 1958] N.R. Hanson. *Patterns of Discovery*. Cambridge: Cambridge University Press, 1958.

[Heim, 1983] I. Heim. On the projection problem for presuppositions. In *Proceedings of the Second West Coast Conference on Formal Linguistics, WCCFL*, volume 2, pages 114–126, 1983.

[Hintikka, 1955] Jaakko Hintikka. *Two Papers on Symbolic Logic*. Helsinki: Akateeminen Kirjak, 1955.

[Hitchcock, 2002] David Hitchcock. A note on implicit premisses. *Informal Logic*, 22:158–159, 2002.

[Hobbs *et al.*, 1990] J.R. Hobbs, M. Stickel, D. Appelt, and P. Martin. Interpretation on abduction. Technical Report 499, SRI International, Artificial Intelligence Center, Computing and Engineering Sciences Division, Menlo Park, CA, 1990.

[Jáskowski, 1948] S. Jáskowski. Rachunek zdak dla systemów dedukcyzjych sprzecznych. *Studia Societatis Torunesis*, 1:55–77, 1948. Sectio A.

[Jennings and Schotch, 1981] R.E. Jennings and P.K. Schotch. Some remarks on (weakly) weak modal logics. *Notre Dame Journal of Formal Logic*, 22:309–314, 1981.

[Josephson and Josephson, 1994] J.R. Josephson and S.G. Josephson, editors. *Abductive Inference: Computation, Philosophy, Technology*. Cambridge: Cambridge University Press, 1994.

[Kakas and Mancarella, 1990] A. Kakas and P. Mancarella. Knowledge assimilation and abduction. In *Proceedings of the European Conference on Artificial Intelligence, ECAI*, Berlin: Springer-Verlag, 1990.

[Kakas *et al.*, 1995] A. Kakas, R.A. Kowalski, and F. Toni. Abductive logic programming. *Journal of Logic and Computation*, 2:719–770, 1995.

[Kapitan, 1992] Tomis Kapitan. Peirce and the autonomy of abductive reasoning. *Erkenntnis*, 37:1–26, 1992.

[Kelly, 1989] G.M. Kelly. Elementary observations on 2-categorical limits. *Bull. Aust. Math. Soc.*, 39(2):301–317, Apr 1989.

[Konolege, 1996] K. Konolege. A general theory of abduction. In G. Brewska, editor, *Principles of Knowledge Representation*, pages 129–152. Stanford, CA: CLSI, 1996.

[Kowalski, 1979] R.A. Kowalski. *Logic for Problem Solving*. New York: Elsevier, 1979.

[Kraus, 2003] Manfred Kraus. Charles Peirce's theory of abduction and the Aristotelian enthymeme from signs. In Frans H. van Eemeren, J. Anthony Blair, Charles A. Willard, and A. Francisca Snoeck Henkemans, editors, *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation*. Amsterdam: Sic Sat, 2003.

[Kuipers, 1999] Theo A.F. Kuipers. Abduction aiming at empirical progress of even truth approximation leading to a challenge for computational modelling. *Foundations of Science*, 4:307–323, 1999.

[Kuipers, 2000] Theo A.F. Kuipers. *From Instrumentalism to Constructive Realism: On Some Relations Between Confirmation, Empirical Progress and Truth Approximation*. Dordrecht and Boston: Kluwer, 2000.

[Laplace, 1904] Pierre Laplace. Mémoires. In *Oeuvres completes de Laplace*, volume 13/14, Gauthier-Villars 1878–1912, 1904.

[Laudan, 1980] Larry Laudan. Why was the logic of discovery abandoned? In Thomas Nickles, editor, *Scientific Discovery, Logic and Rationality*, pages 173–183. Dordrecht: Reidel, 1980.

[Lloyd, 1987] J.W. Lloyd. *Foundations of Logic Programming*. Berlin: Springer-Verlag, 2nd edition, 1987.

[Magnani, 2001] Lorenzo Magnani, ed. *Philosophy and Geometry: Theoretical and Historical Issues*, Dordrecht: Kluwer, 2000.

[Magnani, 2001] Lorenzo Magnani. *Abduction, Reason and Science: Processes of Discovery and Explanation*. New York: Kluwer, Plenum, 2001.

[Mayer and Pirri, 1993] M. C. Mayer and F. Pirri. First order abduction via tableau and sequent calculi. *Bulletin of the IGPL*, 1:99–117, 1993.

[McAllister, 1996] J. McAllister. *Beauty and Revolution in Science*. Ithaca, NY: Cornell University Press, 1996.

[Meheus *et al.*, forthcoming] Joke Meheus, Liza Verhoeven, Maarten Van Dyck, and Dagmar Provijn. Ampliative adaptive logics and the foundation of logic-based approaches to abduction. In Lorenzo Magnani, Nancy J. Nersessian, and Claudio Puzzi, editors, *Logical and Computational Aspects of Model-Based Reasoning*. Dordrecht and Boston: Kluwer, forthcoming.

[Musgrave, 1989] Alan Musgrave. Deductive heuristics. In Kostas Gavroglu, Yorgos Goudaroulis, and Pantelis Nicolacopoulos, editors, *Imre Lakatos and Theories of Scientific Change*, pages 15–32. Dordrecht: Kluwer, 1989.

[Nickles, 1980] Thomas Nickles. Introductory essay: Scientific discovery and the future of philosophy of science. In Thomas Nickles, editor, *Scientific Discovery, Logic and Rationality*, pages 1–59. Dordrecht: Reidel, 1980.

[Paul, 1993] G. Paul. Approaches to abductive reasoning: An overview. *Artificial Intelligence Review*, 7:109–152, 1993.

[Peirce, 1931–1958] C.S. Peirce. *Collected Works*. Cambridge, MA: Harvard University Press, 1931–1958. A series of volumes, the first appearing in 1931.

[Peng and Reggia, 1990] Yun Peng and James A. Reggia. *Abductive Inference Models for Diagnostic Problem-solving*. New York and Berlin: Springer-Verlag, 1990.

[Polya, 1945] G. Polya. *How to Solve it. A New Aspect of Mathematical Method*. Princeton, NJ: Princeton University Press, 1945.

[Polya, 1954] George Polya. *Induction and Analogy in Mathematics*, volume 1 of *Mathematics and Plausible Reasoning*. Princeton, NJ: Princeton University Press, 1954.

[Polya, 1962] G. Polya. *Mathematical Discovery. On Understanding, Learning, and Teaching Problem Solving*, volume I. New York: Wiley, 1962.

[Poole, 1989] David Poole. Explanation and prediction: An architecture for default and abductive Reasoning. *Computational Intelligence*, 5:97–110, 1989.

[Priest, 2002] Graham Priest. Paraconsistent logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 6, pages 287–393. Dordrecht and Boston: Kluwer, 2nd edition, 2002.

[Pople, 1973] H. E. Pople. On the Mechanization of Abductive Logic. In *Proceedings of the Third International Joint Conference on Artificial Intelligence, IJCAAI-73*. Stanford, CA: Morgan Kauffmann, 1973.

[Ramoni *et al.*, 1992] M. Ramoni, M. Stefanelli, L. Magnani, and G. Barosi. An epistemological framework for medical knowledge-based systems. In *IEEE Transactions on Systems, Man and Cybernetics*, volume 22, pages 1361–1375. 1992.

[Reichenbach, 1938] Hans Reichenbach. *Experience and Prediction*. Chicago, IL: University of Chicago Press, 1938.

[Rescher, 1976] Nicholas Rescher. Peirce and the economy of research. *Philosophy of Science*, 43:71–98, 1976.

[Rescher, 1996] Nicholas Rescher. *Priceless Knowledge? Natural Science in Economic Perspective*. Lanham, MD: Rowman and Littlefield, 1996.

[Russell, 1907] Bertrand Russell. The regressive method of discovering the premises of mathematics. In Douglas Lackey, editor, *Essays in Analysis*. London: George Allen and Unwin, 1907. Published in 1973.

[Savary, 1995] C. Savary. Discovery and its logic: Popper and the friends of discovery. *Philosophy of the Social Sciences*, 25:318–344, 1995.

[Smullyan, 1968] R.M. Smullyan. *First Order Logic*. Amsterdam: Springer-Verlag, 1968.

[Tanaka, 2003] Koji Tanaka. Three schools of paraconsistency. *Australasian Journal of Logic*, 1:28–42, 2003.

[Thagard, 1989] Paul Thagard. Explanatory Coherence. *Behavioral and Brain Sciences*, 12:381–433, 1989.

[Woods, 2004a] John Woods. *The Death of Argument: Fallacies in Agent-Based Reasoning*. Dordrecht and Boston: Kluwer, 2004.

[Woods, 2004b] John Woods. The economics of paradox. *Australasian Journal of Philosophy*, 2004. to appear.

[Woods, 2005, to appear] John Woods. Dialectical Considerations on the Logic of Contradiction. *Logic Journal of the IGPL*, 2005, to appear.