Book Reviews

THE REACH OF ABDUCTION: INSIGHT AND TRIAL
By Dov M. Gabbay and John Woods. Amsterdam: Elsevier, 2005.
Pp. xviii, 1-476. ISBN 0-444-51791-X. Hardbound, US$160.00

The Reach of Abduction is the second volume in a planned three-volume set—*A Practical Logic of Cognitive Systems*—spanning relevance, abduction, and fallacious reasoning. Despite this fact, Gabbay and Woods write in the preface that "we have written the individual volumes with a view to their being read either as stand-alone works or as linked and somewhat overlapping items in the series" (p. xvii). The aim of *The Reach of Abduction* is to embed abduction "within a practical logic of cognitive systems" and in so doing, provide "an adequate stand-alone characterization of abduction itself" (p. 9). At the same time, Gabbay and Woods admit this work is only meant to be "an enterprise of first words rather than last" (p. xviii). Indeed, this tentativeness gets expressed again and again throughout the book—here is but one example: "Much of what we will have to say for ourselves here and in the book's succeeding chapters will be fragmentary, tentative, programmatic and promissory" (p. 71).

The book is divided into three sections. Part I, comprised of the first two chapters, is a general introduction to the framework within which they will be studying the problem of abduction, i.e., their general theory of a practical logic of cognitive systems—hence we get a brief articulation of what the authors mean by 'logic', 'practical' and 'cognitive system,' etc. Part II, comprised of chapters 3 – 9 and more than half of the book, is devoted to articulating and discussing various conceptual models of abduction. Part III, comprised of chapters 10 – 13, is devoted to presenting and discussing formal models of abduction. (The third part of the book, like the second part, is titled "Conceptual Models of Abduction"—just one of a frustratingly high number of typographical and copy-editing errors.)

In Chapter 1, which is a seven-page introduction, Gabbay and Woods provide a brief articulation and justification of the presuppositions that will ground their approach to abduction. Most significantly, they construe 'logic' quite broadly "as the disciplined description of the behaviour of real-life logical agents" (p. 1) such that they "propose to absorb the logic of linguistic structures into a more comprehensive logic of agency" (p. 2). Hence, while they acknowledge that their main technical thrust will be via formal models rather than computational implementation, they point out that the reach of their "models extends beyond sets of properties of agent-independent systems" (p. 4).

In Chapter 2, Gabbay and Woods, provide a brief discussion of what they mean by 'practical logic' and 'cognitive system'. (They also point the reader to Chapters 2 and 3 of the first book in the series, *Agenda Relevance*, for a more detailed account.) The core of the discussion surrounds the following two definitions:

> Definition 2.1 (Cognitive Systems) *A cognitive system CS is a triple X, R, A of a cognitive agent X, cognitive resources R, and a cognitive agenda A executed in real time.*

> Definition 2.2 (Practical logics, a first pass) *A practical logic is a systematic account of aspects of the behaviour of a cognitive system in which X is a practical agent.* (p. 10)

According to Gabbay and Woods, "an agent is a practical agent to the extent that it commands comparatively few cognitive resources in relation to comparatively modest cognitive goals," (p. 12) and "practical reasoning is reasoning done by a practical agent" (p. 12). Note that what Gabbay and Woods mean by practical reasoning or practical logic is not what many philosophically oriented readers might first think of—i.e., as reasoning concerning action as opposed to reasoning concerning knowledge (see, for example, Audi, *Practical Reasoning*). According to Gabbay and Woods, it is the type of agent, not the task to which the reasoning is devoted, that makes reasoning or logic practical rather than theoretical. "Our interpretation of the practical–theoretical dichotomy may strike the reader as nonstandard, if not eccentric.... We are prepared to put up with the nonstandardness in return for conceptual yield" (p. 13). The chapter is then devoted to discussions and examples of how practical agents differ from theoretical agents, with particular attention to the sorts of inference strategies practical agents adopt in order to minimize the expenditure of their limited cognitive assets.

Chapter 3 is the most crucial chapter of the book and is devoted to an initial articulation of the authors' conception of abduction. According to Gabbay and Woods, abduction is the finding and engaging of a hypothesis (H) that, when combined with what one already knows (K), enables one to presumptively attain a cognitive target (T) that one could not attain via K (or a ready expansion of K) alone. (p. 47) Engaging an hypothesis is a two-part task comprised of (a) holding that H is worthy of conjecture and (b) releasing H for possible use as premise in future reasoning (p. 69). Finding an hypothesis involves two general tasks, (a) originating the candidate space of potential hypotheses and (b) cutting the candidate space down to, hopefully, a successful H. The authors do not separate out the finding of H quite this explicitly, but tend to weave the two tasks together—most of the emphasis in this chapter is placed on the cutting down task and the logical resources that will need to be brought to bear on describing the behaviour involved in this task—hence Gabbay and Woods argue that an abductive logician will need to deploy the resources of the appropriate logics of relevance and plausibility (p. 57). A crucial condition on H, according to Gabbay and Woods, is that H *only* allows the presumptive attainment of T. Since T cannot be achieved via what one knows or what one can readily add to what one knows, H has a lesser epistemic status than the abducer may have originally hoped for when trying to attain T. Hence, the authors claim, "the fundamental conceptual fact about abduction is that abduction is ignorance-preserving reasoning" (p. 40). Fully generalized, the claim is that abduction is cognitive-deficit preserving (p. 45).

Intertwined with the presentation of their own schema of abduction is a discussion of "the dominant schematic representation of abduction in the current literature," what they call the AKM-schema. Gabbay and Woods object to the AKM-schema on the grounds that it focuses on explanationist/consequentialist abduction, which is at best part of abduction and at worst distorts the true nature of abduction. While this discussion is focused on differences of opinion regarding the conjecturing of H (part of the engagement process), the rest of the chapter is

primarily devoted to a discussion of the cutting down of the candidate space in order to arrive at a (hopefully) successful H.

In Chapter 4, Gabbay and Woods discuss explanationist abduction, i.e., abduction in which the cognitive target requires an explanatory H in order to be attained. (The AKM-schema is meant to model explanationist abduction.) After a presentation and evaluation of Pierce's views on abduction, Gabbay and Woods present various competing models of explanation, e.g., covering law, rational, teleological and conclude that if inference to the best explanation is evidentially clinching, then it is not abductive (since this would violate the ignorance condition)—but Gabbay and Woods leave it as an open question whether inference to the best explanation is evidentially clinching or not (though they argue that, at the very least, the D-N sense of explanation must be subjunctive if it is to have a chance of being abductive).

Chapter 5 is devoted to non-plausibilistic abduction, i.e., abductions in which H is a propositionally implausible hypothesis. The main method of explication here is via examples from science and logic, e.g., from Newton, Plank, Russell and Gödel, superstring theory, Lakatos, etc. However, intertwined with these examples is a general discussion of such issues as how to understand the consequence relation in abductions (it depends on the target), whether semantic tableaux abduction à la Aliseda-Llera is really abduction (the authors are skeptical), whether Bayesian inference is abductive (it isn't). The main upshot of this chapter seems to be that H can be (i) epistemically hopeless (Newton's view on action-at-a-distance) or (ii) self-inconsistent or (iii) inconsistent with K. The upshot of (ii) and (iii) is that abductive reasoning requires a paraconsistent base logic with a dialethic component.

In Chapter 6, "Diagnostic Abduction in AI," Gabbay and Woods begin by comparing two "representative treatments of diagnostic abduction"—the explanationist diagnostic logic AP developed by Josephsen and Josephsen and "the parsimonious covering theory" of Peng and Reggia. They then rely on arguments of Thagard's, via modeling abduction using probabilistic networks, to suggest that probabilistic methods will be inadequate, at least for modeling abductive reasoning in practical agents. They finish the chapter with a brief discussion of modeling abduction using neural-symbolic networks. This chapter is much more formal than any other chapter in this section and generally does not mesh well with the other chapters on conceptual models of abduction.

In Chapters 7 and 8, Gabbay and Woods fill out in somewhat more detail their claim from Chapter 3 that cutting down the candidate space of potential hypotheses will rely on relevance and plausibility. In particular, they discuss the role that a suitable relevance logic plays in reducing the candidate space of solutions for an abduction problem to a more manageable set of 'relevant' solutions. They also discuss how this set of relevant solutions can be adjudicated via plausibility/implausibility constraints. Once some initial examples are made, to get the (at least) intuitive role of these two concepts clear, Gabbay and Woods focus on plausibility throughout the rest of the Chapter 7 and turn to relevance in Chapter 8.

By the end of Chapter 7, Gabbay and Woods postulate a connection between what is 'plausible' and what is 'characteristic'—"plausibility then falls out as what conforms to what is characteristic, with implausibility contradicting the characteristic" (p. 238). They compare their reliance on what is characteristic with Rescher's plausibility logic and while admitting that in some respects their approach

fares no better, they also claim that their approach will avoid problems of circularity and extension to impersonal contexts that bedevil Rescher's approach.

In the first half of Chapter 8, Gabbay and Woods argue that while standard logical treatments of relevance, which treat mostly of propositional relevance (topical relevance, full-use abduction, irredundancy-relevance), are too coarse-grained for the work that needs to be done in producing a theory of abduction, they also suggest that such logics can play a subsidiary role. Ultimately, they argue in favor of the view of relevance expounded in the first volume of this series, *viz.*, agenda relevance. (I is relevant for X with regard to A to the extent that in processing I, X is affected in ways that conduce toward the advancement or closure of A (p. 240).) In the second half of the chapter, Gabbay and Woods, discuss legal relevance and legal presumption and their ties to abduction. Indeed, the authors present many comparisons with legal situations and abduction, without linking the discussion back to relevance—hence the chapter does not hang together well as a whole. They end the chapter with a discussion of the role of analogy in abduction.

In Chapter 9, "Interpretation Abduction," Gabbay and Woods focus on interpretation problems, especially linguistic interpretation problems, and the degree to which these problems are also abduction problems. Their primary cases are enthymeme resolution and principles of charity, and they conclude that while text interpretation is not intrinsically abductive, there certainly are abductive cases.

Chapters 10–13 are devoted to setting the groundwork for formal models of abduction. Chapter 10 provides some examples and suggests what sort of logical apparatus will be needed in order to accommodate the examples. For instance, Gabbay and Woods suggest that the logic will have to be defeasible and require a labeled deductive system. Chapter 11 provides a general theory of logical systems, which, as the authors point out, is "a serious departure from current conceptual practice" (p. 366). In particular, they hold that logical systems are not just the sets of formulas provable under a given consequence relation, but also the algorithmic system used to generate the set of formulas. Hence, a semantic tableaux, a natural deduction system, a truth table system for classical propositional logic, would all count as different logical systems according to Gabbay and Woods, for even though the set of formulas provable is identical in each case, the method of proving the formulas is different. Gabbay and Woods go on to modify the set of formulas to be structured data such that again the same consequence relation could be involved in different logical systems because the systems structure the provable formulas differently. They also add mechanisms that "make use of data and algorithms to extend data" (p. 380). They maintain that data can also be mechanisms, i.e., operations on data, so data need not just have declarative content. Ultimately, they maintain that a logical system should be presented as a quadruple, <Consequence Relation, Algorithm, Relevance, Mechanisms>.

In Chapters 12 and 13, Gabbay and Woods turn to formal abduction. After a brief introduction to a Labelled Deductive System (see Gabbay, 1996 for more detail), Gabbay and Woods show how to begin constructing abduction algorithms for such systems, along with numerous examples (though most in the context of programming and database interaction). In section 13.2, they define one particular abduction algorithm and sketch a proof for its soundness. They conclude with some examples of abduction for intuitionistic implications and for relevance logic.

Having put forward the mere beginnings of formal models for abduction, they note that "proper formal modeling is better done at the end of this series of volumes since all the concepts and mechanisms are interdependent"(p. 441).

The preceding summary provides only the barest outline of much of the book and still leaves numerous sections of the book unmentioned. One reason for this tremendous breadth can be found in the facts that, as the authors point out right from the beginning, (a) there is no dominant theory of abduction and (b) there is "a large literature on abduction ... from philosophy, cognitive psychology, computer science, artificial intelligence and, of course, logic" (p. xvii-xviii). As a result, there is quite a variety of approaches and issues that loosely fall under the abduction umbrella—approaches and issues the authors cover, briefly at least, through the course of the book.

Unfortunately, the sheer breadth of discussion is a hindrance to the reader. Firstly, the scope of the discussions makes following a coherent path through the book extremely difficult, if not impossible. In their bid to be thorough, Gabbay and Woods take the reader down many side paths without making it clear to the reader what the upshot is for their own theory of abduction. Indeed, precisely picking their theory out from the clutter is quite a challenge for the reader.

Secondly, the result is too many 'first words' where in some cases it would have been nice to have gotten some second or third words, especially concerning their own proposal. For example, there is no argument for the ignorance-condition (or cognitive-deficit condition) on abduction—it is stipulated (repeatedly) and then used to argue that many cases of what we thought were abduction are (or perhaps might) not really be abduction at all. For example, non-subjunctive deductive-nomological explanationism and evidentially clinching inference to the best explanation are ruled non-abductive because they violate the ignorance-condition. But given that Gabbay and Woods accuse the AKM-schema's minimality condition on the union of K (the reasoner's knowledge base) and H (the hypothesis put forward to achieve the reasoner's goal) of "a contingency elevated to the status of a logically necessary condition" (p. 55), one might wonder whether Gabbay and Woods are guilty of the same charge with regards to the ignorance-condition. Given that so much of what is taken to be examples of abductive behavior is getting thrown out, one might doubt the logical necessity of the ignorance condition on abduction. Indeed, since Gabbay's and Woods' conception of logic is fundamentally a description of the behavior of the agent, why is the fundamental nature of abduction resting on the epistemic status of H relative to K and T rather than the agent's behavior? At the same time, some may be skeptical of their proposal on the grounds that they have so redefined logic and practical agent that they are not really solving the original problem of abduction at all.

Gabbay and Woods claim that part of their project is to get abduction right. Evaluating whether they succeed is made even more difficult by the very tentative nature of their claims. They call this book their "prologomenon to the logic of abduction" (p. 61) and admit that their "own account of abduction is itself abductive" (p. 42), i.e., they are putting forth a hypothesis concerning the nature of abduction that is a mere conjecture, though a conjecture they are sufficiently confident in to release for use in future argumentation. Have they really cut down the candidate space to a single proposal? Even if they have, is the hypothesis that remains

robust enough to warrant being worthy of conjecture? What standards must be met in order for the hypothesis to be worthy of conjecture anyway? Without answers (or the details necessary to ground the answers) to these questions, whether they ought to be confident in their conjecture remains an open question and, given their minimal discussion of adjudicating proper from improper abductions, it is impossible to tell whether or not they are warranted in their conjecture.

Finally, unsurprisingly, the conceptual discussion far outstrips the modeling of the formal discussion. For example, relative to the issue of 'discovery,' Gabbay and Woods point out that "because of the algorithmic nature of the proof theory, it is always clear what is locally needed to carry on" (pp. 383-384). In other words, for the formal models they consider there is no issue of generating a candidate set—that is determined by the proof theory. For the formal models, the issue is to provide a mechanism for cutting down the candidate space. But in 'real-life' cases there is, in many instances, the issue of generating the candidate space in the first place. I am surprised to find the door to my office open on a weekend. Wanting to know why it is open, I may generate a candidate space that includes, 'I forgot to lock my door when I left on Friday,' 'Someone broke into my office,' 'The cleaning staff left my door open.' But if, through lack of imagination, I leave out 'The physical plant staff left my door open,' I may be led astray in what I hypothesize as the explanation. Indeed, though Gabbay and Woods talk about producing a *bona fide* logic of discovery, they limit most of their discussion (both conceptually and formally) to the problem of cutting down the candidate space—yet surely a true 'logic of discovery' will need to include mechanisms for discovering a candidate space in the first place.

While there is much that is frustrating in their broad ranging discussion of abduction, there is also much to ponder from a variety of perspectives. Gabbay and Woods have offered a comprehensive discussion of abduction and made an intriguing case for embedding abduction within their quite promising practical logic of cognitive systems. Indeed, I strongly suspect that the true merits of their approach will be much more evident against the backdrop of a more comprehensive theory of practical cognitive systems—but that theory must wait for, at least, the third volume in the series. Gabbay and Woods describe this book at one point as "a call to arms to the research community" (p. 61). In this they are surely correct. Clearly, there is much more work to be done.

References

Audi, Robert. (1989). *Practical Reasoning*. London: Routlege.

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

Gabbay, Dov. (1996). *Labelled Deductive Systems*. Oxford: Oxford University Press.

G. C. Goddu *University of Richmond*