Abduction at the interface of Logic and Philosophy of Science

Johan VAN BENTHEM

ABSTRACT: Abduction is a typical theme where logic and philosophy of science meet today: occasionally, with computer science as a go-between. This is just one instance of a broader study of 'styles of reasoning', dating back to Bolzano and Peirce. The resulting concern with 'logical architecture' moves us closer to cognitive science, and the dynamics of reasoning intertwined with learning and belief revision. The crucial process of self-correction involved here is usually triggered by others, and hence a shared target of logic and philosophy of science should be the phenomenon of 'intelligent interaction' between rational agents.

Keywords: logic, philosophy of science, artificial intelligence, cognitive science, abduction, dynamics, interaction.

It is a pleasure to say, and write a few words in connection with Atocha Aliseda's recent book *Abductive Reasoning: Logical Investigations into Discovery and Explanation* —which is based on her doctoral dissertation which I supervised during her years at Stanford as a first-wave student in the pioneering Symbolic Systems Program.

As is clear from the title, the book brings together two fields of research whose relations through history have varied. Clearly, great philosophers like Bolzano, Mill, Peirce, and Carnap were masters of both logic and methodology of science, and so was my predecessor Beth on the general logic chair in Amsterdam. Indeed, when Beth and Tarski realized their planetary ambitions after much plotting and scheming (Van Ulsen 2000), they created the "International Division of Logic, Methodology and Philosophy of Science" around 1960. But soon after, the fields went largely their own way —partly under the influence of the Suppes-Sneed view that set-theoretic structures alone are the privileged representation of scientific theories. This leaves little room for the logical languages needed to start up the model-theoretic machinery of logicians. And as we all know, frictions lead to partner change. History and sociology started wooing the philosophy of science, while logicians made new friends in linguistics and computer science. For these and other reasons, our two fields are largely disjoint today, with their own journals, jargon, and intellectual attitudes.

But now, looking at many other signs around us, it seems that a new spring is in the air. Can old love be re-kindled, after many decades of separation? Well, one reason why the two fields meet today is a shared modern environment, namely, developments in *computer science and artificial intelligence*. In particular, since the 1980s, AI has developed accounts of 'common sense reasoning' doing justice to a variety of inferential tasks far beyond standard logical deduction, and closer to the rich repertoire of activities of confirmation, explanation, and prediction found in the philosophy of science. But computer science uses language and logic essentially, and hence one is together again. Indeed, there are interesting parallels between common sense reasoning and the variety of reasoning styles studied by Bolzano (van Benthem 2006), who advocated a wide-ranging awareness of different logical properties of different styles. And the

same liberal conception is found, of course, in the innovative work of Peirce on de-
duction, induction, and abduction. Such studies can be undertaken in at least two
ways, both of them represented in Atocha's book. One is 'top-down', in terms of
structural properties such as Monotonicity (persistence of conclusions when new
premises are added), or the lack of it. This is in line with many other themes in logic
since the 1980s, culminating in the research area of 'substructural logics'. But there is
also a 'bottom-up' approach looking at specific computational mechanisms for per-
forming different styles of reasoning. There are many of these, but it is a pleasure to
note that Atocha's book uses Beth's semantic tableaux, born and bred in Amsterdam.

Of course, this is just the beginning of a field of study. In particular, attractive and
liberal as it sounds, the 'pluralist' methodology raises some issues. If humans have all
these reasoning abilities, presumably these *meet and combine* in fruitful manners, as we
integrate information from various sources. But then, we should definitely describe
mixtures of styles, including classical deduction, abduction, and what have you. And
this is not at all a simple matter of throwing everything into one combined system. In
various areas of modal logic, it has been discovered that natural combined systems
with decidable components can even become *undecidable* —sometimes by making
some apparently harmless assumptions about agents' abilities, such as their having
perfect memory. So, how harmless are combined abductive-deductive systems for
perfect reasoners? I guess we do not know.

Moreover, let us also put the above developments, encouraging as they are, in a
broader perspective. Should the current meeting of logic and philosophy of science be
cast in terms of consequence relations like abduction? I myself would prefer a much
broader view of *cognitive actions* that are involved in intelligent tasks. E.g., I have argued
that maybe the main phenomenon in non-monotonic common sense reasoning is not
consequence, but acts of *belief revision* (van Benthem 2007). Sticking to mathematical
proof is hard in some ways, but it is still an easy form of cumulation. But in real life,
we are proved wrong all the time, and our intelligence shows in how we deal with *that*.
I myself think that it might be very fruitful to critically compare, and perhaps merge
currently disjoint work on abduction, circumscription, belief revision, and dynamic
logics. And once on this road of actions of information-processing and self-
correction, the longer-term perspective would be one of forming and modifying hy-
potheses over time, as in formal *learning theory* (Kelly 1996). These matters are of joint
interest to logicians and philosophers of science, and they can be studied in dynamic
and temporal logics of information flow (van Benthem 2006).

But let me go one step further than individual action. When Kuhnian influences
entered the philosophy of science, a shudder could be felt, as real people started walk-
ing in the Platonic Museum of abstract scientific theories, with their shouts and steps
reverberating on the immaculate tiles. But nowadays, *multi-agent interaction* is coming to
be seen as crucial to intelligent behaviour (ESF 2007). Logicians think of proofs as
strategies for successful interaction, computer science is all about distributed systems
and interacting 'agents', linguists cast meanings as equilibria in coordinate games be-
tween speakers and hearers, and many philosophers find the test of 'true knowledge'
in being able to survive the grindstones of controversy. In line with this, logic, com-

puter science, economic game theory, and formal sociology are rapidly developing new interfaces today, including studies of strategies, belief change, and preference merge. I think the same would be wise for the methodology of science. Kuhn and even the social critics of traditional neo-positivism had a serious point, and indeed much of what they said makes sense in this new formal setting.

Finally, the realities. For testing their theories, formal philosophers of science and logicians often appeal to intuitions and carefully stream-lined historical facts. But nowadays, empirical *cognitive science* is knocking on the door of such theoretical research. If there are all these inference mechanisms like deduction and abduction, surely, we should find some trace of them in the workings of the human brain? The last few years have seen some exciting discoveries on this score, e.g., proofs by various authors that neural nets may be viewed as computing conclusions in non-monotonic default logics (Leitgeb, to appear). Thus, stale polemics between paradigms in cognitive science come to a fitting end. Indeed, some young logicians and philosophers of science at this interface see all this as leading to a possibly new life for the Vienna Circle Program, but now in a new guise, informed by new logics and new access to empirical facts. I find this broad-band revivalist program very appealing, and it would reflect my own view of logic today: not as an absolute guardian of truth and consistency, but as a way of dealing with the continuous surprises that intellectual life has in store for us, delivered either by Nature of by our interactions with other people. If you wish (and I do…): *logic is the immune system of the mind!*

REFERENCES

European Science Foundation (2007). *LogiCCC. Modeling Intelligent Interaction: Logic in the Humanities, Social and Computational Sciences*. Strassbourg: Eurocores Program Document.

Kelly, K. (1996). *The Logic of Reliable Enquiry*. Oxford: Oxford University Press.

Leitgeb, H. (ed.)(to appear). *Logic and the New Psychologism*, special issue of *Studia Logica*.

van Benthem, J. (1990). "Artificial Intelligence, a Continuation of Philosophy by Other Means", *Algemeen Nederlands Tijdschrift voor Wijsbegeerte* 82, 83-100.

——— (2000). "Reasoning in Reverse", Preface to *Abduction and Induction, their Relation and Integration*, P. Flach & A. Kakas (eds.). Dordrecht: Kluwer, ix-xi.

——— (2006). "Logic in Philosophy", in D. Jacquette (ed.), *Handbook of the Philosophy of Logic*. Amsterdam: Elsevier.

——— (2007). "Modal Logic Meets Situation Calculus", submitted to special issue of *Artificial Intelligence*.

van Ulsen, P. (2000). *Evert Willem Beth als Logicus*. Dissertation, Institute for Logic, Language and Computation, University of Amsterdam.

Johan van Benthem is a university professor of logic at the University of Amsterdam, and professor of philosophy at Stanford University. He was the founding director of the 'Institute for Logic, Language, and Computation' (ILLC). He has published in modal logic, temporal and spatial reasoning, logical structures in natural language, foundations of computation, and interfaces between logic and philosophy. His main current interest are dynamic logics of information flow, belief revision games, and intelligent interaction.

Address: Institute for Logic, Language and Computation (ILLC), University of Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, The Netherlands. E-mail: johan@science.uva.nl. Web: http://staff.science.uva.nl/~johan/.