WIDER STILL AND WIDER...
RESETTING THE BOUNDS OF LOGIC

an essay

Johan van Benthem

ILLC, University of Amsterdam
September 1997

Abstract

Modern logic is often defined in terms of specific formal languages, rules, and calculi. Such architectural decisions about a field form a pervasive implicit definition which determines professional practice – through the structure of textbooks, as well as the research agenda that determines 'interest', and hence acceptance and academic status. Such a practice may come to contain a lot of historical accident, or force of habit. Therefore, it seems worth thinking about the defining agenda of a field once in a while. In this brief essay, we explore alternative views of logic, locating the nature of the field in more abstract themes, concerns and attitudes. The new definition does not remove the need for the old agenda, but we advocate a shift in emphasis, toward greater generality and range of application. The outcome is a conception of logic as a broad methodological stance, looking for invariants in (information) structures and processes.

in A. Varzi, ed., 1999,
The European Review of Philosophy, vol. IV
CSLI Publications, Stanford, pp. 21–44.

1 The Agenda of Logic

The original attraction of logic for many people is the discovery that, in the fast stream of unreflected talking, thinking and planning that we all do, there are stable recurring patterns, which can be brought out and studied as such. This is the same Platonic insight of an unchanging true reality behind the chaotic world of appearances, that underlies all of science and philosophy. Another attraction nowadays to some people might be a more manipulative element. Using logical patterns, we can design machines, or teach humans, to perform these, as well as extrapolated skills, with higher speed and accuracy.

But what precisely is the subject matter of logic? Looking at most textbooks, the implicit answer seems to be that logic is a science of formal systems. This is the usual textbook sequence: propositional logic, predicate logic, modal logic, higher-order logic, etcetera. These systems include a formal language, a semantics, and a proof calculus, as one package. Of course, the professional picture is more diverse. Developing these systems, other dimensions of classification arise. For instance, one could also see logic as a bunch of recurrent techniques: compactness arguments, semantic tableaux, recursive functions. And finally, some of the highlights of the field are theorems, like Gödel's completeness and incompleteness theorems, sometimes formulated for one formal system, sometimes with a much wider scope. Even so, all this is all about formal tools, most of them recent. One might expect logic rather to have natural chapters like "negation", "recursion", and other major ubiquitous structures in reasoning – not tied to any specific formal system. But no professional consensus seems to exist on what such a division would be.

Does logic have a more specific subject matter, and should it? I see two kinds of answer. One is methodological, saying roughly that logic is about ubiquitous patterns in information representation and reasoning, which arise in any subject matter whatsoever. The other is more content-oriented, saying that logic is some kind of cognitive science, be it with a somewhat peculiar relationship to its empirical domain. (When people do not reason the logical way, we can always switch to a normative stance and say they *should*. Or nowadays, we can threaten to design an intelligent machine which *does* behave...) The literature on this issue is diverse. Amongst others, logic has been viewed as a science studying 'the laws of thought', the 'universal grammar' of language and meaning, the Platonic universe of all possible ontologies, the foundations of human cognition, or the abstract structure of information processing by humans or machines. All these answers lead to border crossings with neighbouring disciplines like mathematics, linguistics, computer science, or cognitive psychology. I cannot offer a definite criterion of demarcation here, valid for all times. I am inclined to say that logic is most properly viewed as a methodological stance which applies to most domains of rational enquiry. But if a more specific 'proper province' were requested, I would opt for logic as a foundational science of *information*, both its structures and its transformation processes. This makes it well-suited as an intermediary between the above-mentioned disciplines, which all represent aspects of an emerging 'information science'. Incidentally, a search for definitions is often a sign of an identity crisis, which spells ill for a field. By contrast, the temperament of this essay is optimistic. Logic in the 20th century has achieved much impressive and undeniable progress. But the art is to see *in which direction*.

Given this broader background, here is what I hope to achieve at the least. This essay concerns the architecture of logic. The agenda of our field is the outcome of historical choices that could have been, and can be, made differently. No science escapes from this flux. There is constant renegotiation of what is important, and reinterpretation of the value of major contributions. My aim is to shake up some 'received views' on the business of logic, proposing some constructive alternative views and challenges.

2 Accidents of Formulation

2.1 Language and System Imprisonment

Classical logic has been immensely successful. But this very success has enshrined certain formats and procedures, that also have drawbacks. For instance, many themes suffer from what may be called 'system imprisonment'. We have to discuss the behaviour of negation inside specific formal systems, such as propositional or predicate logic – even though these systems do not correspond to meaningful distinctions in the 'open space' of actual reasoning. Many inference patterns formulated in predicate logic are completely general, and have nothing to do with tool-related issues like 'first-order' versus 'higher-order', or axiomatizable versus non-axiomatizable. A good example is the ubiquitous inference rule called Monotonicity: "positive occurrence sanctions upward predicate replacement", which has no preferred linguistic home (van Benthem 1987). Note the tentative formulation. We even seem to lack a suitable vocabulary for formulating these principles in their true generality. Thus, users of logic must buy into specific formal languages, which spell out much more than may be needed for their needs. Contrast this with mathematics, where a user of the calculus need not buy a fully specified formal language and proof system at all. Thus, we have a problem of achieving the appropriate generality. Another drawback of the formal systems approach is that we are forced to spell out complete formal languages and deductive calculi to their full mathematical extent, including convoluted forms of statement and argument which have no counterparts in natural reasoning. This may generate theoretical issues that are non-starters from the point of view of capturing essences of reasoning. [1]

One might think that system imprisonment is really a severe case of a more general 'language imprisonment'. Logicians always introduce 'formal languages', a habit which surprises (and vexes) many of its potential users. Why bother with the details of syntax, when we are after contents? But this is a mode of proceeding which I think is inevitable – provided that we see it in its proper light. It is not important that we are dealing with 'languages' in any received sense of that word, let alone any specific formal notation. But what is essential is that logic must bring out the interplay between cognitive content and its *representations* in order to get a handle on the relevant phenomena.

2.2 From Proper Names to Common Nouns: Getting at the Logical Phenomena

Lifting specific results to broader contexts occurs all the time in logic. Consider Gödel's Completeness Theorem: originally a specific statement about first-order predicate logic, now a type of theorem. Proper names become common nouns. In logic, losing one's capital is an honour. This is standard generalisation, found in any science. But one can take this trend much further, and seek general phenomena behind even standard logical items. For instance, 'logical constants' are usually taken to be members of a conventional list of Boolean connectives, first-order quantifiers, and perhaps the identity predicate. But the real issue is the general *phenomenon of logicality*, and the degree to which it can occur in arbitrary expressions. Why is the business of logic with just these distinguished items, rather than with 'logical behaviour' – broadly understood – which may be found to some degree in all linguistic categories? For so-called 'generalised quantifiers', this point is becoming acknowledged nowadays – but it holds equally well for prepositions ("in", "with", "to"), or discourse particles ("if", "but", "although"). The same consideration applies to the other traditional focus of the field, namely the 'laws of logic'. Defining features of the field should not be sought in specific rules like Modus Ponens, or other boundary stones of some privileged domain of logical reasoning. In particular, no fixed set of core laws demarcates 'logical' from non-logical reasoning. It is the spirit, not the letter of these laws that defines the field. What logic should be about is rather capturing

---

[1] 'Full specification' does seem necessary for doing natural language processing, or programming of other kinds – but it is not the way in which people experience and structure reasoning. And even computationally, it may be an overly costly architecture, closing our eyes to alternatives. Language understanding and real-time reasoning might consist in massive parallel repetition of simple patterns, rather than elaborate recursive hierarchies.

the *phenomenon of inference* in general, as it occurs in the various connections (deductive, inductive, etcetera) that humans use to drive the flow of information.

2.3 The Architecture of Diversity

Existing logical systems can hide presuppositions, which we do not see just because they are so successful. First-order predicate logic is a good example. It has done amazingly well as a focus for making modern logic more rigorous, and as a 'laboratory' where the most important metatheorems of the field were developed. But as a paradigm for reasoning, it also has some features which one might question. In particular, stepping back a little, its very uniformity seems suspect. It would really be most surprising if one single linguistic formula representation could efficiently serve the purposes of such different activities as interpretation and deduction. Cognitive psychology suggests that memory and performance of humans over time require an interplay of different successive representations for the same information. Therefore, no such uniformity need be assumed for actual cognition, and a much more complex modular architecture might be needed, where semantics and deduction do not work on the same representations, and where information is passed between modules. And the same moral applies elsewhere. With information processing by machines, modularity is at the heart of performance. And to mention one further source, issues of architecture become unavoidable when we ponder the logical meso- and macro-structures of discourse and argumentation.

2.4 From Products to Activities

The previous concerns still follow traditional logic in the following important respect, which again hides a prevalent presupposition concerning the focus of logical research. They show an emphasis on static *products* of logical activities, such as statements or proofs, instead of those activities *themselves.* But interestingly, in natural language, the words "statement" and "proof" are both ambiguous between a reading as a product and as an activity. And rightly so. Indeed, in much current literature there is a 'Dynamic Turn' putting (both physical and cognitive) activities at centre stage as a primary target of research. This move is highly relevant to our conception of the logical agenda. Reasoning is a logical activity, and its process structure is a legitimate research topic. Moreover, so are related intelligent activities like understanding the information in what is said, formulating intelligent questions, more general many-person communication, or playing other cognitive games. What we want to convey to students are the logical patterns that drive these skills. Thus, one alternative way of thinking about the architecture of the discipline is in terms of the cognitive skills that it contains. Its impact lies in this 'know-how', rather than 'know-that': the capital of accumulated truths.

3 The Major Parts of Logic

In order to introduce some alternative viewpoints, we need a brief sketch of established achievements of modern logic. These have been recorded in a number of well-known Handbooks, including volumes on Mathematical Logic (Barwise 1977), Philosophical Logic (Gabbay & Guenthner 1983–1989), Logic in Computer Science (Abramsky, Gabbay & Maibaum 1992–199*), Logic and Artificial Intelligence (Gabbay, Hogger & Robinson 1991–1995), and Logic and Language (van Benthem & Ter Meulen 1997). Our business here is not at all to belittle these accumulated achievements, but rather to change our perspective on what has been, and still can be, achieved.

Expressive Power Logic studies formal languages, as well as types of ontology for them, and 'logical semantics' studies connections between the two realms, dealing with expressive power of languages. This dual aspect makes logic so strange to working scientists – who tend to consider linguistic formulation as a necessary, but trivial evil. The logical stance reveals subtleties of interaction between linguistic form and structural behaviour, which one simply would not be aware of otherwise. Thus, logic has become a paradigm of meaning in the semantics of natural languages in linguistics, and that of computer languages in computer science. The 'strong arm' of logical semantics is *model theory*, which also has purely mathematical applications. One of the most striking benefits of this dual perspective is the creation of *new* unintended models for existing deductive practices. Examples are the infinitesimals used in non-standard analysis.

Deductive Power Logic studies proof calculi for various species of valid reasoning and their combinatorial properties, as well as the proof-structure of deductively organised theories. *Proof theory* started as a typical tool in the foundations of mathematics, studying proof structures and their effective transformations – but its patterns also occur in computer science, artificial intelligence, and also linguistics and even law. It is this aspect of logic which has most permeated general culture, as an ideal of precision and intellectual organisation. Despite the connotations of the term "proof", logic is not necessarily on the side of authority here. Proof need not always mean knock-down argumentation, ending in cognitive surrender of one's opponent. It is also a model for setting forth one's reasons and assumptions, putting them up for public scrutiny.

Computational Power Logic has provided the first, and still most successful general analysis of computation, and its associated notion of complexity. Moreover, it contains precise theorems about the limits of effective computability, when Gödel discovered essentially 'undecidable' problems. On the positive side, computer science has realised old ideals of 'mechanising' logical tasks in automated deduction and logic programming. The scope of computational concerns covers both deduction and semantics. One can provide algorithms for valid inferences, but also for semantic queries. Finally, computational models have again led to new interpretations for old phenomena. An example is the treatment of randomness and learning in Kolmogorov complexity.

With many of these themes, boundaries are fuzzy with surrounding disciplines. E.g., some exciting recent advances in the study of reasoning go under the heading of Artificial Intelligence. This short enumeration shows that modern logic has developed a large body of insights and techniques in a relatively short historical period. These results have set patterns of rigour and coverage that have changed the traditional field beyond recognition, and irreversibly. Every subsequent critical turn in the field must eventually live up to these quality standards. Our coming discussion therefore falls under the above three headings. In each case, we discuss one trend within current research, and one which seems to require a more drastic shift.

4 Expressive Power

4.1 Logicality and Invariance

There is an asymmetry in logical theory, in that most results take some vocabulary for granted, and then develop various deductive and other systems on this basis. Most major results are about semantic consequence and computability given some language. The question how we chose its vocabulary in the first place does not arise. Modern logic has much more to say on derivability than on definability. And yet, we do have intuitions concerning the choice of logical operators. These form the 'glue' of discourse, which does not itself carry any specific information about individuals or issues. Logical notions are 'topic-neutral', as is sometimes said. One well-known way of making this idea precise is through *semantic invariances*. Consider any permutation of a domain of individuals: that is, a function shaking up their identity, leaving only their 'patterns'. Any such permutation induces permuted forms of sets of individuals, and likewise all the way up to objects in higher types. Now call a logical operator O of any arity and type *permutation-invariant* if it commutes with such permutations:

$$\pi\,[O\,(X, Y, ...)] \quad = \quad O\,(\pi\,[X], \pi\,[Y], ...)$$

A useful reformulation recasts this identity as invariance under isomorphisms:

$$(x, ...) \in O(R, S, ...) \quad \text{iff} \quad (\pi(x), ...) \in O(\pi\,[R], \pi\,[S], ...)$$

Permutation invariance has been proposed as a general criterion for logicality by many authors independently (including Tarski 1986), gradually covering the algebra of sets and relations, generalised quantifiers, and eventually the whole range of possible categorial operators (van Benthem 1989A). In general, any operation defined in a logical language that does not refer to specific objects will be permutation-invariant. This is actually a very old perspective. Dual view-points on definability and invariance under suitable transformations date back far into the 19th century. For instance, in the ground-breaking work by Helmholtz and Heymans on perception, invariants of movement provided natural primitives for geometrical theories. Klein made this viewpoint the foundation of his Erlanger Program for the set-up of any mathematical theory. Similar ideas have been proposed by Weyl in this century for analysing the primitives of physical space-time (classical and relativistic), and they have also surfaced in so-called 'ecological psychology'.

Semantic invariances have two virtues. First, they provide a *general* formulation of logicality, which applies to operators in different syntactic categories. Thus, we see how Boolean operations really are logical in the same way as quantifiers, or less obviously 'logical' items, such as the reflexivizer "self". Second, semantic invariance naturally admits of *gradations*. This makes pure logicality one extreme on a spectrum. The standard mathematical examples do not consider all permutations of their base individuals, but only special transformation groups preserving the relevant structure (geometric, topological, ...). Thus, logicality is not an all-or-nothing matter – since arbitrary expressions can be classified for the amount of invariance that they support. (Logic is very robust, mathematics a bit less, while ordinary expressions will only show invariance under transformations respecting a lot of further structure.) Interestingly, similar views have been proposed for linguistic expressions. For instance, prepositions seem to lie in between logical operations with a completely fixed meaning, and lexical items whose interpretation is completely free, like adjectives or verbs. They are somewhat constrained, by being items that are stable under those shifts in orientation that tend to occur in our ordinary spatial movements.

Do semantic invariances 'justify' types of expression, or do types of expression generate the appropriate structure-preserving transformations? Our analysis does not say, and indeed, one can look in either direction. Given the vocabulary, there are corresponding automorphisms, and one can show that further defined expressions will be invariant. One can also ask, conversely, for definability of invariant items (this question dates back to Weyl, who thought it was unsolvable). There are many 'functional completeness theorems' classifying logical invariants of various kinds (van Benthem 1986A, 1991B).

An interesting recent interpretation of the invariance approach is more computational (cf. Section 6.1 below). One can think of logical constants as defining simple evaluation processes (conjunction composes subroutines, disjunctions prompt a choice, etcetera). But then logicality is no longer an absolute notion. It depends on one's choice of process. A semantic equivalence gives us a notion of 'simulation' between instances of the same procedure, running on different models. Invariance expresses the latter fact. Thus, 'logical constants' reflect an implicit procedural notion of semantic evaluation.

4.2 Generality and Linguistic Form

Our second example merely elaborates a limitation that was already pointed out earlier. Sensitivity to syntax is a hall-mark of logical analysis. In particular, the fine-structure of formal languages contains a lot of semantic and computational information. For instance, when comparing truth of statements across different (or changing) situations, one may observe that certain statements remain true whenever the universe of discourse contracts. The Los-Tarski Theorem then tells us that (in first-order logic), these are precisely the statements that can be defined employing only *universal* quantifiers over objects. Likewise, we mentioned the inference rule of monotonicity (the 'Dictum de Omni' of traditional logic) whose range of applicability depends on *positive* occurrence of its key predicate. As an algorithmic example, in proof search, one finds that some statements have a natural procedural interpretation as simple search instructions – and identify *Horn clauses* as a vehicle for efficient logic programming. Nevertheless, there is also a problem here. Many of these observations seem to have a much wider scope than any given formal language, and they crop up all the time. Then, what is their true generality? Traditional logicians wrestled with the Dictum de Omni, and never managed to find a satisfactory general formulation (Sanchez 1991). As a result, we cannot even say what the true range of the classical syllogistic was. But a similar problem afflicts its modern successors... Thus, attention to linguistic form is both a very typical, and powerful source of logical insights, and a not always desirable constraint on their formulation.

5 Deductive Power

Deduction is connected to meaning, but it is also sui generis. Before broaching more specific issues illustrating this point, let us illustrate the 'dangers of success'. Consider the standard completeness theorem for predicate logic, one of the central results of the field. It is taken to say that formal deduction in Frege's or Hilbert's style indeed captures valid logical consequence. The latter semantic notion (due to Bolzano and Tarski) says that, whenever the premises of an inference are true in a model, so is its conclusion. Thus, proving completeness theorems has become something of an industry. Now, this result and this interpretation hide as many questions as they solve. For one thing, *why* should deductive systems be complete in this sense? (As Kreisel once observed, soundness says that the proof system will tell us nothing new, and completeness that the semantics will tell us nothing new...) And even if they are, completeness is only an 'extensional' statement, saying that the sequents "premises-conclusion" forming the output by the deductive machinery are precisely those that satisfy some abstract semantic criterion for validity. But the most important features of a proof system are clearly its 'intensional' ones, having to do with its account of inference, and with the deductive organisation of information which it supports or suggests. A completeness theorem tells us nothing about these intrinsic deductive virtues. Finally, more insidiously, the completeness theorem suggests various things ex silentio. One is that there is one unique notion of logical validity that forms the intuitive norm – something which is not obvious. The other suggestion, still more hidden, is that the comparison between the semantic and the deductive approach to logical validity is the crucial one, and that these two perspectives are the only players. But historically, neither of these two may be the original intuition of logical validity – which seems closer to having *winning strategies* in debate and discussion. The latter, more pragmatic account can be stated in more computational game-theoretical terms, which are close to actual argumentation. Thus, the agenda of standard textbooks (which tend to ignore or marginalise these other viewpoints) can be highly misleading – which again engenders a certain poverty in philosophical studies of logical consequence that do not question their presuppositions.

5.1 Styles of Reasoning

Is there just one notion of logical validity, to which all deduction must aspire? In the history of logic, other views have existed. Notably, 'Bolzano's Program' in the early 19th century rather viewed logic as the science of styles of reasoning. Bolzano 1837 observed that human reasoning comes in varieties, with stricter or looser criteria for validity, ranging from domestic deliberation, through scientific reasoning, to what he considered the acme of thought: philosophical strict entailment. Bolzano's agenda for logic requires the identification of major styles of reasoning (deductive, inductive, and others), and the determination of their logical properties. (Similar broad agendas are still found with C.S. Peirce, towards the end of the 19th century.) Interestingly, logical constants play less of a central role here. On Bolzano's view, the basic distinction in inference is between those parts of speech whose meaning we keep constant (these will form the 'logical skeleton') and those whose denotation is variable. These distinctions can be made in different ways, depending on context, and hence the main emphasis is on proof rules that do not involve the behaviour of special expressions like "not", "and" or "all". Rules of the latter kind are nowadays called *structural rules* of inference. And indeed, in logic and AI, various packages of structural rules have been used to identify major types of inference (van Benthem 1989B, Makinson 1994). Classical reasoning is characterised by structural rules that allow us to treat the premises as a 'sufficient set'. We can permute them, copy them, contract copies, and add things, without losing a conclusion once established. By contrast, for instance, default logics are non-monotonic. If we allow conclusions from ignorance, then further premises may invalidate the grounds on which an earlier conclusion rested. Thus, practical reasoning is deeply tied up with withdrawing conclusions, and *revision* of our theories. Yet another example is probabilistic reasoning, which need not be transitive: a chain of probable inferences may lose force along the way, and yield implausible conclusions. Finally, dynamic inference pays close attention to the sequential presentation of premises, and hence its conclusions may even be sensitive to permutations or contractions of premises.

Although this perspective on inference is not generally accepted, it falls within standard logic in the following sense. Varieties of non-standard inference can be developed using standard techniques, including formal semantics, representation and completeness theorems. Indeed, many combinatorial techniques of classical proof theory have been found to apply to logics with structural bases different from the classical one (cf. Dosen & Schröder-Heister 1994). And yet, this abstract idea of 'inference' seems alien to many logicians. Part of the reason is again the earlier-mentioned emphasis on products. Alternative inference systems often violate classical 'laws of logic', and this seems like a deep loss if the latter are the heart of logic. A better view of this plurality is that these systems are not competitors in any sense. They just describe various 'connections' between statements that humans find it useful to observe and develop. Also, the 'yield' of these systems in terms of their valid sequents is in some sense a secondary output, which should not be at the focus of evaluation. What matters are the underlying ideas, such as the view of default logic where *'ignorance'* supplies conclusions, or the qualitative form of probabilistic reasoning where one considers only those models of the premises that best fit with one's *preferences.*

There is more to be said here. Radicalism usually only goes so far. Proponents of alternative reasoning styles have often stuck to standard methods for developing their systems, either model-theoretically or proof-theoretically. But of course, some other format might be much more appropriate, say, game-theoretic or yet different. Also, in line with our preceding section, these authors have usually taken the classical logical *vocabulary* for granted. But of course, expressive power and deductive power are not independent. If one sets up an alternative logic, then the question should come up what is its most appropriate logical vocabulary. Exceptions to this neglect are relevant logic and linear logic, where new proof systems have come with new views of logical constants. And indeed, Bolzano himself stated structural rules that involve explicit changes in the boundary between 'fixed' and 'variable' vocabulary.

5.2 Modular Architecture

Classical proof systems are uniform for one particular formal language, handling all its potential inferences. They do not contain further information about architecture. Now in natural language, this view is not quite plausible. For instance, many inferences seem to be shallow, involving only a few operators at a time. And more importantly, inferences seem to come in modules, small packages of rules for a special purpose (van Benthem 1986B). The clusters of this 'natural logic' need not fit very well with the distinctions introduced in, say, first-order logic. For instance, nothing in natural reasoning corresponds to the standard emphasis on prenex forms. Examples of natural clusters are monotonicity reasoning, conservativity reasoning, and simple algebraic rules (by no means all of propositional logic). Again, the 'phenomena' of reasoning do not fit the mathematics of formal syntax very well, and the latter may not be the most appropriate medium for bringing them out. What we need is a better account of logical architecture, which tells us what modules are, perhaps even based on different logics, and how they can efficiently interact and pass information.

The need for 'logical architecture' becomes even clearer with other recent developments. For instance, current dynamic logics of reasoning suggest a natural distinction between short-term presentation-dependent and long-term memory-oriented inference, where the latter works on more abstract representations. Again, the architecture of such more complex reasoning systems is at present beyond the scope of logical theory. Putting this more generally, deductive logic so far has little to say about the meso- and macro-levels of reasoning, which is where most of our more strategic thinking takes place. This is not to say that no formal approach could work here. Philosophers of science, rather than logicians, have had many interesting things to say on these higher aggregation levels of both reasoning and theories, witness the influential works of Nagel, Sneed, or Lakatos. Likewise, the best formal accounts of modularity in the representation of information come from the theory of abstract data types in computer science, rather than logic itself. Nevertheless, 'combination of logics' is becoming an acknowledged research area, facilitated by recent more flexible formats of 'labeled deductive systems' (Gabbay 1992). Also relevant here are various newer approaches to 'logical systems', collected in the anthology Gabbay 1994. The latter also illuminates other themes in this essay.

6 Computing Power

The theory of computability is a field in its own right, with concerns of its own that fall outside the scope of this discussion. We raise some issues here concerning interpretation and inference, viewed as cognitive activities that involve some sort of computation.

6.1 Cognitive Dynamics

As we observed before, classical logic is oriented towards the products of activities (mostly cognitive, but also more general ones), rather than the structure of those activities themselves. Logic is about reasoning, confirming, refuting, denying, etcetera. But it mainly talks about reasons, proofs, propositions, and other static traces of these activities. But contemporary dynamic logics are trying to put the latter on the map, thereby making the intuitions explicit that underlie much of classical theory. [2] In this way, logic becomes more of a general theory of information structures plus the processes that modify these. Prominent examples of this trend are the logics of update, contraction and revision in the theory of belief revision (cf. Gärdenfors & Rott 1995), as well as various logics of 'dynamic semantics' for natural and formal reasoning that are currently being developed in analogy with the semantics of imperative programming languages (cf. Muskens, van Benthem & Visser 1997). This 'Dynamic Turn' makes traditional logical languages more like cognitive programming languages whose programs are instructions for modifying the information states of parties in communication.

Again, this is a broader research program for logic, pursued by classical means. Dynamic logics employ mathematical models of information states (both for factual information, and for epistemic information concerning other people's knowledge and ignorance), and their logical constants are much like program constructors of well-known kinds, familiar from relational algebra and process algebra. This process view of logical constants fits very well with our invariance analysis of Section 4.1 (cf. van Benthem 1996). Of course, what are the appropriate logical constants will depend on one's view of the relevant cognitive processes. These need not be the same for all purposes: evaluation may indeed require different information states, and atomic moves transforming these, from the structures manipulated in inference. Here is a concrete example. Think of propositions dynamically as transition relations on information states (standard propositional updates will even be deterministic functions, but propositional revisions may be indeterministic).

---

[2] For readers which have undergone the dynamic Gestalt Switch, implicit dynamic features are ubiquitous in even the most traditional textbook presentations of logic, clamouring for proper attention. For case studies in first-order and modal logics, cf. the chapter on 'dynamification' in van Benthem 1996.

In this move from Boolean set algebra to relational algebra, logical operators come to define algebraic combinations of binary relations. For instance, here are natural dynamic counterparts of the three classical Boolean operations • (and), + (or), − (not) , resp. (for their motivation, cf. Groenendijk & Stokhof 1991, van Benthem 1996):

composition	$R \circ S$ =	$\{ (x, y) \mid \text{for some } x : Rxz \text{ and } Szy \}$
choice	$R \cup S$ =	$\{ (x, y) \mid Rxy \text{ or } Sxy \}$
impossibility	$\sim(R)$ =	$\{ (x, x) \mid \text{there is no } y \text{ with } Rxy \}$

One way of defining logical process operators employs the *permutation invariance* of Section 4.1. Notice that permutations of individual objects lift naturally to permutations of binary relations, taking them to relations having essentially the same abstract 'arrow patterns'. Now it is easy to show that the preceding relational operations of composition, union and impossibility are invariant for permutations of their argument relations, whereas an operation referring to (say) the presence of a specific individual ('king Louis Napoleon I') or some specific relation ('being the founder of') would not be invariant. But the range of permutation-invariant relational operators is vast, and does not match known computational repertoires. So, we need a better approximation of processes. Here is one way of achieving that, which suggests a generalization of invariance to 'safety'.

In many theories of computation, a specific notion of process arises by first choosing some representation of states and transitions, and then imposing an equivalence relation. One common choice are *labeled transition systems* ('process graphs', 'Kripke models') consisting of states with labeled arrows, identified under so-called *bisimulation* (cf. van Benthem 1996, Barwise & Moss 1997). Bisimulations are back-and-forth relations between states in two models which allow for running the same process with similar local states and similar available choices at each state. [3] Now, given this process model, we may require that natural operations $O(R, S, ...)$ on binary relations 'stay inside it'. That is, if we have a bisimulation between two models that have atomic actions $R, S, ...$

---

[3] More precisely, a bisimulation between two rooted Kripke models $\mathbf{M}, \mathbf{N}$ is a binary relation Z connecting states in $\mathbf{M}$ to states in $\mathbf{N}$ subject to the following conditions. (1) ('Same Start') Z connects the two roots. (2) ('Local Harmony') Z only connects states which satisfy the same atomic propositions. (3) ('Back-and-Forth') For every atomic relation R , if $s Z t$ and $s R u$ in $\mathbf{M}$, then there exists some v in $\mathbf{N}$ such that $t R v$ and $u Z v$. And vice versa. This notion is well-known for matching the standard *modal* language: the same propositional modal formulas are true at $\mathbf{M}, x$ and $\mathbf{N}, y$ when $x Z y$. Various converses are discussed in van Benthem 1996. Safety for bisimulation may be viewed as a natural extension of this invariance for Modal Logic to program operations in so-called Dynamic Logic.

for the argument relations, then this *same* bisimulation should still bisimulate with respect to the newly defined relation O(R, S, ...) . This new requirement is known as *safety for bisimulation*. (But safety also makes sense for other processs equivalences.) It is easy to see that safety for bisimulation has permutation invariance as a special case. But it is much stronger. Van Benthem 1996 (chapter 5) shows that the only first-order definable relational operators which are safe for bisimulation are precisely those definable using the above dynamic propositional operators $\mathbf{o}$, $\cup$, $\sim$. This is a functional completeness result for dynamified propositional logic, characterizing its operators in semantic terms. But it also demonstrates a more general analysis of dynamic 'logicality'.

Here as elsewhere, the outcome of the new program is not in conflict with the tenets of classical logic. There rather seems to be a 'Correspondence Principle' at work, guiding the design of the newer systems. Passing to a suitable limit (disregarding sequential phenomena), the latter should 'reduce' to the standard ones, in some reasonable sense. The same principle seems to work, incidentally, for varieties of inference. So far, all non-standard logics reduce to classical logic when one makes appropriate (consistent) additional assumptions. This observation explains a perhaps surprising empirical fact. Most alternative logics are weaker than classical logic, but they never *contradict* it.

6.2 Church's Thesis Revisited

It has sometimes been claimed that the basic notion of computation has been found by Turing in the 30s, bolstered by a bunch of equivalences with alternative frameworks. This then led to *Church's Thesis* stating that all reasonable models of computation amount to the same thing, namely the (general) recursive functions. But in the light of the preceding discussion, things are less clear-cut. Church's thesis is 'extensional', in that it talks about the functions (the input-output graphs) of quite diverse algorithms and programming styles. The latter may differ considerably in their fine-structure. For instance, Turing machine programming is about the least perspicuous style of defining algorithms that has ever been invented. At a more intensional level, it is still unclear what a canonical notion of 'algorithm' would be – so that the theory of computation still has its fundamental open questions. A strong version of Church's Thesis would give a uniform model for algorithmic structure. And what the latter would involve, presumably, is an account of natural 'logical' or cognitive programming structure. In particular, what are the natural logical constants in the above dynamic setting? It seems safe to say that we do not know yet, even though a number of approaches exists.

Indeed, despite the success of Turing's model in AI as a paradigm for cognition, it is not even clear that machine models are the most appropriate mathematization of cognitive activities. One alternative which is gaining ground in the logical community is logical *game theory*. Cognition involves moves in a social setting, guided by higher-level strategies that we can follow to achieve our goals. Logical games exist for argumentation (Lorenzen dialogues), interpretation (Hintikka game semantics), model comparison (Ehrenfeucht-Fraïssé back-and-forth games), and for many other purposes. [4] Moreover, games provide genuinely novel perspectives on logical constants (e.g., negation becomes a 'role switch', and disjunction a 'choice') as well as valid inference (which becomes the existence of a winning strategy for the proponent of the conclusion against an opponent who has granted the premises). Games also seem to lend themselves more easily to analyses of the above-mentioned cognitive group behaviour, as well as the importation of non-deductive probabilistic considerations. But the jury is still out on what should be the fundamental game model for logical activities, and thus, this intriguing alternative approach still awaits its Turing.

7 General Laws

The view of logic developed in the preceding sections takes existing logical notions to much broader formulations. But I feel that there is still something missing in our general conception of the field, which may be brought out through an analogy with the sciences. In physics, in addition to laws for specific domains of phenomena, or facts concerning particular applications, certain broad principles regulate our thinking about the world – which embody some of the most fundamental insights of the field. Examples are the principle of the Least Effort or Shortest Path, or the famous Conservation Laws of Energy or Momentum. Does logic have similar broad principles, that reveal something essential about information or cognition? Thinking about the issues in this essay so far, the following points come to mind, which I put up here, diffidently, for discussion only. I do not pretend to have any definite answers, but I do think we should try harder to crystallize some combined wisdom of the discipline at this general level.

Balance of Language and Ontology   Logical systems should have 'the right' expressive power vis-a-vis their intended semantics. This balance can be measured by the existence of a characteristic semantic equivalence for the language, telling us when two models are indistinguishable. (Examples are back-and-forth relations like potential isomorphisms in first-order logic, or bisimulations in modal logic.) Some others have suggested that a measure for this balance is the existence of an Interpolation Theorem for the language.

---

[4] A classical reference is Hintikka 1973. For a more recent survey and discussion of logical games, cf. van Benthem 1988, 1993.

<u>Compositionality</u> This was the only constraint that Montague brought to light in his 'Universal Semantics' unifying the semantics of natural and formal languages. Dummett has widely defended it as Frege's key insight that distinguishes modern from traditional logic. Logical syntax is given by recursive constructions, and our business is not just the lexical semantics of logical constants, but also the compositional semantics of these constructions. There is a law here: all good semantics can be made compositional, so that syntactic structure harmonises with semantic process structure.

<u>Conservation of Complexity</u> The following reflects a pervasive experience in logical research. What one gains in expressive power, one loses in complexity of inference. In other words, the balance of expressive and algorithmic complexity seems constant. What is the appropriate measure of complexity for this conservation law? It might involve some suitably abstract notion of 'information' – which we lack at present. [5]

The principles so far mainly concern the design of logical systems by themselves. But there are also general experiences concerning connections between different systems.

<u>Translation Thesis</u> Church's Thesis tells us that all reasonable models of computation can be simulated on Turing machines. Many people feel that something similar holds for expressive power. Many major logics are equivalent under effective translation, provided that we take the latter term in a suitably liberal sense. For instance, intuitionistic logic is not faithfully embeddable into classical logic when we keep its language fixed – but it is embeddable when we transcribe its Kripke semantics with explicit quantification over information states. (This equivalence might even follow from Church's Thesis plus Conservation of Complexity.) So, is first-order logic a universal expressive medium?

<u>Duality of Viewpoints</u> Related, though somewhat different in thrust, are certain broad dualities. In mathematics, there is a constant historical interplay between algebraic and geometrical viewpoints. Both seem to correspond to basic intuitions that we have, and there is constant re-encoding of one stance into the other. Likewise, logic has natural recurring stances. What can be said model-theoretically ('geometrically') has natural

---

[5] Indeed, we even lack a good bridge between well-developed more algorithmic notions of information (as found in Shannon's information theory; Ash 1965, or in Kolmogorov complexity; Ming Li & Vitanyi 1993) and the more qualitative notions of information that guide logical semantics (Barwise & Perry 1983, Barwise & Seligman 1996, Veltman 1996).

counterparts proof-theoretically ('algebraically'), and vice versa. Another instance of the same duality might be the above-mentioned relation (once articulated) between 'algorithmic' and 'semantic' information. And perhaps, when the time comes, this duality extends to a triangle affair with the game-theoretic stance.

Once again, these are musings, not definite outcomes. The more ambitious goal behind such principles, just as in the sciences, would be to find logical theories that *explain*, rather than merely *describe* the workings of information and cognition.

8 Summary

The main purpose of this paper has been to present a plea for a broader agenda of logic – partly inspired by the future, but partly also by the past of the field, before Frege gave the steering wheel a momentous push toward the foundations of mathematics. One part of our effort was an attempt to place known notions and results in proper generality. Varieties of 'logical constants' and 'deductive styles' were key examples. Another part was an openness to altogether new topics, such as the 'logical architecture' of complex reasoning systems, or most conspicuously, the inclusion of dynamic 'process structure' as a logical concern on a par with the traditional business of representational structure. These examples were driven by a grand purpose for the discipline. What best advances logic as a general science of information structures and processes, both in human cognition and artificial settings? This broad concern is not a threat to established logic. It rather adds to its impact, and refines some of its concerns (witness our discussion of Church's Thesis). Moreover, our broader canvas raises some, perhaps vague, but at least inspirational themes, like the above quest for fundamental laws of logic, far beyond specific rules of inference, or specific formal systems. These moves may have an intellectual interest. But our boundary discussions also hold a practical interest, at least for those logicians who are concerned with the academic position of their field.

9 APPENDIX Actual Reasoning

A common reaction to deviations from the classical logical agenda classifies them as being concerned with 'applications'. [6] Application has not been the focus of our essay – and our more general view of logic does not automatically make it more 'applicable'. Indeed, there are many misconceptions about what it means to apply logic, or other

---

[6] This reflex makes quite theoretical fields like deontic logic 'applied logic in philosophy', or generalized quantifier theory 'applied logic in linguistics'.

methodological disciplines. These misconceptions form a rich subject which deserves a paper by itself. Here, we conclude by pointing out a number of ways in which actual reasoning differs from current logical systems. (This is why papers from non-logicians are often so much richer than ours when they dig into some specific reasoning task.)

9.1 Real-Time Performance

Complexity awareness is a potential new item on the logical agenda. The known worst-case complexity of current logical systems does not come even close to the speed of human real-time performance on cognitive tasks. Here is one possible explanation for this mismatch. Humans may be fast, but not very sound or complete reasoners. Their procedures do not guarantee absolute certainty, but they are very good at revision when things go wrong. In this sense, our intelligence has to do, not with steady accumulation of eternal truths, but with quick rational response to challenges. Rationality is repair... This is a diagnosis outside of standard logical theory. Alternatively, one can also 'internalize' the challenge of real-time performance, and redesign logical systems to achieve this effect. For instance, the undecidable systems of relational algebra and first-order logic have been remodelled in so-called Arrow Logic and Modal First-Order Logic (Venema 1996, van Benthem 1994, 1996). This program can be extended, towards lowering complexity of decidable (but still complex) to realistically feasible logics.

9.2 Heterogeneous Information

Information is certainly not all symbolic. Graphic information is important, too – and indeed about every physical sign can be a carrier of information that can be processed logically. Perhaps, one good thing about the (sometimes deplored) abstraction of logical systems is that they can handle this generality. Even so, good paradigms for integrating information from these various sources are as yet unavailable.

9.3 Packaging: Representation and Computation

Another take on human real-time performance might be this. It is the package which counts. Humans exploit advantageous representations that help keep complexity down, and try to perform minimal computation over these. This might reflect cognitive principles of 'least effort', in line with the general principles of the preceding section, that explain something about human intelligent behaviour.

9.4 Architecture

This issue was raised before. Actual cognitive strategies seem ad hoc, calling many reasoning modules, and looking up information from all available sources. (For instance, even working logicians will often use a hodge-podge of semantic and proof-theoretic reasoning.) From this perspective, most logical systems are puristic: too clean and uniform. What is the architecture of opportunism, that allows us to get by so fast?

9.5 Mixing Deduction with Observation

No purism holds also for our information gathering. We deduce when we must, but we *observe*, or merely *ask* when we can, and this is shorter. Even scientists often prefer the outcome of an experiment to the result of laborious reasoning. (The happy mixture of a priori axioms and a posteriori experiments in Newton's work is a good example.) Conclusions are like queries, and we use whatever the situation offers.

9.6 The Genesis of Vocabulary

In practice, the formation of good vocabulary is essential to successful cognition. Some of this is local to discussion, or problem solving. Some of it is so generally useful that it gets encoded into our natural language. Many expressions then turn out to have a functional inferential role. For instance, it pays to reflect on the uses of a dispositional expression like "friendly, which encapsulates a whole default rule of inference. Moreover, again a conservation law seems at work here. We create new definitions to keep the complexity of interpreting our discourse, and reasoning, more or less constant.

Should the complexities noted in this Appendix become part of logical theory itself? There is a general danger lurking here, namely, that the true scientific account of anything in the world will turn out as complex as the reality it is intended to describe. Our preliminary answer is therefore: we do not know...

10 References

S. Abramsky, D. Gabbay & T. Maibaum, eds., 1991–199*, *Handboook of Logic in Computer Science*, Oxford University Press, Oxford.

R. Ash, 1965, *Information Theory*, Interscience Publishers, New York.

J. Barwise, ed., 1977, *Handbook of Mathematical Logic*, North-Holland, Amsterdam.

J. Barwise & J. Etchemendy, 1995, 'Heterogeneous Logic', in J. Glasgow, N. Narayanan & B. Chandrasekaran, eds, *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, AAAI Press (Menlo Park) and MIT Press (Cambridge (Mass.)), 209-232..

J. Barwise & L. Moss, 1996, *Vicious Circles. On the Mathematics of Non-Well-Founded Phenomena*, CSLI Publications, Stanford & Cambridge University Press.

J. Barwise & J. Perry, 1983, *Situations and Attitudes*, MIT Press, Cambridge (Mass.).

J. Barwise & J. Seligman, 1996, *Information Flow in Distributed Systems*, to appear in Cambridge Tracts in Theoretical Computer Science, Cambridge University Press.

J. van Benthem, 1986A, *Essays in Logical Semantics*, Reidel, Dordrecht.

J. van Benthem, 1986B, 'The Ubiquity of Logic in Natural Language', in W. Leinfellner & F. Wuketits, eds., 1986, *The Tasks of Contemporary Philosophy*, Schriftenreihe der Wittgenstein Gesellschaft, Verlag Hölder-Pichler-Tempsky, Wien, 177-186.

J. van Benthem, 1987, 'Meaning: Interpretation and Inference', *Synthese* 73:3, 451-470.

J. van Benthem, 1988, 'Games in Logic', in J. Hoepelman, ed., *Representation and Reasoning*, Niemeyer Verlag, Tübingen, 3-15, 165-168.

J. van Benthem, 1989A, 'Logical Constants across Varying Types', *Notre Dame Journal of Formal Logic* 30:3, 315-342.

J. van Benthem, 1989B, 'Semantic Parallels in Natural Language and Computation', in M. Garrido et al., eds., 1989, *Logic Colloquium. Granada 1987*, North-Holland, Amsterdam, 331-375.

J. van Benthem, 1991A, Editorial, *Journal of Logic and Computation* 1:3, 1-4.

J. van Benthem, 1991B, *Language in Action: Categories, Lambdas and Dynamic Logic*, North-Holland, Amsterdam. [Paperback with Addenda, The MIT Press, 1995.]

J. van Benthem, 1993, 'Modeling the Kinematics of Meaning', *Proceedings Aristotelean Society* , 105-122.

J. van Benthem, 1994, 'Dynamic Arrow Logic', in J. van Eyck & A. Visser, eds., *Logic and Information Flow*, MIT Press, Cambridge (Mass.), 15-29.

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford & Cambridge University Press.

J. van Benthem, 1997, 'Logic, Language and Information: The Makings of a New Science?', guest editorial, *Journal of Logic, Language and Information* 6:1, 1–3.

J. van Benthem & A. Ter Meulen, eds., 1997, *Handbook of Logic and Language*, Elsevier Science Publishers, Amsterdam.

B. Bolzano, 1837, *Wissenschaftslehre*, Sulzbach. (Translated by R. George as *Theory of Science*, University of California Press, Berkeley and Los Angeles, 1972.)

K. Dosen & P. Schröder-Heister, eds., 1994, *Substructural Logics*, Clarendon Press, Oxford.

D. Gabbay, 1992, *Labeled Deductive Systems*, Department of Computing, Imperial College, London. To appear with Oxford University Press.

D. Gabbay, ed., 1994,*What is a Logical System?*, Oxford University Press, Oxford, 107-139.

D. Gabbay & F. Guenthner, eds., 1983–1989, *Handbook of Philosophical Logic,* Reidel, Dordrecht.

D. Gabbay, C. Hogger & J. Robinson, eds., 1991–1995, *Handbook of Logic in Artificial Intelligence and Logic Programming*, Oxford University Press.

P. Gärdenfors', 1988, *Knowledge in Flux. The Dynamics of Epistemic States*, MIT Press, Cambridge (Mass.).

P. Gärdenfors & H. Rott, 1995, 'Belief Revision', a chapter in D. Gabbay, C. Hogger & J. Robinson, eds., *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. IV, Oxford University Press, 35-132.

J. Groenendijk & M. Stokhof, 1991, 'Dynamic Predicate Logic', *Linguistics and Philosophy* 14, 39-100.

J. Hintikka, 1973, Logic, *Language Games and Information*, Clarendon Press, Oxford.

I. Lakatos, 1976, *Proofs and Refutations*, Cambridge University Press, Cambridge.

Ming Li & P. Vitanyi, 1993, *An Introduction to Kolmogorov Complexity and its Applications*, Springer Verlag, New York.

D. Makinson, 1994, 'General Non-Monotonic Logic', in D. Gabbay, C. Hogger & J. Robinson, eds., 1994, *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. III, Oxford University Press, Oxford, 35-110.

Y. Moschovakis, 1991, 'Sense and Reference as Algorithm and Value', Department of Mathematics, University of California, Los Angeles.

R. Muskens, A. Visser & J. van Benthem, 1997, 'Dynamics', a chapter in J. van Benthem & A. ter Meulen, eds., *Handbook of Logic and Language*, Elsevier Science Publishers, Amsterdam, 587-648.

E. Nagel, 1961, *The Structure of Science*, Harcourt, Brace & World, New York.

V. Sanchez Valencia, 1991, *Studies on Natural Logic and Categorial Grammar*, Ph.D. thesis, Institute for Logic, Language and Information, University of Amsterdam.

J. Sneed, 1971, *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht.

A. Tarski, '1986, What are Logical Notions?', in J. Corcoran (red.), *History and Philosophy of Logic* 7, 143 - 154.

F. Veltman, 1996, 'Defaults in Update Semantics', *Journal of Philosophical Logic* 25, 221-261.

Y. Venema, 1996, 'A Crash Course in Arrow Logic', in M. Marx, M. Masuch & L. Pólos, eds, *Arrow Logic and Multi-Modal Logic*, CSLI Publications, Stanford & Cambridge University Press, 3-34.