

Johan van Benthem

ILLC

University of Amsterdam

j.vanbenthem@uva.nl

Valedictory Lecture A personal view of logic

Fanning the flames

On 26 September 2014, Johan van Benthem, Professor of Pure and Applied Logic, delivered his valedictory lecture in the auditorium of the University of Amsterdam. Modern logic is the broad, exact science of information directed behaviour that plays an important role in connecting the humanities, exact sciences and social sciences at the university, and which forms the bedrock of our information society. In his valedictory lecture, van Benthem illustrates the creation of this scientific field and the most important future challenges facing it.

When people hear that you are a logician, their eyes start looking for the nearest exit. But now and then, there are creatures like me, whose encounter with logic changed their life. Already in high school, I liked rules and formulas, which I jotted down in a special notebook. And when I arrived at the University of Amsterdam in its revolutionary Sixties, I loved to engage in argumentation about the future shape of Dutch society — to be determined of course by us, the student elite. My argumentative skills were not particularly successful, however, and I well remember the sinking feeling of slowly, but inexorably talking myself into a corner with an indefensible claim. A kind fellow student advised

me to buy a book about logic, and that is what I did. (In Figure 1 I add three pictures of what the species of logicians looked like to me at the time. The third picture is of my inspiring first teacher in Amsterdam, Else Barth.)

Reasoning patterns

What I bought was a pocket edition of a nineteenth century classic by William Stanley Jevons, available on the internet today. And what I learnt was that reasoning has *patterns*, and that some patterns are correct, while others are not.

Consider the following statement, where capital letters stand for arbitrary sentences:

If B, then not-H,

and ask yourself if it follows logically that

If not-B, then H.

You might be tempted to agree. But what if the doctor on duty in the emergency room where you have just been rushed in after a collapse, is trying to figure out what is wrong with you? Let us say that, if the cause is in your brain, it is not in your heart — but nothing is wrong with your brain scan. Is it correct to conclude that something is the matter with your heart? I do not think so, and neither should you. For myself, I would prefer a doctor whose reasoning goes differently, using a correct reasoning pattern like this:

from if B, then not-H to if H, then not-B.

Of course there is more to proper diagnosis than just these few rules — but at least, on the lighter side, you now see that logical correctness may be a matter of life or death.

Explaining and teaching

In general, logical reasoning patterns lie behind how we find the truth, and explain it to others, and ourselves. I always felt a need to explain things I think I understood as a private test. As a student, I would walk at night in the Lairessestraat in Amsterdam talking to myself to explain what I had learnt from the day's classes. Again, people around me quickened their step and increased their distance — and this time, even without my having told them that I was a logician.

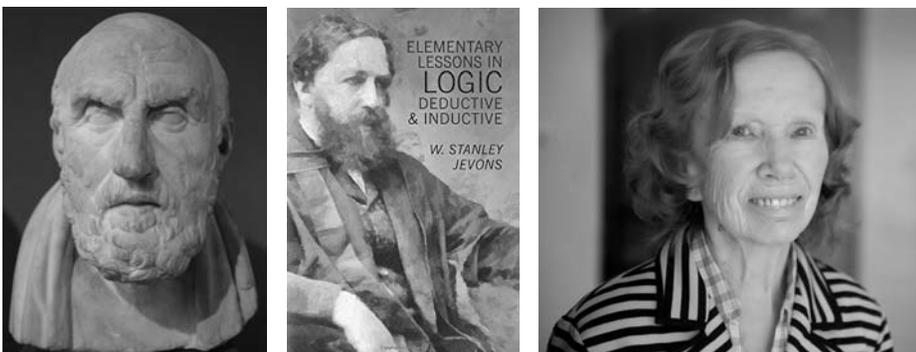


Figure 1 Chrysippos of Soli, W. Stanley Jevons and Else Barth.



Johan van Benthem

of reason

Logic across cultures

In line with the thread of my own personal experiences, most of this lecture will follow the Western mainstream history of the discipline of logic. However, it is important to emphasize that a sensitivity to reasoning patterns and their valid laws is not a Western preoccupation: it is a talent that has occurred quite early across different cultures. Indeed, we live in exciting times where many new historical sources for our field are coming to light, as archives yield their secrets. Logic stands in a worldwide movement including ancient Greece, China, India, and the Islamic tradition, as shown in the efforts of a growing number of scholars today.

To complement the dominant facebook of the field shown earlier, I add in Figure 2 some pictures of major logicians in these other cultural streams. One day, Mōzǐ, Dignāga, and Ibn-Sīnā may be names that are as familiar to the intellectual public as Aristotle.¹

Reasoning and information

But logic is more than seeing patterns and improving reasoning practice. With the study of correctness and incorrectness, we enter a much broader world of exact scientific ideas. One of these ideas is the ubiquitous notion of *information*.

In principle, there are exactly four ways the world might be given the above statements *B* and *H*:

$$+B + H, +B - H, -B + H, -B - H.$$

If we then learn the above rule *if B, then not H*,

we lose the possibility that $+B + H$. That is, we eliminate one out of the four options to obtain the information

$$+B - H, -B + H, -B - H.$$

If we now learn that *not-B*, we lose the further possibility $+B - H$, but two still remain:

$$-B + H, -B - H.$$

This final state does not have enough information to determine that *H* is the case.

You can check for yourself how this scenario of successive information updates differs from that for the logically valid inference from *if B, then not H* to *if H, then not B*.

Here the information does force the conclusion *not-B* through the following stages:

- i. $+B + H, +B - H, -B + H, -B - H$;
- ii. $+B - H, -B + H, -B - H$
(after learning that *if B, then not H*);
- iii. $-B + H$ (after learning that *H*).

So, reasoning is tied up with information, the lubricant of our whole society, and logic suddenly starts looking like a much bigger field than you might have thought.

Logical systems

And there is still more than reasoning patterns and their underlying ideas. Quickly, I learnt that patterns form larger *systems*, that can be studied on their own intrinsic merits — all the way back to Aristotle’s system of syllogistic inference. Logical systems have mathematical structure, which leads to new avenues of its own.

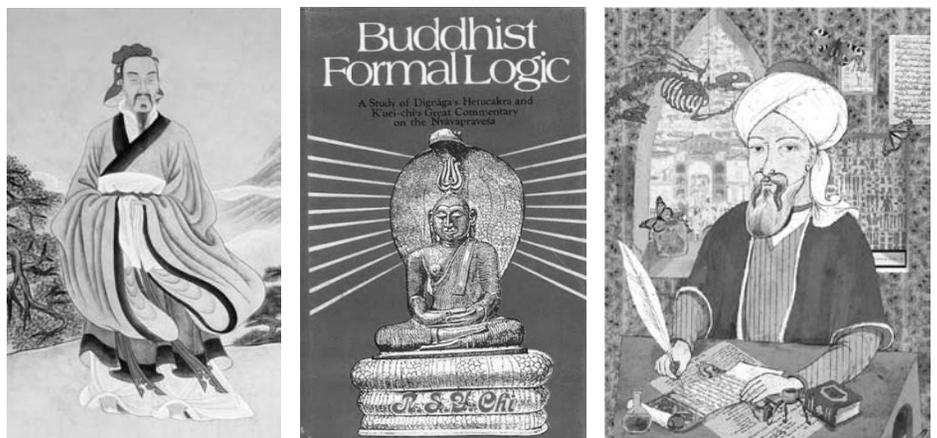


Figure 2 Mōzǐ, Dignāga and Ibn-Sīnā.

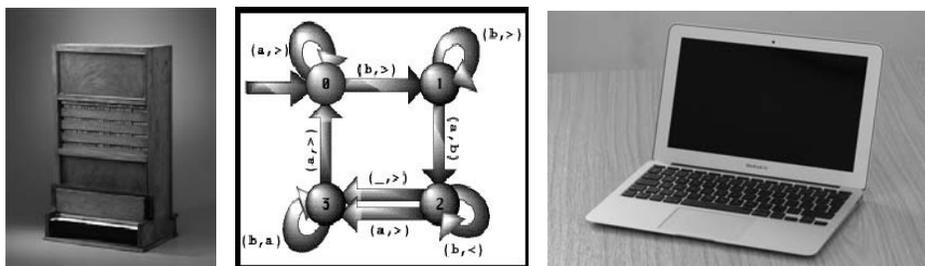


Figure 3 Logical piano, Turing machine and Macbook Air.

Reasoning and computation

The criterion for correctness or incorrectness of simple rules like those I gave for implication and negation is manipulating the ‘truth tables’ of basic logic, and thus reasoning gets tied to mathematical structures in *computation*. This is a general thread in the history of ideas: reasoning is connected to binary arithmetic, or as Hobbes said: “Reasoning is reckoning.” And once we see this computational angle in its general light, the idea emerges naturally, already in the Middle Ages, that reasoning might be done by machines, provided we find suitable logical languages for them to work with. The first picture in Figure 3 is Stanley Jevons’ ‘logical piano’, one of the pioneering attempts at creating mechanical computing devices.²

The second picture in Figure 3 is Turing’s fundamental model of computation, and the third is an object that has swept the world in recent years. There is a straight historical road from logic to computer science, and the information technology pervading our lives.

This mixture of reasoning, information, and computation is the thriving conglomerate of logic, philosophy, mathematics, linguistics, and computer science that has put logic firmly at the crossroads of the modern university, and our modern information society. This, indeed, is the world of the Amsterdam Institute for Logic, Language and Computation where I have spent so much of my academic career.

Logical systems and their theory

As we have seen by now, logical patterns and systems capture major ideas that can be analysed and developed as such. And the great engine for pursuing that analysis and development has been the *mathematical turn* of modern logic since the nineteenth century, turning it into an exact discipline that unleashes the power of mathematical precision to study systems of reasoning and everything that comes in their wake. This is how modern logic works, and nothing that I will say in the remainder of this lecture about

broadening the scope of the discipline is intended to depart from this methodology.

Paradoxes and foundational theorems

In particular, the power of mathematical thinking allowed the great logicians of the Thirties to reflect on the basic nature of reasoning itself, both its strengths and its limitations. A celebrated instance of this style of ‘confronting reason with itself’, to use a sonorous philosophical phrase, are Gödel’s Incompleteness Theorems that capture surprising essential features of the mathematical method par excellence, namely, *deductive proof*.

Consider one of the oldest instances of logical thinking that has come down to us in various cultures, from ancient Greece to ancient China, ‘self-referential’ statements that talk about themselves. A famous case is the Liar Paradox from Antiquity where someone asserts that the very statement she is making is false. But for here, think of a related but subtly different statement *L* that says about itself that it cannot be proved:

L if and only if *L* is not provable,

or in Amsterdam style formula notation:³

$$L \leftrightarrow \neg \Box L.$$

Languages have several ways of achieving such self-reference. If you want a dramatic historical instance, the Hebrew God told the

people seeking his essence:

“I am the unknowable one.”⁴

However, our logical intellect does not stop at such a statement, and the liar sentence *L* admits mathematical reasoning with startling consequences.

Here is a piece of this reasoning at the edge of reason. It is well within your capacities to appreciate — assuming the famous motto of John Perry and Ken Taylor’s radio show Philosophy Talk: “the only thing that we do *not* question is the hearers’ intelligence.”

Let us draw some quick consequences in a first, rather simple pass:

Gödel lite. Is *L* true or false? If we assume that *L* is false, then *L* is provable.⁵ But if *L* is provable, then *L* is true: and we have contradicted our assumption. We conclude from this piece of reasoning that *L* cannot be false. In other words: *L* is true. But given what *L* says, it follows that *L* is a *true statement that is unprovable*.

The pictures in Figure 4 show some resounding names in this world of thought.

Living at the edge of reason

This simple piece of reasoning lies at the heart of Gödel’s famous First Incompleteness Theorem. But it is clearly just a very first pass, which invites questions and refinement at once. Indeed, the art of ‘living dangerously’ displayed here is another major attraction of logic, especially to minds of a somewhat absolutist bent, and it involves a continuous practice of refining our styles of thinking.

To turn the screws of exactness a little bit more, how reasonable, really, was the above step where we assumed that provable statements *L* are true? Does not that represent an unwarranted confidence in the reliability of our proof methods? The great logicians working in the foundations of mathematics in

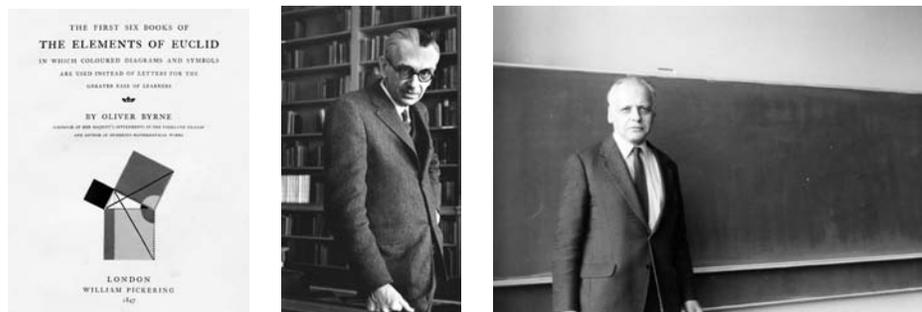


Figure 4 Euclid’s *Elements*, Kurt Gödel and Martin Löb.



Figure 5 Frege, C.S. Peirce, Brouwer and Beth.

the Thirties such as Gödel, Hilbert and Gentzen thought so, and dropped this assumption in their analysis of the power of proof. Let me sketch a more sophisticated version of Gödel's argument for you then:

Assume we have $\Box(L \leftrightarrow \neg\Box L)$ (*). Next, assume that $\Box L$. Then using (*), we can prove that $\Box\neg\Box L$. But at the same time, inspecting the proof for L that exists by $\Box L$, we can prove $\Box\Box L$. Combining the two, a contradiction $\neg\Box L \& \Box L$ is provable. So, our proof system is inconsistent, and unreliable.

This time our conclusion gets more sophisticated. *If our mathematical proof methods are reliable*, then there have got to be true unprovable statements beyond their reach.

The process of refining our understanding of truth, proof, and reliability is still going on. Masters of the art included my thesis supervisor Martin Löb, Sergei Artemov, Dick de Jongh, and Albert Visser — and still last year, I read surprising delightful new perspectives on the mechanics of Gödel's reasoning discovered by Lev Beklemishev at the Steklov Institute in Moscow. We can keep reflecting on our understanding of mathematical proof, and go deeper and deeper, until the mind reels.

Science, religion, or common sense?

With all this, we are at the heart of classical logic and foundations of mathematics, and its deep results that keep inspiring religious awe in every new generation of students worldwide that gets exposed to logic. Yet, this depth comes with a somewhat narrow focus on mathematical proof, a very specialized human activity. The methods are mathematical, the topic is mathematics.

But I started this lecture in our daily reasoning practice, and want to return there eventually. As a student, I had to memorize a list of differences between scientific and common sense reasoning — with the former supposed to be much deeper and better than the latter. But the striking thing to me is the very opposite. The logical steps for Gödel's Theorem are the same ones that we employ in daily life, and this observation fits the facts of human cognitive evolution, where bands of hunters with survival skills eventually produced the greatest art and science, all using the same brain.

The true story of modern logic then is as follows. The methods remain mathematical, but the topic becomes much broader: in principle, any aspect of the human intellect. The true facebook of logic, then, consists of both mathematicians and others looking beyond, forming pairs as in the pictures of contemporaries, see Figure 5.

Frege was contemporaneous with C.S. Peirce, an independent inventor of first-order logic with a much broader agenda in philosophy, semiotics, psychology, and so on, which we are only beginning to carry out today. Our national mathematical icon L.E.J. Brouwer was a contemporary of E.W. Beth, a broad logician and philosopher, and it is their tandem that created the Dutch School in logic that we know today. It would be easy to multiply examples of such facebook pairs all over the history of logic.

So, here is where this lecture is going to go, charting further aspects of what our discipline is, or can become. Eventually, I want to understand the laws of what actual *reasoning agents* can and should do: in terms of knowledge, action, learning, making mistakes, and

all of that in mutual social dependencies. And while setting our bounds that widely, I still hope to convince you that we are doing logic.

But before we enter this enticing Realm of Agency, let me first talk about the personal intellectual development that took me there. I will take things slowly, introducing new ideas one by one.

Proof as seeing and as doing

I start with two essential aspects of mathematical proofs that have occurred entangled ever since Euclid's *Elements* in Antiquity. A proof is *evidence* that a theorem is true: that is, it provides information — but at the same time, it is a *method* for computing solutions and creating situations where the theorem holds. The same duality pervades general learning and knowledge, where 'knowing that' and 'knowing how' are deeply intertwined. For instance, we have truly learnt a language when we recognize when given sentences are correct, but are *also* able to produce discourse that makes sense, and the same is true for logical systems. Thus, studying information also means studying action, and vice versa. The two notions are two sides of the same coin. How can logic do justice to this deep duality?

Reasoning about information and action

Logicians have long studied reasoning about both information and action — though they may come to these pursuits from different academic fields. Logical systems analysing information often come from the philosophical study of knowledge and belief, but in recent years they have entered computer science and economics. 'Dynamic logics' of action come from computer science, in the

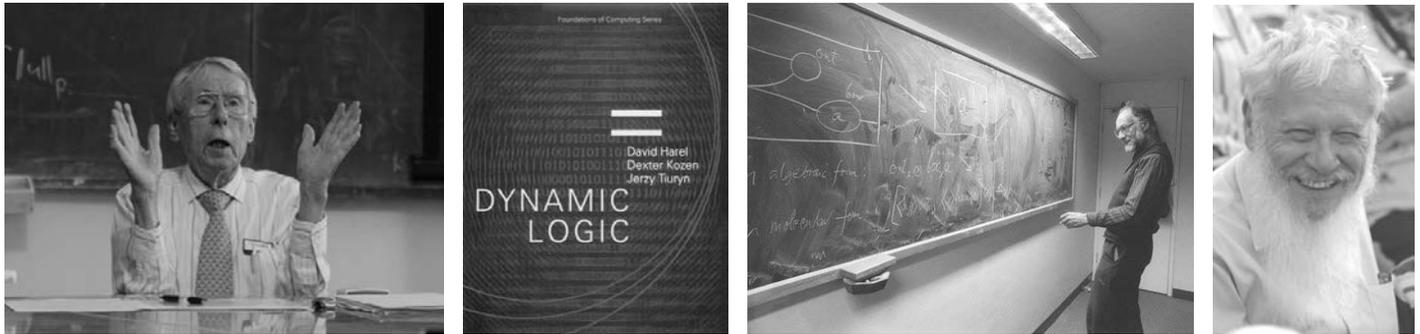


Figure 6 Hintikka, Dynamic Logic, Milner and Aumann.

analysis of programs and processes, but have spread to philosophy and linguistics. These contacts have blossomed over the years, and taken their practitioners to Nobel Prizes, Turing Awards, or their equivalents in philosophy. The four pictures in Figure 6 present a gallery of relevant names and topics ranging from philosophy through computer science to game theory.

Modal logic

The abstract common structure of information and action that underlies large parts of philosophy, computer science, or game theory is studied in a technical field called *modal logic*. A modal notion states how things might be, insofar as not ruled out by our information – or how things might become, given the actions at our disposal. Very similar laws govern such notions. One example is this:

$$\text{Distribution Law } \Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$$

In terms of information, this simple principle expresses that

if I have evidence for $A \rightarrow B$
as well as evidence for A ,
then I also have evidence for B .⁶

In terms of actions, the same principle says that

if the outcomes of an action satisfy $A \rightarrow B$,
and they also satisfy A ,
then all outcomes of that action satisfy B .

There are many other, and technically much more sophisticated, laws of modal logic, but the simple example of the Distribution Law will have to do here.⁷

Modal reasoning patterns have been a focus in my work over the years, which offers a systematic study of modal laws, but also of connections between modal patterns and other basic logical systems, measured by semantic invariances such as ‘bisimulation’. Some sources are depicted in the sequence of four books in Figure 7.

These matters can be found in any good textbook, but let me stress one point. While modal logic is ‘applied’ as a theory of information and action, it is at the same time ‘pure’ as a study of basic aspects of reasoning that had not come to light before, and thus it falls squarely within the mathematical methodology that I have emphasized.⁸

Information-driven agency

But let us now move beyond the agenda stated so far. For a start, what about the actors getting the information and engaging in actions? What do they do, and why? My next step toward the study of agency is an arena where knowledge and action meet naturally, the basic *informational acts* themselves.

Basic informational acts

While reasoning may seem the only informational action of interest to logicians, in reality, inference steps usually occur together with other equally fundamental acts that make information flow. For instance, the following scenario plays out day by day in Amsterdam cafés.

The Restaurant. Three people order three drinks: water, beer, and wine. A new waiter comes to the table carrying three glasses. There are six ways these drinks could be distributed over the customers, and here is what you will see happen.

The waiter *asks* who has the beer, say, and after hearing the answer, puts down that glass in the right place. This reduces his uncertainty from the initial six options to two, much as in our example of diagnostic rules. The waiter then asks a second question, and puts, say, the glass of water. Now his options are down to one, and he does not ask any more, but can *infer* who has the wine, putting it down without a question.

Thus, basic informational acts include, at least, both questions and inferences. But there is more. As it happens, a broad view of informational action was already present with the ancient Chinese logicians in the Mohist school, active from around 500 BC. It stressed the interplay of three sources that come together in producing information:

知闻说亲

(zhi wen shuo qin)

an elegant compact statement which says roughly, in a more modern phrasing, that “knowledge comes from three sources: hearing from others, proof, or experience”.

Here the third category of ‘experience’ is one more major source of information, in daily life as well as in the natural sciences: viz. just

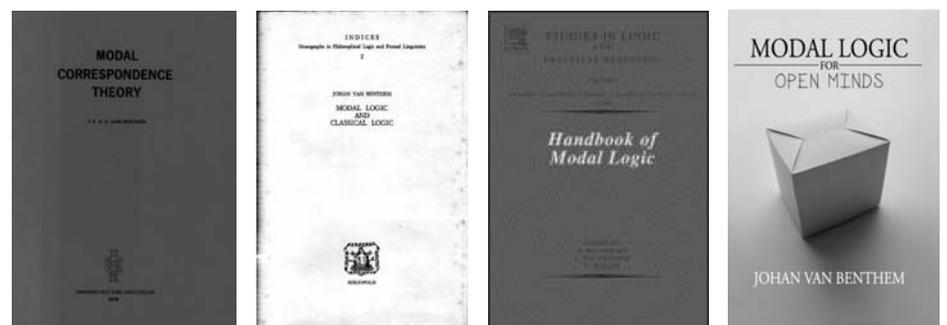


Figure 7 Dissertation, Monograph, Handbook and Textbook.

observing things.⁹ Observation is what a logical detective like Sherlock Holmes needs in addition to pure deduction, and it is the well-spring of the empirical sciences, on a par with mathematical proof.

The pictures in Figure 8 represent a restaurant, a Mohist dialogue, Sherlock Holmes at work, and finally, a card puzzle with social reasoning that is often used in my circles.

Logic of information change

My own view of logic has come to resemble this ancient paradigm. We must study all basic informational acts on a par: inference, observation, and communication, and their interplay in how we deliberate, plan, and decide. A major test for such a venture is a search for formal laws governing these additional informational actions that retain the precision logicians are used to. Here is one:

$$[!\varphi]K_i\psi \leftrightarrow (\varphi \rightarrow K_i(\varphi \rightarrow [!\varphi]\psi))$$

This law describes the new knowledge an agent *i* will possess after having received the information that φ is the case.¹⁰ It does so in terms of the ‘conditional knowledge’ that the agent had before the new information $!\varphi$ came in. Typically, it is ‘recursion equations’ like this law that drive the logical study of information-driven agency.

As an aside, please note how the above principle shows the cooperation of ideas from different fields: the modality $K_i\psi$ for knowledge in this law comes from philosophical logic, while the dynamic modality $[\varphi]$ for action comes from logics of computation.

Still, we are only at the start of a logical study of agency. The actors in the above could still be instinctive devices recording information from the environment, like the bees sucking up nectar in the flowers in my garden. So let me turn to further intelligent activities that

rational agents engage in, of which I will mention a few in this lecture.

Consistency?

The first new aspect of agency concerns the quest for certainty we saw earlier. The founding fathers of modern logic were concerned with the power of proof, and its reliability. Is *consistency* our ultimate goal? Gottlob Frege thought so, saying that mathematics would ‘crumble like a house of cards’ if a contradiction were to arise.¹¹ But there can be logical grandeur in fragility, being wrong, and making a volte face.

Learning is correction

As I said before, proofs give reasons, and a rational agent can be described as someone driven by reasons.¹² But at this point, I part company with the original foundational enterprise. Our drivers should be *good* reasons — but there is no guarantee that they are correct. We make mistakes and often form incorrect *beliefs*, and the task of logic cannot be to prevent that from happening once and for all. Indeed, our beliefs and expectations, even when wrong, represent a creative human talent in navigating a world where knowledge is a currency that is often unavailable.

But this non-eliminable, and perhaps even desirable, potential for error comes with another creative talent, our ability to *correct ourselves* when beliefs are shown wrong by an observation, or when theories explode after a contradiction comes to light. There is no house of cards that crumbles: indeed, human intelligence shows at its finest in a clever recovery after we have been shown to be incorrect. So, a true logic of rational agency tolerates errors and highlights methods for revising beliefs. What this points at more generally is an intimate connection between reasoning and *learning*.

While all this may be true, are beliefs and their revision a fit theme for logic? As it happens, thanks to the work of pioneers like Peter Gärdenfors and others, we know that this more delicate error-tolerant process of adjusting to new information still satisfies logical laws, that now involve operators for belief and acts of belief revision. Of such laws, I just display two at this point, merely in order to show you that, once again, we need not sacrifice the mathematical precision of our foundationalist founding fathers:

$$[!\varphi]B_i\psi \leftrightarrow (\varphi \rightarrow B_i\varphi [!\varphi]\psi),$$

$$[\uparrow\varphi]B^{\psi}\chi \leftrightarrow (E(\varphi \wedge [\uparrow\varphi]\psi) \wedge B^{\varphi \wedge [\uparrow\varphi]\psi} [\uparrow\varphi]\chi) \vee (\neg E(\varphi \wedge [\uparrow\varphi]\psi) \wedge B^{[\uparrow\varphi]\psi} [\uparrow\varphi]\chi).$$

Such laws describe which new beliefs arise as new information comes in. This can now happen in two forms: witness the syntax of our two formulas. Changes can be triggered by a ‘hard’ totally reliable event $!\varphi$ removing all $\neg\varphi$ -situations for good, but there is also a more ‘soft’ plausibility-changing event $\uparrow\varphi$ that keeps all situations, while putting those with $\neg\varphi$ on top of those with φ . I am not going to try to explain to you any further what the above formulas say. Their typography may well scare you. These laws of learning are clearly more complex than the simple rules of reasoning that I started with. But then I never said our broader logical skills are simple!

Let me just state the upshot. Correctness forms a tandem with correction — and as some philosophers have said long ago, logic is also about learning from errors.¹³

My favourite analogy for this comes again from health. Creating a World free from disease seems a sterile and unappealing goal: much more impressive is the dynamics of the human *immune system* that deals with chal-



The Cards: Learning by Update

Children 1, 2, 3 get each one card

red [red square] white [white square] blue [blue square]

1 Red, 2 White, 3 Blue

See their own card only, not that of the others.

2 asks 1: “Do you have blue?”

Does anyone know the cards now?

1 answers “No”

Who knows what now?

The children in NEMO (age 8) solved this!

Figure 8 A restaurant, a Mohist dialogue, Sherlock Holmes and a card puzzle with social reasoning.

lenges as they occur. To me, logic is not the guardian of consistency proofs, but rather,

“the immune system of the mind”.¹⁴

In terms of our initial example, maybe logicians themselves are doctors of the mind.

Even so, I should not play up the difference with foundational research too much. There might well be very interesting extensions of the classical theorems by Gödel or Turing when, in addition to what is known with total certainty, we also explicitly represent agent’s reason-based, but fallible beliefs about what is true and provable.¹⁵

Many minds and social interaction

Beliefs may need correction for all three reasons in the earlier picture of information dynamics: a surprising observation, a paradoxical inference, or also, being contradicted by somebody else. This leads me to my next fundamental theme in the logical study of agency, which can be summarized as:

“intelligence seldom comes alone”.

My example of argumentation was a social scenario where different agents interact. This is significant. Rational behaviour is a *many-mind talent* rather than an isolated skill — just as in modern science, the key is many-body interactions, not single bodies in their natural place, as in Aristotle’s physics. This social theme of deliberation and argumentation runs throughout the history of logic, in both Greek and non-Western traditions. Logic is also about how we convince, refute, or otherwise influence others.

Just as earlier, this social dimension has two dual aspects: *information*, and *action*.

Theory of mind

The social aspect of information involves a crucial human ability often called ‘Theory of Mind’ where we reason about the facts, but also about other people’s information, what they know about the facts and about us, and so on. Theory of Mind underlies many informational acts. My asking you a *question* normally conveys I do not know the answer, and also that I believe you might know the answer.¹⁶

Social behaviour is kept in place by a web of such complex iterated information about others that can be studied as before. Again, it has been found that crucial properties of rational social equilibrium then show up in logical laws. For instance, a proposition φ is

‘common knowledge’ in a group of agents G , i.e., it is known up to any depth of iteration (written as $C_G\varphi$), if and only if the following equivalence holds:

$$C_G\varphi \leftrightarrow (\varphi \wedge E_G C_G\varphi),$$

where an operator $E_G\psi$ says that everybody in the group G knows that ψ is the case. This looks like a vicious circle that foolishly tries to define common knowledge $C_G\varphi$ in terms of that notion itself, but in fact, the above equivalence is a ‘luscious circle’ where the notion at issue is well-defined, and supports precise reasoning.^{17 18}

The program of logical dynamics

The preceding samples suggest a more systematic logical study of information-driven social agency. Over the past decades it has become clear that the methods of logic are well up to this agenda extension, creating systems with beautiful laws, not just for inferences, but also observations, questions, correction, learning, and social interaction. And in doing so, the grand style of our founding fathers does not lose its lustre. Its thinking about strengths and limitations also applies to what diverse sources of information do for rational agents.

In what follows, I discuss one current highlight of this ‘Logical Dynamics’ research program, as exemplified in the depicted three books in Figure 9, from 1996, 2011, and 2014, respectively.

Games

Social action is at the heart of the fast-developing area of logic and strategic behaviour in *games*. Our actions toward each other congregate to form longer-term *strategies*, where what I choose to do next depends

on what you do now, and so on. For examples close to logic, just think of the timing of playing your cards in argumentation, or your points in giving a valedictory lecture. The logical laws of strategies or plans form a natural continuation of the earlier-mentioned logics of actions, and they often concern situations where strategies are in some sort of rational *equilibrium*, meaning that no agent has an incentive for deviating from her current plan.

Only one such law will have to serve as an example here, again mostly just in display. It is very similar qua form to the earlier equivalence for common knowledge, and it describes when someone has a strategy for playing a game G that always achieves the effect φ (this is written here as $\{G^*\}\varphi$) in terms of φ ’s being true and what the player can achieve by her currently available moves (the latter ability is written as $\{G\}$):

$$\{G^*\}\varphi \leftrightarrow (\varphi \wedge \{G\}\{G^*\}\varphi)$$

The analogy with the earlier law for knowledge again shows the fundamental duality of information and action that governs logic of agency. They are two sides of the same coin, and often, the shape of their mathematical laws reflects their similarities.

Harmony of information and desire

However, in the fundamental notion of game-theoretic equilibrium in behaviour between different agents, a new notion comes in sight that you might have thought far outside of the province of logic, namely, our preferences, intentions, and *desires*. We are not purely informational agents attuned to how the world is by knowledge or belief: everything we do is coloured by what we want, hope, fear, or regret. While some people view the latter as a realm of blind impulses and irrationality,

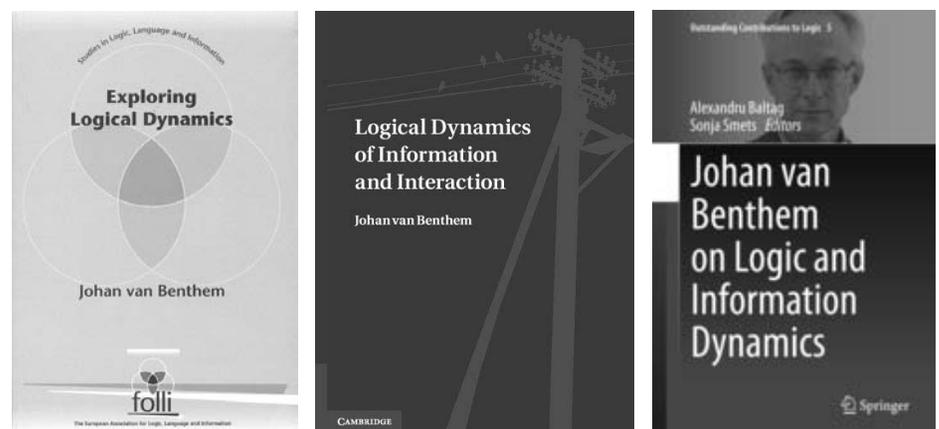


Figure 9 Three books by (and about) the author on logical dynamics.

nothing could be further from the truth. What we want to know explains how we seek information, and what we want to achieve determines our choices of actions. Instead of a separating barrier between these two realms, a logical perspective on rationality involves understanding their delicate *balance*.^{19 20}

A reasonable person keeps her information in harmony with her goals and desires, plus those of others. That quality shows in the activity of taking *decisions*, a growing interest among logicians. Incidentally, it is that broader balance of information, desires, and decisions that we should also seek to instil in a true academic education.

Logic in games

Noble causes indeed, but can logic offer any illumination in such a broad arena? What I can say for now is found in my new book *Logic in Games*. It covers two major lines of research, that have been developing over the past decades.

The first line is *logic of games*, being the study of the structure of games and everything that happens before, during and after their play by rational agents. There is a lot of intricate reasoning involved in plotting one's course in games, including the effects of one's actions, beliefs, preferences, and thinking about the kind of opponent (or fellow player) one is engaging with. This reasoning shows logical patterns that can be studied by the tools I have outlined in the above, including fixed-point laws for social equilibria where all players do the best they can. This line of research may be considered a joint offspring of logic, game theory, and modern computer science in its interactive-systems mode. Logic of games is now engendering a joint research program that might be called 'Theory of Play', where we analyse the reasoning of players in games and related social scenarios — and sometimes, more ambitiously, try to design new games that implement desirable forms of social interaction.

But at the same time, there is a second intimate connection here, that may be called *logic as games*. As we have hinted at in the beginning of this lecture, paradigmatic logical phenomena such as argumentation *themselves* have a game-like character, and this metaphor can be made very precise. For instance, mathematical proofs have turned out to be very much like 'winning strategies' in suitably defined dialogue games. This game-theoretic view of logic itself also fits very naturally with modern views of computation, where the math-

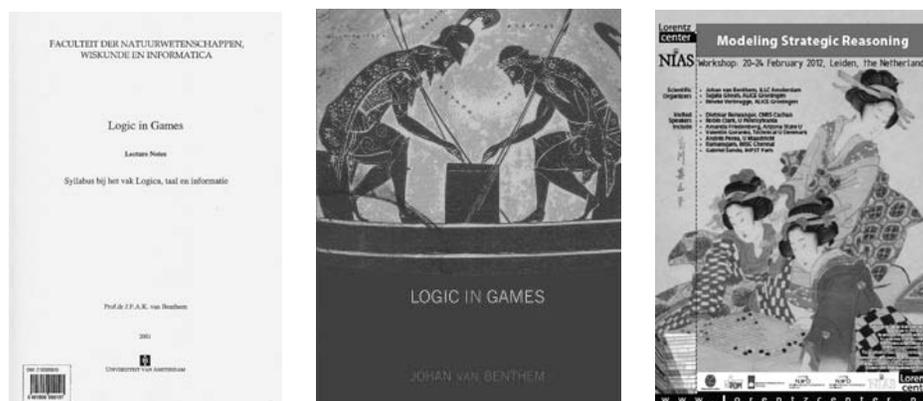


Figure 10 Lecture notes and a monograph by the author on logic in games and a workshop on logics for strategic reasoning.

ematical paradigm is no longer isolated Turing machines, but networks of distributed processors that may cooperate or compete for available resources, and that may differ widely in their abilities and intentions.

Thus, once more, logical laws acquire the dual character we have seen several times. Viewed in an informational perspective, they tell us what the world is like. But in a dual dynamic procedural perspective, they express fundamental facts about strategies and equilibria in basic scenarios of interactive computation or social behaviour. This duality is not an annoying ambiguity, but rather the true face of logic as I see it.

The first two pictures in Figure 10 are lecture notes and a monograph of mine on both of the above interfaces of logic and games. The third picture points to a forthcoming book on logics for strategic reasoning, bringing together many communities, that I have edited with Sujata Ghosh and Rineke Verbrugge. This area is in full development.

What is next?

Ladies and gentlemen, I have shown you glimpses and glimmers of my dynamic program for logic as a broad study of information-driven social agency. As one of my readers put it nicely: "Classical logic handled the straight lines, now you want logic to also do the curves." Many of those curves are found in my own work and that of my colleagues and students. What is left for *me* to do?

The empirical facts

In my declining years, I sometimes stray from the glorious path of theory, and feel the pull of the brute *facts*. What I told you about information and agency can be seen as theory-bound and *normative*: this is the information that is truly available in the world, these are the conclusions that agents *should* draw, these are the game-theoretic equilibria they *should* at-

tain. In this classical perspective, the laws of logical dynamics have no descriptive content, and how people really act, presumably stupid and ill-mannered, can be left to cognitive or behavioural scientists.

But I fear that this common cosy analytic separation of the normative and descriptive reflects intellectual poverty rather than safety. Logical theory that ignores actual cognitive behaviour seems dangerously empty, lacking focus. On the other hand, I do not want to lose the potential of normative thinking either, that can help us *improve* performance, or design better ways of dealing with the world and with one another. For instance, my earlier emphasis on *correction* rather than guaranteed correctness of intelligent behaviour mixes these two motives. It moves closer to real human behaviour, but what we call successful 'correction' will still depend on norms. In other words, things are subtle and complex here, and I have no definitive line to offer.²¹

Let me give an example where these two pressures in my current life come to a head. You might feel that my broad program for logic is becoming a huge machinery of high mathematical complexity. Now there is a mathematical unity in logical methods that makes this complexity less daunting. But it is also good to realize that you are in fact good at everything that I have discussed in this lecture. Information flow, inference, correction, desires, and strategic behaviour occur together in that basic phenomenon that we all engage in all the time: *conversation*. Therefore, a natural empirical focus for my view of logic is natural language, the vehicle of conversation.

Natural languages and real brains

In the Eighties, I was deeply engaged in the logical analysis of natural language, and people sometimes ask where this love has gone. My work then was driven by static perspectives on language as reflecting the world, with

hardly a trace of dynamic agency. But natural language is a major interface where logical intuitions meet empirical facts. One project on my desk is the *language of agency*, studied in ways that will be familiar to you after my discussion of logic and games. One prong is a study of the patterns of agency that constitute language use, the other prong is the ‘language of agency’, and all the subtle ways that we use for describing what we do, plus the laws of the natural logic governing this natural language.²²

At the same time, I am involved with empirical research on language in the brain in our national Gravity Project ‘Language in Interaction’ where we will use logic-related games to see what really happens when people communicate, and how new language arises in that process. The four pictures displayed in Figure 11 show some of my earlier work on logic and natural language in the Eighties and Nineties, a Handbook on this interface, and a logo for the new project that also brings in empirical cognitive science.

Large numbers and probability

The second project waiting on my desk concerns another aspect of the facts, namely, that there are *so many* of them. Logical reasoning and rational agency play in a thin zone of conscious deliberation where what we think and decide matters. This zone is like the thin realm of our physical abilities, hemmed in between objects that are too large, or too small, for our hands to manipulate. Above us lies the statistical world of social behaviour and public

opinion, where each of our well-considered views is just a drop in the ocean. Below us is the statistical neural machinery of our brains, busily clicking away as an audience is listening to a speaker, but not under one’s conscious control. Here logic meets *probability theory*, as two major mathematical paradigms with different views of what are relevant patterns. My own interest in this area today concerns interfaces between logical and dynamical systems theory as providing two views of these interfaces. There are many more good reasons for combining logic and probability, but the above may suffice for here.

The interface of logic and probability has a long history: even Stanley Jevons included it in his book. It is also very much in evidence at our institute the ILLC, and I believe that it poses many fundamental problems that go far beyond what we know so far.

Foundational results after all

Factual interests are by no means opposed to the classical *modus operandi* in logic. Some neuroscience intriguingly repeats ideals from the foundations of mathematics, such as a ‘Universal Scanner’ that tells us exactly what an experimental subject is thinking, a goal with a sinister whiff of Inquisition. Fortunately, one can modify famous arguments in logic to show that universal scanners are as impossible as universal decision methods for reasoning or computation.

Spreading the word in education

The best test of a proposed scientific perspec-

tive may not be whether it pleases one’s senior colleagues, but whether it is *teachable* to audiences with fresh open minds, such as students, or the general public.

I wish I could jump into things here and start teaching you the topics of this lecture, but there are other resources for that, such as our on-line course *Logic in Action* that can be found at www.logicinaction.org.

I recommend a tour of this educational world, which we are still creating and testing out at several sites worldwide, including Amsterdam, Stanford, Beijing, and Seville. The following pictures illustrate part of this world, as well as its outliers into the Open University and on-line education efforts at several pre-university stages.

This reflects the earlier ties between logic and learning. People can get better at argumentation, and many other themes in this lecture point at skills that can be taught.

What is logic?

Let me return to the beginning of this lecture. What subject did I choose as a student, and was it the right choice? My general view of decision is quite unlike the paradigm in decision theory of choosing optimal actions given one’s present expectations. This is too passive. In my own life, decisions turned out to be good, not because of superior a priori reasoning, but because *I made* them good afterwards by committing to the actions that I had taken. This is also true for choosing logic.

Perhaps the field as it was when I chose it in the Sixties was not an optimal choice for someone like me.²³ But I have told you what logic can *become* if we take it to its full potential as a science of information in action, performed by agents engaged in meaningful activities, doing justice to both the classical ideals and a modern future of the field. Indeed, this view is at the heart of my Logical Dynamics program:

becoming is as important as *being*.

Deep thoughts and noble sentiments: but at the end of this long lecture, you may well prefer a lighter take home message. Let me try. Throughout my presentation, I have extolled the virtues of *reason*, as the light that, sometimes, relieves the boredom and darkness all around us. But that light has to come at the right time, in the right style.

One autumn night, fifteen years ago, I had a heart to heart conversation with my 85-year-old mother — and I put the delicate question I had never dared ask before: why she had ever

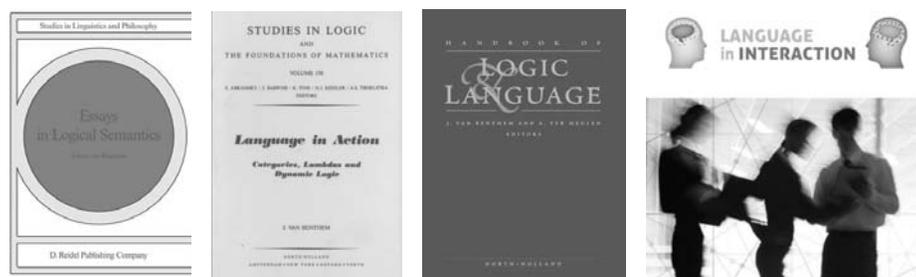


Figure 11 Some work by the author on logic and natural language.

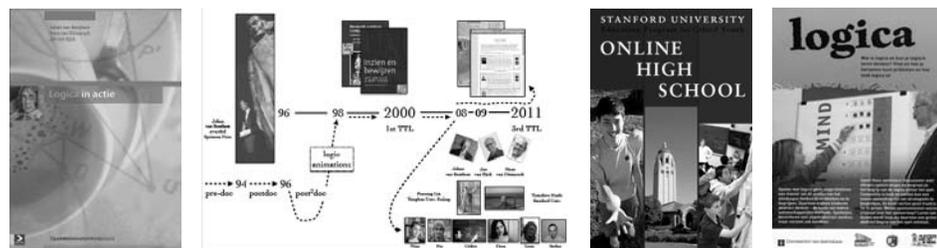


Figure 12 Initiatives in logic education.



Figure 13 Four of the ‘habitats’ of the author.

decided to marry my father. Her response was as follows:

“because you were not there to advise against it”.

There is a logical paradox lurking here, not unlike the ones I discussed in this lecture — but my mother’s response also gives me my final definition of logic for you today:

logic is the science of *wit*.

People and institutions

In line with what I have said about logic and probability, my own views and life’s decisions have ridden on a statistical sea of behaviour by many others. My formative interactions have included many colleagues and students, so much so, that I often wonder where my own professional persona ends and theirs starts. My work and ambitions in academic life also naturally led to helping shape new organizations in our field whose lifespan may well exceed our own.

The four pictures in Figure 13 show four of my habitats over the years: the Institute for Logic, Language and Computation in Amsterdam, the European Association for Logic, Language and Information, the Center for the Study of Language and Information at Stanford, and as a most recent initiative, a new Joint Research Center in Logic between the UvA and Tsinghua University in Beijing.²⁴ These are sites where modern logic in the style I have explained is flourishing, as well as much more.

But the realm of logic is much larger than this personal environment. I warmly invite you to open a door into the actual international world of logicians — or just at home, step through Alice’s mirror into the internet, and see its attractions for yourself. ←

Noten

- 1 With a group of dedicated colleagues, I am currently involved in exploring Chinese strands in all of this, in the forthcoming *Handbook of the History of Logic in China*.
- 2 Here I am not endorsing the old AI mantra that all reasoning can be done by machines. The ultimate goal of logic to me is not a world where reasoning and judgement are left to technology, but one where we understand ‘agency’, in the broad sense to come in this lecture.
- 3 The strange box \square is not a pdf conversion error, but the logical symbol for ‘provability’.
- 4 Yes, there is much more to the ‘unknowability’ in this example than mere formal logic.
- 5 Here we use earlier rules: we have $\neg \square L \rightarrow L$ by the meaning of L , and hence also $\neg L \rightarrow \square L$.
- 6 This use of the box notation is naturally related to our earlier formal analysis of provability.
- 7 In particular, if the actions considered can be different, then Distribution must be modified to an indexed form $[x](A \rightarrow B) \rightarrow ([y]A \rightarrow [x\#y]B)$, where the variables x, y stand for actions and $\#$ denotes some suitable operation of combining actions.
- 8 Modal logic is not the only field that does justice to the information–action duality. Other such paradigms that I admire include process algebra and category-theoretic approaches.
- 9 The Mohist example was a Dark Room, with an object inside whose colour you do not know. However, you see a white object outside of the room, and someone now tells you that the object in the room has the same colour. You then conclude that the object in the room is white.
- 10 These formulas are mainly for display: I ruthlessly omit technical details in this lecture. Just for once: the stated law describes knowledge after update as conditional knowledge in a recursive manner that lends itself well to logical analysis of the relevant modal reasoning.
- 11 In daily life we also value ‘consistency’, a more subtle virtue of applying the same standards to similar matters, without self-serving shifts, a delicate form of *harmony in judgement*.
- 12 The common idea that rationality means ‘serving one’s own interests’ seems an aberration. Good reasons involve much more, such as an ability to put oneself in *someone else’s* place.
- 13 This tandem view again illustrates the earlier point of the duality of information and action.
- 14 Students ask why I still grade exams for correct answers, instead of their responses to their errors. For lack of a good answer, let me just step aside, and say that exams are an ancient Chinese invention that has conquered the world, on a par with paper, gunpowder, and silk.
- 15 E.g., while it might be provably inconsistent to *know* that a given statement is true and unknowable, it could well be consistent for an agent to *believe* that it has these two properties. Technical examples of this line are found in recent work on probabilistic versions of Gödel’s theorems.
- 16 Of course, both of these presuppositions are switched off when a teacher asks a question to a student, or a valedictory lecturer to his audience.
- 17 In mathematical terms, logical laws like this behave like ‘fixed-point equations’.
- 18 The status of common knowledge is widely debated: it may represent a useful fiction.
- 19 Public audiences often say that emotion rules the world, not logic. My response is finding lipstick on the collar of your spouse: observation, inference, and emotion entangle in a flash.
- 20 I sometimes wonder whether this balance also explains why mathematicians keep using natural language in their work and publications, be it enriched with symbols. It has often been asked why this is so. One answer might be that even pure mathematical research requires a constant undercurrent of signals for cooperation, competition, goals, and desires.
- 21 This has not prevented me from publishing papers on the interface of logic and psychology, a fascinating border region where much is happening, also at the University of Amsterdam.
- 22 More concretely, I am currently looking at simple reasoning patterns concerning the main drivers of our actions: hope, and fear — as well as what are in fact major cognitive abilities such as switching perspectives in going from first-person to third-person reports, and back. Throughout, the fundamental logical harmony of information and action reveals itself.
- 23 My sons occasionally wonder if I would not have made a much more productive member of Dutch society as a physicist, the field of study that I started with at the UvA.
- 24 Perhaps not accidentally, counting time zones, this is a realm where the sun never sets.