Rules Regularities Randomness

Festschrift for Michiel van Lambalgen

Rules, Regularities, Randomness

Festschrift for Michiel van Lambalgen

Rules, Regularities, Randomness

Festschrift for Michiel van Lambalgen

Edited by:

Keith Stenning

Martin Stokhof

Institute for Logic, Language and Computation

University of Amsterdam

Cover: Jaap van der Does
Layout: Martin Stokhof

Preface

This Festschrift has been a most enjoyable happening. 'Conversation' with Michiel is an endless inspiration. Amsterdam has been a conducive second home to several of us. It is wonderful to be included as an honorary Amsterdammer, though the credit as editor goes entirely to Martin. The credit for the coherence of this varied crew's contributions goes, of course, to Michiel, but also to Immanuel. If I had realised how crucial he was in establishing the coherence I read here, then he would certainly have been invited to contribute. Many of the conversations with Michiel were only possible through Immanuel. Sometimes an amateur but serious enthusiasm: Michiel's one altogether more serious.

An anecdote may illustrate two different attitudes to this interdisciplinarity. Oxford received in the late 40's demand for a psychology degree, or so the story goes. This was discussed by the University fathers with some concern. There was a strong feeling that a psychology degree might attract the mentally unstable. This worry was resolved by a committee member?s remark that if Psychology were combined with Philosophy, this would keep the unstable element well grounded – philosophy being the staple diet of the neighbourhood, and besides, the two subjects had absolutely no connections. Hence was created PPP – psychology, philosophy and physiology – with the option of avoiding psychology entirely.

The analogous tale for Amsterdam? A rumour suggests that Amsterdam had the greatest portion of its industry attributable to philosophy of any city in the world? So far perhaps two similar cities, and maybe Michiel is the lone catalyst for its reaching out? I doubt it. Amsterdam has a star-studded history of outward looking philosophy and linguistics, and Michiel's work is an illustrious continuation.

Keith Stenning

Contributors

Theodora Achourioti
University of Amsterdam

Natascha Alechina
University of Utrecht

Giosuè Baggio
Norwegen University of Science and Technology

Johan van Benthem
University of Amsterdam & Tsinghua University

Kees van Berkel
Technical University Vienna

Marian Counihan
University of Groningen

Jaap van der Does
Radboud University Nijmegen

Bart Geurts
Radboud University Nijmegen

Peter Hagoort
Radboud University Nijmegen

Machiel Keestra
University of Amsterdam

Sylvia Pauw
University of Amsterdam & Ghent University

Riccardo Pinosio
Avenade & Amsterdam University of Applied Sciences

Keith Stenning
University of Edinburgh

Martin Stokhof
University of Amsterdam & Tsinghua University

Leon van der Torre
University of Luxembourg

Chapter 1

The Kantian turn

Theodora Achourioti

On the occasion of Michiel van Lambalgen's Festschrift

Michiel has devoted a substantial part of his later research to studying Kant's philosophy and I am privileged to have collaborated with him in this extraordinary intellectual journey. After publishing an impressive record of highly original contributions in different areas from logic and mathematics to cognitive science and natural language semantics, becoming a Kant scholar was perhaps not the most predictable career step. Or was it? Looking closer, I conjecture that Kant was already there, even if implicitly, in Michiel's work. To show this, I will travel back to the years before 'A Formalisation of Kant's Transcendental Logic' (Achourioti & van Lambalgen (2011)). As my reference points I will take the two books, *The Proper Treatment of Events* (PTE) (van Lambalgen & Hamm (2005)) and *Human Reasoning and Cognitive Science* (HRCS) (Stenning & van Lambalgen (2008)) that Michiel co-authored with Fritz Hamm and Keith Stenning respectively.[1]

On its very first page, PTE introduces its reader to its topic by making what is very much a Kantian turn. It considers what is said to be a standard question that motivates research on the semantics of tense and aspect, namely, 'what must **the world** be like in order for tensed talk to make sense?'; only to replace it by 'what must **our minds** be like for tensed talk to make sense?' (emphasis mine). This new question is then announced as the point of departure for a book that argues tensed talk to be more complex than a reflection of a single 'earlier than' relation. The complexity, we are told, comes from our minds, not the world.

The reader cannot fail here to recall the famous passage in the preface of the B edition of the *Critique of Pure Reason* (Kant (1998)), when Kant compares his philosophical approach to Copernicus' heliocentric revolution: 'Up to now it has

1. Kant is not explicitly mentioned in PTE, and only briefly in HRCS.

been assumed that all our cognition must conform to the objects [...] let us once try [...] by assuming that the objects must conform to our cognition' (Bxvi).

Let us call this turn, the *cognitive turn*. In PTE it finds its expression in the philosophical position that time itself is a cognitive construct. This is an admittedly loaded position in the philosophical debate about the reality of time. It is also foundational for what is argued in the book. Because the same cognitive mechanisms employed for constructing time are also used to explain the semantics of tensed talk. - And yet, a hint on the irreality of time by Bertrand Russell used as a motto at the beginning of the book is said to be 'the last nod of philosophy in this book'. It is not.

On the perceptual level, the constructive element is evident in the transformations that temporal aspects, such as succession or duration, undergo, depending on empirical conditions (e.g. difference on so-called Stimulus Onset Asynchrony). But our experience of time, as consisting of past, present and future, goes well beyond the perception of changes and events in our immediate environment. PTE fleshes out a cognitive mechanism, *planning*, in order to explain temporal experience, for example, the way we conceptualise the future. This is a truly constructive mechanism: it is proposed not as a way to recover some temporal structure that is already in place, outside experience, but as a way to create one. Planning provides a semantics of tensed talk by giving an algorithm which constructs discourse models. And further than that, on the lexical level, concepts of planning are to be found grammaticalised, for example, in forms of the future tense. This cognitive grounding argues for a way of doing semantics that is not descriptive or corpus-based in a narrow sense but explanatory in its essence.

The cognitive turn is similarly prominent in HRCS which goes as far as reestablishing the relevance of logic to cognition and does so by redefining logic itself on the way. The position we find in PTE is repeated here, namely, that semantics is not given but constructed and this construction process now becomes part of logic itself. The notion of *logical form* is enlarged to capture the dynamics of an interpretation process that fixes the several parameters on which reasoning depends. This cognitive turn has then direct consequences for the normative status of logic. As the various case-studies discussed in HRCS nicely show, validity rules for inferences are no longer given. Making interpretation relevant and even necessary for logic distinguishes the *constitutive* part of 'reasoning *to* an interpretation' from its *regulative* pair of 'reasoning *from* an interpretation'; validity relies on those.[2]

It is an obviously robust Kantian theme running through PTE and HRCS alike, that data, whether sensory input in experience or linguistic expressions, need cognitive processing and interpretation before they can be said to have meaning. Just like our experience of the past, present and future is not the outcome of

2. I explored this theme further in my Master of Logic thesis 'Logic, Normativity, and the A Priori', under Michiel's supervision.

recovering some pre-existing structure but stems out of cognitive processes such as planning, reasoning itself relies on the interpretation of data which carry no (normative) meaning in and by themselves. 'In Kantian terms, we may think of the activity of imposing logical form and integrating the premises in a single representation as **synthesis**; this synthesis is **a priori** since the logical form imposed is not determined by experience, but a constraint contributed by cognition. One needs logical form in order to be able to extract information, but it is as little given in the data as an edge is given in the retinal array' (HRCS p.351, emphasis mine). Interpretation is underdetermined by the data; the 'myth of the given' is dispelled.[3]

Right at the centre, laying the basis for PTE and HRCS, is the emphasis of focus on the continuous processes of active engagement and interaction of the cognitive subject with its environment, an emphasis that echoes Kant's dynamic and intimate relation between *spontaneity* and *receptivity*. Kant is careful to neatly separate the different faculties involved in cognition. He writes of *sensibility* and *understanding* that 'one must not mix up their roles, rather one has great cause to separate them carefully from each other and distinguish them' (B76/A52). However, Kant's famous 'thoughts without content are empty, intuitions without concepts are blind' (B75/A51) is better understood not as a conjunction of two statements but as expressing the reciprocal relation of continuous interdependence of the cognitive (sub)faculties involved, with all the complexity that this entails.[4]

The formalisms proposed in both PTE and HRCS are well equipped to account for the continuous and dynamic processing of input as this is, or *becomes*, available. They can do this by licensing non-monotonic inferences based on closing the space of possibilities ('the world') to include only the contingencies that one anticipates. Since information cannot be presented all at once, being able to revise one's conclusions simply makes language comprehension and reasoning possible. PTE's *planning* is an intrinsically non-monotonic process because one typically has to rewrite a scenario of how to achieve a certain goal and adjust to new data. HRCS's *reasoning to an interpretation* is similarly non-monotonic, as one sees illustrated in reasoning tasks such as the suppression task. But it is *not* the non-monotonicity by itself of the formalisms (of which there are plenty) that distinguishes this work for its Kantian merits, it is grounding these particular formalisms on how human cognition and psychology of reasoning work.[5]

---

3. For the origins of the phrase, see Wilfrid Sellars' famous 'Empiricism and the Philosophy of Mind' paper of 1956.

4. The chapter on Schematism is the place where this interdependence and synergy becomes most clear.

5. Characteristic of both PTE and HRCS is that the proposed underlying cognitive processes correlate with a computational theory that is tractable, hence appropriate for applications in AI. Whether this correlation reflects a more general computational view of the mind, or to which extent, we leave here as an open question.

The cognitive correlates of formal principles such as the closed world assumption or semantic ones such as the immediacy of minimal models are manifestations of what Kant calls *spontaneity*.[6] Spontaneity plays a central role in Kant's theoretical philosophy, as well as his moral philosophy; it is intimately related to the notion of *causality* which can be thought of 'according to nature or from freedom' (A530/B559). Kant defines spontaneity as 'the faculty for bringing forth representations **itself** (B75/A51)' and as able to 'begin a series of occurrences **entirely from itself**' (emphasis mine) (A534/B562).

PTE and HRCS can largely be read as defying and dissolving traditional paradigms of posing *externally* set boundaries as the conditions for having a semantics or reasoning at large. For example, the sentence is no longer the natural unit of computing a semantics, rather the computation is a continuous process that informs and readjusts the representation on the go. Similarly, processing information and drawing inferences does not wait for some natural point to start; it happens all the time. In Kant's words, 'But since [...] no absolute totality of conditions in causal relations is forthcoming, reason creates the idea of spontaneity, which could start to act from itself, without needing to be preceded by any other cause that in turn determines it to action according to the law of causal connection' (A533/B561).

We would need a longer discussion to elaborate and carefully establish the connection between Kant's notion of causality and the cognitive grounds for semantics and reasoning. But what is especially remarkable, and not an anachronism I believe, is that in his Antinomy of Reason chapter, Kant anticipates what has to be explained once cognition is given this primary role. Taken out of context, Kant's words could be used to describe the kind of non-monotonic logics that we see employed in PTE and HRCS. We end this brief journey with the critical passage that speaks for itself: '[...] reason does not give in to those grounds which are empirically given, it does not follow the order of things as they are presented in intuition, but with complete spontaneity it makes its own order according to ideas, to which it fits the empirical conditions and according to which it **even declares actions to be necessary that yet have not occurred and perhaps will not occur**, nevertheless presupposing of all such actions that reason could have causality in relation to them; **for without that, it would not expect its ideas to have effects in experience**' (A548/B576, emphasis mine).

To conclude, nothing said here is really new, or so I hope. My reference did not go beyond PTE and HRCS, and even in these two books there are many more Kantian connections to be made than the ones noted here. HRCS hides a

6. There is a discussion in the Kantian scholarship of whether *spontaneity* resides in understanding alone or is present in intuition as well. I side with the latter reading for reasons having to do with the intricate interaction of the various faculties involved in cognition which I think gives a more faithful representation of Kant's views.

promise of another book which 'would be required to explain the many affinities with Kant' (p.9). Retirement must be the time to hold Michiel to it.

Chapter 2

Quantifiers, sequents, events

Natasha Alechina

Michiel van Lambalgen worked, and continues working, in many areas of mathematics, logic, cognitive science and philosophy. I would like to write about two of the topics I worked on together with Michiel. These two topics are very different, and are only a small fraction of the areas he worked in and obtained beautiful results. The first topic is on sequent calculi for generalised quantifiers, and the second one is explicating semantics of natural language (Russian verbs) in event calculus.

Reasoning with generalised quantifiers made easy

I start with a brief introduction to generalised quantifiers. A generalised quantifier **Q**, as defined by Mostowski in Mostowski (1957), is a class of subsets of the universe, so that a model M satisfies $Qx\varphi$ if the set of elements $\{e : M \models \varphi[x/e]\}$ is in **Q**. Examples of generalised quantifiers are: the ordinary existential quantifier (interpreted as the set of all non-empty subsets of the universe); the quantifier 'there are exactly 2'; a free filter quantifier (where **Q** is a free filter, that is, it is closed under finite intersections, upward closed for set inclusion, and the intersection of all its members is empty, which intuitively means the sets in **Q** are infinite and 'large'), 'there are uncountably many' (where the domain is uncountable, and **Q** contains all uncountable subsets), etc. Mostowski's original definition required that the generalised quantifier is invariant under permutations of the universe, thus restricting attention to quantifiers related to cardinality. Subsequently, other generalised quantifiers were considered which do not have the property of permutation invariance, such as topological quantifiers or measure quantifiers. An example of the latter is the quantifier 'for almost all', which contains all subsets of measure 1 of $[0, 1]$. An overview of generalised quantifiers can be found in a

collection 'Model-Theoretic Logics', edited by Barwise and Feferman, Barwise & Feferman (1985), and in Westerståhl's book Westerståhl (1989).

Michiel van Lambalgen became interested in 'for almost all' quantifier in the course of his work on axiomatisation of randomness van Lambalgen (1990). In van Lambalgen (1990), he obtained an axiomatisation of von Mises' notion of a random sequence and as a side effect, quantifier elimination for the 'for almost all' quantifier. The technique he used there (which I will explain in a moment) promised a way to make reasoning about certain kind of generalised quantifiers easier, by providing a natural deduction style reasoning formalism. The quantifiers in question are $Qx\varphi(x)$ that, intuitively speaking, say that the set of elements satisfying φ is 'big', such as 'for almost all', 'for all but countably many', etc. They have the flavour of a universal quantifier $\forall$ in first order logic, but there is an important difference. In first order logic, we can eliminate x in $\forall x\varphi(x)$ on any domain element: if $\forall x\varphi(x)$ is true, then $\varphi(e)$ holds for any e in the domain. This clearly does not hold for a generalised quantifier that says that most elements satisfy $\varphi(x)$; if $Qx\varphi(x)$ holds, $\varphi(e)$ may still be false for a small number of e's. Reasoning in first order logic is made much easier by being able to eliminate quantifiers, both $\forall$ (on any element) and $\exists$ (using Skolem functions). If a similar technique can be developed for generalised quantifiers, reasoning with them is also made much easier.

In 1991, Michiel van Lambalgen wrote a paper on natural deduction for generalised quantifiers van Lambalgen (1991) (it only appeared in formal proceedings in 1996). The idea which came from his earlier work van Lambalgen (1990) was to look at the generalised quantifiers $Qx\varphi(x,\bar{y})$ as binding a special kind of variables, whose range is restricted by the values of the free variables of the formula, $\bar{y}$. Michiel proposed to model this restriction as an *independence relation* R of indefinite -arity between the quantified variable x and the rest of the free variables of the formula $\bar{y}$. He defined a satisfiability preserving translation $*$ from the language with a generalised quantifier $L(Q)$ into first order language $L(R)$ are follows:

i) $*$ is the identity on formulas not containing Q

ii) $*$ commutes with $\vee$, $\wedge$, $\neg$, $\rightarrow$, $\forall$, $\exists$

iii) $(Qx\varphi(x,\bar{y}))^* :\equiv \forall x(R(x,\bar{y}) \rightarrow \varphi(x,\bar{y})^*)$

where $\bar{y}$ are exactly all the free variables in $\varphi(x,\bar{y})$ besides x. Note that the range of this translation corresponds roughly to the guarded fragment of first order logic defined by Andréka, van Benthem and Németi (Andréka et al. (1995)). [1] It also resembles modal logic, which was famously shown by Johan van Benthem (van Benthem (1983)) to be translatable into first order logic with a binary accessibility relation R:

$$ST_x(\Box\varphi) :\equiv \forall x(R(x,y) \rightarrow ST_y(\varphi))$$

1. In the guarded fragment, formulas are of the form $\forall x(R(x,\bar{z}) \rightarrow \varphi(x,\bar{y})^*)$ where $\bar{y} \subseteq \bar{z}$.

This idea of Michiel's inspired the topic of my thesis, supervised by Michiel van Lambalgen and Johan van Benthem, which was an investigation of 'modal quantifiers' with the following semantics:

$$M, s \models \Box_x \varphi(x, \bar{y}) \Leftrightarrow \forall d(R(d, s(\bar{y})) \rightarrow M, s^x_d \models \varphi(x, \bar{y}))$$

where s is a variable assignment, and $\bar{y}$ are precisely the free variables of $\Box_x \varphi$. The dual $\Diamond_x$ of $\Box_x$ is defined as $\Diamond_x \varphi :\equiv \neg \Box_x \neg \varphi$ and has the following semantics:

$$M, s \models \Diamond_x \varphi(x, \bar{y}) \Leftrightarrow \exists d(R(d, s(\bar{y})) \wedge M, s^x_d \models \varphi(x, \bar{y}))$$

The minimal logic of such quantifiers, corresponding to a subset of the guarded fragment, was proved decidable in Alechina (1995b).

Natural deduction rules for $Qx\varphi(x, \bar{y})$ essentially eliminated x on variables independent from $\bar{y}$, denoted in the syntax by $x_{\bar{y}}$, and called *indexed variables*. Indexed variables have the form $x_{\bar{y}}$, where $\bar{y}$ may also be indexed variables. The variable $x_{\bar{y}}$ ranges over objects in relation R to $\bar{y}$.

Michiel initiated work on a Gentzen-style sequent calculus with left- and right-introduction rules for generalised quantifiers with indexed variables van Lambalgen (1993), which we continued together in Alechina & van Lambalgen (1996, 1995). Briefly, in a Gentzen-style sequent calculus, $\Longrightarrow$ is an entailment relation, and a sequent $A_1, \ldots, A_n \Longrightarrow B_1, \ldots, B_k$ means that if all A_i are true, then at least one of B_j is true. Inference rules show how to introduce logical connectives on the left and on the right of $\Longrightarrow$. For example, the following rules for negation, where Γ and Δ are set of formulas, say that if from Γ and φ it is derivable that Δ, then from Γ is derivable that $\neg\varphi$ or Δ (right introduction rule for $\neg$), and if from Γ it is derivable that φ or Δ, then from Γ and $\neg\varphi$ it is derivable that Δ:

$$\frac{\Gamma, \varphi \Longrightarrow \Delta}{\Gamma \Longrightarrow \neg\varphi, \Delta} \quad \neg r \qquad \frac{\Gamma \Longrightarrow \varphi, \Delta}{\Gamma, \neg\varphi \Longrightarrow \Delta} \quad \neg l$$

The rules for the existential quantifier are:

$$\frac{\Gamma \Longrightarrow \psi(x), \Delta}{\Gamma \Longrightarrow \exists x \psi(x), \Delta} \quad \exists r \qquad \frac{\Gamma, \psi(y) \Longrightarrow \Delta}{\Gamma, \exists x \psi(x) \Longrightarrow \Delta} \quad \exists l$$

where in $\exists l$ y does not occur free in Γ and Δ. There are additional rules that are called *structural rules* that are not concerned with logical connectives, but with the structure of the sequents, for example that repeated occurrences of a formula do not matter:

$$\frac{\Gamma \Longrightarrow \varphi, \varphi, \Delta}{\Gamma \Longrightarrow \varphi, \Delta} \quad CONr \qquad \frac{\Gamma, \varphi, \varphi \Longrightarrow \Delta}{\Gamma, \varphi \Longrightarrow \Delta} \quad CONl$$

Logics defined in terms of absence of structural rules, such as linear logic, are called *substructural logics*.

In Michiel's approach to proof theory for generalised quantifiers, the logics of generalised quantifiers are viewed as substructural logics, where the left and right introduction rules for the quantifiers remain constant, and the structural rules which deal with the indexed variables are modified. For the details, and for discussion of the full logic of indexed variables, the reader is referred to Alechina & van Lambalgen (1996) (where the independence relation was renamed to be the dependence relation, for reasons that are shrouded in the mists of time). Below are the left and right introduction rules for $\Diamond_x$ of the proof system for generalised quantifiers:

$$\frac{\Gamma \Longrightarrow \psi(x_{\bar{z}}, \bar{z}), \Delta}{\Gamma \Longrightarrow \Diamond_x \psi(x, \bar{z}), \Delta} \ \Diamond r \qquad \frac{\Gamma, \psi(x_{\bar{z}}, \bar{z}) \Longrightarrow \Delta}{\Gamma, \Diamond_x \psi(x, \bar{z}) \Longrightarrow \Delta} \ \Diamond l$$

where in $\Diamond l$ $x_{\bar{z}}$ does not occur free in Γ and Δ, also not in indices, and in $\Diamond r$ $x_{\bar{z}}$ should occur free in Γ or Δ.

The corresponding calculus was called L_{triv} in Alechina (1995a), also for reasons that escape me at the moment. It is shown in Alechina & van Lambalgen (1996) that to make $\Diamond$ in the calculus with indexed variables to behave as the ordinary existential quantifier, one needs to add the following *substitution rule*:

$$\frac{\Gamma \Longrightarrow \psi(t), \Delta}{\Gamma \Longrightarrow \psi(s), \Delta} \ \mathit{SUB}$$

where s and t are any variables; the restriction on *SUB* is that t does not occur free in Γ and Δ. (Note that

$$\frac{\Gamma, \ \psi(t) \Longrightarrow \Delta}{\Gamma, \ \psi(s) \Longrightarrow \Delta}$$

with the same restriction on t, is derivable from *SUB* and the rules for negation.)

SUB is a structural rule since it does not involve any logical connectives.

In between L_{triv} and L_{triv} + *SUB*, there is a whole class of substructural logics with respect to the substitution rule. The examples below show that modifications of the substitution rule are made possible by the fact that the variables have internal structure.

The weakest system considered in Alechina & van Lambalgen (1996) contains the following substitution rule:

$$\frac{\Gamma \Longrightarrow \psi(x_{\bar{z}}), \Delta}{\Gamma \Longrightarrow \psi(x'_{\bar{z}}), \Delta} \ \mathit{SUB}_{av}$$

(given that $x_{\bar{z}}$ does not occur in Γ, Δ).

SUB_{av} corresponds to the principle of renaming bound variables (taking alphabetic variants). $L_{triv} + SUB_{av}$ formalises the minimal logic of $\Box_x$.

This system is still rather weak. For example,

$$\Diamond_x \varphi(x, \bar{y}) \to \Diamond_x(\varphi(x, \bar{y}) \vee z = z)$$

(monotonicity of $\Diamond$) is not derivable in $L_{triv} + SUB_{av}$.

The first standard generalised quantifier, namely the filter quantifier $\neg\Diamond_x\neg$ is obtained by strengthening SUB_{av} to

$$\frac{\Gamma \Longrightarrow \psi(x_{\bar{u}\bar{z}}, \bar{z}), \Delta}{\Gamma \Longrightarrow \psi(x'_{\bar{v}\bar{z}}, \bar{z}), \Delta} \quad SUB_{ext}$$

(given that $x_{\bar{u}\bar{z}}$ does not occur free in Γ, Δ). Observe that this rule allows to prove the monotonicity principle. The rule means that only the free variables $\bar{z}$ matter in the index, while other variables $\bar{u}, \bar{v}$ can be added or removed.

The characteristic axiom of the 'for almost all' quantifier

$$\Box_x \Box_y \varphi \to \Box_y \Box_x \varphi$$

corresponds to the following substitution rule:

$$\frac{\Gamma \Longrightarrow \varphi(y_{\bar{z}}, x_{y_{\bar{z}}\bar{z}}, \bar{z}),\ \Delta}{\Gamma \Longrightarrow \varphi(y_{x_{\bar{z}}\bar{z}}, x_{\bar{z}}, \bar{z}),\ \Delta}$$

where both $y_{\bar{z}}$ and $x_{y_{\bar{z}}\bar{z}}$ do not occur free in Γ or Δ.

For determining such substitution rules, and proving their interderivability with the axioms of the generalised quantifiers, one can benefit from knowing to what condition on the relation R the axiom corresponds. For example,

$$\Box_x \Box_y \varphi \to \Box_y \Box_x \varphi$$

corresponds to

$$R(y, \bar{z}) \wedge R(x, y\bar{z}) \to R(x, \bar{z}) \wedge R(y, x\bar{z}).$$

The correspondence theory of generalised quantifiers (correspondence between axioms and the properties of R) is studied systematically in Alechina (1995a), but was already initiated in van Lambalgen (1991). Michiel also proved one of the crucial theorems needed for the correspondence and completeness theory of $\Box_x$, Theorem 4.2.10 in Alechina (1995a) (where it is attributed to van Lambalgen 1994). I wish I could reproduce here the original calligraphic proof of this theorem that Michiel gave me for my birthday in 1994, but unfortunately I do not have access to it because of the pandemic.

Event calculus and Russian verbs

Another topic I was fortunate to work on together with Michiel was using event calculus to provide semantics for *aspectual pairs* of Russian verbs in event calculus, along the lines of the approach developed by Michiel and Fritz Hamm in van Lambalgen & Hamm (2005). Since the work is joint, and so far unpublished (we had to stop working on it for health and other reasons in late 2003-2004), I am not going to present any details in this single-authored contribution, and just give some highlights and motivation.

The event calculus was developed by Murray Shanahan (Shanahan (1990)). It requires a many-sorted first order logic with sorts for the following:

1. individual objects
2. real numbers, to represent time and variable quantities
3. time-dependent properties, such as states and activities
4. variable quantities, such as position
5. event types, whose instantiations mark the beginning and end of time-dependent properties.

Time-dependent properties are called *fluents*. A fluent is a function which may contain variables for individuals and reals, and which is interpreted in a model as a set of time points. Fluents are initiated and terminated by events, and may hold at various time points. If f is a variable over fluents, e a variable over events, and t a variable over time points, the following predicates can be used to talk about this:

Initially(f): f holds at the beginning of time
Happens(e, t): e happens at time t
Initiates(e, f, t): f is initiated by e and begins to hold after (but not at) t
Terminates(e, f, t): f is terminated by e (holds at t but does not hold after t)
HoldsAt(f, t): fluent f is true at time t

In van Lambalgen & Hamm (2005), event calculus was used to give a computational semantics to natural language phenomena, in particular tense. We attempted to apply it to semantics of Russian verbs. There is a considerable controversy surrounding the meaning of the corresponding perfective and imperfective verbs in Russian (Comrie (1976); Forsyth (1970); Zaliznyak & Shmelev (2000)). Aspect plays a very important role in Russian grammar. Each verb is either perfective (perf) or imperfective (imp); it can be classified as being perfective or imperfective without any context, just by looking at its infinitive. Children at school are taught to check whether the verb can be used to answer the question 'что делать?' ('to do what?') in which case it is imperfective, or 'что сделать?' (perfective form of 'do'; roughly, 'to have done what?' or 'to achieve what?'). This classification of verbs into perfective and imperfective is very easy for a native speaker. However, the formal definition of the difference in meaning between perfective and imper-

fective verbs, is still a difficult and contested topic in Russian linguistics. It is also rather difficult to define formally when two verbs constitute an aspectual pair, i.e. a perfective/imperfective pair of verbs with the same meaning (or, perfective and imperfective forms of the same verb, depending on who you are talking to).

In many texts (including the textbooks I used when at primary school) perfective meaning is more or less identified with perfect tense: perfective verb denotes an action which has a clear result, while imperfective verb denotes a process or a state. Oxford Russian Grammar and Verbs Wade (2002), p.111 gives the following commonly used definition of imperfective and perfective aspects in Russian:

> The **imperfective** aspect denotes:
> (a) an action that was, is, or will be in progress ('he was, is, will be ringing'), or
> (b) a **repeated** or **habitual** action ('he used to ring, rings, will ring').
>
> The **perfective** aspect indicates **completion** of an action in the past ('he made, has made, had made a call') or **intention to complete** an action in the future ('he will make a call, will have made a call'). A **result** is often implied (e.g. a message has been passed on, information is now available, etc.).

However, there are problems with this definition. Perfective verbs are used to refer to actions in the future which by definition have not produced any result yet. They can also refer to activities in the past which do not have any clear 'result', for example 'Yesterday we dined at 6' will be translated using a perfective verb 'пообедали' for 'dined'. It is possible to use a perfective verb to describe actions which are no longer relevant in the present: 'Мы пообедали (perf) в шесть, потом пошли (perf) погулять и вернулись (perf) домой в девять.': 'We had dinner at six, went for a walk and came back home at nine'. It is difficult to render something like this in present or past perfect.

In 1946, Maslov suggested that verb1 (imp) and verb2 (perf) form an aspectual pair if for any sentence we can replace verb2 in past tense with verb1 in present tense and get the description of the same event (the latter sentence describes it in present historical tense), for example:

- Однажды он позвонил (perf) и сказал (perf) ... (One day he called me and said ...)
- ? Однажды он звонит (imp) и говорит (imp) мне ... (One day he calls me and says ...)

Since the present historic tense plays a crucial role in Maslov's criterion of aspectual pair, we attempted to formalise its meaning. First of all it should be noted that 'present historic tense' is not a special form of a verb; it is one of the

meanings of the imperfective present. Forsyth (Forsyth, 1970, p.150-151) explains the meaning of the present historic as follows:

> The 'historic' use of present tense forms to express actions which took place in the past is metaphoric, its aim being to present past events as if they were being witnessed at the 'moment of speaking'. It is therefore essentially the effect of the 'real' present which the speaker or writer attempts to produce when he switches from the perspective view of past tense narration to the 'immediate' viewpoint of the historic present. The narrator as it were 'scans' past events from a moving observation point, so that, as in 'live' running commentary, the sequence of verbs implies the sequence of events reported...

Forsyth illustrates this by a quote from a story by Bulat Okudzhava:

> Вдруг наш полуторка останавливает ся (imp). Впереди дорога пуста. Только далеко-далеко какой-то одинокий маленький солдатик стоит (imp) и смотрит (imp) в нашу сторону. Старшина спит (imp). Мы с Сашкой соскакиваем (imp) на дорогу. ... А солдатик бежит (imp) к нам... Вот он подбегает (imp) к нам, и я вижу (imp), что это девоцхка. –"Подвезите, ребята..." – Я помогаю (imp) ей взобраться в кузов... Наш газик наконец трогает ся (imp). (Окуджава: Будь здоров, школяр.) Suddenly our one-and-a-half tonner stops. The road in front is empty. Only away in the distance there is a lonely little soldier standing looking towards us. The sergeant is asleep. Sashka and I jump down on to the road... And the soldier runs towards us... He comes running up to us and I see it's a girl. 'Give me a lift, boys.' I help her to climb into the truck... At last our GAZ truck starts moving.

The essence of the present historic tense is thus the existence of a moving observation point. To model the present historic tense in the event calculus, we introduced a variable, moving reference point or interval $now(x)$ (x a positive integer), satisfying the following conditions, where $?HoldsAt(f,t)$ denotes a query (whether $HoldsAt(f,t)$ holds):

1. $Initially(now(0))$
2. for each natural number x, if $?HoldsAt(now(x), r),\ r < now$ succeeds, and
   $?HoldsAt(now(x), t),\ t > now$ fails, then $?HoldsAt(now(x+1), s),\ s < now$ succeeds
3. $?HoldsAt(now(x), t),\ HoldsAt(now(y), s),\ t > s,\ x < y$ fails
4. a fluent $\overline{now}$ is defined by the clause

$$HoldsAt(now(x), t) \ \rightarrow \ HoldsAt(\overline{now}, t).$$

The point of this definition is to generate a supply of *now*'s, represented as fluents. Each *now*(n) (for positive integers n) determines a (possibly empty) half-open interval, not a point, and no interval of the sequence lies completely to the right of the true *now*. The fluent $\overline{now}$ then collects the intervals corresponding to the *now*(n). We associate imperfective verbs, represented by fluents, with the *now*'s. The present historic differs from the present tense in that the *now*'s are temporally extended and situated in the past, even though each fluent is presented as holding *now*. This modelling was used to prove coercion from process to event in present historic time, giving an event calculus semantics for Maslov's definition of aspectual pair.

Aspectual pairs are classified into different types. For example,

- proleptic pairs: imperfective verb denotes a *state* which results in the event denoted by the perfective verb: опаздывать (imp) - опоздать (perf) (to be late). The semantics of proleptic pairs can be defined in the event calculus semantics as follows. Imperfective denotes a parametrised fluent $f(x)$ such that

$$HoldsAt(f(x), t) \wedge x > c \rightarrow Happens(e, t)$$

  where e is the denotation of the perfective verb and c a constant (such as the deadline in 'being late').
- gradation pairs: the imperfective verb denotes a process of change, while the perfective verb denotes the confirmation of the fact that the process of change took place: стареть (imp) - постареть (perf) (to get older), повышаться (imp) - повыситься (perf) (to rise, about e.g. temperature); The semantics of gradation pairs can be defined in the event calculus semantics as follows. Imperfective denotes parametrised fluent $f(x)$ such that

$$HoldsAt(f(x), s) \wedge HoldsAt(f(y), t) \wedge y - x > c \rightarrow Happens(e, t)$$

  where e is the denotation of the perfective verb.

An unusual feature of Russian aspect is the way perfective verbs are produced from imperfective ones using prefixes and (occasionally) suffixes. The meaning of prefixes can be summarised in several broad categories, which we gave event calculus semantics for. This part of the project was particularly fascinating for me. Some examples of prefixes, their meaning, and event calculus semantics are below.

- inceptives (за-, вос(воз)-, вс(вз)-), referring to the beginning of action, for example зазвонить to start ringing (a bell). In event calculus, $Initiates(begin(f), f, t)$ (where $begin(f)$ is the event denoted by the perfective verb, and f the fluent denoted by the imperfective verb).

- terminatives (от-, до-), referring to the termination of action, for example отзвонить, to finish ringing, to come to an end of ringing, догореть, to burn out. In event calculus, $Terminates(end(f), f, t)$.

Conclusion

I hope that the couple of examples I gave above illustrate how wide ranging Michiel's interests and scientific contributions are. It remains to say that it was always a great pleasure to work with Michiel, and I am very lucky that he was my PhD supervisor.

Chapter 3

Logic in psycholinguistics?

Giosuè Baggio

"Formal semantics is entirely unhelpful for understanding the psychology of language." (Adele Goldberg)

"Formal semantics is pure externalist description. It has nothing whatsoever to say about mental states. […] (It could be in principle married to anything, but it usually chooses to remain separate)." (Gillian Ramchand)

"There's an incompatibility between a referential semantics and gen[erative] gr[ammar], but if you see the models as mind internal, at least some variants of formal semantics survive (though not possible worlds as total ontological units [. . .])." (David Adger)

These are three fragments of a Twitter conversation that, coincidentally, took place on the day Martin and Keith wrote us to share their plan for a Festschrift for Michiel. This coincidence, and Goldberg's pungent tweet in particular, brought my memory back to 2007, when she gave the Nijmegen Lectures at the Max Planck Institute for Psycholinguistics. Michiel was one of the discussants of Goldberg's first lecture on 'The constructionist approach to language'. In his commentary talk, he argued that the meaning of some constructions (e.g., Ditransitive and Caused Motion) could be captured using the computational semantics of tense and aspect he had developed with Fritz Hamm. In my recollection, Goldberg's response was similar to her recent tweet: formal semantics has nothing to say about the psychology of language, so it makes little sense to try to wed it to theories of syntax that strive to be cognitively relevant and plausible, like Construction Grammar. Or indeed Generative Grammar. The tweets by Ramchand and Adger reveal similar suspicions about the possibility of effectively bringing formal semantics (back?) under the auspices of philosophical mentalism and cognitive psychology.

And they are right, at least for some values of 'formal semantics'. Theories of meaning based on possible worlds, truth conditions etc. were designed as "pure externalist description". No wonder that now we cannot fit simple or complex (compositional) meanings 'in the head'. Ramchand and Adger, however, seem to leave the door open: perhaps formal semantics could be "married to anything"; perhaps "variants of formal semantics survive", if we view "models as mind internal". My two cents to that recent Twitter thread were a link to the Hamm, van Lambalgen, & Kamp article, 'There is no opposition between formal and cognitive semantics' (Hamm et al. (2006)), and a reminder that *we compute meanings*, so there must be a formal theory that models the process and result—a theory that might not look like anything we like to call 'Formal Semantics' right now, one that might not be uniform and coherent (much like current formal semantics isn't), but that will still be formal (computational, algorithmic) and still semantics (a theory of linguistic meaning).

Fifteen years have passed since 2006/7, eighteen since I first met Michiel. I came to the ILLC in early 2003 as a Master of Logic student, fresh out of my philosophy degree in Pavia, one term late (the MoL year had started in the fall 2002), and with a letter of presentation by Daniele Mundici that read: "His weakest points are possibly due to a lack of systematic training in mathematics, as is usually the case in Italy, even among the best students in philosophy". I had taken two introductory Logic courses in Pavia, but only in the second, taught by Mundici at Collegio Ghislieri in 2001/2, I was steadily coming to grips with it, thanks to Daniele's limpid classes and lecture notes (we were guinea pigs for his 'metodo breve', now published as 'Logic: A Brief Course', (Mundici (2012))). That feeling of progress, even understanding, was shattered shortly after my arrival in Amsterdam. I dropped out of Yde Venema's Modal Logic classes after just a couple of weeks ("I agree that this may be a bit too hard for you: it assumes a lot of maths"), and I started to gravitate towards Nieuwe Doelenstraat. I binged on Amsterdam-style formal semantics and pragmatics, in courses by Veltman, Stokhof, Dekker, and van Rooij, and I also took Michiel's classes on Logic and Cognition and the Psychology of Reasoning. His lectures had just enough logic, and logic that was familiar to me (I had learned resolution for the propositional calculus with Mundici, and it was comforting to find it applied in Michiel's works with both Keith and Fritz), yet used in new and exciting ways as a cognitive modelling tool. For our term papers, we had to pick an experimental study in the psychology of language or reasoning and try to do what Michiel and Keith had done with Wason's selection task and the suppression task: deconstruct it using logic or formal semantics, spot confounds in the experimental design, find alternative explanations. My assigned article ('When temporal terms belie conceptual order', an ERP study by Muñte, Schiltz, and Kutas, Münte et al. (1998)) had an obvious confound that was, however, only visible through the lens of formal semantics.

If even a *Nature* paper yielded to the logician's tools, Michiel's research program surely had huge potential, I thought.

Then the time came to pick a thesis topic. Michiel was in touch with Peter Hagoort ("He's a psycholinguist, truly world class") and plans for collaboration were taking shape between them. Would I want to go to Nijmegen and write a thesis at the new F.C. Donders Centre for Cognitive Neuroimaging? Sure. Was I "going to show it's all in the head?", Frank Veltman joked at the MoL's end-of-year drinks. At that time and for a long while afterwards, it just seemed convenient to bracket off such questions and what I had learned about the hazards of mentalism in semantics while writing my thesis on Wittgenstein in Pavia. Meanwhile, I would trust Michiel's arguments: one section of the last chapter of 'Human Reasoning and Cognitive Science' (Stenning & van Lambalgen (2008))is titled 'Some Information-Processing is Best Viewed as Logic'. In a more recent interview, he says, with a doubly anti-realist connotation: "We apply higher level descriptions in order to organise our sensory data. Some of these higher level descriptions can be profitably taken to be logical formalisms, but not all of them." So, my ILLC thesis was not only a thesis: I was on a mission to test whether a little logic really did go a long way, whether it could be of any use in designing psycholinguistic experiments, deriving predictions, and interpreting data. Retrospectively, the thesis was very far from the integrative success I had hoped it would be, but fortunately I had a second chance with a PhD with the same team of supervisors: Michiel the logician and Peter the psycholinguist. The intellectual and logistical back-and-forth between Nijmegen and Amsterdam lasted until 2009, but 2003 was the year the parameters were fixed for me: Michiel set me on a 'collision course' with some of the most fascinating and difficult problems in cognitive science, and I am still grateful for that.

One problem is what role logic and formal semantics can play in psycholinguistics, particularly in theories of language processing. Language users compute meanings. What is the relation between those 'live' meanings and the meanings that a formal semantic theory assigns to expressions of a language (recursively, compositionally etc.)? Barbara Partee (e.g., in Partee (2018), and earlier work) asked whether semantics may be reconciled with internalism—whether we can view 'semantics as psychology'. A compositional theory of meaning is finitely stateable, like theories of phonology and syntax. In principle, there are no obstacles to asking how composition operators are algorithmically and neurally implemented and to pursuing research programs that can answer those questions. Perhaps unsurprisingly, there is no single neural event that corresponds to (or implements) syntax-driven, logico-semantic composition à la Functional Application (Pylkkänen (2020), Baggio (2021b)). Composition seems to decompose into a small set of operations, carried out by different cortical networks, each engaged

depending on (yet unknown) properties of the input, the context, and ongoing internal states ((Olstad et al., 2020, p19), Baggio (2021a)).

This seemingly fragmentary picture of meaning composition may reveal its unity in light of a Parallel Architecture in which complex meanings can be computed under the constraining force of linguistic grammar (yielding compositional structures) or autonomously (Culicover & Jackendoff (2006)): these two types of computation may unfold simultaneously, in parallel. The computational role of logic in an architecture of this sort would seem to be restricted to cases in which grammar fully determines composition: imagine cognitively plausible versions of Chierchia & McConnell-Ginet (Chierchia & McConnell-Ginet (2000)) or Heim & Kratzer (Heim & Kratzer (1998)), if that is possible. But logic can provide principles for computing meanings that do *not* mirror the grammar. This possibility resonates well with Michiel's own work—for example, the rich system of constraints of Event Calculus, which yields a non-strictly-compositional theory of tense and aspect, and the idea of 'reasoning to an interpretation', or as Michiel once put it to me: "Logical reasoning is a kind of discourse integration".

Perhaps the most valuable lesson of Michiel's work for the psycholinguist concerns logic as part of a theory of semantic representation, whether or not we also believe that logic has something to say about semantic processing. Language understanding is essentially the construction of a *model* that makes discourse true. This insight, as such, is not new ("In order to understand what another person is saying, you must assume it is true and try to imagine what it could be true of": G.A. Miller's Law), but logic can render it precise, via notions of minimality and partiality. This is a starting point for viewing "models as mind internal". It also has testable implications, such as that minimal models may be both extended and recomputed, depending on what new information becomes available. I recall discussions with Michiel during my PhD about this: what are the processing consequences of minimal models for particular linguistic constructions (e.g., the progressive), and how can one test them? What do other theories predict, such as dynamic theories based on partial models (e.g., DRT), or static theories based on classical total models?

At that time, we thought—like Goldberg—that truth conditional semantics may not have much to say about the internal process of natural language interpretation qua model construction. Specifying the truth conditions of a sentence or discourse is not a plausible "definition of the information processing problem, whose solution is the goal of the computation", nor does it give a correct "characterisation of the abstract properties of the computation" (Poggio (1981)). I now suspect there is a possible role for formal semantics, even in its more traditional incarnations as "pure externalist description", in a neurocognitive theory of meaning. According to Marr and Poggio, an additional goal of a computational-level theory is the description of "properties of the (...) world that constrain the compu-

tational problem" (Poggio (1981)). Models may be 'mind internal', but they draw from and can be embedded in larger reference structures that are not entirely arbitrary cognitive constructions. It is one task of a computational-level theory of meaning to state what elements of those structures (e.g., space, time, situations etc.) are made available for cognitive computation, how they are exploited in particular processing circumstances, and how they constrain the computational problem and its solution. Michiel's work has shown us new ways of thinking clearly about these fundamental questions in the study of meaning.

Chapter 4

Close encounters with the Van Lambalgen world

Johan van Benthem

On this festive occasion, I offer a few reminiscences on my encounters with Michiel's interests and insights concerning logic, probability, quantifiers, and cognition.

Meeting a remarkable student

When Michiel arrived in Amsterdam, he stood out. Here was a youngster with interests ranging from the history and foundations of mathematics to several branches of philosophy, mathematically inclined and gifted, and what is more, somebody who thought on his own. All these virtues show in Michiel's career displaying a unity of philosophical and mathematical thought, resulting in an original research agenda and a reputation that reached far and wide. When starting my annual spring cycles at Stanford in the early 1990s, I still remember the high esteem in which a local grandee like Patrick Suppes held Michiel's work.

The invitation by Keith Stenning and Martin Stokhof for this Festschrift was a good occasion for reflecting on my encounters with Michiel's ideas over the years. I have selected just a few topics, mostly on logic and probability broadly conceived, and describe some of my own subsequent interests. In doing so, I make no claim that what follows is central to Michiel's own agenda, or that he will agree with the directions where I am taking my themes. Anyway, idiosyncratic or not, my offering to this Festschrift will be brief.

Algorithmic randomness

First chance

I had the pleasure of acting as a supervisor, together with Joop Doorman, for Michiel's UvA dissertation *Random Sequences* (van Lambalgen (1987)). The topic of this seminal work was an old line in the foundations of probability, namely, Richard von Mises' account of infinite random sequences as the basis of probability, where no effective selection of subsequences can lead to effective betting strategies guaranteeing a profit. van Lambalgen (1996) is a good historical overview of the debates in the 1930s that led to the adoption of the now dominant measure-theoretic Kolmogorov paradigm. In contrast, Michiel's dissertation develops the von Mises program, and connects it in new ways to the theory of algorithmic randomness as it emerged in the line of Church, Martin-Löf, and others. In subsequent work, he took all this much further, linking up with Kolmogorov complexity, and using his view of random sequences to explore new conceptions of the set-theoretic universe.

These achievements are very much alive. Some years ago, I met a teacher at an ESSLLI Summer School who told me that he would not be standing in front of his classroom if it had not been for Michiel's pioneering work. With this acclaim still in the future, the verb 'supervising' took on a new meaning when interacting with Michiel. Probability was conspicuously missing from my education up to the PhD: the Institute for Mathematics at the UvA did not offer a compulsory core course in the topic [it was considered not fundamental enough], and the little I knew had been learnt through a physics course in Statistical Mechanics plus teaching a course on Carnap's Inductive Logic to educate myself. Moreover, Michiel's dazzling speed in producing new mathematical results on his topic made my preparing for our meetings like studying for an exam, rather than advising an apprentice.

Despite this forcible immersion course with Michiel, random sequences still intimidate me with their tension between two aspects. Their behaviour is patternless, so we cannot use them to predict next events or play profitable games against Nature. But at the same time, that very unpredictability is the basis for their having stable statistical properties like the Law of Large Numbers. Of course, this tension is not vicious or contradictory: it is deep.

Second chance

Fate sometimes offers second chances. Some years ago, Francesca Zaffora Blando asked me to be her supervisor at Stanford. Her dissertation "Patterns and Probabilities. A Study in Algorithmic Randomness and Computable Learning" (Blando (2020)) starts with Hume's Problem of Induction and shows how algorithmic randomness notions throw new light on classical issues in Bayesian epistemology. In

this way, I came to appreciate relations with formal learning theory, itself a natural continuation of my work on dynamic-epistemic logics of information update and belief change (van Benthem 2011). By the way, this connection was made clear in Nina Gierasimczuk's dissertation "Knowing One's Limits. Logical Analysis of Inductive Inference" (Gierasimczuk (2010)), co-supervised with Dick de Jongh: students keep lifting us to broader views. I will not state Francesca's results here, but you can check for yourself, as her thesis will appear in the ILLC dissertation series.

Randomness is just one of the many interfaces of logic and probability, whose variety is sometimes overwhelming, without a unifying narrative. Few people have a complete picture of all the ways in which the two fields interact, and Michiel may be one of the few who do. In what follows, I describe some encounters with his ideas in this broader area.

Generalised quantifiers, modality, guards, and (in)dependence

Abstract logics of independence relations

While the theory of randomness uses a rich array of mathematical tools, in logic, often, small is beautiful. Michiel is no stranger to such a more austere style of thinking. In his paper "Independence, Randomness and the Axiom of Choice" (van Lambalgen (1992)), he isolates a number of formal properties of the fundamental probabilistic notion of *independence*. These form a small first-order theory of an abstract independence relation $R\,x\,Y$ saying that x is independent from the finite set of variables Y. I will not list all the axioms here, but note that they combine some evident logical properties (say, downward monotonicity w.r.t. Y) with mathematical ones, in particular, Steinitz Exchange. Here is the latter as an illustration, transposed to dependence (here: the negation of the relation R) where it may be more easily understood: if $D\,x\,Y \cup \{z\}$, then either $D\,x\,Y$ or $D\,z\,Y \cup \{x\}$. For dependence among vectors, this is the basic abstract principle governing bases and dimension in Linear Algebra.

I consider this paper a precursor to the current field of dependence logics, cf. the survey (Galliani (2021)). Yet, Michiel's core ideas attracted less attention than they deserved because of their embedding in a new set theory including random objects, while the base logic for independence also contained a set-oriented reflection principle plus an axiom for ordinals. Even so, I was immediately interested, since an abstract reading of the relation R suggested analogies with accessibility relations in modal logic. In 1994, Natasha Alechina and I wrote a paper "Modal Quantification over Structured Domains" (van Benthem & Alechina (1994)), where we introduced a logic with the usual first-order quan-

tifiers plus the following modal-style quantifier: $\mathfrak{M}, [\mathbf{d}/\mathbf{y}] \models \Diamond_x \varphi(x, \mathbf{y})$ iff $\exists d \in D : R(d, \mathbf{d})$ & $\mathfrak{M}, [d/x, \mathbf{d}/\mathbf{y}] \models \varphi(x, \mathbf{y})$. We showed how classical modal issues and techniques lift to this logic, such as bisimulation analysis and frame correspondence. But we also noted how axiomatisation raised subtle issues, including new incompleteness phenomena with a still intriguing example pointed out by Michiel. The paper also states a prescient open problem: might the complexity of the system be decidable if one drops the standard first-order quantifiers altogether, leaving only the pure logic of $\Diamond_x$? More on that below. Further details can be found in Natasha's ILLC dissertation "Modal Quantifiers" (Alechina (1995a)), which I had the pleasure of supervising.

Guarded fragments and generalised semantics

In the same early 1990s, I started a collaboration with Hajnal Andréka and Ístvan Németi. On my side the trigger was an interest in what makes modal logic tick in terms of its nice model theory and low complexity, looking for explanatory general syntax patterns. (The latter line reflected my long-standing interest in syntactic translations of modal languages into standard ones.) What they brought was a match with a line in algebraic logic toward relativized semantics for relational algebra, dropping its complexity from undecidable to decidable, the CRS tradition of 'cylindric relativized set algebra', (Németi (1985)). The most prominent outcome of our joint work was the paper "Modal Languages and Bounded Fragments of Predicate Logic" (Andréka et al. (1998)) in which we introduced the *Guarded Fragment* GF, a large decidable chunk of the first-order language that generalises many existing modal languages. GF drops the dedicated guard predicates R of modal languages, or of the above modal quantifier (any atomic predicate will do), while the quantification is polyadic: quantifier syntax must be of the form $\exists \mathbf{y} : G(\mathbf{x}, \mathbf{y}) \wedge \varphi(\mathbf{x}, \mathbf{y})$, where atomic formulas involving the predicate G can occur in any further position in φ. This choice turned out to be natural: in particular, bisimulation analysis is still possible, and decidability is provable by means of a generalised filtration technique using syntactic types. Subsequently, I noticed that these results also worked for 'pairwise clique guarding' instead of single guard atoms, shifting the syntactic border-line with undecidability still further. Another noteworthy result is that, unlike FOL, the Guarded Fragment can carry the weight of higher-order apparatus gracefully, witness the decidability of the fixed-point version μGF proven in (Grädel & Walukiewicz (1999)). The JoLLI issue "Guarded and Decidable Fragments" (*Journal of Logic, Language and Information*, 14(3), 2005) edited by Natasha Alechina collected the first phase of research on the Guarded Fragment, but new results continue to appear.

However, another perspective is of equal interest to me. The CRS remodelling tradition suggests taking a fresh look at the semantics for FOL, generalising it in a modal style, as explained in my book "Exploring Logical Dynamics" (van

Benthem (1996a)). Here models are no longer classical Tarskian, but they come with a range of *available assignments* (maps from variables to objects: think 'possible system states'), and quantifiers refer to varying only between such states. The gaps in such generalised models encode *dependencies* between variables, dropping the tacit independence assumption of first-order logic that values of any variable can be changed at will while keeping those for all others the same. (This move makes particular sense in probabilistic reasoning, where variables may, or may not be independent.) The new semantics validates a decidable sublogic of FOL which retains its important subsystem of monotonicity reasoning. However, further quantifier laws of first-order logic such as $\exists x \forall y \varphi \rightarrow \forall y \exists x \varphi$ express Church-Rosser style existence conditions on the space of available assignments that show in grid patterns of transitions. And crucially, such grid patterns can be used for encoding undecidable geometrical tiling problems as first-order SAT problems. Thus, we have identified those features of Tarski's semantics that induce the undecidability: here and in other logics, usually the tacit inclusion of some rich mathematical object not necessarily needed for the core purposes of the logic. We now replace these decisions by a parameter that can be varied, and find the decidable first-order base logic of mixed independent and dependent quantification. More instances of remodelling strategies neutralising sources of 'imported complexity' in a broad variety of logical systems can be found in Andréka et al. (2017).

The two perspectives presented here might seem to raise a little 'paradoxette' of logical analysis. Guarded fragments *restrict* the language of first-order logic to achieve decidability on standard models. CRS-style logics *extend* the class of models of FOL, but then make the whole language decidable. The two moves, one 'down' and one 'up', seem at odds, but in fact, my paper in Natasha's 2005 JoLLI volume showed how they are essentially equivalent.

Modal dependence logic

Just recently, there has been a further twist to this story. CRS-models encode dependencies, but they only do so implicitly. Given the importance of dependence as a pervasive notion (just think of its role in analysing the notion of causality in AI), it makes sense to introduce explicit syntax and find explicit laws. This was done in (Väänänen 2007), with the introduction of *dependence atoms* $D_X y$ saying that variable y functionally depends on the simultaneous values of the variables in the set X. The resulting dependence logic in its Finnish version is non-classical and second-order, where the latter feature makes it undecidable and non-axiomatisable. However, one can also start the analysis at a much simpler base level. My recent paper "A Minimal Logic of Functional Dependence" with Alexandru Baltag (Baltag & van Benthem (2021)) presents a modal semantics where the basic notion is *local* dependence of y on X at some given state, and then finds a decidable and simply axiomatisable core logic LFD with dependence modal-

ities and CRS-style quantifiers. On this basis, one can then explore the further laws governing notions of dependence in richer concrete domains such as dynamical systems or games. As a concrete example, the Steinitz Exchange law of Michiel's 1992 paper is not valid in LFD, but it is an interesting open problem whether adding it to the modal base logic provides a complete axiomatisation for the notion of dependence in the realm of vector spaces. Our minimal modal dependence logic is under development now, and we are currently into topological and domain-theoretic versions, bringing in more mathematics as we go, though in small doses.

But functional dependence is just one notion. What about Michiel's emphasis on independence, rather than dependence? Actually, independence is not definable in LFD. If we see independence as absence of significant information flow from X to y (a thought already found in Michiel's 1992 paper), then we must, and can, introduce a new modality $I_X\, y$ for independence, and here is something that we discovered. The modal base logic of D and I is *undecidable*: and the reason is that it can encode the three-variable fragment of first-order logic. This may look surprising, since independence assumptions usually simplify calculations in probabilistic reasoning. How can the logic then get complex? To resolve this new paradoxette, it suffices to realise that a reasoning practice and a logic as the *theory* of that practice are not the same. The very mathematical structures that make reasoning simple may be a source of undecidable complexity for the logical theory of that practice. Does our finding show that Michiel's start with independence was unfortunate in a modal perspective? The jury is out on this. Suppose that we only consider the independence modality by itself, then there is an interesting open problem of the decidability of its modal logic.

In any case, these are not yet the key issues. The more important general phenomenon is information flow and *degrees of correlation* between variables. Functional dependence $D_X\, y$ leaves only one value for y given those of X, independence leaves y free to take on any value it could have in the state space: obviously, there is a lot to be studied in between.

Generalised quantifiers

There was also another general angle to Michiel's thinking about independence and probabilistic reasoning generally in the early 1990s. He was interested in logics of *generalised quantifiers* encoding probabilistic notions such as the Friedman quantifier $Q^A\, B$ with the intended interpretation that "all A except for a set of measure zero are B". He also thought of his independence relation, viewed abstractly, as a general device for creating new generalised quantifiers. This, too, resonated with me immediately, since generalised quantifiers had been one of my main interests at the interface of logic and natural language semantics (van Benthem (1986)). Triggered by this probabilistic interest of Michiel's, (van Benthem

(1996b)) considers the laws for the Friedman quantifier which include some evident principles that I will not list here plus the intriguing non-trivial Fubini Property saying that $Q^A x.\, Q^A y.\, Rxy$ is equivalent to $Q^A y.\, Q^A x.\, Rxy$. This commutation law is a crucial principle of probabilistic reasoning that features in much of Michiel's work. My contribution was a negative finding: I proved that no permutation-invariant logical quantifier can satisfy the Friedman axioms. At the same time, this is a positive insight. To have consistent models for the Friedman logic, additional structure on the domain is needed: measures, (in-)dependence relations, or what have you. Stated yet more positively, the introduction of probability allows for more design freedom in new consistent systems of reasoning.

Incidentally, generalised quantifiers also fit well with the above theme of generalising the semantics of first-order logic. A case in point is Aldo Antonelli's recent highly minimal quantifier semantics, discussed in depth in Andréka et al. (2017), including connections to the Guarded Fragment. But for a conclusion to this section, let me point out another unifying insight. In the final analysis, the generalised quantifier perspective is also close to the above modal approach, but then not, as in the above, in its usual relational format, but rather in terms of generalised *neighbourhood semantics*. The latter connection is explained and investigated in more detail in van Benthem & Westerståhl (2012).

Logic and probability: quantitative meets qualitative

Logic and cognition

Now fast-forward to 2001 when Michiel, together with Frank Veltman, became a holder of the chair of logic and cognitive science at the ILLC. Michiel has built up a broad oeuvre at this interface, with highlights such as the joint book "Human Reasoning and Cognitive Science" with Keith Stenning (Stenning & van Lambalgen (2008)). I found this move intriguing because around 2000, I, too, was beginning to feel the pull of the empirical facts knocking on the door of Frege's anti-psychologistic paradise (or if you wish, reservation). In 2008, I published a paper "Logic and Psychology: Do the Facts Matter?" (van Benthem (2008)), which records all the to-ing and fro-ing in my thinking, but which also records with some satisfaction how much concrete interesting work was already going on at the border between logic and cognitive science.

Michiel's work at the empirical cognition interface contains many topics that I find appealing, such as his analysis of neural networks in terms of non-monotonic logics (also studied by Gärdenfors and Leitgeb, and still alive, now in much more detailed connections coming to light between types of machine learning systems and conditional logics, (Icard & Ibeling (2020)). I also like his innovative use of logic programs for modelling brains and minds. Logic programs were

a case study in my interest in logical dynamics since the 1980s, because of their (despite official ideology) unique mix of declarative and procedural features, (van Benthem 1996A). Axiomatising this mix is a key theme in Marianne Kalsbeek's dissertation "Meta-Logics for Logic Programming" (Kalsbeek (1995)), which I supervised yet again in those lively 1990s. From neural nets, it is just one step to logical analysis of dynamical systems, another shared interest of Michiel and mine, be it that we approach them with different mathematical tools. Michiel and his student Levin Hornischer favour domain theory, (Hornischer (2021)), Grisha Mints has worked with modal logics, (Mints & Kremer (2007)), I myself and others with dynamic-epistemic logics, and recently also with modal dependence logics, (Baltag and van Benthem, in progress). This diversity of logical approaches is all to the good, of course: Let a Hundred Flowers Bloom, Let a Hundred Schools Contend.

However, my cue for what follows here is a fundamental distinction made by van Lambalgen and Stenning in empirical cognitive scenarios, between 'reasoning to' and 'reasoning from' an interpretation. Comparisons between logical systems and human practice are bound to be off if we do not acknowledge the major role of how agents *represent* a task at hand, how they form these representations, and how they work with them. I could not agree more. Once we see this distinction, facile judgments of people's lack of logic or rationality fall by the wayside, and we create room for a view that "the facts matter, and they speak in favour of logic". Indeed, to me, creation and maintenance of task-oriented representations is an independent cognitive skill, and the resulting goal is one of understanding the *logical dynamics* of an array of intelligent activities on a par: inference, observation, information update, and a lot more. This is the view in my monograph "Logical Dynamics of Information and Interaction" (van Benthem (2011)), implemented in the framework of dynamic-epistemic logics. However, I will only address this dynamics theme in the specific setting of our running theme of probabilistic reasoning, an area where human subjects have been claimed to be particularly error-prone.

Logic and probability

I said that the total landscape of active interfaces between logic and probability is hard to comprehend for a single person. My paper "Against All Odds. A Logician Looks at Probability" (van Benthem (2017)) is a quick survey of various strands in two directions. One can use qualitative logics of comparative probability statements as an underpinning for quantitative probability, as in the seminal (de Finetti (1937)), continued by Kraft-Pratt-Seidenberg and Scott around 1960, and taken further recently in a body of work on 'imprecise probabilities' surveyed and expanded in Ding et al. (to appear). In the opposite direction lie attempts at deriving qualitative logical theories of belief from quantitative probabilities, in a

tradition going back to Locke. Leitgeb (2017) contains a well-known account of acceptance rules for belief in terms of quantitative probabilistic 'stability', but it also presents a general perspective harmonising the two directions.

Now, in all this, there are issues of dynamics. How do we update probability judgments as new information comes in? I see the task of getting clear on this mixture of representation dynamics and inference as essential to understanding the cognitive behaviour that is sometimes lumped together under the single heading of 'probabilistic reasoning'. By the way, update has always been a staple of probabilistic frameworks such as Bayesian epistemology, and it may be of some interest to note historically that the 'dynamic turn' in language and logic of the 1980s, novel there at the time, just reflected common practice in neighbouring fields.

Probabilistic update: quantitative and qualitative

One technical answer to the preceding challenge comes from my work on dynamic-epistemic logics. In the joint paper "Dynamic Update with Probabilities" with Jelle Gerbrandy and Barteld Kooi (van Benthem et al. (2009)), we analyse the various roles of probability that come together in updates triggered by a new observation. There is the *prior probability* of our current information state (in our approach, an epistemic probabilistic model), representing our experience so far. Next, there is also the *occurrence probability* of the observed event in each world of the model, representing information we have about the general process we are in (sometimes called 'protocol information'). And finally, there is *observation probability* representing our judgments about which event we have actually witnessed, since observation is often just partial. These three components collaborate to form the new probability space after update, and for this purpose we propose a general 'product update rule' and axiomatise the resulting dynamic logic PDEL. The key to its proof system is finding the right 'recursion axioms' analysing truth of probability statements that hold after update in terms of probabilities in the initial model $\mathfrak{M}$ and in the 'event model' $\mathfrak{E}$ collecting all information about which event we are witnessing. The format here is equivalences reducing modal statements of the form $[\mathfrak{E}, e]P(\varphi) \leq q$ to formulas where the dynamic event modality has been pushed inside.

Incidentally, returning to an earlier theme, one can see such equivalences as a transition function of a dynamical system for iterated logical update (Klein & Rendsvig (2021)).

The full system PDEL is admittedly a bit baroque, and simpler subsystems often suffice in particular scenarios. For instance, when just making a public observation of a fact, standard Bayesian conditionalisation suffices, and when analysing an evergreen of probabilistic reasoning like the Quizmaster (also: 'Monty Hall'), which only involves public observations, a simple tree picture with prior and oc-

currence probabilities is all we need for a perspicuous updatable representation. There are also open problems about the computational complexity of PDEL, and perhaps we have sacrificed fit to reality for logical sophistication.

But there is also a deeper challenge here. As is well-known, many probabilistic scenarios do not supply precise quantitative information, and the decisions to be made do not depend on these either. For instance, in the Quizmaster, the decision is a qualitative one between switching doors or not, and this choice would be justified by the qualitative information whether it is more likely that the car is behind the door we chose, or the other door. So, can we make the above product update setting, or other quantitative update rules, qualitative? Can we start, say, with De Finetti style comparative judgments for prior, occurrence and observation probabilities, and then produce a new ordering? In this task, we encounter another major role of mathematics in probabilistic reasoning: the usual formulas that we employ 'glue together' probability values of various kinds (it is actually amazing that this works, as one more instance of the 'unreasonable effectiveness of mathematics'), and the update challenge is to replace this gluing function by some qualitative mechanism of, say, *order merge* to produce the new qualitative ordering after update. As it happens, I am not aware of a solution to this particular problem, which can be viewed as the dynamic counterpart to De Finetti's program, perhaps again working toward a representation theorem.

In this connection, it is interesting to note that a possibly simpler alternative has just been proposed in Ding et al. (to appear). Here a simple DEL-style superstructure is added to a propositional language with comparative probability statements, and the recursion axioms become much simpler thanks to the introduction of some further modal vocabulary for propositions of (thinking in quantitative terms) measure greater than 0. Another novelty is the use of dynamic operators for extending the language with new proposition letters, a device whose technical effect is encoding some more arithmetic of solving systems of inequalities into the logic. While some major questions about this alternative solution remain unsettled (there is no axiomatisation yet, and decidability is open), to me, at least, this shows that the above qualitative update issue is alive and well.

Rethinking qualitative versus quantitative

But perhaps we have been trapped in the wrong narrative here. Are we attempting to reduce quantitative probabilistic reasoning to qualitative reasoning, or in the other direction mentioned above: reducing the qualitative to the quantitative? This may well be a lingering attitude of the foundational era when logicist constructions tried to build everything from a logical ground up. It also seems the attitude of attempts to do "Science Without Numbers" and the like. My own current thinking on this has changed considerably, partly under pressure of cognitive facts. It seems that children develop reasoning and counting abilities at

about the same time, and no order can be detected. But also empirically, when looking at scientific practice, it is very hard to maintain that logical inferences are necessarily simpler than something as evident as the distribution of multiplication over addition, or the numerical Pigeon Hole Principle.

Accordingly, I am now working with Thomas Icard on combined formal systems of logic and counting, treating them as basic abilities on a par. This calls the relevance of qualitative/quantitive divide into question, but of course, not anything goes: some combined systems are much simpler and more natural than others. We are finding a lot of interesting issues when shifting perspectives in this way, that I cannot do justice here, such as the deep entanglement of logical syntax and counting, or the relevance of themes from generalised quantifier theory such as 'semantic automata'. Our results are reported in a paper with the working title "Logic and Counting" which should be available soon.

In a way, this final theme closes a circle in this article. I started by noting that Michiel worked with sometimes quite complex combined systems of logic and mathematics, and then advocated logical parsimony, perhaps even blaming insidious mathematical components for undecidability and high system complexity. However, what I am interested in now might be described as looking into the *fine-structure* of such powerful systems, starting at the bottom end corresponding to simple things that children do. In this perspective, the qualitative–quantitative divide as commonly construed loses much of its significance.

Conclusion

I have lightly described some interests at interfaces of logic, dependence and probability which I see as triggered by, or at least related to, Michiel's work. Of course, as I have made clear at the start, I am not claiming blessing or endorsement, I have just been describing topics and trends as they appear to me. Moreover, in doing so, I have not been systematic at all. For instance, it has been suggested recently that a more sophisticated account of dynamic-epistemic (product) update would have to take dependencies and independence between events into account (Aucher (2020)), and I agree. And I am sure that there will be many further interactions between the topics that I have presented separately here.

My choice of topics is also not exhaustive. There are many further strands in Michiel's work that I can relate to. One is the analysis of vision in terms of scene refinement and inverse limits in van der Does & van Lambalgen (2000) , a topic that we discussed extensively at one time, and which I see as relevant to creating more sophisticated versions of hyper-intensional semantics: a story by itself. Another is Michiel's interest in Immanuel Kant, who seems to be making great strides in logic under Michiel's tutorship. One of my most vivid student memories is wrestling with Kant's "Critique of Pure Reason" in the rose garden

of the Vondelpark in Amsterdam, and one of my earliest published papers was about the analytic/synthetic distinction and Hintikka's fascinating attempts to locate this somewhere inside first-order logic. I see this theme as relevant to the above discussion of discovering decidable parts inside FOL, but again, explaining that would require a separate story.

But even the topic of logic and probability that I did raise, however modestly, seems important to me. My metaphor for intellectual life is as a thin layer of logical thinking forming a delicate interface between two great probabilistic realms: the working of neural networks in our brain 'below' us, and the workings of mass behaviour over time in society 'above' us. Interfacing logic and probability, to me, is not just a filling in of white spots on the research map, but also an attempt at trying to understand the human condition.

Retirement is somewhat bitter-sweet, though term limits are also a sign of civilisation. Moreover, retirement can be a liberation from the repetitive and mundane. I am sure that Michiel will use the great escape well and spread his wings even wider in clearer skies.

Chapter 5

Kant's logic: letting go of a historical assumption

Kees van Berkel

It is commonly assumed that Kant's logic can be adequately represented in classical logic. Up to the present day, this assumption is a source of criticism for Kant's transcendental philosophy. Most notably, there is the criticism voiced by Strawson. The present article advocates to let go of this historical assumption. I provide an analysis of Strawson's argument and its relation to the aforementioned assumption. A possible source of the historical assumption is discussed and traced back to Kant's own work. In investigating the correctness of the assumption, I discuss Kant's only essay on logic: *The False Subtlety of the Four Syllogistic Figures*. In particular, I will argue that the criticism presented by Kant in this work is strikingly akin to Strawson's criticism of Kant, thus raising the question: did Kant not anticipate Strawson's objection? The article is inspired by the work of Achourioti and Van Lambalgen, who demonstrate that novel insights into Kant's transcendental philosophy can be gained by adopting a different perspective on Kant's logic.

Introduction

Logic plays a fundamental role in Immanuel Kant's critical philosophy. [1] This is particularly the case in the *Critique of Pure Reason* (1781/87), where Kant aims

1. On a personal note, I am grateful for the possibility to contribute to the celebration of Professor Michiel van Lambalgen's academic life and his contributions to philosophy, logic, linguistics, cognitive science, and especially those works that lie in the various intersections of these disciplines. I am even more grateful for the invaluable guidance and inspiration which I received from Michiel during my Master's in Logic at the University of Amsterdam between 2012 and 2015. Apart from reaffirming in me an imperishable fascination and admiration for the philosophy of Immanuel Kant, Michiel introduced me to the rich, complex, and beautiful nature of interdisciplinary research and to what, I have come to believe, are its quintessential virtues: patience and creativity. The present article deals with the *discredited view* of Kant's logic. My interest in this topic was roused many years ago during Michiel

to (i) strictly demarcate the domain of metaphysics and (ii) determine the possibility of synthetic a priori cognition, i.e., knowledge of objects independent of experience. For Kant, such a priori knowledge is possible and it is the task of transcendental philosophy to investigate the "mode of cognition of objects insofar as this is to be possible a priori" (B25).[2] The table of categories (A80/B106) consists of those a priori concepts of the understanding without which no object can be cognised (at all).

This is the domain of transcendental logic: the logic which "has to do merely with the laws of the understanding and reason, but solely insofar as they are related to objects a priori" (A57/B81-2).[3]

Kant shows how the pure concepts of the understanding, i.e., the categories, can be derived from a table of pure logical forms of judgment (A70/B95). A judgment is, broadly, a cognitive binding function providing a "conscious mental representation of an object" Hanna (2018). Each judgment has an intrinsic logical form, and since Kant argues that each pure logical form has a corresponding pure concept of the understanding, i.e., category, each concept is grounded in logic Hanna (2018). For Kant, there are twelve pure logical forms of judgment which, in sets of three, can be subsumed under four titles: *quantity, quality, relation,* and *modality*. The table is provided in Figure 5.1.
[4]

Furthermore, Kant claims that the table of judgments is exhaustively complete, and uses this completeness to demonstrate the completeness of the table of categories. In short, it is this very idea of providing a complete table of categories that contributes to the desired demarcation of metaphysics and accounts for the

---

van Lambalgen's course 'Kant and Cognition' at the University of Amsterdam. The present article does not deal mathematically with Kant's logic, but addresses the maintainability of this discredited view. I hope this work will bring pleasure to the one to which this Festschrift is dedicated.

2. I adopt standard reference to the Critique of Pure Reason by referring to its A- and B-edition (1781, respectively 1787). The translation used is that by Guyer and Wood (Kant (1998)).

3. For Kant, judgments arise in two types of logic: general logic and transcendental logic. The former is more abstract, since it "abstracts from all the contents of the cognition of the understanding *and* of the difference of its objects, and has to do with nothing but the mere form of thinking" (A54/B78) (emphasis my own), whereas to their objects and, for that reason, transcendental logic is a logic of truth (A62-3/B87). Consequently, general logic cannot produce any knowledge, but functions as a mere negative criterion for knowledge through the prevention of formal contradiction. For the purpose of this article, this brief comparison suffices. The reader is referred to Achourioti & van Lambalgen (2011) and Tiles (2004) for an in-depth discussion of the two logics. See Hanna (2018) and Tiles (2004) for an introduction to Kant's theory of judgment in relation to logic.

4. Tiles (2004) argues how through the involvement of objects the table of judgments of general logic (composed of eight forms of judgment) gives rise to the "transcendental table of all moments of thinking in judgments" (A73/B98) of transcendental logic, the latter which is given in Figure 5.1. See also Tiles (2004) for a discussion of how the three judgment forms subsumed under each title relate to each other.

Quantity	Universal, Particular, Singular
Quality	Affirmative, Negative, Infinite
Relation	Categorical, Hypothetical, Disjunctive
Modality	Problematic, Assertoric, Apodictic

Figure 5.1: The table of judgments (A70/B95)

possibility of synthetic a priori cognition.[5] The above recapitulation contextualises the following observation: Kant's transcendental enterprise—expressed by (i) and (ii) above—depends on the claimed completeness of the table of pure logical forms of judgment. In fact, as Tiles (2004) points out, "it is perhaps this claim, more than any other that has led modern logicians to dismiss Kant's significance as a logician" (p.106). What is more, this dependence has been a source of serious criticisms of Kant's transcendental philosophy. As observed by Achourioti and Van Lambalgen (2011):

> The *communis opinio* is that the Table of Judgements rests on a thoroughly discredited view of logic, and this has led to many reductive readings of CPR [Critique of Pure Reason] in which logic hardly figures, if at all. (p.255)

At the heart of this discredited view lies the following assumption: Kant's views on logic concern an (insubstantial) extension of Aristotelian logic and are, for that reason, adequately representable in and compatible with classical first-order logic (henceforth, classical logic), as a subsystem thereof. In particular, as Achourioti and Van Lambalgen point out: "It has been taken for granted that Kant's judgement forms can be translated in classical predicate logic" (p.259). The most well-known criticism of Kant's transcendental philosophy, rooted in this assumption, is the criticism voiced by Peter Strawson in his work *The Bounds of Sense* (2018). In what follows, I refer to the above assumption about Kant's logic as *the historical assumption*.

In order to avoid misunderstanding, the idea of 'being compatible with classical logic' must be made more precise: Although both Aristotelian logic and, for instance, intuitionistic logic are subsystems of classical logic with respect to theorem-hood, only the former is philosophically compatible with classical logic. Aristotelian logic, taken as a monadic first-order logic without identity, is a restricted subsystem of classical logic in harmony with the latter's fundamental principles, including the Law of Non-Contradiction (LNC) and the Law of Excluded

5. The above brief introduction to Kant's transcendental enterprise serves merely to provide the appropriate context of this paper. For an extensive discussion we refer to the works by Achourioti and Van Lambalgen (2011), Hanna (2018), Longueness (1998), and Guyer and Wood's introduction to the Critique of Pure Reason (1998).

Middle (LEM).[6] Conversely, from the viewpoint of intuitionism, intuitionistic and classical logic are mutually exclusive approaches to logic: e.g., intuitionism strictly *rejects* the endorsement of a generalised LEM. Intuitionistic and classical logic are in that respect incompatible alternatives. It is the assumption that Kant's logic is compatible with classical logic in the same sense as Aristotelian logic is, that leads to the discredited view.

The historical assumption can be challenged. In this paper, I will pose three guiding questions about the nature of the assumption:

Q1 Where does the assumption come from?
Q2 Are there reasons to think that the assumption is mistaken?
Q3 What happens if we let go of this assumption?

The exact formal nature of Kant's logic may be hard to assess, but still, answers to questions Q1-Q3 may provide good indicators as to whether to abandon the assumption, as an assumption *simpliciter*. In this article, I will briefly address the first question, focus primarily on the second question, and discuss the third in the light of earlier results obtained by Achourioti and Van Lambalgen (2011).

The contributions of the article are as follows: First, in section 'Strawson's criticism and *The Bounds of Sense*' I will provide an analysis of Strawson's criticism of Kant's logic. In particular, I will discuss how the assumption impels Strawson to impugn some of the foundations of the first Critique. In section 'A Kantian source of confusion' , question Q1 is addressed in the light of Kant's own assertions concerning Aristotelian logic, e.g., seemingly unequivocal assertions which can be found in the preface to the Critique of Pure Reason. Subsequently, in section 'Kant's criticism of the Aristotelian tradition', I will address Q2 by discussing Kant's only published work on logic, the essay *On the False Subtlety of the Four Syllogistic Figures* (1762). As I will argue, this surprisingly often-overlooked work provides valuable insight regarding the correctness of the historical assumption and Strawson's criticism. In fact, I will argue that Kant's own criticism of the syllogistic tradition, presented in the aforementioned essay, demonstrates a striking similarity with Strawson's criticism of Kant's conception of logic, thus raising the question: did Kant not anticipate objections such as those voiced by Strawson? Last, in section 'A different perspective: letting go', Q3 will be addressed using

6. In fact, both LNC and LEM can be retraced to Aristotle himself who, in the *Metaphysics* 2014, writes: "since it is impossible that contradictories should be at the same time true of the same thing, obviously contraries also cannot belong at the same time to the same thing" (LNC) (Book IV, CH 6, p.531) and "there cannot be an intermediate between contradictories, but of one subject we must either affirm or deny any one predicate" (LEM) (Book IV, CH 7, p.531). Note that Aristotelian logic must not be confused with Aristotle's logic, the former which arose out of the latter over subsequent millennia; see Lagerlund (2021) and Łukasiewicz (1957). In this paper, I will be concerned with Aristotelian logic only. See Smith (2020) for the essential differences between Aristotelian logic and classical logic, and how the limitations of the former eventually lead to the development of the latter in Frege's *Begriffsschrift* (1879) (e.g., think of the absence of relations and multiple quantification).

the work by Achourioti and Van Lambalgen (2011) as an example. There, I will discuss how their results likewise question the historical assumption. The main contribution of the present article is an analysis of Kant's own criticism of the science of logic and how it motivates reconsideration of the formal nature of his logic. In short, I advocate letting go of the historical assumption that Kant's views on logic can be adequately represented in classical first-order logic.

Strawson's criticism and *The Bounds of Sense*

In 1966, Peter Strawson published *The Bounds of Sense* Strawson (2018). It is in this work that Strawson famously criticised the logic in Kant's Critique of Pure Reason. Almost directly after its publication, "The Bounds" had a major impact on Kant scholarship (see Snowdon & Gomes (2021)). On the one hand, it incited thorough investigation of *transcendental arguments* as a field in its own right[7], where such arguments are designed to refute/reduce skepticism by deriving necessary grounds for statements that are, so to say, trivially acceptable to the 'reasonable' skeptic (e.g., think of elementary statements such as 'I experience change' in Kant's analogies of experience (see Sacks (2006)). On the other hand, "The Bounds"' reaffirmed the believe that Kant's conception of logic was not satisfactorily developed and too narrow, particularly in the light of subsequent developments in mathematical logic (starting with Frege's *Begriffsschrift*). Specifically, Strawson challenged Kant's claims that (i) the table of judgments is complete and (ii) contains only primitive, that is, irreducible logical forms. In this section, I will examine Strawson's criticism and discuss its relation to the historical assumption.

Strawson's argument

Strawson embraces Kant's general standpoint that knowledge depends on judgment, but points out that this does not imply that each logical form in which knowledge may be expressed is necessary. Kant claimed that the table of judgments consists of *primitive* logical forms of judgment only. Furthermore, he claimed the table to be complete. As stressed by Strawson, the challenge for Kant is not to show that such logical forms of judgement are possible, but to demonstrate that these are *essential* features of the form of judgment. Only then Kant may conclude that the categories, rooted in these judgment forms, are indispensable with respect to the possibility of knowledge (including experience as well as synthetic a priori cognition). Strawson is skeptical about this idea, relating Kant's claims to the development of mathematical logic in the twentieth century: formal logic shows that many (if not all) logical functions are interdefinable in terms of alternative logical functions. As an example, Strawson (2018) points out that:

7. See, for instance, Stern (2003), Stroud (1968), and Taylor (1978).

> For instance, that list [the table] includes the hypothetical and disjunctive forms, the analogues of which in modern logic are interdefinable with the help of negation. (p.45)[8]

From this observation, Strawson infers that a minimal system of primitive logical forms is thus rather a matter of a logician's *choice*. Moreover, this basic observation leads him to conclude that instead of thinking of fundamental logical forms, it is more suitable to talk of *fundamental logical ideas* underlying these forms: ideas which may be expressed through various logical forms. Therefore, a logical system primarily reflects the logician's choices in defining the system in one way, rather than another (yet equivalent way).

Strawson further pursues this line of thought and observes that the two fundamental ideas of modern logic are *truth-functional composition* and *quantification*. The former relates to propositional logic and its connectives, the latter to predicate logic and its quantifiers. Just as for the connectives (cf. the quote on hypothetical and disjunctive forms above), Strawson stresses that for quantification we are likewise left with a fundamental idea instead of necessary primitive forms, since it amounts to the logician's choice to take either universal or existential quantification as primitive (I will return to this below).

Sympathetic to Kant's general aims in the Critique of Pure Reason, Strawson concludes that his proposed refinement to move from fundamental forms to fundamental ideas, unfortunately, cannot contribute to a better understanding of the possibility of experience and synthetic a priori cognition. Namely, Strawson argues, the generality of the above two fundamental ideas leaves us with a reformulation of Kant's starting question, given prior to the introduction of the two tables in the Critique of Pure Reason. The reformulation being,

> [h]ow in general must we conceive of objects if we are to make empirical judgements, determinable as true or false [i.e., truth-functional composition], in which we predicate concepts of identified objects of reference [i.e., quantification]? (Strawson, 2018, p.46)[9]

According to Strawson, this refinement brings Kant's endeavour back to its original starting point: namely, he argues that the proposed refinement to consider logical ideas, instead of forms, as fundamental results in two fundamental ideas that are already implicitly present in Kant's initial question concerning the possibility of synthetic a priori cognition.

I take the gist of Strawson's criticism to be twofold: (i) interdefinability of logical forms of judgment motivates a move from fundamental—i.e., primitive—logical forms to fundamental logical ideas and (ii) the two fundamental ideas of modern logic—i.e., truth-functionality and quantification—do not provide us with the required additional insight for understanding the possibility of know-

8. Strawson equates 'modern logic' with classical logic. I will come back to this.
9. The quote furthermore suggests that Strawson has bivalent truth-functionality in mind.

ledge through objective judgment. Since interdefinability makes the idea of a primitive form problematic, it consequently problematises the claimed completeness of the table of judgments. Given the above, Strawson's take on Kant's logic leads to problems in Kant's transcendental philosophy: the table of judgments seems neither complete nor does it seem to consist of primitive notions. However, as I will argue next, the argument—underlying (i) and (ii)—is strongly dependent on Strawson's conception of 'modern logic'.

A criticism of Strawson's criticism

The argument put forward in *The Bounds* relies on an assumption about the underlying logic in terms of which Strawson defines 'interdefinability'. In fact, this dependence reveals a crucial weakness of the argument. Strawson grounds his criticism in what he calls 'modern logic'. Although he does not specify it formally, we can infer from Strawson's remarks that he conceives of modern logic as classical logic. In a footnote to *A Formalization of Kant's Transcendental Logic* (2011), Achourioti and Van Lambalgen make a similar observation:

> At least in classical logic, it is impossible to talk about a distinguished set of primitive logical forms, since judgement forms are interdefinable (e.g., the hypothetical in terms of the disjunctive and negation). (p.258)

In fact, the collection of primitive logical forms could be drastically reduced from the perspective of classical logic. For example, 'negation' ($\neg$) and 'disjunction' ($\vee$) are sufficient to classically define the other connectives 'conjunction' ($\wedge$), 'material implication' ($\rightarrow$), and 'equivalence' ($\leftrightarrow$) (etcetera). Even then, there is no need for embracing negation and disjunction as primitives since other logical connectives would equally suffice (e.g., the 'Sheffer stroke' as a single connective). However, the interdefinability observed by Strawson would fail from the point of view of intuitionistic logic.[10] The often-claimed interdefinability of logical forms depends on a classical logic reading of these logical forms.[11]

---

10. That the observed interdefinability does not hold from the viewpoint of intuitionist logic, is due to the restricted use of negation in this logic, which is expressed through the absence of the double negation elimination axiom $\neg\neg\phi \rightarrow \phi$ (which is classically equivalent to LEM). Consequently, in intuitionistic logic one must axiomatize each connective individually in order to capture their intended meaning.

11. Hanna (2018) presents another 'interdefinability argument': from the viewpoint of contemporary logic, he states, Kant's logic appears limited "since Kant's list of propositional relations leaves out conjunction, even his propositional logic of truth-functions is apparently incomplete". However, since conjunction can be systematically defined in terms of other logical relations (cf. the remark on Sheffer above), "at least implicitly, Kant's propositional logic of truth-functions is complete". Here, interdefinability may be taken as a reason in favour of concluding Kant's logic to be implicitly complete. However, as is the case for Strawson, the argument depends on the assumption that the logical connectives behave classically. Also, see Geach (1972) for a discussion on the presence/absence of conjunctive properties in Aristotle's works on logic (thanks to Nils Kürbis for pointing this out).

Another indicator for Strawson's endorsement of the historical assumption can be found in his analysis of the fundamental idea of quantification. Here, Strawson (2018) emphasises that "it is a matter of choice whether we introduce the existential quantifier [$\exists$] without a formal definition and define the universal quantifier [$\forall$] in terms of it, or vice versa" (p.45-46). The fact that a logician may choose to adopt either universal or existential quantification as a primitive notion, is due to the interdefinability of these two notions. Here too, Strawson's observation assumes a classical logic interpretation of quantification: the equivalence $\forall x P(x) \equiv \neg\exists x \neg P(x)$ is classically valid, but for instance not valid from an intuitionistic point of view.[12]

A refinement of Strawson's objection is thus required: the quantifiers $\forall$ and $\exists$ express the same fundamental logical ideas (with the help of negation), *provided* the underlying logic allows for their interdefinability. The alleged choice available to the logician is thus a consequence of the *endorsed* underlying logic. From the point of view of intuitionistic logic, $\forall$ and $\exists$ represent, so to say, two mutually irreducible logical ideas, even though they share the general idea of quantification. There, the logician is not presented with such a choice.

Recall the two central pillars of Strawson criticism: (i) interdefinabilty of logical forms motivates a move from fundamental logical forms to fundamental logical ideas and (ii) the two fundamental logic ideas of modern logic—i.e., truth-functionality and quantification—do not advance Kant's transcendental enterprise. As I argued for above, (i) and (ii) cause serious issues for Kant's logic *under the assumption that* the logic can be adequately represented in classical logic.[13]

It is the historical assumption that leads Strawson in (i) to the idea of fundamental ideas (since forms are interdefinable). It is again the historical assumption that leads Strawson in (ii) to the conclusion that the most primitive fundamental ideas are truth-functionality and quantification (since no other refinements are considered primitive in a classical setting). However, in both cases one may consider a different, non-classical logic for which the claimed interdefinability does not hold (e.g., intuitionistic logic), which may in fact give rise to a variety of fundamental logical forms in contrast to mere logical ideas. The central question that must be answered is therefore:

Is Strawson justified in adopting the historical assumption?

The present article seeks to answer this question.

---

12. Failure of this equivalence is due to the absence of double negation elimination in intuitionistic logic.

13. I consider Strawson's criticism to indicate his endorsement of the historical assumption. Alternatively, one could take Strawson's arguments to suggest that Kant adopts, so to say, 'the wrong logic' and should have instead grounded his transcendental philosophy, and table of categories, in classical first-order logic (which arose through Frege's *Begriffsschrift* a century later). Nonetheless, my aim is to lay bare the dependence of Strawson's criticism on the assumption that Kant's views on logic can be represented in classical logic in the first place.

Concluding this section, I would like to provide two reflections on Strawson's analysis. First, I believe that Strawson's observation that logical systems depend on choices made by logicians is valuable in its own right. In designing logical systems, one is often confronted with design choices, alternatives which may be justified from different perspectives (e.g., considerations of complexity versus expressivity). For that reason, it is beneficial to think in terms of fundamental ideas. Take for instance Von Wright's article *Deontic Logic* (1951). The modal logic presented in this seminal work (regarded as the first formal system of deontic logic) enables reasoning with the modal concepts 'obligation', 'permission', and 'prohibition'. In this work, only permission is taken as primitive and the remaining two modalities are defined in terms of it, using negation. Von Wright later remarked (see Hansson (2013)), that one could consider 'obligation' as a more fundamental concept in deontic reasoning, thus making it a more suitable candidate to serve as a primitive operator. The reason for initially taking 'permission' as primitive was motivated by Von Wright's (1963) previous choices for taking 'possibility' as the primitive modality of alethic logic (defining 'necessary φ' as 'not possible that not φ'), a modality which has a striking logical resemblance to permission. Interestingly, this focal shift to obligations as the more central notion of deontic logic later led to criticism concerning the overlooked complexity of the notion of permission: Hansson (2013) argues that permission must be considered as primitive, i.e., irreducible to obligation.

The second remark concerns determining the completeness of Kant's table of categories. Since, according to Kant, the table of categories is derived from the claimed completeness of the table of judgments we need to determine the correctness of the latter. In order for a judgment in the table to classify as primitive, Strawson (2018) argues, one must prove that,

> [it] must be an essential form or feature, one which exhibits, as no other can, some part of those fundamental and indispensable resources [of judgment] themselves. [...] But it is by no means clear that this condition is satisfied by all the items in Kant's list [table]. (p.45)

To substantiate this claim, Strawson makes his famous claim about (classical) interdefinability in Kant's table of judgments. Nevertheless, irrespective of the exact logical nature of the logical forms of judgment, the assessment of Kant's claims concerning the completeness of both tables remains a central challenge in Kant scholarship. A fundamental question here is: are these concerns preserved when one adopts a different stance on Kant's logic (Q3)? Before addressing this question in section 'A different perspective: letting go', I will first discuss possible sources of the historical assumption (Q1).

A Kantian source of confusion

Strawson is not the only philosopher who takes Kant's logic to be representable in classical logic. As observed, this assumption has led to quite some negative verdicts concerning Kant's transcendental philosophy. How do we historically situate Kant's ideas on logic? As Young (1992b) points out in the introduction to *The Lectures on Logic*: "Kant's approach to logic falls within what can broadly be called the Aristotelian tradition" (p.xv). The aim of this section is to highlight how Kant's remarks on Aristotelian logic may have contributed to the received view expressed in the historical assumption (Q1).[14] That is, Kant himself may have created the expectation that his conception of logic is little different from the Aristotelian tradition and, for that reason, at least compatible with classical logic. One thing is certain, Kant is well-known for *praising* Aristotle's work on logic.

In the introduction to the Critique of Pure Reason, Kant sets out his general aspiration to establish a rigorous scientific foundation for philosophy (metaphysics in particular). Inspired by well-grounded sciences such as mathematics and physics, he claims that "[the] same spirit would also prove itself effective in other species of cognition if only care had first been taken to correct their principles" (Axi). Kant aims at bringing this rigour into philosophy, thus defending it from what he calls "complaints about the superficiality of our age's way of thinking" (Axi). Kant appears to have a similar high regard of logic as a well-grounded science. On the first page of the preface to the second edition of the Critique of Pure Reason, he lauds Aristotelian logic accordingly:

> That from the earliest times logic has travelled this secure course [of a science] can be seen from the fact that since the time of Aristotle it has not had to go a single step backwards [...]. What is further remarkable about logic is that until now it has also been unable to take a single step forward, and therefore seems to all appearance to be finished and complete. (Bviii)

Kant's remark that logic, conceived of as Aristotelian, seems finished and complete leaves little room for thinking that Kant's own ideas on logic—pertaining to general and transcendental logic—should not follow this tradition too.[15]

Similar remarks can be found for instance in the Jäsche Logic Kant (1992b), a work of logic composed on Kant's behalf by Gottlob Benjamin Jäsche in 1800, and which is solely based on Kant's own lecture notes. (The Jäsche logic is therefore assumed to reflect Kant's later ideas on logic.) Kant appears to endorse the idea

14. The aim here is not to reconstruct the actual origin of the assumption. I refer to Tiles (2004) for a historical situation of Kant's ideas on logic.

15. Strawson (2018) quotes the following passage from Kant's *Prolegomena*: "The work of the logicians lay before me, finished" (p.42-43).

that logic, as a science[16], is already complete in the work of Aristotle, safe for some lack of clarity. In the Jäsche logic, Kant says:

> From Aristotle's time on, logic has not gained much in content, [...] but it can surely gain in regard to exactness, determinateness, and distinctness. [...] Aristotle had not omitted any moment of the understanding; we are only more exact, methodical, and orderly in this. (9:20)[17]

Given these remarks, it seems justified to take Kant's ideas on (general) logic as merely more precise and determinate representations of Aristotelian logic. I believe that Kant's overt contentment with Aristotelian logic may be acknowledged as one of the possible contributions to the received view that Kant's logic is mathematically trivial, not significantly diverging from the Aristotelian tradition, and insubstantial to the development of contemporary formal logic.[18] As Young (1992b) phrases it:

> Kant is not a major contributor to the development of formal logic. He fails, too, in his most conspicuous efforts to build his transcendental logic on clues provided by formal logic. (pxvi)

Kant's remarks about Aristotelian logic, suggesting that his own logic may not be radically different from the former, together with the observation that Aristotelian logic is compatible with (an expansion to) classical logic (cf., introduction), may be regarded as partly responsible for (the persistence of) the historical assumption. However, the fact that Kant frequently emphasises that Aristotelian logic can still gain in exactness and distinctness also shows that Kant was not uncritical of the Aristotelian tradition.

Kant's criticism of the Aristotelian tradition

As an answer to Q2—are there reasons to think that the assumption is mistaken?—I provide an analysis of Kant's only published essay on logic: *The False Subtlety of the Four Syllogistic Figures* (1762).[19] In this often-overlooked essay, Kant voices his

16. For Kant (general) logic is a formal science; e.g., see Hanna (2018).

17. I adopt standard reference to the Jäsche Logic referring to the volume, respectively page number of the Academy edition of Kant's works. The translation used is that by Young (1992b).

18. Tiles (2004) discusses three essential contributions made by Kant to the development of modern logic: "the distinction between concept and object, the primacy of the proposition (or sentence) as the unit of logical analysis, and the conception of logic as investigating the structure of logical systems" (p.85). Achourioti and Van Lambalgen 2011 show that Kant's logic is not mathematically trivial, and that more favourable views may be acquired, if one abandons the historical assumption. I will discuss this in section 'A different perspective: letting go'.

19. Kant's *A New Elucidation of the First Principles of Metaphysical Cognition* (1755) also addresses the topic of logic. In this work, Kant argues that all principles employed in deductive reasoning are rooted in the principles of identity and contradiction. Still, as a work of logic in itself, "The False Subtlety" remains the sole candidate.

criticism of dividing Aristotelian syllogistic reasoning into four distinct primitive figures. A criticism which runs contrary to common practice of his time.

The essay was published in 1762 and predates the so-called 'silent decade' in which the Critique of Pure Reason was conceived (from 1770 to 1781) (Kant, 1998, p.24). "The False Subtlety" was written in the same period as Kant's influential *The Only Possible Basis of a Demonstration of the Existence of God*, published in 1763. Following Cassirer (1981),both works were most likely completed in the Autumn of 1762. Next to the publication of Kant's inaugural dissertation *On the Form and Principles of the Sensible and Intelligible World* (1770), this period between 1762-1764 is often seen as marking an important change in the development of Kant's ideas. In this period, we find the first hints of the fundamental ideas that will later shape the Critique of Pure Reason. For instance, as Cassirer (1981) points out:

> Here the first sharp dualism in the Kantian system emerges. The view that logic in its traditional form, as syllogistic, could suffice to "construct" the system of actuality crumbles once and for all, since it and its supreme principle, the principle of contradiction, are inadequate to express the peculiarity of even the simplest real relation, that of cause and effect. (pp.75-76)[20]

As Cassirer suggests, in this period Kant becomes aware of the limitations of the traditional approaches to logic. In particular, in "The False Subtlety" we find a discussion of what turns out to be one of the fundamental building blocks of Kant's Critique of Pure Reason: that is, we find "the claim that the fundamental notion in formal logic and in the analysis of the powers of human capacity for cognition is the [procedural] notion of judgment [instead of concept]" (Kant, 1998, pp.28–29, introduction by Guyer and Wood).

The argument of *The False Subtlety*

The essay starts with a simple definition of syllogism: "Every judgment which is made by means of a mediate characteristic mark is a syllogism" (2:48).[21] Such a judgment expresses a relation between a characteristic mark and, what Kant calls, "the thing in the judgment" (2:48). One can think of a characteristic mark as a predicate of a subject. In a syllogism, the ascription of a characteristic mark occurs mediated by another applicable characteristic mark. The presence of an intermediate mark implies that a syllogism consists of at least three judgments; two serving as premises, one acting as the conclusion. To illustrate this, consider the classic syllogism 'all humans are mortal, all Greeks are human, therefore, all Greeks

20. Cf., footnote 6.

21. In what follows, I refer to the volume, respectively page number of the Academy edition of "The False Subtlety". The translation used is by Walford and Meerboote (1992a).

are mortal' in which the characteristic mark 'human' mediates the application of the characteristic mark 'mortal' to the subject 'Greek'.[22]

The essay's focus is on the four traditional syllogistic figures.[23] We briefly recall them here. A syllogism consists of three statements linking three marks: a major, a minor, and a middle term. For instance, in the above example 'mortal' is the major term (Predicate, henceforth P), 'Greek' is the minor term (Subject, henceforth S), and 'human' is the middle term (Middle, henceforth M). Arranging permutations of these different terms gives 256 distinct syllogisms, of which 24 are considered valid (see Lagerlund (2021)). The arrangement of the terms of these 24 valid syllogisms can, subsequently, be ordered into 4 general syllogistic *figures* (with the structure S-P as default conclusion). The figures are presented in Figure 5.2.[24]

For Kant, the mediation occurring in a syllogism can follow two fundamental principles, one affirmative and one negative principle. The first, affirmative, principle is called *dictum de omni*: "[a characteristic mark] which belongs to the concept belongs to the lower concepts which are subsumed under it" (2:49), namely, "that which belongs to a concept is a characteristic mark of a characteristic mark" (2:49) and thus also of those subordinate characteristic marks. The second, negative, principle is the *dictum de nullo*: "that which contradicts the higher concepts, must also conflict with the lower concepts which are subsumed under it" (2:49). For Kant, all rules applicable in the context of syllogisms have their ground in these two rules.[25]

On the basis of the above, Kant describes two types of syllogisms: If a syllogism is composed of three judgments solely combined through the above two fundamental principles, it is a *pure* syllogism. If a syllogism can only be construc-

22. As argued for by Łukasiewicz (1957), Aristotle does not consider singular terms suitable for syllogistic reasoning—e.g., 'Socrates' instead of 'Greek'—but only those terms that are capable of both inclusion and exclusion. The reason is that terms occurring in syllogisms must be applicable both as a predicate and a subject, which is for instance not the case for the term 'Socrates' (cf. 'some humans are Socrates').

23. Aristotle's analysis of the possible combinations of premises led to three figures of syllogistic reasoning, i.e., figures I-III of Figure 5.2. The additional Figure IV has been present since ancient times and was thoroughly investigated throughout the Middle Ages. Lagerlund (2021) points out that all instances of the fourth figure are already encountered as examples in Aristotle's *Prior Analytics*, although not introduced as a figure proper. Also, see Łukasiewicz (1957) for an extensive discussion of the differences between Aristotle's syllogistic and traditional, here called Aristotelian, syllogistic.

24. See Smith (2020) for an introduction to Aristotelian syllogisms.

25. Kant points out that these principles are ultimate grounds which cannot be proven themselves: "For a proof is only possible by means of one or more syllogisms, so that attempting to prove the supreme formula of all syllogisms would involve arguing in a circle" (2:49). To demonstrate that the two principles are the only two ultimate grounds from which syllogisms are constructed, is according to Kant a straightforward but tedious task. Also, see Aristotle's distinction between every assertion expressing either the affirmation or the denial of a predicate of a subject in Smith (2020).

	Figure I	Figure II	Figure III	Figure IV
Major	M c P	P c M	M c P	P c M
Minor	S c M	S c M	M c S	M c S
Conclusion	S c P	S c P	S c P	S c P

Figure 5.2: The four basic syllogistic figures, with 'c' for copula.

ted through using more than three judgments, it is called *mixed*. The two types are exhaustive for syllogistic reasoning.

In "The False Subtlety", Kant argues that only figure I is pure, and the remaining three II-IV are, despite being valid, mixed syllogisms *reducible to the first*. Kant's main conclusion is that figures II-IV are falsely considered 'simple', i.e., primitive. The argument revolves around the observation that all syllogisms belonging to figures II-IV make use of *implicit premises* that are strictly speaking necessary in order to make the reasoning valid. In all of these cases, Kant states:

> the conclusion could only be drawn from these judgments [major and minor premise] by means of legitimate logical conversion, or by contraposition or some other logical transformation of one of these premises. (2:50)

The validity of these inferences, hence, strictly depends on the presence of additional premises obtained through logically valid transformations of the involved premises, transformations which according to Kant must "be present if only in thought" (2:51).

Before proving his claim, Kant first observes that syllogisms of the first figure are always pure. That such syllogisms are constructed from the two supreme rules only, a quick examination should make clear. Reconsider the example 'all humans are mortal, all Greeks are human, therefore all Greeks are mortal' (which is an instance of figure I), here we find a single application of the *dictum de omni* rule. The fact that the syllogisms of figure I can be directly constructed through applications of the two rules shows that no intermediate premises are required and that such syllogisms are *pure*. That this is not the case for the remaining three figures II-IV, Kant discusses in the middle part of the essay.

I will merely present the gist of Kant's reduction through an example of the second figure. Kant uses the following example: 'no mind is divisible, all matter is divisible, so no matter is a mind'. Although the inference is valid, it is so by virtue of a valid but implicit conversion of the first premise, i.e., 'hence, nothing divisible is a mind'. Kant remarks:

> This conversion must, therefore, be tacitly thought in making the inference, for otherwise my propositions do not form a valid inference. (2:52)

One may take syllogisms of the second figure as, so to say, *lemmata* of syllogistic logic, but they are nevertheless not primitive. The full inference would have the following form:

P1.	No mind is divisible;	premise
P2.	Hence, nothing divisible is a mind;	contraposition P1
P3.	All matter is divisible;	premise
C1.	So, no matter is mind.	conclusion

The conclusion C1 is an inference obtained directly from P2 and P3, through application of a syllogism of the first figure. P2 itself is a contraposition of P1. The reason why the above syllogism is invalid without the implicit P2, is that neither the *dictum de omni* nor *dictum de nullo* is directly applicable to P1 and P3. Consequently, that P1 requires a logical transformation is the reason why the second figure is not pure.

The reason why figure II (to which the above syllogistic instance belongs) is a 'false subtlety' is that, once the logical transformation is applied, we end up with a syllogism of the first figure I: namely, first we change the order of P2 and P3, then we observe that P3, P2, C1 has the structure M c P (P3), S c M (P2), S c P (C1), which is valid by virtue of the *dictum de nullo*. Put differently, 'divisible' is the higher concept of 'matter' and since 'mind' contradicts the concept 'divisible' by *dictum de nullo* it must conflict the lower 'matter'. The same argument applies to all other instances of the second figure, and thus syllogisms of the second figure are mixed syllogisms reducible to syllogisms of the first figure. Kant presents a similar reduction for figures III and IV.

Kant's analysis shows that:

> [I]t is only ever the first figure which, concealed in a syllogism by means of covert inferences, has the power to generate the conclusion. [...] [O]ne can instantly convert any syllogism belonging to one of the other figures into the first simple figure. (2:58)

In other words, the alleged 'subtlety' expressed in the traditional division of syllogistic figures is 'false'.

A similar distinction between figure I and figures II and III can already be found in Aristotle's *Prior Analytics*.[26] Aristotle calls the former complete (perfect), whereas he calls the latter incomplete (imperfect). Perfect syllogisms need no external premise and are self-evident, i.e., they do not require additional proof of their validity (see Smith (2020)). Figures II and III are imperfect and require additional reasoning (premises) in order to demonstrate their validity. Aristotle proves the validity of these imperfect syllogisms through reducing each instance to an instance of perfect syllogistic reasoning (of figure I). Like Kant's arguments, Aristotle's reductions are obtained through logical conversion of some of the involved

26. See the translation by Smith (1989).

premises. In particular, a central result of Aristotle's analysis is that all valid syllogisms are reducible to two instances of the first figure. We find similar results in "The False Subtlety".

Kant's analysis is quite meticulous and his findings do not change Aristotelian logic at all: that is, the same class of syllogistic inference and principles remain logically valid. As Kant puts it, "the three other figures [II-IV] would, at worst, be useless, but not actually false" (2:55). So, one may justifiably wonder, what is the point of Kant's criticism? The answer is, I believe, that Kant's argument is of a scientific, methodological nature: his criticism advocates certain fundamental scientific values, values which are according to Kant absent in the Aristotelian tradition of logic.

The scientific implications of Kant's criticism

In "The False Subtlety" Kant does not deviate from the Aristotelian tradition when it comes to the validity of syllogistic inferences, but deviates only from a generally accepted *classification* by problematising the claimed purity of certain inference schemes. Why all this effort for such a minor result? I believe the answer is to be found in Kant's vision of the fundamental tasks of science and the core duties of a scientist, in particular, that of a logician. Kant's analysis discloses a violation of one of the central tasks of the science of logic:

> The purpose of logic, however, is not to confuse but to clarify; its aim is not to obscure but clearly to reveal. Hence, these four modes of inference [figures] ought to be simple, unmixed, and free from concealed supplementary inferences. (2:56)

What the analysis shows is that the latter three figures do not, paraphrasing Kant, *deserve* to appear in the science of logic as those syllogisms representing syllogistic reasoning in its clearest form. The mistake reflects critically the work of logicians in the past:

> It is also certain that hitherto all logicians have regarded them [syllogisms of figures II-IV] as simple syllogisms, not requiring the interpolation of additional judgments. (2:56)

In the above quotes, Kant seems to suggest that those logicians before him failed to serve the central purpose of logic, and this despite the fact that, following Lagerlund (2021), the reduction of incomplete to complete syllogisms through additional logical conversion was known since the rediscovery of Aristotle's *prior analytics* in Western Europe (in the middle of the twelfth century).

My view is that the above analysis shows that from a methodological perspective, Kant has a clear idea about his core duties as a logician (an idea reaffirmed in the first Critique and during his later life through the Jäsche Logic): a logician's duty is to reduce confusion and to promote clarity. In fact, the complaint in "The False Subtlety" generalises to Kant's criticism of the history of

metaphysics voiced in the first Critique (e.g., see the quote in the introduction). But what is more, the distinctive purpose of logic consists not just of providing clarity and reducing obscurity, but consists, as Kant puts it, first and foremost

> of reducing everything to the *simplest mode of cognition*. (2:56) (emphasis my own)

The reader may have an idea of where this is going: what Kant identifies as the logician's foremost duty—a duty he suggests violated by those logicians before him—is exactly that task which, according to Strawson, Kant fails to fulfil himself: reducing everything to the simplest mode of cognition. The criticism presented by Kant in "The False Subtlety" is strikingly akin to Strawson's criticism of Kant. It raises the question: did Kant not anticipate Strawson's objection? If he did, it may suggest that Kant's logic is exactly that logic which makes the table of judgments pure and complete; which, consequently, would be at odds with what is entailed by the historical assumption.

What can we conclude from the above? At a minimum, "the False Subtlety" demonstrates that Kant was very much aware of the potential errors—such as the false ascription of primitives—that may occur in the field of logic. In fact, we saw that for Kant the central task of logic is exactly to avoid such false ascriptions. For that reason, it seems rather odd that Kant did not take this into account during the twelve year voyage that it took him to finish the Critique of Pure Reason. This, I would say, is especially odd when taking into account the pivotal position assigned to logic (i.e., general and transcendental logic) in Kant's first Critique. We must be cautious with drawing overambitious conclusions here, but still the above investigation suggests that when analysing Kant's transcendental philosophy, we should at least give Kant the benefit of the doubt when he says that the table of judgements consists of only pure judgement forms (which is ruled out by the historical assumption). Kant's position in "The False Subtlety" provides us with good reasons for abandoning the historical assumption, as an assumption *simpliciter*.[27]

A different perspective: letting go

In letting go of the historical assumption, a fruitful approach would be to reconstruct Kant's conception of logic (including both general and transcendental logic) and develop a formal system corresponding to it. This approach is for instance taken up by Achourioti and Van Lambalgen (2011), Van Berkel (2015), and Kovač (2008). What these approaches have in common, is that instead of imposing a logical framework on Kant's thought, they aim at deriving the logical apparatus

27. Kant's criticism is of a methodological nature and the argument presented in this section is independent of the exact formal nature of Kant's views on logic.

underlying Kant's reasoning via the process of *logical formalisation*. In this section, I briefly recapitulate the findings of Achourioti and Van Lambalgen as presented in *A Formalization of Kant's Transcendental Logic* (2011)—arguably the first work to rigorously adopt this approach—and point out how their approach auspiciously answers question Q3: what happens if we let go of the historical assumption?

Achourioti and Van Lambalgen

The aim of "A Formalization" is to investigate Kant's logic from a mathematical point of view and to restore a more favourable view of Kant's logic accordingly.[28] Its motivation lies in the discredited view of Kant's conception of logic (a consequence of the historical assumption). The result of "A Formalization" is threefold: (i) the discredited view of Kant's logic only arises within a classical first order semantics; (ii) the semantics derived from Kant's first Critique has a mathematical correspondent; (iii) the derived semantics is, contrary to common belief, a fragment of *intuitionistic* logic, i.e., geometric logic. The advantage of providing a mathematical model of Kant's thinking is that certain mathematical tools can be employed to demonstrate consistency and completeness of the logic underlying the formalisation (I will come back to this).

Achourioti and Van Lambalgen (2011) proceed, so to say, with a *tabula rasa*.[29] They commence with a philosophical analysis of Kant's ideas on cognitive activity, to subsequently arrive at an analysis of judgment as modes for binding sensory input in order to construct objects. Here, their first observation concerning the historical assumption arises:

> The semantics for classical predicate logic is given by models consisting of a domain of objects over which relations are interpreted; thus here the objects are assumed to be given. We have just seen however that for Kant, judgements somehow play a role in constituting objects from sensory material, so that it seems wrong to take objects as given from the outset. (p.259)

With this observation in mind, Achourioti and Van Lambalgen develop a mathematical model—reconstructed from their analysis of Kant's conception of judgment and objects—that ascribes a more central role to objects as the result of synthesis, i.e., judgment. Their result: all judgment forms relating to objects share the same logical feature, namely, they are all mathematically expressible as conjunctions

28. Alternative formal approaches to Kant's conception of logic are presented by Kovač (2008) and Pinosio and Van Lambalgen (2018). See Van Berkel (2015) for an analysis of the logical reasoning underlying Kant's transcendental argumentation in the Critique of Practical Reason.

29. The purpose of this section is to highlight some of the fundamental observations made in "A Formalization" and we refer to the article itself, as well as Achourioti & van Lambalgen (2017), for further details of the analysis and definitions of the obtained logic.

of geometric implications.[30] Interestingly, this logic of geometric implications is *intuitionistic*: for each finite set Γ of geometric implications which imply some geometric implication ϕ, there exists an intuitionistic proof of ϕ from Γ.

The above shows that Kant's conception of judgment is in line with a *constructive* stance on logic. A stance which is closer to an intuitionistic, rather than a classical view on logic.[31] In fact, it is this very constructive stance that breaks down the problematic interdefinability claimed by Strawson, as discussed above. Achourioti and Van Lambalgen (2011) conclude:

> It was Kant's intention to enumerate primitive logical forms which each express a particular function of cognition as it attempts to construct objects out of sensory manifolds. [...] Kant selected 'primitive' logical forms of judgement already with a particular transcendental purpose in mind, a purpose that renders the classical semantics underlying Strawson's objection inapplicable. (p.286)

In other words, through reconstructing Kant's account of judgment and object (fundamental to an understanding of Kant's table of pure logical forms of judgment) Achourioti and Van Lambalgen obtain an alternative semantics for Kant's theory of judgment, a semantics that is (i) inherently constructive and (ii) does not validate Strawson's claims of logical interdefinability.

Concerning question Q3, I want to make two final remarks. First, "A Formalization" shows that certain criticism—e.g., Strawson's—is resolved if one abandons the historical assumption and takes a non-classical stance on Kant's logic. Second, "A Formalization" also provides a favourable answer to one of Strawson's general worries concerning Kant's enterprise: the alleged completeness of the two tables (recall the last paragraph of section 'Strawson's criticism and 'The Bounds of Sense'). Namely, on Achourioti and Van Lambalgen's account the table of judgments is in fact complete. As they rephrased in *Kant's Logic Revisited* (2017):

> It turns out that the objectively valid formulas are exactly the geometric formulas. It follows that no judgement whose logical form is more complex than that allowed by the Table of Judgement can be objectively valid, i.e., this Table is complete. (p.863)

I believe that here we find a good example of how a better understanding of Kant's transcendental philosophy can be acquired when one lets go of the historical assumption.

---

30. For the sake of completion, a geometric implication is a formula $\forall\overline{x}(\phi(\overline{x}) \rightarrow \psi(\overline{x}))$ such that ϕ and ψ are geometric formulae, i.e., formulae constructed from atomic formulae using only $\vee$, $\wedge$, $\exists$, and $\bot$. See Achourioti & van Lambalgen (2017) for details.

31. Also, see intuitionism in the philosophy of mathematics and its relation to the role of space and time in Kant's transcendental aesthetics; Brouwer (1912), Posy (2013), and Tiles (2004). Contrasting Kant's view (and also that of intuitionism), Frege advocated elimination of any appeal to "intuition in the proofs of the basic propositions of arithmetic" (Zalta, 2022, Ch.2.7.1).

Conclusion

In this article, I discussed the historical assumption that Kant's logic is adequately representable in and compatible with classical first-order logic. To highlight the problems for Kant's philosophy arising from this assumption, I analysed Strawson's famous criticism of Kant's logic and discussed its relation to the historical assumption. Three guiding questions were posed with the aim of assessing the correctness of the assumption:

Q1 Where does the assumption come from?
Q2 Are there reasons to think that the assumption is mistaken?
Q3 What happens if we let go of this assumption?

As an answer to Q1, there are good reasons for tracing the confusion about Kant's logic and the origin of the historical assumption to Kant's own remarks about logic. In particular, Kant often claims Aristotelian logic to be a completed science since the days of Aristotle. Kant's remarks may have contributed to the idea that his views on logic do not (significantly) diverge from Aristotelian logic, the latter which is—as a monadic first-order logic without identity—compatible with classical logic. Nevertheless, I argued that from a methodological point of view Kant is simultaneously critical of the Aristotelian tradition, advocating the need to clarify, organise, and reduce obscurity in the science of logic.

In addressing Q2, I discussed Kant's only (but often overlooked) work on Aristotelian logic: *The False Subtlety of the Four Syllogistic Figures* (1762). In this work, Kant's criticism of the traditional division of syllogistic reasoning into four pure figures demonstrates his insistence on distinguishing those elements of logic that are genuinely pure and primitive, thus belonging to the science of logic, from those which are mixed and which merely obscure our understanding of logic (e.g., syllogistic figures II-IV). The discovery of fundamental principles and pure forms in establishing a well-grounded science remains a central theme in Kant's later work (especially in the Critique of Pure Reason). As a central contribution of the present paper, the analysis of "The False Subtlety" demonstrates that Strawson's criticism of Kant's logic is strikingly similar to Kant's own criticism of Aristotelian logic: both Strawson and Kant argue that certain logical forms/figures are falsely considered primitive. Furthermore, the analysis shows that (i) Kant was quite aware of the potential errors and pitfalls pertaining to the study of logic, as well as of the central duties of a logician, and that (ii) we should give Kant the benefit of the doubt when he claims his table of judgments to consist of pure, primitive judgment forms only. Consequently, there are good reasons to think that Kant's claim, concerning the completeness and pureness of the table of judgments, is not based on a logical mistake, but on an altogether different view of logic than the one ascribed to Kant in the historical assumption.

Last, as an illustrative answer to Q3, I recapitulated the results obtained by Achourioti and Van Lambalgen (2011), who demonstrate that the table of judgments can be shown complete if one adopts a different stance on Kant's logic. Their results are obtained through a mathematical reconstruction of Kant's conception of judgments and objects. Contrary to the historical assumption, the resulting logic is non-classical. With respect to the aim of the present paper, their results exemplify that certain problems claimed to belong to Kant's philosophy are in fact problems caused by the historical assumption.

⋆ ⋆ ⋆

As a final remark, with this article I hope to have induced the reader's sympathy for those philosophical investigations that, contrary to the historical assumption, pursue a non-classical conception of Kant's logic. The reconstructive method employed in such works reveals aspects of Kant's transcendental philosophy which are a priori excluded by the historical assumption. Such novel perspectives not only generate interesting questions and answers concerning the nature of Kant's transcendental and general logic, but may additionally clarify the structure of Kant's own logical reasoning in the arguments, proofs, and deductions provided in his critical philosophy (e.g., think of Kant's rewriting of the proofs occurring in The Analogies of Experience in the B-edition of the Critique of Pure Reason). Independent of the actual nature of Kant's logic, I hope to have the reader's support that the least we can do is to let go of the historical assumption as an abiding postulate.

Acknowledgements

For helpful comments and discussions on earlier drafts of this paper, I am very grateful to Rachel Boddy and Nils Kürbis.

Chapter 6

A personal message

Marian Counihan

Michiel:

We have hardly been in touch in the years since I completed my PhD under your supervision, but nevertheless the years working with you have left an indelible impression on me and on my work. What particularly remains is a furious desire to understand concepts or phenomena in a kind of 'total' way, with little interest in operating within disciplinary boundaries for this purpose. This delivers rewards and problems. Rewards because it enables one to be creative, to innovate and to operate more freely in the world, cross-cutting and collaborating across existing paradigms. Problems because nothing fits into an existing mould, established methods and quality standards are ill-fitting, and, well, the world still favours specialists. But having the opportunity to develop this way of working during the years of the PhD was, in hindsight, enormously valuable. In all the work I do since then, I am seeking to again experience the deep satisfaction I felt towards the end of my PhD, the sense of having a fundamental grasp of a phenomenon. We both originally studied mathematics. I have often wondered whether it was the sense of irrefutability that a mathematical proof delivers which we kept searching for, but then out in the messy world of empirical phenomena. I hope you have found it, at least to a sufficient degree, in your academic career, and can look back with a sense of satisfaction. And wishing you a healthy, productive, and peaceful next period in your life.

Warmest regards, Marian Counihan

Chapter 7

To count on... But how?

Jaap van der Does

An invitation to discuss philosophy of time and arithmetic in Kant and Early Wittgenstein

Sunday, 3rd October 2021[1]

Dear Henk,

Martin just e-mailed me that the *Festschrift* for Michiel will be delayed somewhat. Corona. Some hold there is a war going on against a virus raging. I rather think that people, whose bodies are multi-cultures of the little ones that will survive us all, have lost a balance in coping with nature. We shall be forced to cope. Let it be gentle.

Michiel... It has been more than twenty years since we have collaborated, and more than thirteen years since we've last spoke. We corresponded when he contributed to Martin's *Festschrift* (van Lambalgen (2011)). I spent my last years at the University of Amsterdam working with him. It was impressive, to say the least. Michiel was one of the most gifted in an environment full of gifted people. It was also a challenge. It seemed to me Michiel, like T.S. Eliot, wanted to have 'a lifetime burning in every moment'. Instead, I sometimes felt at home in a comfortable living room.

As you know, I grew up in an environment that was full of the human element. There were continuous efforts to come to grips with differences in character, and with traumas of war. There were worries about sufficient income. There was love and care, especially for the postwar generation. In this environ-

1. Me pen pal Henk is fictional, but he does remind me every now and then of Henk Visser, a Math-teacher at secondary school and later professor in Philosophy. Henk suggested I should study in Amsterdam or Groningen. Unfortunately, his part of our correspondence is lost.

ment access to more refined culture left to be desired. No way my family would have corresponded with a conductor about the dynamic signs in a Van Beethoven symphony, as you once did in some spare time between lessons. Still, I'm grateful my mother had a copy of Ella Fitzgerald's *Live in Berlin*. How High the Moon!

Don't you think in such an environment a rather blind talent for admiration can prove a strong force to bootstrap oneself? For sure I was helped copiously by times of social and economic growth.

Then, as now, there was much to admire in Michiel. His singular focus on philosophy and science. The depth and originality of his work in logic and set-theory. His erudition, in philosophy, science and literature. His strength of personality, which showed most profoundly when he had to deal with serious illness. Until today he still seems to resist the pressure of: publish or perish. He only writes what he thinks is worthwhile. In doing so he sets an example to all of us . His list of publications, on a diverse array of subjects, still keeps growing, and there is much to enrich one's insight.

When we collaborated, Michiel invented his so-called substructural quantification, in which explicit management of the dependencies of existential and universal quantifiers, properly interpreted, increases the expressiveness of the resulting system dramatically. When invited for this *Festschrift*, I wondered for a short while whether I should attempt to show that this logic can simplify and correct the original proof of Herbrand's theorem. Herbrand lived too short to notice that so-called spurious dependencies left part of his impressive work to be desired. The way in which Dreben, together with others, amended his proof was ingenious but fairly complex. Given the intricacy of the subject, I feared not to manage in time.

Do you recall *A Logic of Vision*? In 2000, it filled the entire first issue of *Linguistics and Philosophy*. It had its authors in alphabetical order: Van der Does & Van Lambalgen. This was my bad and highly misleading. We once cleared things up, but I still feel lousy about it. The highly original approach to perception reports and, among other things, evidentials was based on Marr (1982). It came with so-called conditional quantification, in which quantifiers were 'filtered', and with a new take on default reasoning. Of course, the logic was mainly due to Michiel. At the time I thought: 'So this is what it must have been like to work with Richard Montague.' Later Michiel refined the logic even more and expanded it to clarify Kant's critical philosophy.

I was a very old postdoc – 42 and counting! – with no prospects on grants or tenure positions, and with a lovely wife and two lovely, promising children at home. Our ways parted. As a result, I could not keep track of Michiel's career. Now that I, like Michiel, am about to retire I hope to catch up.

Wednesday, 6th October 2021

Dear Henk,

I have a better idea for this *Festschrift*. In van Lambalgen (2011), Michiel invited Martin to a discussion on the status of time in Wittgenstein's *Tractatus*. Whether this discussion took place or is still on-going, I do not know. I have found no published traces of it. Did you? As discussing this topic is open to anyone, I should perhaps take the liberty and share my bit.

What do you think? Would it be too ambitious to broaden the scope of the discussion and try to get more clarity on the the notion of time in *Tractatus* via a comparison with the rôle of time in Kant's critical philosophy of arithmetic?

Without doubt this is a tricky route, for several reasons. Kant's philosophy on this topic – of which Michiel is an expert and I am not, – is subtle and complex. Also, Kant's philosophy of mathematics keeps engendering a lively debate. See Shabel (2021) for an overview. But perhaps these objections should be seen as virtues. Some people prefer tricky routes. And isn't it the purpose of my short note to rekindle the exchange Michiel once aimed to start? Let me seek contact with a *Stalker* and anticipate on what it would be like to enter this *Zone*. Please bear with me, for I will share my attempt to get clarity about all this with you.

Saturday, 9th October 2021

Dear Henk,

When I lived in Frankfurt am Main a few years ago, near where Schopenhauer used to live, I sought contact with Ede Zimmerman at *Goethe-Universität*. There I met people who as part of their basic education had to study Kant's *Kritik der reinen Vernunft* (1781, 1787) thoroughly. Would that still happen nowadays? I also recall that for Natasha Alechina, who was at ILLC when I was working with Michiel, this was part of her schooling. During my *kandidaats* in Leiden, a kind of BA, we read *Prolegomena zu einer jeden künftigen Metaphysik* (1783); no more Kant, no less.

In view to this *Festschift* contribution, I re-read some relevant parts of Kant and the overview in Shabel (2021). I also studied Engelhard & Mittelstaedt (2008), who aim to show how Kant's scattered remarks can be used to give a cohesive view on his philosophy of arithmetic. Given the *petite mer à boire* on this subject, this is way too little to come up with something serious. But perhaps it suffices to see what the main issues are to further a comparison with early Wittgenstein.[2]

2. In what follows, the parts on Kant profited much from Engelhard & Mittelstaedt (2008).

I think it would be good to proceed on the assumption that Kant's view on time can be seen as deeply influenced by his view on mathematics. I will focus on the question:

How is pure arithmetic possible?

Kant indeed presumed arithmetical judgments to be synthetic *a priori*. In answering this question, Kant uses major parts of his critical philosophy in subtle and intricate ways. It's a shame I'm unable to do justice to all this.

That arithmetical judgments are *a priori*, and so necessary and universally valid, is rarely contested. But, as you know, that they are synthetic, and so require a conceptual synthesis going beyond the concepts involved in these judgments, has been deeply criticised, among others by Frege and Russell, founding fathers of analytical philosophy. Engelhard & Mittelstaedt (2008) give a nice summary of Kant's position (p. 251):

> His famous example for showing the synthetic character of mathematical judgments is the arithmetical equation $7 + 5 = 12$. This judgment means: "the sum of '7' and '5' equals '12"'. Kant's argument is that in judging the concepts '7', '5' and 'sum' it is thought that '5' and '7' are united to one single number, but not that this number equals '12'.

The crux of Kant's insight, or so I understand it, is that the judgment concerns *concepts* of numbers and addition, and that logical analysis does suffice to yield the concept of this addition having a unique result, but not the concept of the addition's specific result. Instead, establishing the identity requires the constructive synthesis of *calculation* with singular *instances* of these concepts. For me, this still leaves much in the dark.

Sunday, 10th October 2021

Dear Henk,

I think I may have found an anachronistic, perhaps misleading way of understanding Kant's position why arithmetical identities are synthetic, not analytical.

It is open to debate to what extent Kant's notion of *Menge* can be captured as a more recent notion of set,[3] or what notion of set is involved in Kant's take on concepts of specific numbers. Be this as it may, for the sake of clarification why not assume a sufficiently large universe E of numerically different entities given? Then extensional, non-Kantian proxies of the concepts **five**, **seven** and **twelve**

3. See (Engelhard & Mittelstaedt, 2008, footnotes 13, 35) for details.

may be modelled along the scheme:

$$\mathbf{n} := \{X \subseteq E : |X| = n\}.$$

Let us not worry yet about how in this setting $|X| = n$ is to be determined. Isn't it more important to notice that if the concept of addition comes with a way of combining **five** and **seven** into a new number, but a way that does not involve any operation on their instances (elements), that for such an operation it is unlikely something resembling the concept **twelve** is to be found? For then:

$$\mathbf{add}(\mathbf{5}, \mathbf{7}) \subseteq \mathbf{5} \cup \mathbf{7} \text{ and } \mathbf{add}(\mathbf{5}, \mathbf{7}) \cap \mathbf{12} = \emptyset.$$

This is true in particular if rather simple-mindedly **add** is taken to be union. Accordingly, such strange and useless concepts of addition leave $5 + 7 = 12$ not analytic: we can even assume without contradiction that $5 + 7 \neq 12$.

Obviously, something more involved is required. In particular, we seem to need an operation in terms of *instances* – here: *elements*, – of the number concepts:

$$\mathbf{add}(\mathbf{n}, \mathbf{m}) := \{X \cup Y \subseteq E : X \in \mathbf{n} \ \& \ Y \in \mathbf{m} \ \& \ X \cap Y = \emptyset\}.$$

This operation does have $\mathbf{add}(\mathbf{5}, \mathbf{7}) = \mathbf{12}$, and now $5 + 7 \neq 12$ is clearly contradictory. With this concept of addition at hand, what still makes the judgement $5 + 7 = 12$ non-analytical?

What do you think? Is it that the concepts **add**, **5**, **7** as such, without being combined or synthesised, have no trace of **12** yet? Or is it rather the observation that this notion of addition – or of similar such functions, – requires an operation on instances $X \in \mathbf{n}$, *etc.*, instead of on concepts *simpliciter*, and that this allows us to hold that such addition is *not* purely conceptual? If addition is not purely conceptual, the entire judgment is obstructed from being analytical.

As I said, my attempt is anachronistic. Presumably, too anachronistic. . .

Friday, 15th October 2021

Dear Henk,

Two and a half months from now I will be retired. I look forward being able, among other things, to study and write again. Would I still be possible to do something worthwhile? Who knows? Wayne Shorter's favourite saying is: Let's go for it!

In this phase of my life I realise stronger than ever before what a privilege it is to have had a proper schooling, and to have teachers and colleagues, now friends, who still want to collaborate. I realise, too, that each study comes with its own

biases. I think it is fair to say that my education in logic and formal semantics was to a certain extent under the influence of the *Wiener Kreis*. I, for one, followed Carnap thoughtlessly an held Heidegger in low esteem, and this without having read a single word of him. How silly! I still find Heidegger's work hard going, but have learned in the meantime how rich and interesting it is. Perhaps a similar bias kept me from a proper study of Kant. Given the criticism of Frege and Russell, I thought there were better ways to spent the limited time available for philosophy. Here, too, I was mistaken.

Thanks for your e-mail. It is now clear to me the above attempt to clarify that arithmetical judgments are synthetic *a priori* only goes thus far. Kantian concepts are intensional rather than extensional; they capture abstractions, general aspects that objects have in common, aspects involved in their synthesis. Such concepts are not just some kind of sets of objects that have these features. Be this as it may, I will try to retain the core of the clarification in my new attempts.

For Kant judging comes with constructive activity in which our sensibility, understanding and reason interact. With arithmetical judgements synthetic, their truth-conditions follow the line of all synthetic judgements. Engelhard & Mittelstaedt (2008) put it thus (p. 251):

> The truth condition of synthetic judgments is connected to intuition as well as to possible experience (CpR, B 73; 193 ff.). This implies that synthetic judgments cannot be deduced from logic. The truth-claim of a synthetic judgment, which claims objective reality, has to be justified by the object, to which it refers. Since mathematical entities depend on construction, the truth claims of arithmetical judgments have to be justified by construction.

Thus, referents need to be found for numeric concepts, and here time as pure form of our internal sense comes in. Indeed, I started my search for more clarity at the conceptual end, which is abstract and general. Perhaps it is better to focus first on the singular, the specific that is given more immediately.

Dinnertime, I have to cook. More on this later.

Saturday, 30th October 2021

Dear Henk,

Do you know next year it is hundred-fifty years ago that Piet Mondriaan is born, at *De Korte Gracht*, *Amersfoort*? The *Fotobond*, of which *Fotokring Eemland* is a member, is organising an exhibition in *Het Rietveldpaviljoen*. I was unaware that

Mondriaan was a painter *and* a writer. To prepare members of *Fotokring Eemland* who want to submit work, I thought it could be helpful to give a short presentation covering both Mondriaan's work and text. This was gratifying but kept me busy for a while. Back to Kant, now starting and the sensible, singular end.

What about the following 'singular' variant of Kant's argument? Arithmetic in its most basic form comes with the manipulation of fairly simple signs, such as:

$$|,\ ||,\ |||, \ldots .$$

In terms of such signs, an analogue of Kant's observation is: one may recognise the sign '|||||' as an instance of the concept *five*, one may recognise the sign '|||||||' as an instance of the concept *seven*, one may know moreover that in this context *addition* comes with the concatenation of such signs. Thus, one knows the result will be unique. Still, this does not mean that '||||||||||||', i.e., the sign of five and twelve's addition, is recognised to be an instance of the concept *twelve*. Apart from conceptual 'effort', this requires one to count the elements of the resulting sign in a sensible way, and the result of that is not immediate.

Let me try formulating what this part of Kant's philosophy amounts to, for sure disregarding subtleties that cognoscenti find crucial. Should you have any comments or additions, please let me know.

As is all too familiar, signs only flourish in a much wider contexts, and here difference in philosophy is to a large extent difference in view on what that context should be. The beauty of Kant's philosophy, or so I think now, is captured well in what Engelhard and Mittelstaedt baptise the *Construction Thesis*. In this thesis, conceptual, epistemological aspects interact with objective, ontological ones:

> In epistemological terms, mathematical knowledge is gained through construction of mathematical concepts in pure intuition by reason. In ontological terms: mathematical objects, that is mathematical entities with which mathematical concepts and judgments correspond, are products of reason's construction in pure intuition.
>
> *(Engelhard & Mittelstaedt, 2008, 250)*

According to Kant, both the construction of concepts in an arithmetical judgement and the construction of the referents such judgements are about, only involve a self-affection of our cognitive abilities and so remain pure *a priori*. Synthetic *a posteriori* judgements are about objects grounded in the receptivity of our sensibility. By contrast, synthetic *a priori* judgements lack such receptive content.

They refer to entities constructed in intuition, i.e., a constructive activity of the understanding rooted in the pure form of time.

Why time? More in general the notions of time and space are constitutive of our ability to have experiences, respectively as subsequent and internal or – in combination with our ability to construct objects out of experiences – as induced by things that are externally positioned, as different in different places. On Kant's account, time and space are neither entities in and of themselves, in which other entities appear, nor empirical properties of or relations between objects that exist independent of us. Time and space are *a priori*: i.e., inherent to the way in which we represent (*vorstellen*). We cannot imagine there being no time or space, but we can imagine time or space to be empty.

But again, why time, especially in relation to pure arithmetic? That time and thus sensibility is involved in mathematical knowledge is based on Kant's insight that mathematics must be applicable. Engelhard & Mittelstaedt (2008) put it this way (p. 253):

> Mathematical knowledge gains its meaning only through possible or even actual experience (CpR, A 239 f./B 298 f.). That is, it is possible that an empirical object exists, which falls under a certain mathematical concept. As an example, the number '5' has a meaning if it is possible that a group of sheep exists that falls under the concept 'five.'

Also, Kant held arithmetic to be constructive, so it is quite natural to expect time to have some rôle to play, even if limited. As process, 'timeless construction' is an oxymoron. Only in time is it possible to successively construct different entities, e.g., strings:

$$\underbrace{|||\ldots|}_{n-times}\ .$$

This means time must be rich enough to allow for such construction.

In van Lambalgen (2013), Michiel treats of Kant's notion of time without regarding its relationship to philosophy of mathematics. But if Kant had not taken mathematics to be synthetic *a priori*, would he have had reasons to treat of time as a pure form of our sensibility? I take it, Michiel just wanted to concentrate on time *simpliciter*, which is more than challenging enough. Anyway, the properties of time that Kant discerns at different places, and which Michiel shows to be consistent and rich in structure, are obviously sufficient to do arithmetic, and more... This is clear, e.g, from the infinite divisibility of time that Kant mentions in his transcendental aesthetics and that Michiel formalises in van Lambalgen (2013), section 3.1. Yet, it does make me wonder: how does Kant's approach compare, e.g., with

assuming axioms, such as the axiom of infinity? Or is it, on Kant's view, rather the synthetic *a priori* nature of science that prompts the richness of time, and with that makes the explanation of a purely mathematical form of infinity less pressing? Is it a modern bias to expect: Math first!?

In the meantime I'm working on my submission for the Mondriaan exposition. It's far from obvious how to come up with something visually rich, which has an obvious connection with Mondriaan's work, but refrains from being a straightforward analogue. Let's see. The observation that it is the process that counts, deserves to be a cliche.

Thursday, 4th November 2021

Dear Henk,

Philosophy. In the above sketch of numbers, their appearance as a sequence generated over time was most prominent. But how does this lead to quantity and cardinals? It all begins with a form of counting, i.e., iteratively discerning moments in time and assigning numerals to them according to a clearly defined rule:

num1 | is a numeral
num2 If n is a numeral, so is $n^\frown|$.

Here, '$^\frown$' indicates the concatenation of two strings. In this way a linearly ordered series is created, in which the choice of notation is quite arbitrary as long as the notation can be iteratively produced. Qualities of signs or of the moments in time they are assigned to, are irrelevant. Although time enables the production of different numerals, in the end what counts is the resulting series and the positions of the numerals in it. That is, our understanding abstracts the quantitive concept of a number as cardinality from its ordinal position in the generated time-series. The abstractions can only be due to construction in time, but the abstractions themselves are timeless.

Arithmetical judgements and the concepts involved in them require sensibility to enable construction in time. In a way, the signs resulting from the construction are like duck-rabbits in disguise. On the one hand the sign:

$$\underbrace{|||\ldots|}_{n-times}$$

can be seen as a sequence of different entities assigned at different moments in time, on the other hand it can be seen as a unity symbolising the quantity of the number of strokes occurring in it. On Kant's view, this shift in aspect – my term, not his, – is due to our understanding being able to synthesise multiplicity into

unity through the pure categories of quantity: unity, plurality, totality. In case of a specific number this is the unity of set-like objects that represent that specific cardinality.

Now that we have numbers, what about numeric functions, such as addition, multiplication and so on?

Kant observes that basically the same insight can be used in case of addition, for addition can be seen as continued counting. Engelhard & Mittelstaedt (2008) quote from a letter of Kant to Johann Schultz (p. 261): 'Arithmetic has, as Kant emphasises, no axioms but postulates, i.e. instructions for actions:[4]

> If I consider $3+4$ as the expression of a problem, namely to the numbers 3 and 4 to find a third one $= 7$, [...] then the resolution takes place by a simple action, which does not need a special instruction, namely by the successive addition that generates the number 4, merely performed as continuation of counting the number 3.'

In general, identities are established by showing that the instruction for (mental) action on the left leads to the same number – abstracted from its position in the sequence generated, – as the instruction for action on the right. Engelhard & Mittelstaedt (2008) emphasise the rule-based necessity involved (p. 257, 258):

> The schema of quantity – the rule that prescribes, how the successive addition has to be executed – guarantees that whenever the elements are counted, the result of counting invariably will be identical. The rule guarantees, therefore, the strict universality of the results of any counting operation or result of a calculation. This is the basis of strict universality and necessity of arithmetic.

Formally, the rule for addition can take the following shape:

sum1 $n + | = n^\frown|$
sum2 $n + m| = (n + m)|$

At this point, the reconstruction of Kant's arithmetic in Engelhard & Mittelstaedt (2008) stops. Of course, their reconstruction reminds one of the logic-free approach in Skolem (1923), an approach that is hinted at in the *Tractatus*, and that has been developed further by, among others, Hilbert, Bernays, Curry, Robinson, Goodstein, ..., work Engelhard & Mittelstaedt do not refer to. Before the reconstruction can be seen to yield a serious candidate for arithmetic, more detail needs to be given (i) on how to handle zero – nil action? – and make it part of the system,[5] (ii) on basic functions – such as multiplication, exponentiation,

4. Translation: Engelhard & Mittelstaedt.
5. This, I think, is unproblematic, cf. the rule: **num0**: The empty • is a numeral; **num1′**: If n is a numeral, so is $n^\frown|$, and *mut. mut.* the same for the other rules.

tetration,. . . , which as repetitions of previously defined functions will pose no problem, – (iii) on the general rules for iteration or recursion, (iv) on different forms of substitution, (v) on the PRIM uniqueness rule, *etc.*. See, e.g., (Goodstein, 1957, Ch. I, V). For some of this, e.g., substitution, we could perhaps use Kant's approach to *Buchstabenrechnung* (algebra). Cf., (Engelhard & Mittelstaedt, 2008, 262). Given the apparent richness that Kant holds time to have, i.e., the realm of construction, one wonders how strong induction can be? Or is this in Kant's philosophy a conceptual matter, concerning the rules involved in the schema of quantity? To be honest, I have to study Kant's philosophy of arithmetic more deeply before I can work this out myself.

Wednesday, 10th November 2021

Dear Henk,

Let me return to ground that for me is more familiar: Wittgenstein. So, you own a copy of Ostwald's *Annalen der Naturphilosophie*, 14 (1921), in which Wittgenstein's *Logisch-philosophische Abhandlung* had its first 'piratic' publication? Nice! But I must say I prefer the *Kritische Edition* or more dynamic versions, such as `tractatus.nl`, which allow one to read the tree-structured text in different ways.[6]

As I stated earlier, my rambling on Kant is prompted by Michiel's contribution to Martin's *Festschrift*, in which he invites him to discuss the notion of time in the *Tractatus*. Here is how I would respond to his invitation. If you want to use this to compare with Kant, feel free.

It is best, or so I think, to begin with discerning the ethical aspect of the Tractatus. This concerns finding a proper, timeless stance toward that problematic world we find ourselves in.

> Der Tod ist kein Ereignis des Lebens. Den Tod erlebt man nicht.
>
> Wenn man unter Ewigkeit nicht unendliche Zeitdauer, sondern Unzeitlichkeit versteht, dann lebt der ewig, der in der Gegenwart lebt.
>
> Unser Leben ist ebenso endlos, wie unser Gesichtsfeld grenzenlos ist. *6.4311*

Sorry, too much beauty is lost in translation. When Michiel first read these lines, he 'understood them as referring to a mental struggle in which the darker features of the 'lapse of time' – decay and death – had to be overcome by a conscious

6. A modernisation and extension of `tractatus.nl` is planned for 2022.

effort to live in the present; but the struggle presupposed that the lapse of time and change are real, even if only in the relative sense of 6.3611':

> Wir können keinen Vorgang mit dem »Ablauf der Zeit« vergleichen – diesen gibt es nicht –, sondern nur mit einem anderen Vorgang (etwa mit dem Gang des Chronometers). *6.3611a,b*

Cf. (van Lambalgen, 2011, 1). But shouldn't it be stressed the ethical 'present' comes from an instantaneous shift, a mystical insight in the world being.

> Nicht wie die Welt ist, ist das Mystische, sondern daß sie ist. *6.4423*

> Wie die Welt ist, ist hier vollkommen gleichgültig. Gott offenbart sich nicht in der Welt. *6.44231*

How this relates to common experience is left overexposed. The ethical 'present' is not in time, is not preceded by past, is not succeeded by future. It is holistic, timeless (*Unzeitlich*), spaceless, bleak, presumably, and without colour.

> Die Lösung des Rätsels des Lebens in Raum und Zeit liegt außerhalb von Raum und Zeit. *6.4312a3*

Early Wittgenstein's urge for 'space-timelessness' is similar to that of Tolstoy's now in his *Gospel in Brief*, and thus ultimately in a way to that of Schopenhauer.[7] See (Stokhof, 2002, ch. 4).

Apart from the mystical absence of time – the world purely being, – there is the mundane living in the coloured space-time of our world. Although this worldly life is in time and space, Michiel still wonders whether 'space-timelessness', my term, is not built into the tractarian ontology (Van Lambalgen, 2011, 1). Theses 2.0232 and 2.0251 are key to any attempt to understand the place of space, time and colour in the tractarian system.

> Beiläufig gesprochen: Gegenstände sind farblos. *2.0232*

> Raum, Zeit und Farbe (Färbigkeit) sind Formen der Gegenstände. *2.0251*

Let me quickly note an important difference with Kant. For Kant colour is fully subjective, like Locke's secondary qualities. Early Wittgenstein seems to hold time, space and colour by and large on a par. The point being, or so I presume, Kant saw no synthetic *a priori* judgments whose truth-conditions would require colour as a

7. Both Tolstoy and early Wittgenstein were inspired by Schopenhauer.

pure form of our sensibility. Wittgenstein is unconcerned with the representation involved in such judgments:

> Ein a priori wahres Bild gibt es nicht. *2.225*
>
> Die Erforschung der Logik bedeutet die Erforschung aller Gesetzmäßigkeit. Und außerhalb der Logik ist alles Zufall. *6.3*

Concerning time, the *Tractatus* seems to leave narrow space for heuristic manoeuvre. One option is to attribute time solely to configurations of objects (*Sachverhalten*). Then, objects themselves are indeed as time- and spaceless as in a way they are colourless (2.0232). Another option is that time can somehow be attributed to objects themselves. For how can time result from combination of timeless objects? Well, according to Wittgenstein all material properties result from such configuration (2.0231) and colour does (2.0232, 2.0232 above), so why not time and space?[8]

The idea that states of things (*Sachverhalten*) are in time – and the objects occurring in them in time only as epiphenomenon, – is based on the observation that time is a form of objects (2.0251) and so is a way in which objects may combine. This suggests, to me at least, their combination is in time, but not so much the objects themselves. Thesis 2.0232 can be read to support this interpretation: time, space, colour are on a par. Since objects are colourless, they are also time- and spaceless. Therefore, only *Sachverhalten*, states of things, are in time.

Reading states of things as timed, hinges on the translation of *beiläufig gesprochen*. Its translation varies quite a bit. Ramsey and Ogden translate *beiläufig gesprochen* as 'roughly speaking'. Pears and McGuinness have 'in a manner of speaking'. W.F. Hermans has '*tussen haakjes*' (between parentheses). According to one of my dictionaries something like 'noted in passing' would be correct as well. As editor of Martin's *Festschrift* I consulted Joachim Schulte on the matter. He wrote:

> Thesis 2.0232 is (I remember this very well from my own case) difficult and puzzling for a non-Austrian German speaker, perhaps even more so than for a non-native speaker, because one does not dream of looking at a dictionary.
>
> As a matter of fact, in my (and most people's) German *beiläufig* means *en passant* (etc.). But in Austrian German it means 'roughly' (etc.). So both English translations are right ('roughly

8. Thesis 6.3751 suggests the 'logical structure' of colour can be reduced to that of time and space. I agree with Ramsey (1923) that this reduction is without success. If the logical structures of time, space, colour, matter are due to the object configuration of states of things, their famous exclusion problems comes in their wake. As Wittgenstein realised later, this contradicts the assumed logical independence of states of things (2.061). This clearly shows that reducing colour to time or space does not help here.

> speaking', 'in a manner of speaking'). On p. 65 of the Ogden correspondence (1973) you find (under 5.5303) that LW himself had changed the translation of 2.0232 to serve his purposes.
>
> I'm sure that LW was familiar with the standard German use of *beiläufig*, but in his own use he generally seems to have followed the Austrian pattern, if I remember correctly. *E-mail: 4th October 2011*

If only states of things are in time, we are left with objects that are, roughly speaking, timeless. How to understand this *roughly*? That this is roughly so, may also have to do with the idea that objects need not occur in realised states of things. Time is a form of objects, and so a state of things a temporal possibility. States of things are in time, I would like to suggest, only if they are realised.

Now, the realised states of things are not ordered in a flow of time given independently. According to 6.3611 quoted above, there is no such flow. Instead, realised states of things induce a flow of time, but this formally, without material content. Perhaps the order of time can be compared with how logical structures can be ordered purely formally, based on their complexity. On this view, there is no need to assume that realised states of things, as all mutually related in the induced time, turn the world into one unique states of things, as Michiel suggests ((van Lambalgen, 2011, 2, 1st par.)). On this view, we should distinguish between formal and material relations, and hold temporal relations formal, and *mut. mut.* the same for space and colour (cf. 4.122).

Beiläufig gesprochen, 2.0251 seems to admit of objects that lack a temporal form. In this regard, objects are not necessarily sempiternal, as e.g. Hacker (1984) holds.

Friday, 12th November 2021

Dear Henk,

Last Wednesday I shared my thoughts on the nature of time in the *Tractatus*. Clearly, the heuristics was suggestive rather than conclusive. But aren't fixed interpretations a luxury in philosophy, perhaps in a way alien to it? Lately, Martin shared a draft of his article 'Philosophy as Change' with me, which I read as suggesting this.

Did you notice my interpretation doesn't say much about the structure of time? This may seem strange, for a large part of Michiel's note is about searching for a formalisation. Actually, I hold the few theses in the *Tractatus* on time, space and

colour do not allow us to give a very detailed reconstruction of their structure.[9] In this regard, the *Tractatus* seems much poorer than Kant's first Critique. But this poverty may well be deliberate. I surmise early Wittgenstein held that the details of this structure are not for philosophy to specify. Philosophically it suffices to indicate the place of time, space and colour in the overall philosophy. To make clearer why I think so, let me share what I think the nature of the system is that the *Tractatus* indicates.[10]

Other than for modern logics, it can be argued the tractarian system is top-down rather than bottom-up. Like Frege, early Wittgenstein held sense to be determinate. Also, he held sense to be prior to truth. Thus, once a philosophical problem forces one to analyse what is meaningful underneath the clothing of ordinary meaningful sentences, a series of 'decompressing' definitions will result in truth-functions of elementary sentences that shows the form of their sense. Its content, what it is about, comes from projecting the function onto states of things. In a way, both language and what language is about are internal to logic.[11]

In elementary sentences and states of things, names and objects can be discerned. However, due to a strong contextuality principle, the names should not be considered independent from the sentences in which they occur.

> Nur der Satz hat Sinn; nur im Zusammenhange des Satzes hat ein Name Bedeutung. *3.3*

Given projective isomorphism, the same holds for objects and states of things. Although states of things are determined by the combinatorial possibilities of objects, strictly speaking the states of things are not generated from them. The combinatorial possibilities are given holistically, in logical space as a whole. In this regard, it is indeed telling that the *Tractatus* has no variable to generate elementary sentences, as it does for logically complex sentences. In thesis 6, elementary sentences are assumed given.

What does this mean for the rôle of time in the tractarian system? As observed above, time is a form of objects. Time is not itself an object, nor a type of objects. Also, temporal properties and relations or no objects, but come with the form of objects (if the form of an object is indeed temporal, Wittgenstein seems to leave open whether this is always so). In the *Tractatus*, what in modern logic is called, relations, functions, constants, are all on a par. They are all objects, be it with different kinds of form.

---

9. And there is not much more in the documents from which the *Tractatus* is composed. To check this, a search on the powerful `wittfind.cis.uni-muenchen.de` suffices.

10. For details, see also van der Does (2011), which can be read as a more formal sequel to Stokhof (2002). Parts of my book should be updated to bring it in line with van der Does & Stokhof (2020).

11. Somehow time *etc.* will be part of our representation, i.e., a logico-linguistic analogue of Kant.

Since time is a form of objects – so itself without content, – time only shows at the level of states of things. Michiel rightly observes that on this view each state of things is turned into a temporal 'island'. But, I would like to add, these islands are possible islands. As time is purely formal, states of things when realised can be in the order of time. Yet, the notion of time is left so abstract that no further structural properties (linearity, density, continuity, branching, etc.) can be assumed. According to Wittgenstein, such are not a priori, only logic is.

In van der Does & Stokhof (2020), Martin and I shed new light on the idea that the *Tractatus* has sufficiently many indications that Wittgenstein's philosophy can be held continuous in certain respects. That its ontology allows for restricted pragmatic variation is one source to see such continuity, a source that attained abundant flow only later. Now, if the form of objects allows for some degrees of freedom, the structure of time may be partly up to us to construct. This can be read as preparing the way for the kind of philosophy Schatzki (2010) develops, inspired by Heidegger and the later Wittgenstein. In this philosophy a meagre notion of temporal succession is held to be objective, merged with a meagre notion of objective spatial in-betweenness. Richer structures of space-time are left human, depending on social practice and what people aim for in certain contexts.

Thursday, 16th December 2021

Dear Henk

It has been long since I wrote to you. The hurly-burly of my upcoming retirement got hold of me: it is more complicated to arrange for things when times go viral. To keep my peace burning I sometimes returned to my little black cushion and sat. This made me realise it was Michiel who pointed me in this direction. During one of the summer schools we attended, to spread the logic of vision, we came to speak about meditation. Michiel suggested I should read Shunryu Suzuki's *Zen Mind, Beginner's Mind* or Charlotte Joke Beck's *Everyday Zen*. I loved these classics. His suggestions enriched my life and helped me cope with disappointments to come (fortunately not too many, up until now). I guess my letters to you have made it clear that beginner's mind is now quite prominent, be it in the more common sense.

Another thing I realised lately is how strong early Wittgenstein's personality must have been. He was influenced by two of the most prominent logicists – '*den großartigen Werken Freges und den Arbeiten meines Freundes Herrn Bertrand Russell*' (*Vorwort Tractatus*) – but in philosophy of mathematics never seemed to have had

the inclination to become a logicist himself. In Ramsey's copy of the *Tractatus* the following notes on philosophy of mathematics were added near 6.02 and 6.03, a topic developed further in 6.2 and its offspring:[12]

> The fundamental idea of mathematics is the idea of *calculus* represented here by the idea of *operation*.
>
> The beginning of logic presupposes *calculation* and so number.
>
> Number is *the* fundamental idea of calculus and must be introduced as such.

Some of Wittgenstein's reasons for taking this position are much like those of Kant. To give just one indication, mathematics is embedded in human life, but if, when in doubt about $5 + 7 = 12$, humans first need to determine, e.g., whether

$$\exists \mathbf{5} x.Px \wedge \exists \mathbf{7} y.Qy \wedge \neg \exists z[Pz \wedge Qz] :\Leftrightarrow: \exists \mathbf{12} v.[Pv \vee Qv]$$

is a tautology, they must, among other things, count the signs involved in its full definition. Thus, in an important sense, counting is prior to logic, and a philosophy of mathematics should do justice to that.[13]

A direct approach to calculation should be promising. Kant proposed one by postulating our sensibility comes with a pure form of time which enables counting, pure and applied. Wittgenstein points in another direction:

> Die Frage, ob man zur Lösung der mathematischen Probleme die Anschauung brauche, muß dahin beantwortet werden, daß eben die Sprache hier die nötige Anschauung liefert. *6.233*
>
> Der Vorgang des Rechnens vermittelt eben diese Anschauung.
> Die Rechnung ist kein Experiment. *6.2331*

I take Wittgenstein's move toward the logic of language to be a move away from Kant. Numbers become indices of the operators that generate complex logical structure.

> Die Zahl ist der Exponent einer Operation. *6.021*

Given an arbitrary such operation Ω and a now more common notation of application, 6.02 defines numerals much as in the above reconstruction of Kant.

$$\begin{cases} \Omega^{\bullet}(x) & := & x, \\ \Omega^{n|}(x) & := & \Omega(\Omega^{n}(x)). \end{cases}$$

12. See e.g. (Wittgenstein, 1922, 113).
13. Of course, much more is needed to develop this indication into a full-blown criticism of logicism. See (Wittgenstein, 1967, Teil III) and the abundant, very helpful clarification in Mühlhölzer (2010).

Here, x is some structure already available or generated. Similarly for addition:

$$\begin{cases} \Omega^{n+\bullet}(x) & := \quad \Omega^{n}(x), \\ \Omega^{n+m|}(x) & := \quad \Omega^{(n+m)|}(x). \end{cases}$$

Early Wittgenstein's approach raises about the same question as posed with regard to Kant; in short: how powerful is the resulting arithmetic? A strong connection with fragments of arithmetic has been noted by Frascolla (1994), Frascolla (1997), Marion (1998), Potter (2000). In Frascolla (1997) one finds a formalisation of the idea that identities $t = s$ between terms obtained from $0, S, +, \times$ are provable in the equational fragment of Peano Arithmetic, if and only if the corresponding operational identity is provable:

$$\Omega^{t}(x) = \Omega^{s}(x) \Leftrightarrow \vdash_{PE} t = s.$$

Frascolla develops a theory of operators for this purpose that ensures the result of applying operator-sequences is invariant enough under different groupings.

In all this, any 'basic' operator should do, e.g., the operator N of joint negation. That Wittgenstein phrases his approach for operators in general, I take to be related to the observation that the same numeric structure can arise in a myriad of ways, and so should not be tied to a specific choice. One should rather show that all these approaches are formally identical. This is an insight that Wittgenstein captures in the *concept* of number, which he identifies with the general form of all numerals, the rule of succession:

$$[0, \eta, \eta + 1].$$

It would be nice to show more directly that PRA can be had in this way, for instance as formalised in Goodstein (1957), but whether much more can be had is doubtful. Ramsey (1923) wrote about Wittgenstein's approach to mathematics:

> I do not see how this account can be supposed to cover the whole of mathematics, and it is evidently incomplete [...].

This cannot be denied. Wittgenstein's point seems to be the logicist program is untenable even in the more elementary parts of mathematics, and thus it suffices in a way to point toward an alternative only for the part criticised. Yet, even here much remains to be done to develop Wittgenstein's suggestion in more detail. Anyway, that Kant bases his philosophy on the pure form of time, which he assumed to have a rich structure, may be one of the reasons why Kant's approach is presumably more expressive.

This brings us back to the notion of time, about which we corresponded earlier. How do Kant and Wittgenstein compare in this regard? The distance between the

two philosophies will depend largely on how the *Tractatus* is read. If, as Hintikka does, one takes its objects to be like sense data, time as a form of objects is quite similar to time as a form of our sensibility.[14] Instead, I'm in the camp of those who hold objects to be logical in nature, which, as I showed earlier, makes the differences between Kant and Wittgenstein more pronounced.

What is most relevant now is what rôle time plays in both philosophies of arithmetic. Kant's philosophy is based on the assumption that time as a pure form of our sensibility can resolve, among other things, the apparently problematic, synthetic *a priori* nature of arithmetic. The referents involved in arithmetical judgement remain pure *a priori* as they are constructed in the pure form of time. This is very different from Wittgenstein: numerals as indices of operators do not refer, their rôle is just to keep track of iteration. In *Tractatus* the operation is on states of things that are in time only when realised. Its logical nature allows iteration to be potentially countable, thus it is independent of the richness of pure time.

Formally, the arithmetics that Kant and Wittgenstein indicate are at least primitive recursive. Kant's arithmetic is perhaps stronger, due to the rich structure of time and to the interaction of sensibility and understanding, which may also allow for more powerful rules. How much more powerful, I do not know.

Reading Kant, even in part, is a deeply human experience, intellectually as well as emotionally. Intellectual creativity comes with an abundance of subtle distinctions and inspiring moral insight. Yet, I wonder whether this richness would flourish even more when freed from its cognitive prison. To posit that numbers result from an interplay between pure understanding and pure sensibility, is in a way the strongest explanation if the phenomenon of arithmetic is seen as synthetic *a priori*. But an explanation in terms of what? For me, pure time and pure understanding continue to have a non-naturalistic, 'spooky' feel about them.

Can Kant's insights be transported to the open, where nothing is hidden? To human practices with human ends? Then we 'move away from a purely conceptual analysis' (or synthesis, JDO) and rather 'investigate what we are able and willing to do with a concept such as it is. Slogan: 'The later Wittgenstein? Kant, properly naturalised.' See: stokhof.org, ramblings, 10.4.2009.

Mathematics involves and establishes concepts and conceptualisations, which, as early Wittgenstein already stated, are captured in the logic or grammar of language. I already recalled the duck-rabbit nature of:

$$\bullet \underbrace{|\cdots|}_{n},$$

14. See (Stokhof, 2002, ch. 3) for an overview. His overview does not include resolute readings yet. Since such readings deny *Tractatus* much philosophical substance, they seem less relevant here.

which can be seen as a sequence of different elements or as a single group. And what is wrong with the following proof of $2 + 3 = 3 + 2$:

$$((\bullet||)|||) = ((\bullet|||)||)?$$

Is it its specificity, its lack of generality? Then what about:

$$((\bullet\underbrace{|\cdots|}_{n})\underbrace{|\cdots|}_{m}) = ((\bullet\underbrace{|\cdots|}_{m})\underbrace{|\cdots|}_{n})?$$

Given the way in which we learn to count, why is this less apodeictic than the more familiar derivations in PRA that require similar pattern recognition, but now based on how we are trained in formal systems?

As Wittgenstein held later: proofs forge connections between concepts and should somehow be surveyable. No wonder that such a rich book on Wittgenstein's philosophy of mathematics as Mühlhölzer (2010) often discusses *Anschauung*, *Aspekt*, *Begriff*, *übersehbar*, *Übersichtlichkeit*s. Here the use of concepts and of perception intermingle in ways that are roughly analogous to Kant, but naturalised. I suggest to revisit this research and to discuss how it can be applied systematically to the case at hand.

This is it for today. Did you notice the date? Six months from now it is Bloomsday, its centenary even! I will stay at home. Dublin will be over-crowed. But I do look forward to reading *Ulysses* once more. I belong to those we think its complexity is not unlike everyday life, but more poetic, and find its narrative human, at times humorous, and moving. It's as human as math. Yes!

Chapter 8

Pragmatics as a foundation for psychology

Bart Geurts

Psychological experiments are often presented as ways of "taking measurements" on subjects. This usage conjures up images of microscopes and Bunsen burners operated by solemn men and women in lab coats, and although such images may help to give psychology an air of respectability, I can't help feeling that they are incongruous and somewhat off-putting. More to the point, treating people as objects of measurement betrays denial of a rather fundamental fact about psychological experiments, namely, that they are social interactions between people. In particular, they inevitably involve communicative exchanges between experimenters and subjects; exchanges which are *designed* by the former to affect the behaviour of the latter.

I've always had an interest in psychology, and my first forays into psychological territory occurred two decades ago. At the time, I was working on the semantics and pragmatics of quantified statements whose meanings were so elusive that, without quantitative data, I felt no theoretical headway could be made. So I turned to the psychology of reasoning, and was in for a surprise. For what I found there was a bit like a pre-colonial tribe of hunter-gatherers, isolated from the world at large, minding their own business, oblivious to developments in other societies. In large part, that business centred around two topics: syllogistic reasoning and the Wason task. To a newcomer like myself, this seemed rather parochial, and with respect to the first topic I was pretty sure that it *was* parochial. My suspicions about the second topic were confirmed when I read an early version of a paper by Keith Stenning and Michiel van Lambalgen, which was duly published in 2001. To be sure, that paper wasn't intended to criticise the Wason task. But that's how I read it, and to my mind the criticism was entirely convincing and devastating.

Wason's (Wason (1968)) task is about selecting evidence relevant to the truth or falsity of a conditional rule. Subjects are presented with four cards, each of which has a letter on one side and a number on the other; for example:

A B 2 3

Then a rule is presented; for example:

> If there is a vowel on one side of the card, there is an even number on the other.

The subjects' job is to select all and only those cards which must be turned to decide whether the rule is correct for these four cards.

Wason's design has been replicated with countlessly many minor and major variations, and it is agreed that the task has one and only one correct solution: subjects should select the A and the 3. This, then, is what psychologists like to call the "normative response". However, as it turns out, this normative response is extremely rare: many, though by no means all, subjects choose the A and the 2, some choose the B, but very few choose the 3.

What's wrong with these people? This question has elicited a variety of answers. Predictably, there are those who rush to conclude that these findings prove once again that humans are irrational, that formal logic is wrong, or both. Equally predictably, there is a Darwinian school of thought, according to which our species just lacks a cognitive module for solving the Wason task, because none was needed for getting on in the struggle of life. But fortunately, there will always be serious and sober-minded scholars who are prepared to dig deeper and think harder, and Stenning and van Lambalgen are definitely in that class. Based partly on previous research but mainly on their own (which, to their great credit, included a series of tutorial interviews), they document in lavish and sometimes painful detail that it is the Wason task itself that is to be blamed. To mention only some of the issues they raise: subjects disagree with experimenters and between each other about the meanings of key expressions, including "true", "false", and even "the other (side of the card)"; they are sometimes uncertain about the status of the target rule, which they may take as given (i.e. true) even when explicitly instructed that its truth value is to be determined; and they may realise that a given card could falsify the rule, and still decide not to select it. And even if some of this variation is mitigated by tweaking the experimental design, thus far all versions of the Wason task have produced heterogeneous response patterns, and there is good reason to believe that this is an intrinsic defect of the experimental paradigm.

> Quite apart from possible social psychological effects of discomfort, the communication situation in this task is bizarre. The subject is first given one rule to the effect that the cards have letters on one side and numbers on the other. This rule they are supposed to take

> on trust. Then they are given another rule by the same information source and they are supposed *not* to trust it but seek evidence for its falsity. ((Stenning & van Lambalgen, 2004, 507))

Moreover, the first rule is supposed to hold only of the four cards on display, and the experimenter knows whether or not the second rule is correct, all of which makes for a highly artificial design. Put otherwise, the ecological validity of the Wason task is quite poor, and in all likelihood, the quirky behaviours it elicits are the outcome of desperate attempts at sense making more than anything else.

The Wason task was a bad idea all along, simply because it lacks pragmatic sense. The road to understanding is strewn with bad ideas. Many of them were exposed almost as soon as they were conceived. Many enjoyed a measure of success for a while, and some were generally respected until they, too, fell by the wayside. I suspect that the Wason task belongs to the last category, though I'm not sure whether it's dead already. But even if it isn't, Stenning and van Lambalgen deserve applause for hastening its inevitable demise.

Interestingly, Stenning and van Lambalgen disagree, claiming as they do that the Wason task is "one of cognitive science's fruit flies – a laboratory phenomenon of deceptive simplicity which is a potential basis for theory reaching far beyond its confines [and which] has the potential to unite disparate areas of cognitive science." ((Stenning & van Lambalgen, 2001, 273)) This is one of the few passages in their paper that I have my doubts about.

As I observed already, psychological experiments are social interactions, which necessarily involve communicative exchanges between experimenters and subjects. In general, these exchanges will be puzzling enough for unsuspecting subjects, and it is vital not to mystify them even more by pragmatically deviant procedures. Thus, pragmatics is foundational for psychology in a methodological sense. But there may be another sense, as well. There is an old tradition in philosophy and logic to view reasoning and dialogue as closely related topics. On this view, reasoning is first and foremost a form of communicative interaction and we learn to reason by speaking with each other. If this is on the right track, then pragmatics is foundational for psychology, or at least the psychology of reasoning, not only in a methodological, but also in an aetiological sense.

Chapter 9

Reasoning and the brain

Peter Hagoort

A Contribution to the Liber Amicorum for Michiel van Lambalgen

My contact and interactions with Michiel go back to 2003, when a young, brilliant student of Michiel at ILLC approached me for an internship in my research group at the Donders Centre for Cognitive Neuroimaging. His name is Giosuè Baggio, currently professor of psycholinguistics at the Norwegian University of Science and Technology, in Trondheim. At ILLC Giosuè had received a solid training in philosophy and logic, which he wanted to apply to the study of semantic phenomena in language processing and their neural basis. In the meantime Michiel had spent a sabbatical with Keith Stenning at the University of Edinburgh where he studied the psychology of reasoning. From my side, I had a keen interest to put the empirical investigations of language processing on firmer formal grounds. This created a middle ground for Michiel and myself, where we felt some fruitful collaboration could be established.

At about the same time NWO had launched a thematic program for foundational research in cognitive science. Johan van Benthem was the chair of this committee and I was a member. My interactions with Johan and Michiel led to a, by now, long term connection to the ILLC, which has been continued and bears fruit until this very day. When the NWO Cognition Program launched a call Michiel and I teamed up with two other colleagues and submitted a proposal on "Reasoning and the Brain". The central theme of the proposed project focused on defeasible inference. According to deductive logic, reasoning is non-defeasible. That is, if a conclusion follows deductively from a set of premisses P, it cannot become invalid if additional information is obtained. However, this does not characterise our reasoning and inference making in everyday life. Van Lambalgen and Hamm had published their important formal theory on tense and aspect (The proper treatment of events, 2005) in which they treated the lin-

guistic processing of temporal information as a defeasible process. This processing involved a mechanism of computing and re-computing so-called minimal models. Such (re)computations are, among other cases, assumed to take place when a reader/listener processes the default implications of the English progressive, as in the following examples:

(a) Michiel was building a barn, when a neighbour dropped by for a chat.
(b) Michiel was building a barn, when he was struck by lightning.

In the absence of information to the contrary, it is assumed that the barn will be finished, as in (a), but this inference is defeasible, as in (b). In our grant proposal we specified experimental procedures to study the neural substrates of defeasible inference with the help of recording electrical and magnetic brain activity (EEG, MEG). The idea was that if the formal model of Van Lambalgen and Hamm is correct, we should see differential traces in the brain signals if we compare cases as in (a) to those in (b). The second part of the project focused on possible deviation in defeasible inferences in people with an Autistic Spectrum Disorder (ASD). The project proposal convinced the referees and the NWO Program committee. On this basis we could offer our joint master student Giosuè Baggio a PhD position and we attracted another gifted student, Judith Pijnacker, for the ASD part of the project.

I don't have space here to summarise the outcomes of these projects, apart from saying that the results were in line with our predictions from the model. Our NWO-funded project resulted in two PhD theses, *Semantics and the electrophysiology of meaning: Tense, aspect, event structure* (Baggio, 2009), and *Defeasible inference in autism: a behavioural and electrophysiological approach* (Pijnacker, 2010). In addition, it resulted in 7 co-authored publications with Michiel in international journals and as handbook chapters. Certainly, both in quality and quantity a sizeable outcome that we should be proud of.

When writing this piece for Michiel, I just realised that there was a common characteristic between Michiel and our two PhD students. Here is a quote from my laudatio at the doctoral ceremony of Giosuè Baggio: "You are not one of those who enter the room with a lot of decibels. Modesty and integrity are two of your most salient characteristics. You are a deep thinker, and always listen carefully to other people's arguments. You will not say easily 'I think that is nonsense', even if you might think it every now and then. Given your qualities, a little less modesty in your interactions with the rest of mankind would not be unjustified. Of course, the ambition to develop a neurobiology of meaning is far from modest." Mutatis mutandis the same applied to Judith and Michiel. Michiel clearly had his views and ideas, but never presented them in an imposing way. That was also clear in our supervision meetings with the students, where he listened carefully to what the students had to say and gave his advise in a gentle manner. Our meetings with the students took place mostly in Nijmegen, but also sometimes in Amsterdam.

Michiel didn't have his office in the buzzing ILLC headquarters, but in a quiet room close to the Oude Manhuispoort, secluded from the student traffic and the networking scientists. Sometimes he had to cancel our meeting for external reasons, such as shown in the following email exchange from June 2005. Michiel: "Dear all, to my great regret I will not be able to make it tomorrow due to the strike, and, no, I don't have a driver's license ....". My reply: "We will prepare a court case against ProRail for their obstruction to the progress of science, and thus of mankind."

In the same year I received an email from Michiel that characterizes him well (it is in Dutch): "Beste Peter, Martin Stokhof heeft me gevraagd een hoofdstuk te leveren voor een nieuw handboek, nu voor Philosophy of Linguistics. Mijn stuk dient te gaan over taal en cognitie. Ik heb onmiddellijk ja gezegd. Mijn vraag is: wil jij op enigerlei wijze meeschrijven? Het lijkt me dat dit een gestructureerd wijze kan geven om over onze modellen en ideeën van gedachten te wisselen. Ik realiseer me ten volle dat je het vreselijk druk hebt, en dat ik mogelijk het leeuwendeel van het werk zal moeten doen. Om werkverlichting gaat het me ook niet – het lijkt me gewoon erg interessant". As is clear from this email, Michiel is interested in the intellectual challenge, the exchange of ideas. He wasn't interested in taking center stage or in offloading his part on other shoulders. One could say that Michiel is a true intellectual in the classic sense of the word. All the paraphernalia of modern life in academia might not have been seductive for him.

My interactions with Michiel have been very fruitful. In later years they became less frequent, mainly due to health-related issues. The interactions also led to connections with Michiel's own network. As a result, on and off I had the possibility to discuss ideas with Keith Stenning, Fritz Hamm, Oliver Bott. I am grateful for our collaboration. It not only resulted in a very productive project on Reasoning and the Brain, but also, and even more importantly, it has enriched my own thinking, especially on issues to do with the compositionality of meaning.

Chapter 10

Configurations of pluralisms

Machiel Keestra

Navigating polyphony and diversity, in philosophy and beyond

A short prelude: from tragedy to polyphony with Michiel

Supervising a dissertation is in some sense comparable to directing an opera: the supervisor hasn't written the libretto nor the score, is not performing one of the main character roles nor an instrument that accompanies the singers. Nonetheless, the supervisor's role is important in that they are guiding the overall interpretive and rehearsal process such that the end result is a meaningful and consistent whole which does in some sense reflect or breathe their thoughts. When I enjoyed Michiel van Lambalgen and Martin Stokhof's careful and valuable supervision of my *Sculpting the space of actions. Explaining human action by integrating intentions and mechanisms* (Keestra (2014)), I first made use of examples from ancient tragedy to support my defence of a framework for explaining a plurality of processes constituting how humans determine their actions. Tragic actions were used to illustrate how action decisions can occur both in an automatised, habituated mode and in a rationally controlled way, with there being important interactions between the two modes. Being aware that both of us enjoy singing individually as well as choral singing, Michiel suggested that opera singing might better than tragic action illustrate and reflect the different modes. Happily embracing that suggestion, I illustrated my framework with the observation that most singers will have that our initial rationally controlled yet less flexible performance might become habitual and automatised over time and consequently also become more nuanced, flexible and complex. This process does not prevent this automatised performance from complying with musical and interpretive standards and being in harmony with the rest of the score, on the contrary.

This grateful memory has inspired the brief exploration of pluralism below, starting from ancient tragedy and inspired by the musical notions of polyphony and counterpoint leading to reflection on how a pluralism can coexist. As any form of pluralism is specifically opposed to monism, I'm especially interested in considering different configurations of the plurality of options presented by pluralism. This also offered a lens on interesting work from Michiel in which pluralism of sorts figures as well.

Pluralism and diversity in tragedy, disturbing ancient monism

One of the most intriguing moments in western cultural history is when, according to Aristotle's account, the imitations and representations of humans as performed in rituals and arts became more complex and dynamic, especially with the emergence of tragedy. In his innovative and influential description of the 'evolution of tragedy' - as one commentator puts it - Aristotle pays attention to a history spanning several centuries and characterised by the development of novel genres, each of which has its pertinent object, its formal properties and certain means of performance (Aristotle et al. (1968)). If we trace back this history while focusing on the last feature, the means of performance, this history is relatively simple, displaying a few decisive moments. It starts in prehistoric times with dithyrambian and hymnal songs performed mainly during religious rituals, according to Aristotle.

Even though these songs did include exchanges between calls and answers, these were only a shallow precursor to what over time would grow into genuine exchanges between protagonists representing different, sometimes even opposing, positions. The first time an individual person did appear on the scene and performed a role in distinction from the communal singing chorus, was probably when its leader started singing calls which were answered by the group of singers. As such the group and its leader would in their mimetic acts not represent very different, let alone contrasting, roles. This was about to change drastically when actual dialogues were added to the singing parts. "Aeschylus was the first to increase the number of actors from one to two; he also reduced the role of the chorus and made the dialogue the major element in the play. Sophocles increased the number of actors to three", Aristotle writes, crediting the two most celebrated tragedians with the crucial innovation of introducing dialogue into the venerable art of tragedy (Poetics, 1449a 16-19, ((Aristotle et al., 1968, 9))).

With the chorus now being constrained to a more secondary role, often merely commenting on, or responding to, the events on stage, the focus had shifted to two or more interacting protagonists who often no longer belonged to

the same group. On the contrary, instead of the representation of a single voice or perspective, tragedy and comedy would now confront its audiences with clearly distinct individuals, at times differing in norms, histories, social positions, genders, characters[1] and especially in the actions that demonstrated these differences. Even those plays in which members from the same family appear, like Aeschylus' trilogy Oresteia, the Elektra's by Sophocles and Euripides, the Iphigeneia's by Euripides and Sophocles' Orestes, are driven by differences in characters, oppositions in perspectives and oftentimes actions through which family members Agamemnon, Clytemnestra, Iphigeneia, Elektra, Orestes, and others murder and revenge each other or instead try to resolve their conflicts (cf., Keestra (1999)). In so doing, these tragic individuals have to navigate between sometimes contradictory appeals, relying upon their autonomy and their freedom, and can no longer naïvely rely on the gods as their Homeric predecessors did (Snell (1975)).

This short sketch emphasises the emergence of pluralism and even dissonance of voices in ancient tragedy. Below I will briefly pause on pluralism and its implications, yet it is important to first note the importance of recognising this pluralism. Ancient Greek culture and philosophy, from Homer via the Ionian natural philosophers to Plato, is often characterised as having a tendency towards unity and monism, describing, and explaining reality in as few elements or principles as possible. Whether it is a single principle like Thales' water in its different phases, Anaximander's 'apeiron' or indefinite, the Parmenidean and spherical 'One', or even the platonic idea of the Good: dynamics of change and development are hard to explain with such simplicity. Or to use Aristotle's musical metaphor: "when we say [that] the non-musical man becomes a musical man, both what becomes and what it becomes are complex" (Physics, 190 a 3-4, (Aristotle (1984))) for the explanation of which a monism of principles is insufficient. This preference for unity and monism holds even for Greek polytheism, even if that term suggests otherwise. Notwithstanding the presence of multiple gods, the Greek pantheon is unified as the 'Olympian gods appear as a family community' with a 'compactness and clarity of organisation' ((Burkert, 1985, 218)).

Although alternative views did exist before, a decisive rift in this monist tradition occurred with the popularity of the sophists, among others, who would become prominent in the public eye around the time when tragedies demonstrated the important roles that diversity and pluralism play in human affairs.[2] Even more so, the tragedies did present diversity and pluralism while demonstrating that unity and monism are no options for the tragic protagonists, struggling as they are "over

1. The Greek words for 'character' and 'habit' are probably related to each other, as Aristotle points out in Nicomachean ethics II,1.

2. Related to this pluralism is how tragedy portrays the human experience of being torn apart, which is connected to the Dionysian ritual of 'sparagmos': sacrificing an animal by tearing it apart (Storm (1998)).

the meaning of a single value, within a character (dilemmas), between characters (disagreements), between the cultural schemes different characters may represent" ((Apfel, 2011, 11)). Indeed, Aristotle himself affirms pluralisms of sorts, recognising variety in our knowledge with not all bodies of knowledge allowing the same certainty and validity as mathematics, for example. Moreover, mathematics itself, building upon axiomatics, allows for varieties because "if the principle should change, practically all that is proved from it would alter" (Eudemian Ethics, 1222 b 25). In the field of politics, something similar holds according to Aristotle, recognising that it is not necessarily a negative thing there being different political constitutions as politics does not allow the unity or monism that in other domains might be possible, in much the same way as citizens are different (Johnstone & Marienthal-Maschler (1962)). In addition to a pluralism of knowledge Aristotle has embraced a value pluralism, implying that "the goods that a human life appropriately values are plural and incommensurable" each of which might deserve our commitment without there being an overarching measure allowing us to compare and rank them ((Nussbaum, 1999, 182)).

What this brief history shows is that though our tradition may show on average a preference for unity and monism, pluralism has been present from the beginning as well, like in the pluralism of gods and of tragedy's voices, associated with a pluralism of forms of knowledge and of values. I will now briefly attend to this pluralism of pluralisms, which will then be followed by an exposition of a contemporary position that surprisingly and convincingly embraces even a pluralism of logics. This raises the question whether denying monism in the domain of logic might raise the spectre of embracing inconsistency and contradiction when pluralism is accepted. Showing that this is not necessarily the case, I will finally reflect upon the important topic of how different configurations of these voices or positions are enabled by their pluralism.

Pluralism of pluralisms: recognising the value of multiple voices

Defending pluralism, Aristotle criticised the platonic position as interpreted by him, which allegedly holds that irrespective of differences, a thorough reflection on the limitations of human knowledge should convince all rational beings to ascribe to a monistic position. According to this monism, apparent differences in the nature and validity of knowledge and ethics are simply due to flaws in reasoning. Yet Aristotle is also critical of the opposing, skeptical position which maintains that no reliable knowledge or ethical reasoning is at all possible. Aristotle rejects this skeptical position as he maintains that it fails in recognising the 'variety and fluctuations' that are prevalent in multiple domains with which

humans occupy themselves including ethics and science (Johnstone & Marienthal-Maschler (1962)). Indeed, acknowledging variability of human experiences and reflections has led Aristotle to accept pluralism in a similar fashion as Rescher does more than two millennia later: "The experiential diversity of differently situated rational inquirers must mean that they are destined to reach variant conclusions about the nature of things. In a human community of more than trivial size, dissensus rather than consensus is the normal condition" ((Rescher, 1993, 77)). Again, this pluralism is recognised without assuming that they eventually can be replaced by consensus or monism.

Such recognition of its irreducibility is key to pluralism. Take the first sentence of the Routledge Encyclopedia of Philosophy's lemma on 'pluralism', which defines it as follows: " 'Pluralism' is a broad term, applicable to any doctrine which maintains that there are ultimately many things, or many kinds of thing; in both these senses it is opposed to 'monism'." ((Craig, 1998, 463)). There being 'ultimately' many things or kinds of things is fundamental here, as it implies the irreducible plurality of these things - whether forms of knowledge or values or religions or meanings are intended, for example. Embracing pluralism about these things implies accepting that there are multiple correct or valid accounts about these possible which can coexist in some configuration.

It is important to distinguish pluralism from relativism as the two are often confused with each other. With pluralism recognising the validity of multiple accounts of certain things, relativism posits that the validity of any account is relative to some external factor. According to relativism, a form of knowledge or moral value is only valid relative to a certain culture or historical period, for example. Instead of accepting the correctness of multiple accounts in parallel, a relativist points out that each account is only valid in a limited sense. Most people will embrace a certain relativism regarding etiquette, for example, and accept that some behaviours are acceptable in certain situations while not in other contexts. Regarding knowledge such relativism would imply a rejection of the knowledge claim, which is also unnecessary as most phenomena allow epistemic pluralism without different accounts excluding each other as is more common when it comes to our social behaviours (cf., Cook (2010)).

Although monism and pluralism are contrasting positions, they are not unusually combined with each other. What appears to be a pluralist position can sometimes turn out to be a monist position at another level. Such a position entails that underlying the plurality is a hidden systematics, like a hierarchy, that eventually allows the reduction of the acknowledged multiplicity to a single, more foundational unit. An example may clarify this. Tragic conflict depends upon the differences between irreconcilable values that two or more protagonists uphold. In the Antigone, for example, we can observe the conflict between Antigone's familial piety towards her fallen brother Polynices on the one hand and the loy-

alty to the state which makes King Creon forbid the burial of this rebel on the other. Sophocles' tragedy provides the spectators with the embodied experience and reflection that enable them to understand and empathise with both positions, making the values inherent in them appear equally valid and defensible. Indeed, the bleak ending of the tragedy makes it doubtful whether Sophocles himself believed in such conflict's resolution or reconciliation. In contrast, Hegel's interpretation amounts to a rejection of the subjective individuality of the protagonists, with their irreconcilable conflict's resolution being possible once they understand the necessity of another form of sociality - entailing a complex form of monism (Keestra (1999)).

Such reconciliation at another level is an example of *non-foundational pluralism*, with an apparent pluralism of (not fundamental) values being eventually related to another, more fundamental value. This is different from *foundational pluralism,* which does accept there being multiple sets of moral values available for shaping one's life, for example (Mason (2008)). Moreover, such foundational pluralism can be associated with another source of moral variability since it is possible that each value is 'multiply realisable'. So in addition to there being multiple fundamental values - like happiness and equality - a foundational pluralist can also accept that each of these are 'subjectively realised' in different ways by individuals. Happiness may be an important value for both religious persons and for secular political ideologists, for example, but the way they're realising it will be determined partly by their distinct beliefs and reasoning (cf., (Audi, 2007, 27)).

Value or ethical pluralism and the tragic conflicts emanating from it are not only prevalent in the arts but generally close to human experience. Due to our finitude and to the contingency of our position and possibilities, for example, we are unable to realise all possible values in a single lifetime (Ivanhoe (2009)). This limitation can't be resolved by some form of monism, as we've just seen. Compared to this experience of irreconcilable value pluralism, epistemic pluralism or the pluralism of knowledge is strikingly different. Although most people will accept multiple accounts of a factual situation, they will still maintain that reality itself is singular. So how might one subscribe to explanatory pluralism - to focus on a specific form of epistemic pluralism - without succumbing to some form of metaphysical pluralism? Since an explanation entails an answer to a specific question, the fact that we can ask multiple questions about a single phenomenon corresponds with there being several explanations available (Ruben (1992)). Human action as well as understanding human action, for example, allow for explanatory pluralism as neuroscientific, psychological, sociological and hermeneutic explanations each offer valid perspectives which do not exclude each other (Keestra (2014, 2015)).

Explanatory pluralism can imply that we develop theories and laws that help us to understand the same phenomenon at different levels of description which

can not only co-exist but even co-evolve: language processing can be explained both by cognitive psychological and neurobiological theories, each addressing different yet related features of the process (Looren de Jong (2002)). These levels of description refer to the fact that a single phenomenon - like climate change or consciousness - can be described and explained in various ways. Considering it as a complex system composed of components, subcomponents, and their interactions, we can offer multiple non-overlapping 'decompositions' of it. A fruit fly, for example, can be decomposed or described in terms of its physiological systems, with nervous, muscular, and other components organised in a specific way in its body. Entirely different is the representation of water by its physicochemical dispositions, as water is the main ingredient spread all over its body making this representation rather uniform (Wimsatt (2007)).

Here again, pluralism must not be assumed to be reducible to either one or to a more fundamental type or level of explanation. Interdisciplinary integration of different explanations being still a goal, this does not entail some form of monism or reductionism (McCauley & Bechtel (2001)). Instead, explanatory pluralism accepts the current plurality of theories, methods and data as a consequence of the complex dynamic realities of the living and social worlds - and to some extent the (quantum-)physical world, too. This complexity allows for the pluralism of decompositions that was noted above, but another feature of it is the presence of multiple causal relations within a single system which are different in nature. Part of that complex reality is its multi-causality, with multiple causal factors interacting dynamically with each other, which is another reason why some scholars defend pluralism while contending that it is improbable to develop a single comprehensive account of reality (Kellert et al. (2006)).

Much more might be said about the two forms of pluralism treated here. In addition, there are many more forms of philosophical pluralism available, including metaphysical and ontological pluralism, semantic pluralism, aesthetic pluralism, scientific pluralism: indeed, there is a 'plurality of pluralisms' (Wylie (2015)). Whatever object domain is at stake, it allows the development of multiple systematic and consistent perspectives that can figure next to each other and be involved in rational argumentation or determine our actions. Now one may ask whether similar considerations apply to the domain of logic and reasoning: is logical pluralism possible, or should we expect this to be a non-foundational pluralism? Taking up this question, I will consider the logical pluralism that is involved in the insightful work by van Lambalgen and Stenning on human reasoning. Does the logical pluralism they present, eventually give way to logical monism? More generally, I will ask whether there is an alternative to the monistic option for resolving or even dissolving this plurality of pluralisms, by considering different configurations in which a plurality of perspectives might be related to each other. Such configurations will be treated after the next section.

Change of scenery: human reasoning as a pluralist affair

Unity and monism were not just guiding principles in western antiquity but remained in some sense in place until more recent times. With monotheism penetrating most domains of western (and Mediterranean) societies and scientific reductionism motivating the work of most scholars for many centuries, pluralism remained at most a marginal if not suppressed position. This also held for philosophers, probably mostly because of the anxiety that lifting the grip of monism might immediately bring the risk of a position according to which 'Anything goes'. Introducing a volume on philosophy and pluralism, the editor captures this sentiment succinctly: "Those unsympathetic to monism are also anxious to distinguish pluralism from relativism. That no single correct answer can be agreed upon does not mean that each and every answer is true" ((Archard, 1996, 2)). Applying our choral metaphor again we can shed stark light on this sentiment: whenever we want to go beyond simple monophony, we're not immediately handing over to a cacophony without any structure. Instead, the alternative to monophony can be a polyphony with different harmonic structures, some of which might not easily give away the interdependence or relations between the different voices. Further down I will more closely discuss pluralism and the possibilities it offers, yet before doing so I will briefly point out how a similar struggle between monism and pluralism is observable in the field of logic and the interdisciplinary study of human reasoning. For this I will take inspiration from Michiel van Lambalgen's work, especially his co-authored provocative monograph *Human Reasoning and Cognitive Science* (Stenning & van Lambalgen (2008)).

Bringing together insights from logic, semantics, cognitive psychology, and neuroscience, with the addition of original empirical research in human interpretation and reasoning, logician van Lambalgen and cognitive scientist Keith Stenning offer a fresh and convincing argument about human reasoning and the logics employed in it. Logics, in the plural indeed, since they take issue with the generally held, traditional position that there is only a single logic which underlies or governs all valid human reasoning. Indeed, they diagnose how in classical logic and even still for Frege "the normativity of logic seems to be bound up with the uniqueness of logic" ((Stenning & van Lambalgen, 2008, 11)). In contrast to this position, their argument amounts to a very different one which has implications both for the relation between the logic of reasoning and the interpretation of linguistic expressions with the norms involved: "our answer will be that norms apply to instances of reasoning only after the interpretation of the (logical and nonlogical) expressions in the argument has been fixed, and, furthermore, that there are in general multiple natural options for such interpretations, even for interpreting the logical expressions" (*ibid.*).

The authors elsewhere explore the fact that human cognition is extraordinary in that it allows humans to navigate between both domain specific reasoning as well as reason with surprising domain generality. The varieties in reasoning that can be observed in humans goes along with a 'multiplicity of logics', they argue, a multiplicity that is related to the multiple semantics required for distinct domains of reasoning ((Stenning & van Lambalgen, 2005, 2)). The semantics pertinent for a particular domain are open for discussion, as when the concepts and meanings we use to reason about train schedules can vary, as can those we employ in the social domain. Yet whenever we decide about the contents of a domain and its pertinent conceptual relations, a corresponding set of logical rules is determined with different sets of rules not necessarily being reducible to each other. Generally the process consists of two distinct steps that together support a form of logical pluralism: "We therefore view reasoning as consisting of two stages: first one has to establish the domain about which one reasons and its formal properties (what we will call "reasoning to an interpretation") and only after this initial step has been taken can one's reasoning be guided by formal laws (what we will call "reasoning from an interpretation") ((Stenning & van Lambalgen, 2008, 20)).

In addition to the logical pluralism implied by there being multiple domains of reasoning, there is one more reason why humans are so used to this pluralism, albeit more implicitly. In a way, then, humans are reasoning on a daily basis with different logics even within a single domain - a fact that has only more recently been recognised as such. Moreover, these different logics applied to a single domain play out even within a single brain - compare a singer who engages not only with polyphony in an ensemble but also individually by using the overtones she can produce simultaneously. For this, van Lambalgen and Stenning build upon so-called dual system or dual process theories, which occupy an important role in the psychological explanation of reasoning in a wide sense. Psychologist and Nobel prize winner Kahneman famously distinguishes these two systems as either processing information fast, intuitive, and emotional - System 1 - or as processing it rather slow, more deliberative, and more rational - System 2 (Kahneman (2011)).

Typically, automatic system 1 processing is considered not to perform logical reasoning, which van Lambalgen and Stenning reject as being grounded in a flawed understanding of logic. They argue that a different, non-classical logic is being employed, which has largely gone unnoticed in psychological reasoning studies: "We conceptualise the part that logic plays in system 1 as being the foundation of routine discourse interpretation, when a suitable knowledge base already exists in long-term memory" ((Stenning & van Lambalgen, 2008, 124)). Thus logical pluralism depends upon there being distinct domains of reasoning in parallel with there being different kinds of reasoning processes engaged with a particular domain.

This observation of a form of logical pluralism raises the question how the different processes and their logics are related to each other. Instead of considering the two processing systems as being independent with regard to their respective functions and regarding their developmental history in humans, van Lambalgen and Stenning are interested in their interdependence. While we share system 1 with large sections of the animal kingdom, system 2 is probably less common even though it is not per se dependent upon literacy and schooling. The two systems indeed operate according to different underlying logics, with automatic processes performing 'defeasible closed world reasoning, and deliberative processes performing either classical or closed world reasoning. Importantly, the interactions between the systems emerge at an early stage of human development with the rather deliberative system 2 processes starting "as repair processes when a system 1 process meets an impasse and gradually shade into full blown adversarial discourses, perhaps with their underlying logic being classical" ((Stenning & van Lambalgen, 2005, 130)).

Logical pluralism is here defended as a consequence of there being different systems of reasoning, applied to distinct domains. However, this does not rule out the option that this pluralism is associated with logical monism at another level.[3] The authors indeed argue that a multiplicity of logics is possible because multiple choices are possible for setting the parameters that determine the semantics and syntax of a particular language. This apparent pluralism, though, does not imply that inferential or consequence relations are equally flexible. Indeed, their logical pluralism is not a foundational pluralism, as we can learn from their comparison with multiple concrete grammars related to a single underlying universal grammar: "we do not claim that a logic can be seen as a point in a well-behaved many-dimensional space. The use of the term parameter here is analogous to that in generative linguistics, where universal grammar is thought to give rise to concrete grammars by fixing parameters such as word order" ((Stenning & van Lambalgen, 2008, 25)).

Nonetheless, with dual processes employing different logics, their explanations do reveal the production of different outcomes in response to the same task by one and the same person. Apparently that person processes identical information differently, depending upon the activation of one of two different processes, each with its own properties and - as we noted above - logic. Dual process or dual systems theories are applied widely, from the social domain (Chaiken & Trope (1999)) via moral deliberation (Craigie (2011)) and reasoning (Frankish & Evans (2009)) to the explanation and treatment of addiction (Wiers et al., 2007)

3. I'm grateful for Martin Stokhof's comments on an earlier version of this text, which included some pressing questions regarding the logical pluralism I am ascribing here to Michiel. This useful exchange echoed the many inspiring conversations I've enjoyed with Michiel and Martin as my co-supervisors. Remaining misunderstandings in this text are, again, due to me.

and the determination of human action (Keestra (2014)). Going back to a platonic metaphor, these two systems have been referred to also as the rider and its horse, with the horse representing system 1 and the rider with their limited capacities working to control and constrain the animal. Indeed, van Lambalgen and Stenning suggest that system 2 evolved more recently, with the interactions between the two systems contributing to specific human capacities in planning, false belief tasks and others (Stenning & van Lambalgen (2005)). Apparently, it is with the presence of two voices in one mind that these can be performed which then raises the question about their interactions or configurations: given this pluralism of voices, we might ask how they can be related to each other? What counterpoint or other configurations are possible? It is this question that we will focus upon in this chapter.

Pluralism and counterpoint: from isolation to interactions

It required an important step to enable a dialogue and even conflict on the Greek stage, I argued above. The tension between monism and pluralism has partly shaped our tradition and thought. A more recent development is that of polyphony in music which has been accompanied by the unfolding of thoughts on counterpoint: "the combination of simultaneously sounding musical lines according to a system of rules" ((Sachs & Dahlhaus, 2001, 1)). If multiple voices are added to a musical score, how should these be configured such that the total effect is more than a mere addition of notes, is musically meaningful and is aesthetically pleasing as well? Over the centuries, different systems have been laid out, offering examples and rules for the creation of scores that benefit from the availability of more than just a single, melodic, voice. Building upon harmonic ideas - about consonance and dissonance, for example - counterpoint entailed writing a score such that the musical meanings of different voices are dependent upon each other while providing means for musical progression - as when an alternation between consonant and dissonant chords resolves eventually in harmony or when a fugue offers variation and repetition simultaneously. The rules underlying counterpoint have been constantly in flux, in many ways gradually offering more freedom and possibilities to composers, with contemporary atonal composition techniques often still involving counterpoint. Interestingly, counterpoint and the configurations between 'Leitmotifs' allowed composers like Wagner and Strauss even to express literary ideas, representing dramatis personae, themes and their relations in musical form (Sachs & Dahlhaus (2001)). In this section I will explore some configurations pertaining to the discussion of pluralism above.

What configurations can we observe in the plurality of pluralisms? How are voices, positions or perspectives related to each other such that they suggest a progression or development in which these merge into a single one, or remain a

pluralism? Above, we already mentioned the monism to which pluralism is often opposed and observed that there are several ways in which apparent pluralism might resolve in monism. Such resolution of a pluralism of voices or perspectives might occur in at least two ways: one of the voices will emerge as the dominant voice, into which others are dissolved. Alternatively, the resolution of a - perhaps dissonant - chord of voices leads to a third, hitherto absent voice. In explanatory pluralism such a resolution typically involves the reduction of different levels of explanation to a more fundamental level that refers to fundamental particles, neurophysiology and the like. Such a resolution after a phase of pluralism is called *moderate* or *temporary pluralism* in Van Bouwel's account of explanatory pluralisms in psychiatry (van Bouwel (2014)).

Genuine pluralism, however, would not permit such reduction to a monist position. In contrast to monism, *antagonistic pluralism* maintains that we're sometimes forced to choose between alternative concepts or explanations as they mutually exclude each other (Currie & Killin (2016)). Such antagonism plays out differently, depending upon the domain at stake. Indeed, this might imply *incompatible pluralism*, which especially applies to normative or moral positions. Tragedy offers us many examples of this, as when Agamemnon cannot both implement the values of a war hero and those of a father, since the former requires the sacrifice of his daughter Iphigeneia (Apfel (2011)). It might be argued that this conflict is not just a matter of the incompatibility of both values, but that it is also impossible to compare or order them as they apply non-overlapping measures, which makes the conflict an example of *incommensurable pluralism* (Mason (2008)). Conflicting values force a person to make a choice, even if it is impossible to compare these. Such a choice is not always necessary in the context of scientific pluralism, as this allows for the presence of incompatible and incommensurable alternatives, even for *'Anything goes' pluralism* which amounts to "retaining all, possibly inconsistent, theories that emerge from a community of investigators." ((Mitchell, 2003, 186)). Monism and 'Anything goes' pluralism can be considered two extremes on a continuum of forms of pluralism, which at both extremes implies the absence of a specific configuration and relation or interaction between options involved: for monism implies singularity and 'Anything goes pluralism' entails indefinite or absent relations between available options (Mitchell (2003), van Bouwel (2014)).

Between monism and 'Anything goes pluralism' we can distinguish several forms of *complementary pluralism*, involving some relation between the perspectives at stake. In the case of complementary pluralism regarding music, for example, multiple concepts of music can coexist and even complement each other. Whether taken as a form of communication or an art form, each perspective presents an equally valid perspective on music by highlighting different aspects of music or its function across times (Currie & Killin (2016)). The challenge facing

us now, is whether the perspectives are not only complementary to each other but can be related to each other in a more productive sense.

Focusing on scientific pluralism, Mitchell defends *integrative pluralism* as biologists typically offer an integrated explanation of a multi-causal and contingent phenomenon, while employing theories and models that remain relatively independent although being compatible with each other (Mitchell (2003)). Such pluralism is also at stake in most forms of interdisciplinary explanations, integrating theories, methods and/or results from multiple disciplines each of which alone can explain partly a phenomenon's variability whereas integrated a more comprehensive explanation is possible (Keestra et al. (In press)). Van Lambalgen offers original examples of such interdisciplinary explanations, for example integrating logical analysis with cognitive psychological and neuroscientific investigations of reasoning and interpretation in normal and autistic subjects: logical and explanatory pluralism being involved in integrative pluralist results (Baggio et al. (2008), Pijnacker et al. (2009)).

However, integration or synthesis should not be expected to be the end result of all such interdisciplinary endeavours. Insisting on the possibility that not all partial explanations might be integrated with each other, van Bouwel is not satisfied with this integrative pluralism as scientific telos. Instead, he adds *interactive pluralism* to the continuum or list of options. Leaving open the ir/reconcilability of pluralism, it also allows for the interaction with heterodox perspectives (van Bouwel (2014)). Allowing such non-mainstream perspectives to play a role in pluralism is relevant, given a history of science in which these have repeatedly contributed to scientific revolutions and progress.

Final chord: pluralism and diversity

Irrespective of whether pluralism is found in the domain of values of science or elsewhere, the encounter with different configurations shows how some forms of pluralism are likely to be productive, whereas others are less so. There is, I think, an interesting relation between the dual system pluralism in human reasoning according to Stenning and van Lambalgen's account, and the interactive pluralism presented by van Bouwel. With regard to human reasoning, the authors contend that the plurality of processes interacting with each other such that one process repairs the other process's flaws improve on what a single process might accomplish on its own. Similarly, van Bouwel presents a set of norms - borrowed from Longino (Longino (2002))- that structure a productive interaction or dialogue between perspectives. Although both arguments apply to quite different phenomena - cognitive processes versus scientific perspectives - they both in some sense defend interactive pluralism's contribution to our epistemic progress. This concurs with research on metacognition and reflection, which shows that if performed not

individually but with others, the explication and articulation of implicit assumptions and norms underlying our cognition is enhanced. It is the interaction and confrontation with a diversity of perspectives, norms, and positions that help us to recognise our own, which isolated self-reflection might not give away (Keestra (2017)). Van Lambalgen and Stenning as well argue that a diversity in reasoning styles is only to be expected given the contributions of genetic, environmental and experiential factors to human development. Interactional pluralism is implied in their appeal in the book's next to last sentence: "This understanding of why it "takes all types" (of people, to use a vernacular expression) might even contribute some much needed motivation for rubbing along with each other" ((Stenning & van Lambalgen, 2008, 366)). Compare again the pluralism of voices in choral singing: the beauty of a particular voice or melodic line is often enhanced by the polyphony and counterpoint in which it is bound.

Chapter 11

Emeritusbesluit: het eerlijke verhaal

Sylvia Pauw

Afgelopen voorjaar verkeerde het normaal zo vredige departement Wijsbegeerte aan de Universiteit van Amsterdam in crisis. Aanleiding was een memo van het College van Bestuur dat paal en perk moest stellen aan het onbezoldigd uitvoeren van onderwijstaken door gepensioneerden. De maatregel betekent dat verschillende geliefde docenten van Filosofie hun werk na dit collegejaar zullen moeten neerleggen. De gevolgen lieten zich voorspellen: Collega's boos, studenten nog bozer, en in een mum van tijd was een sympathieke briefschrijfactie gekaapt door een stel ongezonde rechtse clubs.

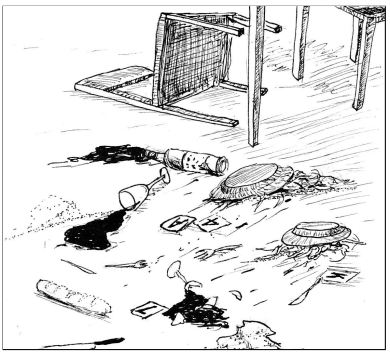
De verontwaardiging die het emeritusbesluit opriep is begrijpelijk. Is het Maagdenhuis in 2015 niet bezet omdat we genoeg hadden van dit soort ondemocratische top-down beslissingen? Is het niet aan de Afdeling om te bepalen wie het onderwijs verzorgt? En wat is er überhaupt mis mee, een paar gepensioneerde vakidioten die hun expertise blijven overbrengen?

En toch.

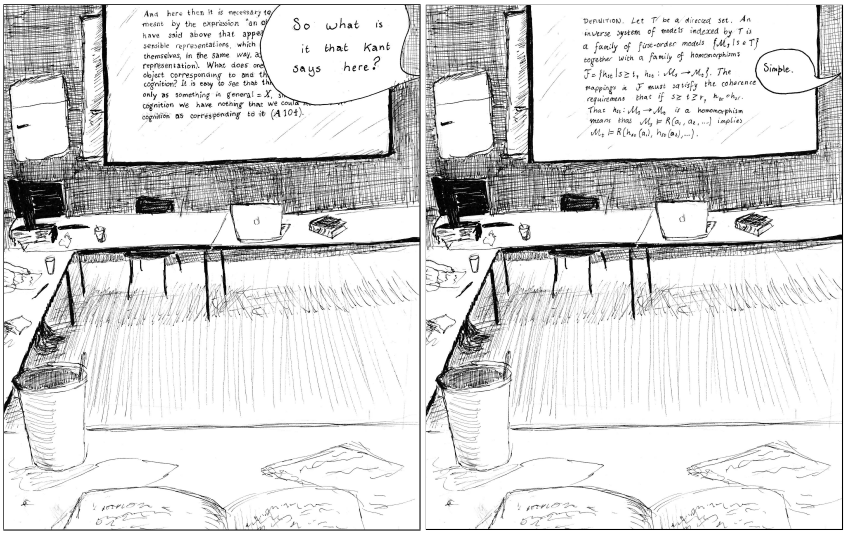
Laten we eerlijk zijn: Er zijn gevallen waarin de maatregel gunstig uit zou kunnen pakken. Om een concrete casus te nemen: Zou het niet verstandig zijn toekomstige generaties studenten te behoeden voor het onderwijs van Michiel van Lambalgen? Laat me dit toelichten.

Tijdens de Master of Logic heb ik twee van Michiels vakken gevolgd. Het eerste, *Rationality, Cognition and Reasoning*, was relatief onschuldig. Het vak gaat over Michiels onderzoeksproject met Keith Stenning, en laat studenten kennismaken met het gebruik van logische modellen voor empirisch onderzoek naar het menselijk redeneren. De cursus roept weliswaar de nodige filosofische vragen op ("had Frege dit niet verboden?"), maar in elk geval was een deel

van de inhoud redelijk te volgen. Niettemin heeft dit vak fikse psychologische schade berokkend bij de betrokkenen. Als deelnemers zelf al geen slachtoffer werden, dan wel hun naasten. Studenten werd gevraagd familie en vrienden te onderwerpen aan redeneertesten als de *Wason selection task*. Lang niet iedereen kan het oordeel dat ze niet klassiek-logisch zouden redeneren verkroppen, zelfs niet als ze wordt uitgelegd dat ze misschien een andersoortige logica volgen. Dit levert onvermijdelijk het nodige drama op. Mij heeft het in elk geval maanden gekost de relaties met mijn proefpersonen enigszìns te herstellen.



De ellende die *Rationality, Cognition and Reasoning* opriep was niets vergeleken bij die van het vak dat Michiel gaf samen met Dora Achourioti: *Kant, Logic and Cognition*. Na een semester redelijk wat tijd en energie aan de cursus te hebben besteed viel wat ik begrepen had als volgt samen te vatten: Eén: we lazen die delen van de *Kritik der reinen Vernunft* die ik tot dan toe met goede redenen vermeden had. Twee: die delen van de *Kritik* zouden met voor mij onbegrijpelijke wiskunde in verband te brengen zijn. Drie: dit alles had ook nog iets te maken met cognitieweten-schap. Kortom: een jaar later moest ik het vak opnieuw volgen.



De tweede keer lukte het om enig inzicht in Dora en Michiels formalisering van Kant te krijgen, maar ik was zo mogelijk nog wanhopiger dan het jaar daarvoor. Maanden van lezen, piekeren en inverse systemen hadden me Kant nauwelijks verder doen begrijpen. Slechts ternauwernood wist ik het vak te halen.

Na de tweede mislukte poging grip te krijgen op Dora en Michiels fascinerende project vertrok ik naar Berlijn. Een jaar van Kant lezen verder begreep ik er nèt genoeg van om Michiel te durven vragen een scriptie over het onderwerp te begeleiden. Andere docenten zouden misschien de noodzaak gevoeld hebben me tegen mezelf te beschermen, maar helaas voor de rendementscijfers van de UvA gold dit niet voor Michiel. Het schrijven van de scriptie kostte ruim twee jaar, maar ook daarna vond Michiel het onnodig me te adviseren mijn heil elders te zoeken. Hij hielp me een promotieplek vinden, met als gevolg dat ik ten tijde van dit schrijven, ruim tien jaar na mijn eerste poging het Kant-vak te volgen, nog altijd verstrikt ben in onbegrijpelijke Duitse teksten.

De ellende die me bespaard had kunnen blijven als ik Michiel niet had leren kennen! De ellende die anderen bespaard kan worden...! UvA: Stop deze waanzin, desnoods via het emeritusbesluit!

Maar dan: Het èchte eerlijke verhaal.

De doorwaakte nachten, zenuwinzinkingen en therapiekosten ten spijt: Ik prijs mezelf enorm gelukkig dat ik Michiels fantastische colleges heb mogen volgen, en dat hij me zowel tijdens mijn master als daarna heeft willen begeleiden. Michiel was (samen met Dora) misschien wel de meest inspirerende docent van mijn studietijd. Hij geeft les over extreem ingewikkelde onderwerpen, en toont ze in hun volle complexiteit. Dit kan enorm frustrerend zijn, maar het roept ook nieuwsgierigheid op op een manier waarop andere benaderingen dit niet kunnen. Je verdiepen in een werk als de *Kritik der reinen Vernunft* is als het beklimmen van een enorme berg. Iemand kan je uitleggen wat Kant grofweg wil zeggen, maar dit is alsof je de Mount Everest leert kennen via een landkaart. Michiels vakken leiden je via de steilste maar mooiste routes de berg zelf op.

Michiel, ik hoop dat je, ondanks de tegenwerking van de UvA, nog lang les wil blijven geven, en dat we nog heel lang heel veel van je mogen blijven leren. Heel veel dank, en van harte gefeliciteerd!!


Immanuel Kant
Kritik der reinen Vernunft

Chapter 12

Some considerations on thinking machines

Riccardo Pinosio

If an artificial intelligence researcher from the early seventies was transported, by some wondrous accident, to our contemporary times, he would be struck - lacking insight into the evolution of the field in the last forty odd years - by how different it now is. To be sure, the fundamental underpinnings of the connectionist methods that are now state-of-the-art Goodfellow et al. (2016) would be familiar to him Rosenblatt (1958), though he would no doubt be dazzled by the sophistication, optimization, and engineering maturity they eventually achieved. He would also be delighted at the progress that these methods made possible on a host of core AI problems, in particular in computer vision (e.g., image recognition, object segmentation and detection) and natural language understanding (from machine translation to information extraction, intent recognition, and question answering); and at the pervasive and ever growing application of such AI techniques to tackle pressing problems in the industry and civil society.

Still, he would be, if not concerned, at least disoriented at some of these developments. The goal of building a McCarthy-style artificial intelligence agent whose central capabilities are symbolic reasoning and planning, which had been the holy grail of both early AI research (i.e., before the first AI winter triggered by the Lighthill report) and of the expert systems movement in the late eighties, is now gone from centre stage. Moreover, the goal of building an agent with human-level intelligence is now sidelined under the name of "artificial general intelligence" (AGI), replaced mainly by a divide-and-conquer approach where domain-specific, and often application-specific, problems are tackled by means of carefully tailored neural network architectures that are trained (or pre-trained) on enormous datasets. It is true that the latter approach - only made possible at scale by the spectacular advances of microprocessor technology in the early noughties - yielded neural network frameworks which can accurately perform

multiple related tasks, and in particular tasks that the model was not necessarily trained for Brown et al. (2020), thus exhibiting a level of flexibility that had never been reached before in artificial intelligence. Nevertheless, approaches of this kind (such as the transformer architectures of GPT-3 and Bert) can still be considered domain-specific in two important ways: first, they are able to deal with a single modality (natural language, as opposed to vision and planning); second, and more importantly, they lack the ability - which we all recognise in ourselves and that I hold to be one of the fundamental features of human cognition - to encode, or *crystallize*, what has been learned into organised and communicable knowledge - which can be used to plan for the future, generalise inference to unseen but analogous situations, and explain its predictions to humans.

Furthermore, the sidelining of the goal of building a truly reasoning artificial intelligence brought with it a move away from the general logical frameworks and symbolic algorithms that our hypothetical researcher would have recognised as the central techniques of the field, whether syntactic or semantic in nature. These techniques, which dominated the first period of modern artificial intelligence research, would later be branded - rather disparagingly - as GOFAI (or, good old fashioned AI) Haugeland (1989). It lies beyond the scope of this short piece of writing to provide a detailed historical examination of why logical AI, as is generally the opinion of most AI practitioners, failed in its goals. However, while there are certainly historical and sociological reasons that influenced its demise, from a technical perspective the root causes of this failure seem to me quite clear - and can be arranged in plain view rather briefly.

First, symbolic approaches require the discretisation of the input space in learning tasks. This makes them generally unsuitable for supervised learning problems that have unstructured or real-valued (i.e., non-symbolic) inputs - be it sentences, images, or sounds - and where the task is eminently one of pattern-matching, or one that can be framed and dealt with successfully as a pattern-matching task - e.g., machine translation, or named entity recognition. Indeed, note that virtually all major successes of machine learning in the last years have been achieved by methods - such as deep learning and reinforcement learning - which are essentially representation-less propositional learners[1]. In these learners, the central technology is the statistical optimization of cost functions formulated over tensor spaces that encode both the model inputs and its parameters. This technology proved much more effective, robust, and scalable on learning problems involving unstructured data than symbolic manipulation, which formed the

1. It could be debated whether these models are really representation-less, since neural networks can encode representations of higher level concepts (e.g., lines and boundaries in computer vision) in their weights. Still, this is far from symbolic representation as is generally understood in the logical AI tradition.

core of early AI systems - systems that, in any case, emphasised reasoning and representation over learning and perception.

A second, and perhaps more serious, limitation of early AI systems is that their reliance on the symbolic representation of concepts, or, more generally, on symbolic representational languages and the manipulation of their constituent elements, exposes them to a host of problems of syntax and semantics, the most basic one being the symbol grounding problem Harnad (1990) - which we understand here, perhaps abusing the concept, as also comprising the problem of choosing an adequate symbolic representation language in the first place. Despite some research Taddeo & Floridi (2005), and claims that the problem has already been solved Steels (2008), I am not aware of any industrial-grade AI system[2] that provides a fully satisfactory answer to this problem. The challenge here, as I see it, is that the formal semantics of these symbolic systems, which are variations on the semantics of first-order logic, struggle to capture the wealth of meaning of naturally occurring concepts and relations, and thus fail to ground them in the same way in which they are grounded for human agents: through a network or web of meaning that comprises both declarative knowledge and pattern matching on unstructured data, and where facts, analogies, images, sounds, smells, pieces of text all jointly contribute to give meaning to a concept. This semi-structured and perhaps unsystematic approach to meaning seems to me at odds with the crisp approaches of formal semantics. Furthermore, it seems to me fundamentally at odds with the traditional philosophical conception of meaning, as exemplified by Kant's notion of a concept being defined and formalised through intensional rules Achourioti & van Lambalgen (2011); Pinosio & van Lambalgen (2018) in a sort of Porphyrian tree (which now would be termed a deep ontology). Of course, the problem of how to effectively formalise the meaning of concepts to be able to tackle higher level intelligence (such as common sense question answering on pictures or texts) is not exclusive to logical AI - connectionist approaches are not much better at this -, but it is more insidious for these approaches because their reliance on the symbolic logical machinery *makes it seem* like they should be successful at this.

---

2. I emphasise industrial-grade here because I contend that in artificial intelligence a problem has not been fully solved until a system has been built that can address that problem in a real-world situation (i.e., not a toy scenario), yielding concrete value or return on investment. For example, the problem of image recognition was not solved until we had algorithms that could operate reliably, and in the wild, over large image datasets, and could be integrated in commercial products or scientific applications. Hence, the engineering required to transform a proof of concept of an algorithm or system outlined in a scientific paper into a full production system is, in my view, an essential component of that system - and not just an engineering afterthought. This is because in order to ensure scalability one often must engineer adaptations to the original solution radically impact the original design of the system. I believe that if the early AI researchers had adopted this practical mindset before overselling what AI could do with excessive optimism we would not have had the first and second AI winters.

A third limitation of symbolic AI systems is that, until relatively recently, little thought has been given to the difficult problem of how to design symbolic systems that can autonomously invent new concepts and relations, and more generally learn abstract representations and concepts from data[3] (i.e., without requiring human input), and autonomously ground them semantically (whether with a formal intensional semantics or with any other method).

It seems to me that most difficulties faced by symbolic AI systems can be traced back to the above issues. For instance, while early symbolic approaches were quite successful in the building of AI prototypes that could exhibit reasoning and planning abilities in limited[4] and crisp[5] symbolic environments (see, e.g., the blocks world scenario Slaney & Thiébaux (2001)), they ultimately proved inadequate for the designing of AI systems for larger, real-world domains (e.g., the medical domain, or a specific business domain). It suffices to examine the literature on medical ontologies (see for instance El-Sappagh et al. (2018)) using the OWL language McGuinness & van Harmelen *et.al.* (2004) to see that in these larger domains, a host of difficult issues arises that all stem from the above problems, such as:

- what formal language should we use to represent concepts and relations symbolically in the ontology? What primitives should we choose? (problem 2 above)
- how can we ground the meaning of the symbols so that, e.g., we can deal satisfactorily with naturally occurring synonyms, and support, for instance, analogical reasoning? (problem 2 above)
- how can the system be designed so that it automatically extracts entities and relations and higher-order concepts from unstructured data (e.g. text or images), integrating these new symbols with the existing knowledge base, so that the need for human maintenance of the ontology is removed? (problem 1 and 3 above)

Consider, as a seemingly trivial example, the task of formalising the sentence 'Elon Musk is the CEO of Tesla' in a symbolic system. Assuming that we have a relation is_ceo_of and entities e, t representing 'Elon Musk' and 'Tesla', respectively, we might represent the above piece of knowledge with the first order formula is_ceo_of(e, t). However, we also need to specify that the first element of the is_ceo_of relation must be an entity that instantiates the concept 'ceo'; hence, we could introduce a unary predicate is_ceo, and add facts $is\_ceo(e)$, $\forall x, y : is\_ceo\_of(x, y) \rightarrow is\_ceo(x)$ to the knowledge base, to ensure the type

3. The Inductive logic programming community is starting to rise to the challenge with the recent work on predicate invention in ILP systems, see, e.g., Cropper et al. (2020a), although many challenges on this problem remain, in particular how to reduce the space of possible invented concepts to keep only the concepts that are "useful" Kramer (2021).
4. That is, with a small amount of predicates and entities
5. That is, with predicates whose intension and extension are clearly defined

instantiations are correct. However, this does not provide the system with a real grounding for the 'ceo' concept or for the 'Elon Musk' concept, which could be used, for instance, to answer common sense questions about them. In particular, the system would not be able to subsume a new entity under the 'ceo' concept based on, e.g., a description of the professional duties of this new entity at a given company; hence, unless we are to take the 'ceo' concept as primitive, we would need to provide an intensional rule to establish the grounding of the concept 'ceo', which would be formulated in terms of other concepts and relations, and so on, recursively, until it can be fully defined in terms of the primitive concepts and relations of the system. We are then back to the Porphyry-tree approach; but which concepts and relations should be taken primitive? and how would we then ground on the basis of these primitive concepts the concept of 'ceo'? We could attempt to define it along the lines of 'the highest-ranking person in a company or institution, with chief decision-making authority', at which point we need to provide grounding for at least the concepts 'highest-ranking person in a company or institution' and 'chief decision-making authority' (assuming that the concepts 'person', 'company', and 'institution' are taken as primitive); and these are highly abstract concepts, whose meaning is grasped by humans chiefly through usage and concrete examples, rather than through explicit formal definitions that can be difficult to provide. And we have not even taken into consideration in this example the high plasticity of human concepts and relations, which often admit of exceptions to the rule (a bird is a flying animal - except when it's a penguin), and exceptions to the exceptions, and so forth. Very soon, the size and complexity of a formal symbolic system built to deal with common sense question answering balloons, while still being unable to reach anything approaching human-level performance; the cyc system Lenat (1995) is the most spectacular failure of this approach. When reflecting on these issues, one is ultimately reminded of the Buddhist saying where the attempt to capture the richness of reality with human language is compared to trying to empty a lake with a sieve; only, the richness in meaning that human concepts exhibit plays here the role of reality - and the puny formal systems that we set up to capture it is our sieve.

In light of the above considerations, one might wonder whether we should just abandon the road of logical or symbolic AI completely, and stick with the representation-free deep and reinforcement learning methods that have become central to contemporary machine learning. However, such methods have so far avoided the issues discussed above merely by eschewing the goal of building machines that can not only learn, but also *codify* this learning into *knowledge*; knowledge that can be communicated to humans, and that can support inference. Consider, for instance, the following exchange between the user and the machine in the early block-world AI system SHRDLU[6] Winograd (1972):

6. See http://hci.stanford.edu/winograd/shrdlu/ for the full dialogue

Person: What does the box contain?
Computer: THE BLUE PYRAMID AND THE BLUE BLOCK.
. . .
Person: Can the table pick up blocks?
Computer: NO.
Person: Can a pyramid be supported by a block?
Computer: YES.
Person: Can a pyramid support a pyramid?
Computer: I DON'T KNOW.
. . .
Person: Had you touched any pyramid before you put the green one on the little cube?
Computer: YES, THE GREEN ONE.
Person: When did you pick it up?
Computer: WHILE I WAS STACKING UP THE RED CUBE, A LARGE RED BLOCK AND A LARGE GREEN CUBE.
Person: Why?
Computer: TO GET RID OF IT.
Person: Why did you do that?
Computer: TO CLEAN OFF THE RED CUBE.

In this admittedly toy scenario, SHRDLU exhibits the ability to *reason* about the state of the block world and its physical laws, but also the ability to *introspect* on its actions and explain them to a human. Systems like, e.g., GPT-3, while able to deal with a variety of generative NLP tasks, still lack SHRLDU's ability to reason about a system and, most importantly, they lack the ability to provide human-understandable explanations for the natural language text and answers they generate. In short, they lack the ability to encode what they have learned into declarative knowledge. Hence, while statistical machine learning methods have proven to be superior to purely symbolic methods on the problem of learning from (unstructured) data, they have not solved the original problem that early AI researches set out to solve - that of building a thinking machine; rather, they have side-stepped the problem all together.

Of course, one might wonder whether we do need algorithms that support knowledge generation and reasoning at all in order to build thinking machines. Are not statistical learning, pattern matching, and prediction sufficient? The answer to this question must be a resounding: no!

For the lack of knowledge generation and reasoning capabilities does limit current AI systems in various aspects that are crucial for the building of human-level AI. First, it hampers model generalisation, causing decreased performance on input data that was not seen during the training of the model. Second, it severely limits the possibility of achieving *lifelong learning*, i.e., the ability for a

system to repurpose, e.g. via analogical reasoning, what it has learned in solving a given task to solve a new task, potentially from a different domain[7]. Third, knowledge generation and reasoning capabilities are essential for AI systems to be able to explain, justify, and communicate to humans (and potentially other machines) what drives their predictions, thereby guaranteeing the transparency required for their application to sensitive domains such as medicine, drug discovery, and finance; domains where experts need to be able to work side-by-side with, and validate the predictions of, these AI systems. In short, knowledge generation and reasoning are essential components to achieve Michiel's notion of ultra-strong ML Michie (1988) - where a learned hypothesis is required to improve the performance at the given task of a human who is provided with the learned hypothesis itself.

Of course, developing AI systems with knowledge generation and reasoning abilities does not, *per se*, require the formal methods of symbolic logic. It is not to be excluded *a priori* that deep and reinforcement learning algorithms could be developed to fulfil the above *desiderata*, eschewing symbolic representation altogether. Still, the most promising avenue to achieve the goals above seems to me that of finding good ways to augment - or combine - the mainstream statistical machine learning approaches with methods drawn from formal systems, since the latter provide the most refined machinery we have to date for dealing with the knowledge representation and reasoning problem. The hard question, of course, is how to combine the two so as to yield systems which, so to speak, go beyond the sum of their parts: addressing at once and organically as many as possible of the (lifelong) learning, knowledge extraction, and reasoning aspects, while at the same time circumventing the limitations of symbolic representation (in particular the grounding and knowledge base maintenance problems) discussed before. While we are currently still very far from having a satisfactory answer to this question, various communities are working on research agendas that seem likely to hold different parts of the whole puzzle.

First, the literature on neuro-symbolic systems d'Avila Garcez et al. (2019) already contains very relevant attempts to unify statistical and symbolic approaches into hybrid algorithms for the learning and knowledge extraction problems - although most of the methods whose learning process yields human-readable programs still struggle on datasets with a realistic size.

An alternative, but also promising, approach to combine statistical learning and symbolic knowledge extraction comes from the explainable AI literature. Most of the methods proposed by this community to explain the predictions of machine learning models are representation-free, and cannot really produce knowledge from the learning process (see Molnar et al. (2020) and Molnar (2019) for an overview of the methods). However, what if we combined the model-

7. For an exciting example of lifelong learning in the ILP literature, see Lin et al. (2014).

agnostic explainable AI methods, such as LIME Ribeiro et al. (2016) or counterfactual explanations Molnar (2019), with inductive logic programming techniques that can learn symbolic knowledge bases from the output of the explanation algorithms? We would then be able to extract the knowledge learned by any machine learning algorithms, and codify it into a symbolic knowledge base. It is this approach that is explored in Shakerin & Gupta (2019), but further work in this direction is needed, e.g., to be able to deal with regression problems (not just classification), and to evaluate the effectiveness of different explanation algorithms from LIME, such as counterfactual explanations.

On the formal systems side, inductive logic programming based on meta-interpretive learning Muggleton et al. (2015); Cropper et al. (2020b), and the possibility of learning the meta-rules it requires directly from data Patsantzis & Muggleton (2021), seem promising to mitigate the issue of how to choose the linguistic bias for the knowledge induced in the learning process; all the while providing better methods for automated predicate invention and abstraction of higher-order relations.

Still, a workable solution for the symbol grounding problem, which, in my view, is one of the major issues holding back symbolic AI, does not seem forthcoming. Indeed, the above approaches all focus on integrating neural networks and logic for learning, explanation generation, or reasoning, but the underlying issue of how to give meaning to the formal symbols used in these systems still remains, and is likely to hamper hybrid solutions for problems like common sense question answering and text summarisation. Here, I believe we can make progress by leveraging recent work on semantic technologies, and in particular by using large-scale knowledge graphs holding both structured and unstructured data as the substrate that provides meaning to the symbols, and over which the learning, knowledge codification, and question answer tasks can be defined. While the topic of reasoning over knowledge graphs has become popular again in the last years in the NLP community Chen et al. (2020), the work in this space tends to focus on learning unobserved relations between entities, mostly for graph completion. Still, the possibility of using, e.g., hybrid neuro-symbolic techniques to automatically induce (higher-order) concepts and relations (and their definition) from a mix of structured and unstructured data in a knowledge graph remains largely unexplored, but could potentially provide a way to solve the symbol grounding problem at scale.

In conclusion, in order to make progress towards the building of truly thinking machines we will have to face the knowledge representation and reasoning problem - and how it relates to the learning process - once again.[8] The goal here

8. To be sure, the interaction between the developments in artificial intelligence discussed above and cognitive robotics must also play an essential role, since to build a thinking machine it seems likely that we will need to have it learn, acquire knowledge, reason, and evolve organically over a long period

is to give new substance to Valiant's vision of a true *semantics of knowledge*Valiant (2003):

> The aim here is to identify a way of looking at and manipulating common sense knowledge that is consistent with and can support what we consider to be the two most fundamental aspects of intelligent cognitive behaviour: the ability to learn from experience, and the ability to reason from what has been learned. We are therefore seeking a semantics of knowledge that can computationally support the basic phenomena of intelligent behaviour.

We are now equipped with much better tools than the pioneers of artificial intelligence ever had at their disposal to achieve this vision; and the integration of learning, knowledge representation and reasoning into a single new AI paradigm that "leapfrogs" us a step closer to the goal of human-level AI seem reachable in the next few decades. But, as the old adage goes, predictions are very difficult, especially about the future.

---

of time in an environment similar to our own, i.e., in a setting where the amount of data available is limited.

Chapter 13

Interdisciplinary conversations on interpretation

Keith Stenning

Michiel and I have been engaged in an interrupted conversation for nearly 25 years. I think some bits that don't usually get mentioned in print may be interesting to some of this audience, and perhaps a Festschrift permits an unusual topic: the nature of one interdisciplinary communication. I am not here concerned with the truth or significance of the scientific products, though the examples require to be instantiated in some detail to get a feeling for the process. The intent is to highlight some of the features of interdisciplinary conversations, why they are necessary, and how they ramify. For an audience probably slanted toward the logical side, perhaps the examples of how logical input issues in psychological hypothesis and insight may be interesting?

What did we bring to our beginning at the turn of the millennium? Michiel and Fritz Hamm were engaged in an analysis of narrative discourse in Constraint Logic Programming, focussing on temporal relations, eventually issuing in van Lambalgen & Hamm (2005). A masterly application of a non-monotonic logic to narrative discourse. I had started what was to become this thread of the conversation back in 1970, with a PhD argument that narrative reasoning's output could be productively viewed as the construction of a single logical model, and that the conventions of narrative (expressed most particularly by anaphora) conspired toward the singleness of model (eventually summarised as Stenning (1978)). Michiel had made his mark mathematically and logically. I saw my contribution as being to the psychology of reasoning (unlike more or less anyone else, with a few noble exceptions). It contained a rather curious experiment but was essentially a proposal about the representation of narrative, based on systematic intuitions/observations. So we shared a common interest in discourse and cognition, both convinced that logical analysis was a necessary part of interdisciplinary

progress in understanding cognition. And an appreciation of how far apart were our disciplinary backgrounds.

This piece will illustrate with two works in progress (in collaboration with Dora Achourioti and Francisco Vargas): the first an experiment showing that with the right context and instructions (the Syllogistic Dispute task), 'naive logicians' (ones without any logical training) can apply a fully classical logical + refutation procedure interpretation to their reasoning. This strongly suggests that this is *not* the interpretation participants bring to the 'Conventional' tasks normally used to study syllogistic reasoning. The second work in progress takes up the challenge of proposing what interpretation naive logicians do apply to those tasks, and does so by using the comparative method through a new 'patently cooperative' task we will call 'Syllogistic Mind Reading' which we claim has the same interpretation. This second experiment constrains aspects of the interpretation brought, but stops short of a full logical interpretation: much detail remains to be fixed. But it is certainly not detail of a classical logical interpretation. Vargas et al. (2022) and Vargas et al. (2020) will present the full evidence which is summarised here.

Briefly the three tasks: the *Conventional* tasks present syllogisms, with proposed inferences (judgement version) or without (generation version). The instructions are to judge the validity of the inference, or to generate either one of the eight possible conclusions, or a 'no valid conclusion' response. We used the generation version here. These are the standard tasks for assessing 'syllogistic reasoning'. The new *Syllogistic Mind Reading* task makes the same presentation as the Conventional generation version, but explains that the problem is chosen to communicate a model and that the participant's task is to indicate which model on a form for 1- or 2-element models. In the new *Syllogistic Dispute* task, the presentation is the same as the judgement version, with a conclusion. These inferences have been proposed by Harry-the-Snake, a notorious gangster and unreliable source, who is offering bets on their validity. The participant's first sub-task is to judge whether the offered inference is valid (about a third are valid). If they agree with Harry then the problem ends, and they go on to the next. But if they judge the bet invalid, then they bet against Harry and they then have to produce a counterexample (which is explained to them), to justify their bet. The participants who have never done a logic class are assigned randomly to one of these three tasks. Harry's invalid inferences for his bets are chosen from the most popular *invalid* erroneous choices from the Conventional task Khemlani & Johnson-Laird (2012). So these inferences are as plausible as possible.

As a titbit to encourage an audience, the two tasks' (Dispute and Conventional) problem accuracy profiles are *not* significantly correlated, but the latter *is* significantly positively correlated with the novel Syllogistic Mind Reading task which makes it easier to assess what participants are doing. Dispute's classical logic plus refutation procedure interpretation, stretches right down to reproducing

a paradox of material interpretation in the participant's own constructed counter-examples: the very source of the many derisive comments about the implausibility of classical logic for modelling human reasoning. Lot's more psychology flows from the little logic. If we are right, in the Conventional task psychologists of reasoning have been studying 'story' all along: some variety of preferred-model reasoning. Remember the intended goal for this piece is a better understanding how logics and procedures are both necessary for understanding the psychology of reasoning. And how that understanding here emerged from an interdisciplinary conversation.

'Naive' classical logic: under our noses

The first thing to say is that the psychological rewards often radiate out from the logical observations in unexpected ways. Stenning & Oberlander (1995); Stenning & Yule (1997) showed that all valid syllogistic conclusions were valid in respect of at least one single-element premiss-model—one of the eight such. A logical triviality, but one we stumbled on, and suspect it is so trivial as to not appear in logical print. Grown-ups don't mess about with very small finite models much? A simultaneous interest in matching and mismatching middle-terms arrived from the several uses of different meanings of 'mismatching' in psychology Yule (1996) and led to a problem about integrating premisses. It's easy to see some mismatched syllogistic problems do have valid conclusions. If a problem has a positive and a negative B-term in its two premisses, how can they be integrated into a 1-element model? A 1-element model requires a 'trick'. The trick is applicable if one universal premiss has its middle term as subject, because then a negative B in the 1-element model makes the universal premiss true by (empty class generalisation (ECG))—it's antecedent is false so the conditional is true. And the other positive B just has to be part of a true premiss to integrate a mismatched problem in a single element. (i.e. the trick doesn't work if the B in the model is positive, or the B in the universal premiss is negative). So 1-element premiss-models of mismatched problems have to depend on ECG. Ironically, this actually makes countermodelling them *easier*. Their preferred-models make the commonest classically invalid conclusions in the Conventional tasks 'automatically' and desirably false. Whereas the easy tricklessly integrated models of matched premisses make the invalid conclusions true, which means they are *not* countermodels—note the transparently psychological regularity coming from the logical analysis. So this 'implausible', indirect, and arcane logical premiss-shuffling leads to a central psychological regularity about the difficulty of countermodelling invalid matched problem conclusions. It turns out that the ECG requirement for mismatched cases is not a detectable burden in our experiment, but the required, usually cyclical, adjustment of matched problems' unification models to get countermodels

is highly damaging—another psychological regularity. In fact this observation is the kernel of a whole procedure for countermodelling—after a few twists. It rapidly transpired that there are ways of avoiding this damaging 'adjustment cycle' for negative matched problems by ignoring the unification model and using the same element as mismatched problems end up with—end literals both positive and a negative B-term. This method for matched problems also depends on ECG for 1-element models. And after a few more twists there is a 2-element version for problems without a universal which mimics the single element element method but distributes the positive and negative B across two elements (and does not need ECG). Methods for positive problems are different: their 1-element premiss models are all the same ABC element, so 1-element countermodels are a matter of placing the required negation on the right end-term. This works, but the 'placing' can be avoided by adopting a uniform 2-element countermodel with positive B and both end-terms each negative in one of the two elements. This method explains some observed redundancy of 2-element models. This is the outline of an algorithm for countermodelling, all of which flows from the puzzle of how to integrate mismatched problems in a single element. This algorithm, with some extra wrinkles (e.g. for double negative problems) fits rather well the participants' countermodels for the twenty or so invalid inferences selected for our observations. There are plenty of 'calculation' errors, but few successful countermodels that don't use ECG. This latter raises some good questions about the algorithm's role in the observations, not just the facility with, but also the near 'necessity' of, ECG for success.

Here is the production of psychological insight flowing directly from a combination of logical, procedural, and psychological observations. A common psychological question is: "Why is it interesting that all valid problems can be integrated in 1-element models? They have perfectly good 2-element models." A short answer is 'Just wait and see!'. A slightly longer one is that the syllogism is a logic of types. When we stand back, more will follow. Mental modellers have shown some limited ability of naive participants' at counterexample reasoning Bucciarelli & Johnson-Laird (1999), but eschewing reference to logical contributions they have not made the crucial logical connection; this counterexample reasoning *just is* a semantic presentation of this syllogistic fragment of classical logic with a refutation procedure. They think it is a *part* of what is done in the Conventional task—sometimes. We are in the process of showing not. So the next contribution of logic is a large scale map of how different bits of reasoning relate. These naive logicians do this semantic version of classical syllogistic logic right down to producing counterexamples that depend on ECG: one of the paradoxes of material implication. This 'paradox' is famous as the 'implausible' nonsense that classical logic is, as far as human reasoning goes (it is claimed) (Johnson-Laird et al., 2015, Abstract). What is even more relevant here is that participants' performance at

this Dispute task is not significantly correlated with their performance on the 'Conventional' syllogism tasks that define the current state of art in the psychology of syllogistic reasoning. And thirdly, in our data, producing countermodels for problems with at least one universal premiss that depend on ECG is a strong indicator they will be successful proposals: few that do not depend on ECG are correct, though a careful definition of 'dependence on ECG' may be wider than supposed.

So what interpretation does the Conventional task elicit?

Michiel's treatment of narrative involved an adaptation of a non-monotonic logic, Logic Programming (LP), to its task here. An obvious requirement is to understand how story depends on the recruitment of the relevant general knowledge required to integrate each new sentence into the growing 'current model'. Especially the temporal structure. The Bartlett (1932) studies of story processing emphasised 'inference by retrieval': an obvious psychological reference point. Logically, one problem is that the vast majority of long term memory knowledge-base conditionals are irrelevant to the vast majority of narrative conditionals. Yet, in story understanding, information has to be retrieved fast from an unimaginably large heap of knowledge. One organisation is to store generalisations in the form of conditionals. But every interesting universally quantified conditional (perhaps outside of mathematics) has exceptions which then have to be indexed to their relevant conditional as 'abnormalities'. These conditionals cannot be classical ones for the obvious reason that when their antecedents are false (almost always) that does not make the conditional true, nor false. It makes the conditional *inapplicable in the current reasoning episode* ('undefined' in the jargon). But still alive to contribute to other stories. Kleene's logic, designed for studying computer algorithms, has this third U-value. The Kleene reading accords much better with our intuitions about conditionals in 'vanilla' story contexts (outside dispute). Pijnacker et al. (2010) even provides neurophysiological evidence of ERP occurrence timed to where the logic predicts the occurrence of 'abnormalities' becoming active in the discourse. LP offers an explanation how this 'reasoning-by-retrieval' can be 'automatic'—outside awareness. What would life be like if the retrieval process had to be fully consciously accessible? Impossibly slow, and maybe impossible. But the automatic integration of narratives through recruitment of conditionals' consequences in the current model is a central part of cognition. The reasoning involved is among the most elaborate human cognitive achievements upon which much else depends—a keystone of communication. The computational problem has at least two parts: here we are concerned with the retrieval part, but there is also a 'memory creation and indexing' part. There is no reason given here why the latter cannot be nearer to the connectionist systems that dominate at present:

the material here is not about that sort of learning. On-line retrieval in comprehension is still a problem even if learning has done its job. Such symbolic accounts are deeply unfashionable, and there are plenty of questions how this one can be made to work at full scale, but at least it should keep some old questions alive. Here it is an example of the psychology that is kicked off by questions about the implementation of logics. A place where the division between long-term and working memory has an impact on the logic that currently looks most suitable.

One standard psychology of reasoning response has been; 'We are studying *deduction* so we are not interested in story—that's psycholinguistics'. But this won't wash. What they are actually studying in the Conventional task, however ineptly, is whatever interpretation the participants adopt for the materials presented. 'Story' is a pretty good one-word answer. Exams may test whether a student has developed an understanding of an interpretation: experiments have to find out how tasks are actually interpreted. So what are participants trying to do in the Conventional task? We have several times suggested in print that the answer is some variety of *preferred-model* reasoning Shoham (1987). But the psychological Holy Grail of an *experimental* demonstration was not available. We set about an experimental demonstration that would throw new light on the old task. We call it the Syllogistic Mind Reading task: it involves a study of social coordination in reasoning. Our strategy is to show that the extra information that the new task provides by using responses in terms of models helps to show that it has an at least closely related interpretation to that of the Conventional task, which is also interpreted cooperatively.

The presented problems in the new task are syllogistic problems (pairs of premisses without conclusions)—just like those presented in the Conventional generation task, but the standard demand in those tasks is either for the judgement of classical validity of presented conclusions, or their generation. The new cooperative task which produces the second lot of data mentioned above presents problems exactly as for the Conventional generation task, but the responses are to identify the 'intended model', though the 'intended model' phrase is not used in explanations for the participant. So in the trial's instructions a particular syllogistic problem has been chosen to communicate a model, and the participants' task is to decide which model that is. In fact a useful name for the new task is 'Syllogistic Mind Reading': 'Which syllogistic model am I thinking about for this problem?' Just a titbit to engage the reader's permission for this lengthy background, this mind reading task is significantly positively correlated with the data from the Conventional task, but uncorrelated with that of the Dispute task. These correlations are all of mean problem performances of participants randomly sampled from a single large class of 'naive logicians' about to take the same introductory logic course. So, to summarise, if this is what goes on in the Conventional task the field has been studying a particular syllogistic mind-reading task (in logical

terms a variety of preferred-model reasoning), but scoring it on the criteria of classical logical syllogistic reasoning, even while complaining vociferously that logic contributes nothing to our understanding of human reasoning.

To set all this in the context of our conversations, late on in (Stenning & van Lambalgen, 2008, Chapter 10), there is a sketch of an account of how participants interpret the standard syllogistic Conventional task, cast in the non-monotonic logic framework developed and applied to some psychology literature in that book. The book's sketch is the background to the much delayed (12 years, and counting) Syllogistic Mind Reading task chosen here to illustrate our own communications. Michiel and Fritz's book focusses on integrations of the temporal structure of stories. Our use of 'stories' ignores the fact that non-temporal descriptions like syllogisms are not stories or narratives in the usual sense, but they share much with narrative and story: crucially, they require 'integration-by-recruitment-of-knowledge'. We beg forgiveness! 'Story' has here the right connotations.

This recruitment in story is the hallmark distinguishing feature from proof, which is, in classical logic, sealed against recruitment. We have already seen the crucial effect this has on conditionals, requiring Kleene rather than classical conditionals. Another design feature of the Syllogistic Mind Reading task, also strongly correlated with recruitment is the cooperative nature of the task. Dispute has to be confined: imagine how dispute would endlessly disperse if any information could be recruited. But for cooperative story, relevance-to-integration is enough constraint. In the current context, does recruiting this piece of knowledge make possible an intended integration? A species of 'mind reading'. Another property of the task just introduced is important: the task is impossible. That is, a perfect score (choosing exactly the model that the sender had in mind) is impossible to get right all the time in this context, without supreme luck. The mapping from problems to models is many-to-many.

Participants are incredibly forgiving of the vagaries of experimenters, and generally do their best within their interpretation of what they are supposed to do. We pick this rather convoluted example because it encapsulates an absolutely crucial psychological dimension that has been almost completely ignored in the psychology of reasoning, and is perhaps 'old hat' in the logic community. It has been understood that classical logic is adversarial for a long time, even if one of Aristotle's great achievements was to bury this fact, for his good contemporary reasons. Game Theory now provides a more precise definition of both the level at which classical logic is inherently adversarial, and the different prior level at which it rests on a cooperative foundation (agreed interpretations).[1] And it

1. Dutilh Novaes (2021) provides an interesting review of attitudes to adversariality in argument. This review is pitched at a different level than our use of 'adversarial' here (and her related ones elsewhere). Our example of trivial derailment of narrative (page 13) illustrates an apparent contradiction that has to be fixed if narrative is to continue successfully, and this derailment may be signalled by

also encapsulates a topic that has occurred intermittently in our conversations—something that strikes this ageing practitioner as possibly illustrative about interdisciplinary collaboration. The Syllogistic Mind Reading task is blatantly cooperative. A choice is made and a signal sent, and an interpretative choice is made by the participant. The rules are simple: if the second choice is the same as the first, the *team* succeeds; if not it fails. No fault is assigned, but both parties fail. Just like mind reading.

How many participants pick the same model each for the same presented problem? This short cut is harmless. The only aspect of the data that concerns us is to what degree the participant team succeeds, or not. Do they attain this bizarre social coordination? Our research concern is what happens when this data is used to measure participants' differential ability to coordinate on different problems, and thereby, we claim, reveal the basis for response in the Conventional tasks that participants actually adopt. In particular we want to show that there are groups of problems, which should be among the easiest to give the 'there is no classically valid conclusion' (NVC) response to, yet Conventional task participants systematically produce 'invalid' specific conclusion errors. And more interestingly, that there are predictable systematic differences between sub-groups of problems in the frequencies of these so-called 'classical logical errors'. Patterns which can be explained by the differential difficulty of coordinating responses to different problems in the Syllogistic Mind Reading task.

The subset of problems chosen for the final focus after much exploration is on the eight $\exists\exists$ problems which have one $\forall$ and one $\exists$ quantifier/premiss. These are four pairs of problems that are premiss-reorderings of each other—thus controlling several factors in their choice. The quantifiers blatantly cannot give rise to valid specific syllogistic inferences, yet do so in many participants' judgements. And the pair-members are differentially predictable. It was predicted at the design stage that one of each pair would be easier to find mutual agreement for than the other. That is, mutual agreement that they would achieve more mutual agreement. An essence of communication. Remember that none of these problems have classically valid conclusions: a most fundamental syllogistic fact.[2]

The results are as predicted. In two pairs, the Figure 2 member is successfully coordinated more often than the Figure 1 member; in the Figure 3 and 4 pairs the positive-premiss-first problem is coordinated better than the negative-premiss-first one. And this is true for a dimension of choices defined by first ordering the problems by their contributions from the best 'row-models' for problems, and

---

a classical logical contradiction arising. This is not the level at which the issues reviewed by Dutilh Novaes arise. Classical logic's concept of validity itself provides our concept: a valid argument permits no counterexamples.

2. Even the other great regularity that two negative syllogistic quantifiers cannot yield valid conclusions is, in modern terms, simply because a required quantifier is omitted Stenning & Yule (1997).

then assigning the most contributing model for each problem. Up to ten such problems. In plainer English, the contributions of the best contributing model for the first problem in this ranking goes on winning for much longer along the problem dimension. The competitor models take over from the row-model much later down the dimension for the winning member of the pairs (Vargas et al., 2020, Table 8).

So the suggestion is that participants in the Conventional tasks actually interpret the goal they are meant to adopt as cooperative syllogistic mind reading (of the 'preferred model' in some non-monotonic logics), whereas the Dispute task evokes an adversarial stance (at the top level) in the game theoretic sense. A proper logical formulation this is not, but we are here concerned with explaining with as little logical detail as possible, that this is basic psychological stuff which cannot be fully understood without some simple logical/procedural distinctions. What could be more psychological than the goal of the reasoning expressed in the adopted interpretation's concept of validity? And this adoption in the Conventional task is not because they do not know how to adopt the classical logical stance the Experimenter wants. If only they are just given a setting in which it is clear to them that an adversarial interpretation is intended (i.e. in the Dispute task).

What is left for this naive countermodelling of classical logical syllogistic inference to do?

If participants are rather good at countermodelling in a suitably contextualised semantically presented classical logic with refutation procedure, and therefore the Conventional task hasn't been studying participants with a classical logical interpretation, the intriguing question arises what is this small nugget of implicit classical logical refutation skill doing for these participants? Remember this is the tiny syllogistic logic fragment in 1-or 2-element models. And that these are highly educated 'naive logicians. But remember too that their successful countermodels for invalid but attractive inferences are highly dependent on one paradox of material implication—that oddity of classical logic widely reviled by psychologists. They will still exhibit the amazement at the paradoxes of material implication when they first encounter them explicitly presented in logic class. So this tiny fragment elicits implicitly one of the most distinctive features of classical logic.

A preliminary conjecture is that this naive classical logic plays a 'housekeeping' role in breakdowns in the narrative process of reasoning to an interpretation. This is deliberately mundane, but not to belittle: centuries of being cast as the wellspring of eternal truth here need some psychological antidote. The hypothesis is that this kernel of intuitive understanding of classical logical refutation

by counterexample may point to where classical logic gets its foothold in human cognition. No more. It is no accident that this argument emerged in a study of syllogistic reasoning. And no accident that the role is cast as one supporting narrative reasoning.

Syllogisms provide a logic of classification: truth values of binary properties combine to determine *types* of thing in the domain. In the syllogism, here is not even apparatus for identifying or counting the tokens of these types. A consequence is that identifying properties by their extensions in the current domain is much more plausible—Boolean sets. If the current domain is small, and the properties are chosen to be naturally binary *in that domain*, and the background goal is something as mundane but demanding as timetable design, then none of the complexities that stem from exceptions need interfere. As mentioned, all valid conclusions from syllogisms can be modelled in single-element models (the eight of them for three properties). A few countermodels require two elements (28 models available), but that is it. Here Boolean operations are naturally sufficient. So exhaustively agreeing an interpretation is possible, and fundamental base-level disagreements or lack of information made implausible. So the hypothesis suggests that classical logical refutation provides a kernel of dispute-resolution when narrative breaks down. It doesn't even fix disputes. It merely demonstrates a breakdown: deriving $p \wedge \neg p$ (or any other contradiction) signals that there is no model as matters stand, and may provide clues about possible fixes. And narrative often breaks down, especially when knowledge bases diverge or disagreement lurks. Often trivial: 'Oh! I was thinking you meant Shirley Smith, but it must be Shona Smythe? She can't be in two places at the same time.' But sometimes fundamental. Continued communication after whatever breakdown requires the misalignment be identified and fixed, or sometimes worked around. Human narrative discourse (in its widest sense including our non-agentive descriptions) strongly depends on fixing derailments, and in contentious cases, the very recognition of misalignment is crucial to outcomes. Imagine being a member of a species that has just 'invented' narrative, but not having any way of detecting derailments.

If these understandings dependent on logics do not elicit enough psychology, then remember, the participants have not had any logic teaching and we absolutely do not assume that they are repeating to themselves the relevant paradox of material implication when they do the Dispute task. We are looking at a highly contextual phenomenon. They do this in certain heavily loaded contexts, but not in the perhaps 'default' narrative one. Remember also that we do not know exactly which changes to the Conventional task are necessary for the Dispute task to get the differences in responses observed. We simply threw the 'kitchen sink' at getting a classical interpretation. Dispute participants first personally chose whether to reject Harry's inference, and so to countermodel. Requiring participants to countermodel *each problem judged invalid* proved absolutely crucial;

including the general instruction that valid conclusions had no countermodels, but *omitting* this demand for countermodelling of all conclusions judged invalid, made the results significantly positively correlated with the Conventional task. So repeatedly justifying their judgement with countermodels was crucial to their classical logical interpretation. Possibly their personal choice whether to challenge Harry (rather than the experimenter) was also important. And Vargas & Stenning (2020) shows that similar results to those of dispute can be achieved in an exam marking scenario, where the participant 'plays exam marker' presented by supposedly student-produced inference proposals: another adversarial situation well understood by participants. None of these questions arise without logical analysis: none of them are answered without psychological analysis.

To understand how this figures in our conversations, the first observation is that the adversarial/cooperative contrast has surfaced several times in the 20 odd years we have worked together, and I certainly, and I suspect Michiel probably, found other issues to be at the heart of the observations we were at the time trying to understand. Michiel starts out from an awesome command of a wide range of logical, mathematical and philosophical issues that can bear on the cognitive problems. I start out from a very basic grasp of introductory logic (not to forget what I have painfully slowly picked up from Michiel) and we share a severe entanglement in the burning question of how reasoning works. Not to forget a prior firm belief that we will need to use all the different disciplines' insights (and avoid as many of their errors or irrelevancies to our questions) as possible. So I bring conundrums and Michiel brings solutions. I usually start with a glimmer of understanding when he produces his solution, but it usually takes much time before I understand much of what he has done. And the process of my gradual understanding goes on for a long time. I'm afraid I am deeply mathematically obtuse, but I have a nose for interesting problems, and some limited methodological know-how about testing explanations. Once upon a time I was a biology student, and never cease to be amazed at how my 50 years out-of-date acquaintance is still so relevant in the deep background, despite all the revolutions. Most important is that I can trust what Michiel tells me—including when he says he doesn't really know for sure.

A short summary of a long conversation

The demonstration of 'naive classical logical ability' in a Dispute situation was a 'discovery made in the cupboard': countermodelling as classical logic in disguise—under our noses. Turning it into a psychological demonstration through fitting an algorithm, fell out of work that started fifty years ago in an argument that story produced single logical models. Integrating mismatched problems in 1-element models required using ECG. And the rest tumbled out as described above. One

moral is that it is hard to predict what bits of logic will take you to which bits of psychology. And the logic that is needed may or may not be 'logically' interesting. Hard psychological work remains in completing the algorithm and fitting the data more accurately than this demonstration-of-principle, figuring out the individual differences—just for example. When all is done, there will be an account of how another bit of logic relates to another bit of psychology.

The demonstration of the cooperative nature of the Conventional task involves creating a novel task of concluding *models*, and using 'number of participants managing mutual coordination' as the fundamental output variable. We started with two dimensions models-for-problems and problems-for-models: for the focal 'row-problem' ranking the contributions of all the models proposed for it gives its models dimension. The greatest contributor is called the 'target-model' for the problem and all the models making contributions are ranked down from that. Then this target-model defines the dimension of problems that receive responses from it to the ones that receive least. A third 'external' dimension takes the same problems and ranks the most contributing models for each. Basically, the winning problem of each of the four pairs maintains the most contributions being from the 'target-model' much further down the ranked list of problems than its losing problem. At the end there are always some other models that contribute more. This difference is absolutely clear cut.

$\exists\exists$ problems have been a puzzle ever since reading the early psychological papers on syllogistic reasoning (How could participants possibly think there could be classically valid inferences if they had the haziest grasp of what classical logic was? If they understood the question, that is. Which we have just seen they don't—when asked it in the Conventional task. And this fragment of eight (positive/negative premiss problems out of sixteen), produces the most sensitive measure, controlling several other variables. The findings fit in outline the 32 problems in our experiment but certainly should be extended to the full 64. The most important logical clues came from the modelling of the narrative in van Lambalgen & Hamm (2005) and the suppression task in Michiel's tailored LP ((Stenning & van Lambalgen, 2008, Chapter 7/8)). Once Kleene gives the conditional semantics necessary for retrieval (and incidentally explains the vehement rejections of the paradoxes in the classical conditional), the psychologist understands the importance of distinguishing Long Term Memory for knowledge-bases from the Working Memory for episodes of reasoning producing new interpretations: reasoning *to* and reasoning *from* interpretations. And the necessity of a fundamentally cooperative algorithm for retrieval from the former. Note that this retrieval (known as the 'Belief Bias' fallacy in the psychology of reasoning literature) is *obligatory* in the operation of LP in story discourse. So again its the same moral: its hard to predict how, which logical understandings will lead to which deeper psy-

chological understandings. The idea that logical and psychological contributions are in competition is deeply bizarre in modern times.

I doubt there is anything obligatory about this particular sampling from our conversations: and even here Michiel would doubtless emphasise different connections. But yes, it should have been possible to have had this conversation more quickly.

Chapter 14

Episodic problems

Martin Stokhof

In a recent paper, Stokhof (2021), I formulated some thoughts about the possible role(s) of philosophy. I started from the observation that much of modern philosophy – in the analytical tradition but also, increasingly, in other traditions – is almost exclusively concerned with concepts and conceptual analysis, with current 'conceptual engineering' as a very prominent instance. The main factor driving that development is the increasing influence of science on the way in which we view the world and ourselves. The results of science and technology pervade our everyday lives in ways that even in recent history would have stricken many as fiction, science fiction to be sure. But even more pervasive is the underlying picture of what the world is like – material, governed by strict laws – and of what we are like – material as well, and governed by material principles: survival, utility maximisation. Even much of the pseudo-science that some cling to, feeds on that picture – misunderstood, sure, but influential nonetheless.

What place is left for philosophy if science shapes and drives us in such a profound and inescapable way? As a way of knowing, it has relinquished its claim to a domain of its own a long time ago. There is no separate reality that is the subject matter of philosophy. There is not even a distinct set of questions about reality that only philosophy can answer. Rather, philosophy has associated itself with science in an uneasy alliance. Philosophy as conceptual analysis is a necessary prolegomenon to science, according to some, or just a clarifying afterthought, according to others. The distinction seems to be more a matter of temperament than that it is based on a factual difference of opinion. Philosophy's subject matter are concepts, those used by science, but also those used in everyday life. With regard to the latter, more often than not philosophy sets as its goal to show that these are based on mistaken, i.e., non-scientific, views that need correction, if not eradication.

Although analytic philosophy is especially prone to this, other philosophical traditions display similar tendencies. In the paper I mentioned Foucauldian archeology and genealogy as examples. In the work of Pierre Hadot we can find further support. In his *What is Ancient Philosophy?* ((Hadot, 2002, p. 261)) Hadot states the following claim:

> The dominance of Idealism over all university philosophy, from Hegel to the rise of existentialism and subsequently the vogue of structuralism, has done much to foster the idea that the only true philosophy must be theoretical and systematic.
> Such, it seems to me, are the historical factors that have led to the conception of philosophy as pure theory.

'Pure theory' in this tradition is, of course, not exactly the same as conceptual analysis in the analytic tradition, but there is a common core: a move away from the everyday, a philosophy that is modelled on the example of science, resulting in an intellectual endeavour that is struggling to maintain its identity confronted as it is with the ever-extending reach, and successes, of science.

The result is that professional philosophers risk what Hadot calls 'intellectual and moral perdition', and he quotes Jacques Bouveresse's description ((Bouveresse, 1973, p. 74), which I quote from Hadot) of what lies in wait for them (*ibid.*):

> In a sense, there is no servitude more intolerable than that which constrains a man professionally to have an opinion in cases in which he may not necessarily have the least qualification. What is at issue here, from Wittgenstein's point of view, is not by any means the philosopher's 'wisdom' – that is, the stock of theoretical knowledge he has at his disposition – but the personal price he has had to pay for what he believes he is able to think and say. [...] In the last analysis, a philosophy can be nothing other than the expression of an exemplary human experience.

Hadot argues for the need to do philosophy based on 'a choice for a certain way of life', which enables the philosopher to conceive (*ibid.*, p. 270):

> [...] of philosophy not only as a concrete, practical activity but also as a transformation of our way of inhabiting and perceiving the world.

This conception of philosophy, Hadot illustrates with references to the works of Montaigne, Descartes, Kant and others, among whom, remarkably, we also find Wittgenstein (*ibid.*, p. 273):

> For instance, we know from one of Wittgenstein's letters that his Tractatus Logico-Philosophicus, which is apparently, and indeed

> truly is, a theory of the proposition, is nevertheless fundamentally a book of ethics in which 'what pertains to ethics' is not said but shown. Wittgenstein elaborates his theory of the proposition in order to justify this silence concerning ethics, which is foreseen and deliberate from the beginning of the book. What motivates the Tractatus is the will to lead the reader to a certain kind of life, and a certain attitude, which, moreover, is fully analogous to the existential options of ancient philosophy: 'to live within the present,' without regretting, fearing, or hoping for anything.

This conception of 'philosophy as a way of life' weds, we could say, theoretical reflection with 'care of the self', but not as two distinct and independent partners. Theoretical reflection is not an end in itself. It is not something that produces a particular kind of 'philosophical' knowledge, it is a tool. Philosophy serves first and foremost to transform our way of seeing things, things that are important to us in a practical manner.

In the aforementioned paper I discussed Wittgenstein's considerations about aspect seeing and aspect change in precisely such a context. The upshot of that discussion was that aspect seeing and aspect change are not meant to discard one way of viewing something in favour of another, 'better' one. Rather they are tools that open up possibilities in a practical context ((Stokhof, 2021, p. 11)):

> The philosopher's engagement ultimately is normative. It is to open up possibilities not just for the sake of it, but in order to be practical. The guiding idea is that in order for there to be meaning, actual or possible, there has to be a practical point. This is key: in the end, philosophy is a matter of seeing *and* acting, of reflection and practical engagement. And the insistence on the latter introduces a moral perspective.

That then led me to the conception of philosophy as 'philosophie pauvre' (*ibid.*, p. 12):

> a modest, hesitating, critically self-reflecting philosophy, one that suggests, asks, observes; not a philosophy that makes claims, defends theses, projects visions. Rather than carving out a highly specialised, exclusively philosophical domain, it seems it is both more modest and more productive to view philosophy as one way of dealing with the episodic, the everyday.

In his as always sharp but constructive comments Michiel raised questions about this conception of philosophy, specifically about the idea that it should be concerned with episodic problems: what exactly characterises an episodic problem? Although the term as such is conveniently vague, at least that is what I thought

at the time, the question is, of course, pertinent. Without a proper delineation of what episodic problems are, the conception of a philosophy that is concerned with such problems remains up in the air. In what follows I will try to come up with an answer. It is a first attempt, and no doubt Michiel will have further questions to expose its weak points (or gaping holes . . .).

Starting point is the connection with Wittgenstein's idea that a language game, or practice, should 'have a point'. This is something that he does not often discuss explicitly, but that clearly informs the way he thinks of the very concept of a language game. It serves a purpose, it has an intentionality, if one wishes to use that term, that distinguishes it from a mere conglomerate of verbal and non-verbal actions. This purpose need not be conceived in utilitarian terms. After all, Wittgenstein mentions, in *Philosophical Investigations*, 23, a wide variety of language games, which often do not have a purpose that can be framed in such terms.

However, that does not mean that any structured ensemble of verbal and non-verbal actions classifies as a language game. Obviously, Wittgenstein construes 'having a point' in a more restricted way. A case that illustrates this is that of philosophy, more particularly, of the kind of philosophy that Wittgenstein opposes. In *Philosophical Investigations*, 38, in his discussion of names and simples, Wittgenstein referring to the idea that demonstratives are perhaps the only genuine names, makes the following observations (italics in original):

> This is connected with the conception of naming as a process that is, so to speak, occult. Naming seems to be a strange connection of a word with an object. – And such a strange connection really obtains, particularly when a philosopher tries to fathom the relation between name and what is named by staring at an object in front of him and repeating a name, or even the word "this", innumerable times. For philosophical problems arise when language *goes on holiday*. And *then* we may indeed imagine naming to be some remarkable mental act, as it were the baptism of an object. And we can also say the word "this" *to* the object, as it *were* address the object as "this" – a strange use of this word, which perhaps occurs only when philosophising.

Traditional philosophy is an attempt to go beyond how language is actually used. But then 'language goes on holiday', it no longer 'has a point'. The German original of the phrase in question is, perhaps, more telling:

> Denn die philosophischen Probleme entstehen, wenn die Sprache *feiert*.

'Feieren' means 'celebrate', but it also connected with 'Feierabend machen', which denotes the end of the working day, when we quit. So, a more appropriate translation might be:

For philosophical problems arise when language stops doing its job.

What is important to note is that 'language doing its job' is something specific: it is language being used in the way in which we use it in everyday contexts, with reference to everyday, practical concerns. Wittgenstein would not deny that language is being used in philosophy, intentionally and systematically. Philosophical discussion is not a matter of inarticulate mumbling and grumbling. But, according to Wittgenstein, it does lack a point, being removed as it is from our everyday concerns. That is what is the problem with traditional philosophy on his view.

This interpretation of what 'having a point' comes to ties meaningful language use intrinsically to our everyday concerns and that, at least for Wittgenstein, has a distinct moral dimension. Throughout Wittgenstein's life, the question how to live, how to find the proper moral stance to the challenges that world and life present, has been a central concern. This is quite explicit in Wittgenstein's early writings. And although explicit discussion of it in the later work is scarce, it does seem to have this dimension as well.

However, it is one thing to claim that traditional philosophy lacks sufficient connection with everyday concerns, and that a proper philosophy should. But it is quite another thing to somehow argue, let alone show, that this is indeed the case. From a standard philosophical perspective this would mean that we need to come up with a definition of what 'having a point' exactly is, and then show that traditional philosophy lacks such a point, in contradistinction to other types of intellectual activity. Conceived of in this way, we find ourselves in a bind. For coming up with a definition seems to presuppose exactly the kind of philosophical conceptual analysis that we want to do away with.[1] So what are we to do? Perhaps we can take our clue from Wittgenstein's stance with regard to ethics. That is a complicated story, but one thing is clear: ethics is not concerned with the formulation and philosophical grounding of general moral principles, but with attitude and action. That means that moral education does not consist in teaching rules but is a matter of 'leading by example' and of reference to such examples. Applied to our current problem this means the following. We cannot first define with the means of traditional philosophy what the alternative that does away with it will be: that is self-defeating. But we can do philosophy in a different way, and by doing so show what is wrong with doing philosophy in the traditional way. It seems that this is exactly what Wittgenstein did, or at least,

1. The problem appears to be related to the kind of self-reference that plagues philosophy in many guises (radical scepticism, radical relativism, but also their counterparts) and that is arguably connected to its absolute character.

what he tried to do.[2] Wittgenstein did indeed oppose traditional philosophy, but he also engaged in a different type of philosophy, one that is supposed to show, rather than prove, that it is the more meaningful activity.

Notice that the movement of our argument here is circular: in order to understand what this particular kind of philosophy is, we already need to have a grasp of what it is. Or, to put it differently, we cannot convince someone that this is the right way of viewing the matter when they do not already see the difference that we are after. But that is okay: it just means that the insight we are after is not something that is internal to philosophy. It has to come from outside, from a moral stance on what matters in life. In order to bring about such a change in perspective, we need, not an argument, but an aspect change, a different way of looking at what we are doing, and why. Some more discussion of how aspect seeing and aspect change can be used to bring such changes about can be found in the earlier paper.

In the present context we can say that what is at the core of the aspect change that is needed is a turn from viewing a problem as a conceptual one to looking at it as an episodic problem. Or perhaps it is better say that we are looking for a change in how we see a problem, i.e., we need to move from a conceptual way of looking at a problem to the realisation that it is connected with what concerns us practically, i.e., an awareness of the matter in question 'having a point'. If we do that, we are looking at the problem as 'episodic'. So, episodic problems are not a class of problems of their own, distinct from another class of problems, viz., conceptual ones. The distinction is not one that can be characterised in terms of content, rather, it is a matter of function, of how we view things.

But doesn't the association of 'episodic' with 'practical' imply that episodic problems are limited, and that the abstract and theoretical problems that are at the core of philosophy as conceptual analysis are ruled out? On the contrary, the functional nature of the opposition shows that this does not necessarily follow. Taking an episodic attitude towards a problem does not rule out that abstract and theoretical elements are involved. It does mean that these, too, are approached in a particular way, and that might make a difference for how we deal with them. Starting point as well as end point of our investigation should be grounded in practical concerns. But that does not rule out that abstract and theoretical considerations enter into the picture as well.

Let me finish with a first, very concise attempt to illustrate this with an example. It concerns AI. The question how/when/why to treat AI as akin to/on a par with human intelligence is one of the most central philosophical questions

2. This touches on the debate on the nature of Wittgenstein's (later) philosophy, in particular its therapeutic function. According to some, this is the *only* function of Wittgenstein's writings. Obviously, the interpretation that I am using here, does not agree. However, a proper defence of that needs to await a different occasion.

surrounding AI. It has received a lot of attention and many different answers have been proposed. Obviously, it represents something that is intellectually challenging and important. However, despite all the intellectual efforts that have gone into it, the matter is far from satisfactorily resolved. Why is that? The suggestion here is that asking the question in the standard philosophical way, as a question about concepts, human intelligence and artificial intelligence, that needs to be answered by analysis and argument, is the wrong way of approaching it. What is needed, it seems, is a change in perspective. The question is not: Are artificial systems intelligent like humans? That is a purely conceptual way of phrasing it. Rather, it makes more sense to address the matter from the point of view of its importance to us: When will we treat artificial systems as intelligent? In much the same way that the humanity of others is a matter of attitude, as Wittgenstein indicated in *Philosophical Investigations*, 420, II iv, 19, AI and our relation to it is a matter of attitude as well. When we grow up with AI in ways that are sufficiently similar to the ways in which we grow up with humans, the attitudes we have towards them, AI and humans, will be the same. If we look at the initial question from this episodic point of view, our considerations may still touch on abstract and theoretical questions. But the perspective from which we address them and the role that is played by the answers that we come up with, is fundamentally different. One such more abstract consideration revolves around the fact that there are, of course, lots of intermediate cases. AI can play many different roles in our lives that range from pure, tool-like functionality to something that invokes a 'human-like' stance. What is crucial is that these differences do not correlate one-to-one with conceptual differences. And that is why the conceptual way of viewing the question falls short.

These considerations are just a first attempt to get clear what the episodic nature of a 'philosophie pauvre' might entail. Much more detail needs to be added, of course, as well as further justification. But I hope the above goes some way to addressing the concerns that Michiel has voiced, and that by doing it has given him further reason to continue our ongoing discussion about it.

Chapter 15

Reflection on 'Reasoning in non-probabilistic uncertainty'

Leon van der Torre

Contribution to the liber amicorum for Michiel van Lambalgen

To show that probability is not the only way of dealing with uncertainty, we note that William James and other philosophers have used the concept of possibility and necessity to distinguish between what we know for sure and what we can only guess at. Maybe we know for sure that the Earth goes around the Sun, but maybe we only think it goes around the Sun. We cannot know for sure anything else, so we are left with the other possibility: maybe the Earth goes around the Sun, maybe it does not.

There are kinds of uncertainty which are for principled reasons not addressable with probabilistic means. For example, it is not possible to assign a probability to a proposition such as 'The Earth might be destroyed by a meteorite within the next 100 years.' This is because it involves an event which is logically impossible, given the laws of physics. It is also not possible to assign a probability to a proposition such as 'I might wake up tomorrow with a different gender identity,' or 'I might wake up tomorrow as a different person.' This is because such events are not in principle predictable, even though they are in principle possible.

To provide evidence that logic-based methods can well support reasoning with uncertainty, we note that there are several areas of logic with a long history of dealing with uncertainty. Moreover, these areas of logic have in the last decade become the focus of active research which has the potential to transform them into strong and effective formalisms for reasoning with uncertainty. These areas of logic are first-order logic with probability measures, generalized probability theory based on Dempster-Shafer theory, modal logics and relational logics.

Logic Programming with Kleene semantics for modelling reasoning from information in a discourse, and probabilistic logic programming, in which events are interpreted probabilistically. To express uncertainty precisely and to reason with it, we need to introduce a new kind of logic, one which allows for the expression of uncertainty.

Uncertainty in dynamic normative contexts is still a very large research topic. Although it is possible to formulate conditions under which uncertainty can be ignored, it will remain important to consider the role of uncertainty in other situations that might commonly occur.

For example, there might be a number of norms relevant to a decision, and the decision maker may have incomplete information on which norms are relevant and how they should be weighted. In such situations, it may be possible to consider the set of norms as a single problem and to use the decision-making approaches that work for single norms. Also, it might be possible to consider the norms as a single norm with multiple components. For example, if there are a number of norms relevant to a particular decision, but the decision maker knows that certain norms are much more important than others, it may be possible to represent the more important norms as a single norm and then to use the more general decision-making approaches.

A neural-symbolic implementation of Input/Output logic for dealing with uncertainty in dynamic normative contexts can be constructed in a manner similar to the neural-symbolic implementation of Fuzzy Logic discussed. The basic ideas for a neural-symbolic implementation of Fuzzy Logic are similar to the neural-symbolic implementation of Input/Output Logic.

The input to this system is a fuzzy logic network, which is a network that implements a fuzzy reasoning system. The output is a symbolic representation based on fuzzy logic. A fuzzy logic network is a network where every input node is associated with a fuzzy set, and every output node is also associated with a fuzzy set. The weights on the links are fuzzy numbers. The network's output is computed the same way as any other neural network, except that the output is computed as a weighted sum of the fuzzy values of the output nodes. In other words, each node's output is represented by a fuzzy number, and the output of the network is a fuzzy number.

Michiel van Lambalgen's view on logic is that it is a theory of the content of a language. This places him in the anti-realist tradition of the early Wittgenstein and the later Carnap. He argues that there can be no knowledge of the world because all statements about it are about language. This is not a theory about the world, but a theory about knowledge. He does not mean that we have no knowledge of the world, but that we do not have knowledge of the world as it is in itself. This view is hard to reconcile with van Lambalgen's own views about the actual world, since these views imply that we have knowledge of the world.

His view on Reasoning in Non-Probabilistic Uncertainty is also anti-realist and anti-cognitivist. Probability is not the only way of dealing with uncertainty; there is also the option of accepting ignorance when there is no way of making a reasonable prediction.

This text has been generated with the help of OpenAI's GPT-3

Bibliography

Achourioti, Theodora & van Lambalgen, Michiel. 2011. A formalization of Kant's transcendental logic. *The Review of Symbolic Logic*, 4(2), 254–289.

—. 2017. Kant's logic revisited. *IfCoLog Journal of Logics and Their Applications*, 4, 845–865.

Alechina, N. & van Lambalgen, M. 1996. Generalized quantification as substructural logic. *Journal of Symbolic Logic*, 61, 1006 – 1044.

Alechina, Natascha. 1995a. *Modal Quantifiers*. Ph.D. thesis, ILLC, University of Amsterdam.

—. 1995b. On a decidable generalized quantifier logic corresponding to a decidable fragment of first order logic. *Journal of Logic, Language and Information*, 4, 177 – 189.

Alechina, Natascha & van Lambalgen, Michiel. 1995. Correspondence and completeness for generalized quantifiers. *Bulletin of the Interest Group in Pure and Applied Logic*, 3, 167 – 190.

Andréka, Hajnal, van Benthem, Johan, & Németi, István. 1995. Back and forth between modal logic and classical logic. *Bulletin of the Interest Group in Pure and Applied Logic*, 3, 685 – 720.

—. 1998. Modal languages and bounded fragments of predicate logic. *Journal of Philosophical Logic*, 27, 217–274.

—. 2017. On a new semantics for first-order predicate logic. *Journal of Philosophical Logic*, 46(3), 259–267.

Apfel, L. J. 2011. *The Advent of Pluralism: Diversity and Conflict in the Age of Sophocles*. Oxford University Press, Oxford.

Archard, D. 1996. Introduction. In: Archard, D., ed., *Philosophy and Pluralism*, pp. 1–5. Cambridge University Press, Cambridge.

Aristotle. 1984. *The Complete Works of Aristotle: The Revised Oxford translation*. Princeton University Press, Princeton.

—. 1989. *Prior analytics*. Hackett Publishing.

—. 2014. Metaphysics. In: *Complete Works of Aristotle, Volume 2: The Revised Oxford Translation*, pp. 1552–1728. Princeton University Press.

Aristotle, Hardison, B., O., & Golden, L. 1968. *Aristotle's Poetics: A Translation and Commentary for Students of Literature*. Prentice-Hall, Englewood Cliffs, NJ.

Aucher, G., ed. 2020. *Expedition in the Update Universe*. Proceedings Dynamic Logic. New Trends and Applications, Prague.

Audi, Robert. 2007. *Moral Value and Human Diversity*. Oxford University Press, New York.

Baggio, Giosuè. 2021a. Compositionality in a parallel architecture for language processing. *Cognitive Science*, 45, 5.

—. 2021b. Epistemic transfer between linguistics and neuroscience: Problems and prospects. In: Nefdt, Ryan, Klippi, Carita, & Karstens, Bart, eds., *The Philosophy and Science of Language: Interdisciplinary Perspectives*, pp. 275–308. Palgrave Macmillan, London.

Baggio, Giosuè, van Lambalgen, Michiel, & Hagoort, Peter. 2008. Computing and recomputing discourse models: An erp study. *Journal of Memory and Language, 59*, 1, 36–53.

Baltag, Alexandru & van Benthem, Johan. 2021. A minimal logic of functional dependence. *Journal of Philosophical Logic*.

Bartlett, F. 1932. *Remembering*. Cambridge University Press, Cambridge, UK.

Barwise, Jon & Feferman, Solomon, eds. 1985. *Model-Theoretic Logics*. Springer.

van Benthem, Johan. 1983. *Modal logic and classical logic*. Bibliopolis, Naples.

—. 1986. *Essays in Logical Semantics*. Reidel, Dordrecht.

—. 1996a. *Exploring Logical Dynamics*. CSLI, Stanford.

—. 1996b. Inference, methodology and semantics. In: Bystrov, P. & Sadofsky, V., eds., *Philosophical Logic and Logical Philosophy. Essays in Honour of Vladimir Smirnov*, pp. 63–82. Kluwer Academic Publishers, Dordrecht.

—. 2008. Logic and reasoning: Do the facts matter? *Studia Logica*, 88, 67–84.

—. 2011. *Logical Dynamics of Information and Interaction*. Cambridge University Press, Cambridge.

—. 2017. Against all odds: A logician looks at probability. In: Katoen, J., Langerak, R., & Rensink, A., eds., *Festschrift for Ed Brinksma*, pp. 239–253. Springer, Lecture Notes in Computer Science 10500.

van Benthem, Johan & Alechina, Natascha. 1994. Modal quantification over structured domains. In: de Rijke, Maarten, ed., *Advances in Intensional Logic*. Kluwer, Dordrecht.

van Benthem, Johan, Gerbrandy, Jelle, & Kooi, Barteld. 2009. Dynamic update with probabilities. *Studia Logica*, 93, 67–96.

van Benthem, Johan & Westerståhl, Dag. 2012. Generalized quantifiers meet modal neighborhood semantics. In: *Outstanding Logicians: Hájnal Andréka and Ístvan Németi*. Springer, Dordrecht.

van Berkel, Kees. 2015. *Kant's Logic in the Critique of Practical Reason - A Logical Formalization of Kant's Practical Transcendental Argument*. Master's thesis, University of Amsterdam, ILLC Publications, Master of Logic Thesis (MoL) Series. Thesis supervised by Prof. Michiel van Lambalgen.

Blando, Francesca Zaffora. 2020. *Patterns and Probabilities. A Study in Algorithmic Randomness and Computable Learning*. Ph.D. thesis, Stanford University.

Bouveresse, Jacques. 1973. *Wittgenstein: La Rime et la Raison*. Les Editions de Minuit, Paris.

van Bouwel, J. 2014. Pluralists about pluralism? Different versions of explanatory pluralism in psychiatry. In: Galavotti, M. C., et al., eds., *New Directions in the Philosophy of Science*, pp. 105–119. Springer, New York.

Brouwer, Luitzen Egbertus Jan. 1912. Intuitionism and formalism. *Bulletin of the American Mathematical Society*, pp. 81–96.

Brown, Tom, et al. 2020. Language models are few-shot learners. *arXiv preprint arXiv:2005.14165*.

Bucciarelli, M. & Johnson-Laird, P. 1999. Strategies in syllogistic reasoning. *Cognitive Science*, 23(3), 247–303.

Burkert, W. 1985. *Greek Religion*. Harvard University Press, J. Raffan, Trans.). Cambridge, MA.

Cassirer, Ernst. 1981. *Kant's Life and Thought*. Yale University Press.

Chaiken, S. & Trope, Y. (Eds.). 1999. *Dual-process theories in social psychology*. Guilford Press, New York.

Chen, Xiaojun, Jia, Shengbin, & Xiang, Yang. 2020. A review: Knowledge reasoning over knowledge graphs. *Expert Systems with Applications*, 141, 112948.

Chierchia, Gennaro & McConnell-Ginet, Sally. 2000. *Meaning and Grammar: An Introduction to Semantics*. MIT press.

Comrie, Bernard. 1976. *Aspect. An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge University Press.

Cook, R. T. 2010. Let a thousand flowers bloom: A tour of logical pluralism. *Philosophy Compass, 5*, 6, 492–504.

Craig, Edward. 1998. *Pluralism*. Routledge, In *Routledge Encyclopedia of Philosophy*. London.

Craigie, J. 2011. Thinking and feeling: Moral deliberation in a dual-process framework. *Philosophical Psychology, 24*, 1, 53–71.

Cropper, Andrew, Dumančić, Sebastijan, & Muggleton, Stephen H. 2020a. Turning 30: New ideas in inductive logic programming. Preprint. arXiv:2002.11002.

Cropper, Andrew, Morel, Rolf, & Muggleton, Stephen. 2020b. Learning higher-order logic programs. *Machine Learning*, 109(7), 1289–1322.

Culicover, P. W. & Jackendoff, R. 2006. The simpler syntax hypothesis. *Trends in Cognitive Sciences*, 10(9), 413–418.

Currie, A. & Killin, A. 2016. Musical pluralism and the science of music. *European Journal for Philosophy of Science, 6*, 1, 9–30.

Ding, Y., Holliday, W., & Icard, T. to appear. Logics of imprecise comparative probability. *International Journal of Approximate Reasoning*.

van der Does, Jaap. 2011. *Passed over in Silence. On Wittgenstein's Tractatus and its System*, *Studies in Logic*, vol. 28. College Publications, London.

van der Does, Jaap & van Lambalgen, Michiel. 2000. A logic of vision. *Linguistics and Philosophy*, 23, 1–92.

van der Does, Jaap & Stokhof, Martin. 2020. Tractatus, application and use. *Open Philosophy*, 3(1), 770–797.

Dutilh Novaes, Catarina. 2021. Who's afraid of adversariality? Conflict and cooperation in argumentation. *Topoi*, 40(5), 873–886.

El-Sappagh, Shaker, et al. 2018. SNOMED CT. Standard ontology based on the ontology for general medical cience. *BMC Medical Informatics and Decision Making*, 18(1), 1–19.

Engelhard, Kristina & Mittelstaedt, Peter. 2008. Kant's theory of arithmetic: A constructive approach? *Zeitschrift für allgemeine Wissenschaftstheorie*, 39(2), 245–271.

de Finetti, B. 1937. La prévision: Ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 17, 1–68.

Forsyth, James. 1970. *A Grammar of Aspect: Usage and Meaning in the Russian Verb*. Cambridge University Press.

Frankish, K. & Evans, J. S. B. T. 2009. The duality of mind: An historical perspective. In: Evans, J. S. B. T. & Frankish, K., eds., *In Two Minds: Dual Processes and Beyond*, pp. 1–29. Oxford University Press, New York.

Frascolla, P. 1994. *Wittgenstein's Philosophy of Mathematics*. Routledge, London.

Frascolla, Pasquale. 1997. The *Tractatus* system of arithmetic. *Synthese*, 112, 353–378.

Galliani, P. 2021. Dependence logic. The Stanford Encyclopedia of Philosophy.

d'Avila Garcez, Artur, et al. 2019. Neural-symbolic computing: an effective methodology for principled integration of machine learning and reasoning. *arXiv e-prints*, pp. arXiv–1905.

Geach, Peter. 1972. Aristotle on conjunctive propositions. In: *Logic Matters*, pp. 13–27. University of Califonia Press.

Gierasimczuk, Nina. 2010. *Knowing One's Limits. Logical Analysis of Inductive Inference*. Ph.D. thesis, ILLC, University of Amsterdam.

Ginsberg, M. L. 1987. *Readings in Non-Monotonic Reasoning*. Morgan Kaufman Publishers, Inc., Los Altos, CA.

Goodfellow, Ian, Bengio, Yoshua, & Courville, Aaron. 2016. *Deep Learning*. MIT press.

Goodstein, R.L. 1957. *Recursive Number Theory*. Studies in Logic. North-Holland, Amsterdam.

Grädel, E. & Walukiewicz, I. 1999. Guarded fixed-point logic. *Proceedings LICS*, pp. 45–54.

Hacker, Peter. 1984. *Insight and Illusion*. Oxford University Press, Oxford. Second revised edition. First edition 1972.

Hadot, Pierre. 2002. *What is Ancient Philosophy?* Harvard University Press, Cambridge, Mass. Translated by Michael Chase.

Hamm, Fritz, Kamp, Hans, & van Lambalgen, Michiel. 2006. There is no opposition between formal and cognitive semantics. *Theoretical Linguistics*, 32(1), 1–40.

Hanna, Robert. 2018. Kant's Theory of Judgment. In: Zalta, Edward N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2018 edn.

Hansson, Sven Ove. 2013. The varieties of permission. In: Hansson, S.O., et al., eds., *Handbook of Deontic Logic and Normative Systems*, vol. 1, pp. 195–240. College publications London.

Harnad, Stevan. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3), 335–346.

Haugeland, John. 1989. *Artificial Intelligence: The Very Idea*. MIT press.

Heim, I. & Kratzer, A. 1998. *Semantics in Generative Grammar*. Blackwell.

Hornischer, L. 2021. The logic of information in state spaces. *Review of Symbolic Logic*.

Icard, Th. & Ibeling, D. 2020. Probabilistic reasoning across the causal hierarchy. *Proceedings 34th AAAI Conference*.

Ivanhoe, P. J. 2009. Pluralism, toleration, and ethical promiscuity. *Journal of Religious Ethics, 37*, 2, 311–329.

Johnson-Laird, P., Khemlani, S., & Goodwin, G. 2015. Logic, probability, and human reasoning. *Trends in Cognitive Science*, 19(4), 201–214.

Johnstone, H. W. & Marienthal-Maschler, C. 1962. Skepticism and inferior knowledge: A note on Aristotle's pluralism. *Philosophy and Phenomenological Research, 22*, 4, 472–480.

Looren de Jong, H. 2002. Levels of explanation in biological psychology. *Philosophical Psychology, 15*, 4, 441–462.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. Farrar, Straus and Giroux, New York.

Kalsbeek, Marianne. 1995. *Meta-Logics for Logic Pogramming*. Ph.D. thesis, ILLC, University of Amsterdam.

Kant, Immanuel. 1992a. The false subtlety of the four syllogistic figures. In: *Theoretical Philosophy 1755–1770*, The Cambridge Edition of the Works of Immanuel Kant, pp. 87–105. Cambridge University Press, Cambridge. Translated and edited by David Walford & Ralf Meerboote.

—. 1992b. *Lectures on Logic*. The Cambridge Edition of the Works of Immanuel Kant. Cambridge University Press, Cambridge. Translated and edited by J. Michael Young.

—. 1998. *Critique of Pure Reason*. The Cambridge Edition of the Works of Immanuel Kant. Cambridge University Press, Cambridge. Translated and edited by Paul Guyer and Allen W. Wood.

Keestra, Machiel. 1999. "Elektra" und Hegels Unterbewertung der Individualität und öffentlichen Gerechtigkeit auf der antiken Szene. *Hegel-Jahrbuch, 1999*, 1, 116–120.

—. 2014. *Sculpting the Space of Actions. Explaining Human Action by Integrating Intentions and Mechanisms*. Ph.D. thesis, ILLC, University of Amsterdam.

—. 2015. Understanding human action. Integrating meanings, mechanisms, causes and contexts. In: Scholz, R. W., Kiyaschenko, L. P., & Bazhanov, V. A., eds., *Transdisciplinarity in Philosophy and Science: Approaches, Problems and Prospects.*, pp. 201–235. Navigator, Moscow.

—. 2017. Meta-cognition and reflection by interdisciplinary experts: Insights from cognitive science and philosophy. *Issues in Interdisciplinary Studies*, 35, 121–169.

Keestra, Machiel, Uilhoorn, A., & Zandveld, J. In press. *An Introduction to Interdisciplinary Research. Theory and Practice.* Amsterdam University Press, Amsterdam.

Kellert, S. H., Longino, H. E., & Waters, C. K. 2006. Introduction. The pluralist stance. In: Kellert, S. H., Longino, H. E., & Waters, C. K., eds., *Scientific pluralism*. University of Minnesota Press, vii-xxviii, Minneapolis.

Khemlani, S. & Johnson-Laird, P. N. 2012. Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 138(3), 427–457.

Klein, D. & Rendsvig, R. 2021. Convergence, continuity, recurrence and Turing completeness in dynamic epistemic logic. *Journal of Logic and Computation*.

Kovač, Srećko. 2008. In what sense is Kantian principle of contradiction non-classical? *Logic and Logical Philosophy*, 17(3), 251–274.

Kramer, Stefan. 2021. A brief history of learning symbolic higher-level representations from data (and a curious look forward). In: *Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence*, pp. 4868–4876.

Lagerlund, Henrik. 2021. Medieval Theories of the Syllogism. In: Zalta, Edward N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edn.

van Lambalgen, Michiel. 1987. *Random Sequences*. Ph.D. thesis, ILLC, University of Amsterdam.

—. 1990. The axiomatization of randomness. *Journal of Symbolic Logic*, 55, 1143–1167.

—. 1991. Natural deduction for generalized quantifiers. In: van der Does, Jaap & van Eijck, Jan, eds., *Generalized Quantifier Theory and Applications*, pp. 143–154. Dutch Network for Language, Logic and Information, Amsterdam. CSLI Lecture Notes, 54, pp. 225-236.

—. 1992. Independence, randomness and the axiom of choice. *Journal of Symbolic Logic*, 57, 1274–1304.

—. 1993. Lecture at the 5th European summer school on logic, language and information, Lisbon, Portugal.

—. 1996. Randomness and foundations of probability: Von Mises' axiomatisation of random sequences. In: *Statistics, Probability and Game Theory. Essays in Honor of David Blackwell*, pp. 347–367. Institute for Mathematical Statistics.

—. 2011. Tractatus on time. In: van der Does, Jaap & Dutilh Novaes, Catarina, eds., *"This is Not a Festchrift" To/Voor Martin*. ILLC, Amsterdam, `http://festschriften.illc.uva.nl/M60/`.

—. 2013. Time in the *Critique of Pure Reason*. In: Aloni, Maria, Franke, Michael, & Roelofsen, Floris, eds., *The dynamic, inquisitive, and visionary life of ϕ, ?ϕ, and $\diamond\phi$. A festschrift for Jeroen Groenendijk, Martin Stokhof, and Frank Veltman*. ILLC, Amsterdam.

van Lambalgen, Michiel & Hamm, Fritz. 2005. *The Proper Treatment of Events*. Blackwell.

Leitgeb, Hannes. 2017. *The Stability of Belief. How Rational Belief Coheres with Probability*. Oxford University Press, Oxford.

Lenat, Douglas B. 1995. CYC: A large-scale investment in knowledge infrastructure. *Communications of the ACM*, 38(11), 33–38.

Lin, Dianhuan, et al. 2014. Bias reformulation for one-shot function induction. *Frontiers in Artifical Intelligence and Applications*.

Longino, H. E. 2002. *The Fate of Knowledge*. Princeton University Press.

Longuenesse, Béatrice. 1998. *Kant and The Capacity to Judge: Sensibility and Discursivity in the Transcendental Analytic of the Critique of Pure Reason*. Princeton University Press.

Łukasiewicz, Jan. 1957. *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*. Oxford: Clarendon Press, 2nd, enlarged ed.

Marion, Mathieu. 1998. *Wittgenstein, Finitism, and the Foundations of Mathematics*. Oxford Philosophical Monographs. Clarendon Press, Oxford.

Marr, David. 1982. *Vision. A Computational Investigation into the Human Representation and Processing of Visual Information*. W.H. Freeman, San Francisco.

Mason, E. 2008. Value pluralism. In: Zalta, E. N., ed., *Stanford Encyclopedia*. of Philosophy.

McCauley, R. N. & Bechtel, W. 2001. Explanatory pluralism and heuristic identity theory. *Theory & Psychology*, 11(6), 736–760.

McGuinness, Deborah & van Harmelen *et.al.*, Frank. 2004. OWL. Web ontology language overview. *W3C Recommendation*, 10(10), 2004.

Michie, Donald. 1988. Machine learning in the next five years. In: *Proceedings of the 3rd European Conference on European Working Session on Learning*, pp. 107–122.

Mints, G. & Kremer, Ph. 2007. Dynamic topological logic. In: Aiello, M., Pratt-Hartmann, I., & van Benthem, J., eds., *Handbook of Spatial Logics*, pp. 565–606. Springer, Heidelberg.

Mitchell, S. D. 2003. *Biological Complexity and Integrative Pluralism*. Cambridge University Press, Cambridge.

Molnar, Christoph. 2019. *Interpretable Machine Learning*. Leanpub.

Molnar, Christoph, Casalicchio, Giuseppe, & Bischl, Bernd. 2020. Interpretable machine learning–a brief history, state-of-the-art and challenges. In: *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 417–431. Springer.

Mostowski, A. 1957. On a generalization of quantifiers. *Fundamenta Mathematicae*, 44, 12 – 36.

Muggleton, Stephen H, Lin, Dianhuan, & Tamaddoni-Nezhad, Alireza. 2015. Meta-interpretive learning of higher-order dyadic datalog: Predicate invention revisited. *Machine Learning*, 100(1), 49–73.

Mühlhölzer, Felix. 2010. *Braucht die Mathematik eine Grundlegung? Ein Kommentar des Teils III von Wittgensteins Bemerkungen über die Grundlagen der Mathematik*. Vittorio Klostermann, Frankfurt am Main.

Mundici, D. 2012. *Logic: A Brief Course*. Springer.

Münte, T. F., Schiltz, K., & Kutas, M. 1998. When temporal terms belie conceptual order. *Nature*, 395(6697), 71–73.

Németi, I. 1985. Cylindric-relativized set algebras have strong amalgamation. *The Journal of Symbolic Logic*, 50(3), 689–700.

Nussbaum, Martha. 1999. Virtue ethics: A misleading category? *The Journal of Ethics*, 33, 163–201.

Olstad, A. M. H., Fritz, I., & Baggio, G. 2020. Composition decomposed: Distinct neural mechanisms support processing of nouns in modification and predication contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 46(11), 2193–2206.

Partee, B. H. 2018. Changing notions of linguistic competence in the history of formal semantics. In: Ball, D. & Rabern, B., eds., *The Science of Meaning: Essays on the Metatheory of Natural Language Semantics*, pp. 172–196. Oxford University Press, Oxford.

Patsantzis, Stassa & Muggleton, Stephen H. 2021. Meta-interpretive learning as metarule specialisation. *arXiv preprint arXiv:2106.07464*.

Pijnacker, Judith, et al. 2009. Defeasible reasoning in high-functioning adults with autism: Evidence for impaired exception-handling. *Neuropsychologia, 47*, 3, 644–651.

—. 2010. Reasoning with exceptions: An event-related brain potentials study. *Journal of Cognitive Neuroscience*, 23(2), 471–480.

Pinosio, Riccardo & van Lambalgen, Michiel. 2018. The logic and topology of Kant's temporal continuum. *The Review of Symbolic Logic*, 11(1), 160–206.

Poggio, T. 1981. Marr's computational approach to vision. *Trends in Neurosciences*, 4, 258–262.

Posy, Carl J., ed. 2013. *Kant's Philosophy of Mathematics: Modern Essays*, vol. 219. Springer Science & Business Media.

Potter, M. 2000. *Reason's Nearest Kin*. Oxford University Press, Oxford.

Pylkkänen, L. 2020. Neural basis of basic composition: what we have learned from the red–boat studies and their extensions. *Philosophical Transactions of the Royal Society B*, 375(1791), 20190299.

Ramsey, Frank P. 1923. Critical notice: *Tractatus Logico-Philosophicus*. *Mind*, 32(128), 465–478.

Rescher, Nicholas. 1993. *Pluralism: Against the Demand for Consensus*. Clarendon Press, Oxford.

Ribeiro, Marco Tulio, Singh, Sameer, & Guestrin, Carlos. 2016. "Why should I trust you?" Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1135–1144.

Rosenblatt, Frank. 1958. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6), 386.

Ruben, D. H. 1992. *Explaining Explanation*. Routledge, New York.

Sachs, K.-J. & Dahlhaus, C. 2001. Counterpoint. In: *Grove Music Online*. Oxford University Press, Oxford. Retrieved on January 15, 2022.

Sacks, Mark. 2006. Kant's first analogy and the refutation of idealism. In: *Proceedings of the Aristotelian Society*, vol. 106, pp. 115–132. Wiley.

Schatzki, Theodore R. 2010. *The Timespace of Human Activity*. Lexington books, Lanham.

Shabel, Lisa. 2021. Kant's philosophy of mathematics. In: Zalta, Edward N., ed., *The Stanford Encyclopedia of Philosophy (Fall 2021 Edition)*. Standford University, https://plato.stanford.edu/archives/fall2021/entries/kant-mathematics/.

Shakerin, Farhad & Gupta, Gopal. 2019. Induction of non-monotonic logic programs to explain boosted tree models using lime. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 3052–3059.

Shanahan, Murray. 1990. Representing continuous change in the event calculus. In: *9th European Conference on Artificial Intelligence, ECAI 1990*, pp. 598–603.

Shoham, Yoav. 1987. A semantical approach to non-monotonic logics. In: *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, pp. 388–392. Reprinted in Ginsberg 1987.

Skolem, Thoralf Albert. 1923. Begründung der elementaren Arithmetik durch die rekurrierende Denkweise. Videnskapsselskapets Skrifter. I. Mat.naturv. Klasse. 1923. No. 6, Kristiania.

Slaney, John & Thiébaux, Sylvie. 2001. Blocks world revisited. *Artificial Intelligence*, 125(1-2), 119–153.

Smith, Robin. 2020. Aristotle's Logic. In: Zalta, Edward N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2020 edn.

Snell, B. 1975. *Die Entdeckung des Geistes*. Vandenhoeck & Ruprecht, Göttingen.

Snowdon, Paul & Gomes, Anil. 2021. Peter Frederick Strawson. In: Zalta, Edward N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2021 edn.

Steels, Luc. 2008. The symbol grounding problem has been solved. So what's next? In: de Vega, Manuel, Glenber, Arthur, & Graesser, Arthur, eds., *Symbols and Embodiment: Debates on Meaning and Cognition*, pp. 223–244. Oxford University Press, Oxford.

Stenning, K. & Oberlander, J. 1995. A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, 19, 97–140.

Stenning, K. & Yule, P. 1997. Image and language in human reasoning: A syllogistic illustration. *Cognitive Psychology*, 34, 109–159.

Stenning, Keith. 1978. Anaphora as an approach to pragmatics. In: Halle, Moritz, Bresnan, Joan, & Miller, George, eds., *Linguistic Theory and Psychological Reality*. MIT Press.

Stenning, Keith & van Lambalgen, Michiel. 2001. Semantics as a foundation for psychology: a case study of Wason's selection task. *Journal of Logic, Language and Information*, 10, 273–317.

—. 2004. A little logic goes a long way: basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science*, 28, 481–529.

—. 2005. Explaining the domain generality of human cognition. In: Roberts, M., ed., *Integrating the Mind: Domain General Versus Domain Specific Processes in Higher Cognition.*, pp. 179–209. Psychology Press, East Sussex.

—. 2008. *Human Reasoning and Cognitive science*. MIT Press, Cambridge, Mass.

Stern, Robert. 2003. *Transcendental Arguments: Problems and Prospects*. Oxford University Press.

Stokhof, Martin. 2002. *World and Life as One: Ethics and Ontology in Wittgenstein's Early Thought*. Stanford University Press, Stanford.

—. 2021. Philosophy as change. To appear in Wang Zhongjiang (ed), *Philosophy in China Today*.

Storm, W. 1998. *After Dionysus: A Theory of the Tragic*. Ithaca, NY: Cornell University Press.

Strawson, Peter. 2018. *The Bounds of Sense: An Essay on Kant's Critique of Pure Reason*. Methuen, London.

Stroud, Barry. 1968. Transcendental arguments. *The Journal of Philosophy*, 65(9), 241–256.

Taddeo, Mariarosaria & Floridi, Luciano. 2005. Solving the symbol grounding problem: A critical review of fifteen years tf research. *Journal of Experimental & Theoretical Artificial Intelligence*, 17(4), 419–445.

Taylor, Charles. 1978. The validity of transcendental arguments. In: *Proceedings of the Aristotelian Society*, vol. 79, pp. 151–165. Wiley.

Tiles, Mary. 2004. Kant: From general to transcendental logic. In: Gabbay, Dov & Woods, John, eds., *The Rise of Modern Logic: From Leibniz to Frege*, vol. 3, pp. 85–130. Elsevier.

Valiant, Leslie G. 2003. Three problems in computer science. *Journal of the ACM*, 50(1), 96–99.

Vargas, Francisco, Achourioti, Theodora, & Stenning, Keith. 2020. What are subjects really doing when asked to solve syllogisms: preferred-model reasoning? To appear.

—. 2022. Syllogistic dispute: A task to invoke 'naive' classical logical reasoning. To appear.

Vargas, Francisco & Stenning, Keith. 2020. Communication, goals, and counterexamples in syllogistic reasoning. *Frontiers in Education*, 5(28), 1–18.

Wade, Terence. 2002. *The Oxford Russian Grammar and Verbs*. Oxford University Press.

Wason, P. 1968. Reasoning about a rule. *Quartely Journal of Experimental SPychology*, 20, 273–281.

Westerståhl, D. 1989. Quantifiers in formal and natural languages. In: Gabbay, D. & Guenthner, F., eds., *Handbook of Philosophical Logic*, vol. IV, pp. 1–131. D. Reidel Publishing company.

Wimsatt, W. C. 2007. *Re-Engineering Philosophy for Limited Beings. Piecewise Approximations to Reality*. Harvard University Press, Cambridge, Mass.

Winograd, Terry. 1972. Understanding natural language. *Cognitive psychology*, 3(1), 1–191.

Wittgenstein, L. 1922. *Logisch-philosophische Abhandlung*. Suhrkamp Verlag, Frankfurt am Main.

—. 1967. *Bemerkungen über die Grundlagen der Mathematik*. Basil Blackwell, Oxford. Bi-lingual edition. English title: Remarks on the Foundations of Mathematics. Edited by G.H. von Wright, R. Rhees, G.E.M. Anscombe. Translated by G.E.M. Anscombe.

von Wright, Georg Henrik. 1951. Deontic logic. *Mind*, 60(237), 1–15.

—. 1963. *Norm and action*. Routledge and Kegan Paul.

Wylie, A. 2015. A plurality of pluralisms: Collaborative practice in archaeology. In: Padovani, F., Richardson, A., & Tsou, J. Y., eds., *Objectivity in Science: New Perspectives from Science and Technology Studies*, pp. 189–210. Springer International Publishing, Cham.

Yule, P. 1996. *Logic and Implementation in Human Reasoning: The Psychology of Syllogisms*. Ph.D. thesis, The University of Edinburgh.

Zaliznyak, Anna A. & Shmelev, A. D. 2000. *Vvedenie v russkuyu aspektologiyu*. Yazyki Russkoi Kultury. In Russian.

Zalta, Edward N. 2022. Gottlob Frege. In: Zalta, Edward N., ed., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Spring 2022 edn.