Logic, Neuroscience and Phenomenology: In Cahoots?

Ahti-Veikko Pietarinen*

Department of Philosophy, P.O. Box 9, FIN-00014 University of Helsinki
pietarin@cc.helsinki.fi

Abstract. Cognitive sciences, including cognitive neurosciences, have provided important insights into the notions of awareness, implicit/explicit information processing in knowledge, perception, object identification and memory, as well as general information retrieval. Meanwhile, propositional-attitude logics have coped with awareness in terms of symbolic tools, but are lacking the pathways by which to relate the two fields. I argue that empirical findings concerning rare neural dysfunctions (blindsight, unilateral neglect, prosopagnosia, implicit memory) contribute to logical investigations. On the other hand, the early phase on cognitive science, the origins of which coincide with that of pragmaticist philosophy, shared roots with phenomenology. Accordingly, I will identify strands in that early period that have surfaced in logic, AI and computer science. In phenomenology, the significance of the division between implicit and explicit aspects of knowledge in understanding cognition was acknowledged very early on.

1 Introduction

One of the key conceptual tools in cognitive neurosciences is the *implicit/explicit* distinction. It may be drawn in similar ways with respect to a variety of notions such as knowledge, belief, perception, memory and learning. These notions have been, in the mainstream analytic philosophy, subsumed under the concept of *propositional attitudes*, especially epistemic ones. Unfortunately, this perspective masks the processual, active and dynamic character of these notions. Above all, it masks the difference between implicit and explicit methods of knowing, believing, seeing, recalling or learning, a key descriptive division in cognitive approaches to information processing in the human brain.

Empirical and *logical* sides of these manifestly different implicit/explicit distinctions may nevertheless be examined in a parallel fashion [20]. The goal is to provide cognitive neuroscientists, computer scientists and philosophers with integrative tools that smooth progress in mutual understanding of what it means for the mind to be simultaneously both conscious and an aware, and unconscious and unaware processor of information. This amalgamation is expected to provide

* Supported by the Academy of Finland (1103130). My thanks go to the reviewers for comments.

new cognitively-grounded logical tools for both philosophical and computational purposes, and contribute to the drafting of a preliminary agenda of informatics that would not evolve in isolation from neighbouring disciplines.

In this paper, I will delineate some parallels between logical and neuroscientific aspects of awareness and the implicit vs. explicit distinction, with special reference to *phenomenology*.

2 Relating Neuroscience and Logic

2.1 Implicit versus Explicit Knowledge

It is quite striking to become aware of the extent to which neuroscientific research contributes to conceptual and logical approaches to the epistemic concept of knowledge and its ilk. Indeed, in recent years, cognitive neuroscience has provided insights into implicit versus explicit information processing in *perception*, *identification*, *memory*, *information retrieval*, *belief content*, *belief formation*, *iterated knowledge* and *introspection*, to name but a few [6, 27]. Perhaps above all, cognitive neuroscience has brought the notion of awareness to the forefront of human knowledge [28].

Meanwhile, philosophers and cognitive scientists have suggested diverse approaches to problems posed by the mental concept of *awareness* [5]. There are significant logical reflections of neuroscientific phenomena that have remained uncharted. But neuroscientific insights may be carried over to bear on logical theories. A particularly useful way of doing this is by introducing *operators* into epistemic languages aimed at formalising constructions of agent's knowledge and other variants such as memory and belief [8, 10, 16, 17, 20].

This has some inevitable repercussions to general theories addressing the *logic of consciousness*. To date, however, logical theories that aim at incorporating some interpretation of the notion of awareness into the language have merely been addressing the question of how to dispose of the *logical omniscience problem* in the resulting systems. That is to say, they have been baffled by the property of traditional epistemic logics, namely agents coming to know all the logical consequences of known formulas. While the problem of logical omniscience may have been the first job description of the introduction of the notion of awareness to bear on logic [8], this no longer needs to be the case. Logical omniscience may be extirpated by other, much stronger means, for example by augmenting the received *possible-worlds semantics* for knowledge with *impossible possible worlds*, in which not all the classically-valid theorems, such as the law of non-contradiction, hold in the sense of these worlds being epistemically possible — even if they were to accommodate contradictions [24].

In addition, it has been shown that logics of awareness may be embedded in the impossible worlds framework [26]. Because (but not only because) of this fact, vital applications for the task of bringing logical concerns closer to cognitive ones are better to be sought elsewhere than in the omniscience problem.

This said, the logical omniscience problem is not devoid of explanatory power with respect to neuroscientific issues. It provides a logical answer to the neuroscientific problem of what it is that separates the implicit and explicit aspects of knowing from one another. The customary explanation is that implicit knowledge does not exhibit *conscious access to information*. From the logical viewpoint, it may be said that an agent implicitly knows something if and only if that knowledge is *closed under a logical consequence relation*. The phrase 'conscious access to information' recently ingerminated in neuroscientific literature turns out to carry logical content.

Conversely, cognitive neuroscience puts before us a range of data concerning *blindsight*, *unilateral neglect*, *prosopagnosia* or *implicit memory* not prehended in logic yet. In fact, a wealth of conceptual and interpretational issues exists to be assessed for relevance of neuroscience to logical modelling and the analysis of knowledge and cognition in relation to awareness.

A genuine need exists for a unifying language in which these investigations may be carried out. One such candidate is epistemic logic, the *epistemic variant of modal logic* dealing with an agent's propositional attitudes, introduced in [10]. Indeed, both epistemic logic and the associated possible-worlds semantics turn out to be quite versatile in modelling actual cognitive phenomena (cf. some related suggestions to that effect in [17]). These phenomena include the distinction between explicit and implicit knowledge, belief and memory, different senses of awareness, and perception.

Accordingly, new ways of extending the basic epistemic language to cover empirically-supported phenomena in cognitive neuroscience such as blindsight and aspects of amnesia and memory become amenable to a new kind of conceptual analysis. The relevant cognitive phenomena are, in fact, quite overwhelming: [20] parallels only a few logical and neuroscientific facets of the implicit vs. explicit distinction in relation to knowledge, memory, and other propositional attitudes, and there are certainly more.

Hintikka [11] was perhaps the first to recognise the importance of the interrelations between experimental findings in neuroscience on the one hand, and epistemic logic on the other. He discussed these relations with an eye on philosophical insights into the famous cogito argument 'Cogito, ergo sum'. One of the primary insights in that paper was the differentiation between two modes of *identification*, the *perspective* and the *public* one, and to relate that distinction to neuroscientific and cognitive 'where-versus-what' systems that have independently been observed as relevant to higher-level cognitive functions.

2.2 What is the Logic of Awareness?

Without probing into the details of the approach [17, 20], in *logics of awareness*, an agent i can be said to be aware of a proposition p in case in which i knows that $K_i\,(p \vee \neg p)$. In other words, the agent knows that the law of excluded middle holds. This has sometimes been explicated in the sense of *situation semantics*, so that the agent is aware of the proposition precisely in those situations that 'support' or 'preserve' the truth-value of the proposition. In other words, an agent cannot be aware of propositions that have a truth-value of Undefined. In alternative terminology, there are no partial interpretations in such situations for sentences among the agent's set of 'aware' ones.

In [8], awareness is taken to be a syntactic operator, conjoined to an implicit knowledge operator precisely in order to convert the proposition into explicit knowledge. (A related distinction exists in neuroscience, in which one would instead speak of *overt* vs. *covert* knowledge [27, p. 256].) In other words, given implicit knowledge $\mathbf{K}_i\,\varphi$, this is transformed into explicit knowledge by defining $K_i\,\varphi ::= \mathbf{K}_i\,\varphi \wedge A_i\,\varphi$. The operator A_i attached to the proposition φ means that i *is aware of* φ, and is explicated in [8] by the syntactic means of denoting an 'aware' proposition. To this effect, a mapping $\mathcal{A}_i(w)$ from a world w to propositions provides a set of propositions of which agent i is aware.

The suggestion that the awareness operator may be interpreted in different ways depending on the purpose at hand is not motivated by conceptual or cognitive concerns, but by *computational* ones. But it is contestable whether computation is a likely candidate for a general theory of cognition. Such readings are proposed in [8] as 'an agent is aware of p if she is aware of all concepts that it contains', or that 'p's truth-value can be computed within an interval T'. There are other conceptual, cognitive and computational readings of awareness, yielding to independent logical systems of their own [20].

2.3 Three Levels of Investigation

The cognitive workings of awareness and related areas of information processing in the brain may be studied at several levels. The first is the actual, *physiological* and *neural end* of mechanisms of information processing. This level is intimately connected with the details of how to *implement* a certain kind of awareness in a particular hardware configuration, irrespective of whether such hardware consists of neural processors, bio-computers, quantum gate devices, or any of the platforms employed in traditional computing.

The second level looks away from these actual mechanisms and examines those tasks in which an understanding of awareness is sought via *cognitive information processing*. Examples of this *task-oriented level* are problems in *psycholinguistics* such as text or dialogue processing and comprehension, or the role of the cross-modal (perceptual, tactile, auditory) conception of awareness in abstract skills such as problem-solving and reasoning. Memory and information retrieval are also relevant in these tasks.

On the third level, we look away from both actual neural mechanisms and task-oriented analysis and develop *abstract* (say, *logical* or *semantic*) ways of addressing awareness. It is this third level that, in my opinion, sets the most topical agenda for the interplay between philosophy and informatics. This methodological stance is not intended to diminish the role of computation *per se*, but that concerns only a derivative agenda, materialised once the logical workings of cognitive notions are better understood.

This is by no means the only viable standpoint to the overall task of marrying philosophy, cognitive and computational sciences with one another. At the very least, complex co-acts between the three main levels and their sub-levels appear incontestable in any broad-minded cognitive theory. But if so, then the input provided by neuroscience is not to be ignored on the other two levels, either, no matter how abstract (e.g. non-empirical) the level of analysis may be.

The interplay between logical notions of knowledge, belief, perception and memory on the one hand, while setting cognitive neuroscience and neuropsychology on the other, also raises its head in artificial systems. For example, in *systems controlled by* such languages that involve epistemic notions there is the need to create proper naming relations instead of just generic terms in order to attain tasks that involve the conscious processing of information. Logical systems are illustrations of what is generally required of such epistemic languages, and they are for that reason useful for several *knowledge-representation tasks* concerning *multi-agent systems* [22].

However, this does not have to lead to any *reduction*, for in viewing either side of what is much the same matter. Both logical and neuroscientific terms and concepts are needed when approaching intelligence in cognitive systems in order to embed more reliable and realistic epistemological features into such artefacts. There is little substance in those arguments that claim that one day, by looking at subneuronal levels in the brain, we may find logical operations and properties in action at those levels.

2.4 Binding as a Trans-Disciplinary Problem

Although the logical systems of implicit knowing and attitudes outlined in [20] are propositional, the addition of quantifiers would in the end be indispensable, since full notions of knowledge cannot exist independently of how objects are identified, a fact that is intimately connected with the neuroscientific phenomena of how humans actually individuate objects [17]. However, such extensions give rise to new questions: Is *cross-identification* viable between different compartments of sets of possible worlds needed in disgregating viable implicit acts or attitudes from explicit ones? If so, would not, then, cross-identification be related to the neuroscientific *binding problem*, namely the question of how do independent bits of data come to be combined into unitary, coherent percepts?

There is a wider insight motivating the topicality of these questions. In logic as well as in neuroscience, *binding data* is vital. Symbolic, logical systems typically achieve this by reusing the same variables in a formula, in other words arranging variables to become bound by the same quantifiers. In neuroscientific terms, the problem has been to formulate a conjunction of different local areas of the brain capable of producing *coherent experiences.* The traditional method in logic is to rewrite universal (resp. existential) quantifications as infinite conjunctions (resp. disjunctions) of atomic predicates. On the other hand, a richer (albeit historically earlier) way is to consider both the *processes of predication* and *identity* between different occurrences of variables as reflections of the same underlying conceptual process.

The nature of this unifying process may logically be unravelled as a diagrammatic and iconic rather than a symbolic activity. One fallout is that by such a *diagrammatisation*, an alternative and formally rigorous path to conceptualisation processes advocated in *cognitive linguistics* [14] is revoked.

More generally, the affinity between binding data and binding percepts vindicates Lakoff & Johnson's [14] conclusion that "to understand reason we must understand the details of out visual system, our motor system, and the general mechanism of neural binding" (p. 4). In diagrammatic approaches to logic [18, 21], predication, existence and identity are attained via the same underlying sign that marks continuous connections between differently localised predicates. I believe that such *heterogeneous, iconic* and *topological representations* will provide a promising locus for one of the most striking illustrations of the congeniality of two hitherto-isolated areas, logic and cognitive neuroscience.

3 Phenomenological Ramifications

3.1 Phenomenology: an Aide to Artificial Intelligence after all?

Phenomenology has traditionally been inimical to AI [7]. However, Albertazzi [1] has recently revealed a rich conceptual repertoire and thematic complexity in the early Central European theories revolving around the accounts of cognitive science developed in between 1870 and 1930. They include psychophysical, functional and constructivist approaches to cognition. According to Albertazzi, these theories anticipated much of the content of the contemporary cognitive sciences as well as many of the modes of research typical of the current mainstream research in cognitively-oriented sciences and philosophy.

Let me add that the revival of this early endeavour to develop *conceptual categories of the mind and psyche* is currently underway in several junctures of AI, logic and cognitive science. Unfortunately, this historical connection has not featured regularly in these theories. The following points need to be highlighted:

- Scientists and logicians, who instead of *objects* were accustomed to talk of *individuals* during the symbolic era, have been resorting to the more phenomenological terminology. A case in point is the idea of *discourse objects* (or *discourse markers* or *references*) in discourse-representation theory, which are certain intermediate representations or mental models concerning discourse and objective facts, taking place between sentences and the world.
- In lieu of *predicates*, many logical theories, especially the computational ones dealing with issues revolving around 'practical reasoning in rational agenthood' [9] have switched to pictorial, visual, graphical and iconic modes of representation in order to capture different notions of *qualities* associated with assertions and objects [21].
- A related tendency is to dispense with the traditional notion of *logical constants* and substitute it with an assortment of iconic, topological and similar *Gestaltpsychologische* notions in *dynamic theories of action and experimentation* concerning the *relations* involved in the representations (e.g. reasoning about space or conceptual graphs in AI). Similar dissociations are manifest in tendencies to dismiss the division between logical/non-logical constants as any good logical counterpart to the analytic/synthetic division.

Largely as a result of the recent investment in the foundations of computational sciences, the early European cognitive theories are thus, however unintentionally, beginning to be discernible in the concepts of *embodied and enactive minds* [25], *game-theoretic and open-systems approaches* to verification, refinement and composition of concurrent processes and programs [4], *reactive and embedded* computational systems in robotics and AI [15], *interactive, emergent* and *synthetic notions of meaning* entertained in cognitive semantics [14], and in *evolutionary approaches to linguistic meaning, semantics and pragmatics* [23].

This revival is, I believe, due to two principal factors: the decline of the mainstream *logico-formal* and *analytic paradigm* that dominated the better part of the twentieth-century philosophy, and the revitalisation of the *pragmatist stance* in logical philosophy, especially a Peircean one, the origins of which coincide with the origins of the early European contributors to cognitive science. Moreover, the realisation of these paradigms could not have happened without a range of new tools and techniques developed in logic and computation relatively recently.

3.2 Phenomenology and Implicit Knowledge

Finally, let us consider the final question concerning the passage from logic to phenomenology. Husserl's concept of *noema* [7, 12] has predominantly been interpreted transformationally, according to which *unconscious inferences* turn *sense data* into *perceptions*. Alternatively, it has been argued that in noema, mental activity plays a vital role in determining what the object types are that agents intentionally choose among the alternatives presented to the mind [2].

We may think this as mental activity selecting among the *categorial objects* those that are to become objects of the conscious or aware mind. This provides an alternative interpretation for awareness filters formalised in logics of knowledge, in other words filters do not merely turn some raw data into interpreted propositions, but select among all possible propositions those of which the agent is aware.

Representational content is, under this view, a complex of precepts via which other perspectives are synthesised and delineated. Accordingly, one is tempted to think of there being a correlation between, on the one hand, *implicit and explicit aspects of the representational content* to which a subject's mind is attuned to, and on the other, the *character of the kind of mental activity* that determines which object types are selected to become conscious, observational and articulated knowledge.

We note how this phenomenological ramification resonates with the view of consciousness in Peirce's *pragmatist philosophy* and his phenomenology as *phaneroscopy*, developed slightly earlier than Husserl's:

> Thus, all knowledge comes to us by observation, part of it forced upon us from without from Nature's mind and part coming from the depths of that inward aspect of mind, which we egoistically call *ours*; though in truth it is we who float upon its surface and belong to it more than it belongs to us. Nor can we affirm that the inwardly seen mind is altogether independent of the outward mind which is its Creator. (*Collected Papers of C.S. Peirce*, 7.558, c.1893.)

Observations, however implicit or explicit, are both *inward* and *outward*. It would be tempting to learn what light cognitive neuroscience can throw on *that* division.

4 Conclusions

Some initial, overlapping vocabularies in the interfaces of logic, cognitive neuroscience and phenomenology were delineated. Precisely how strong these tentative connections will be, or how concrete the practical relevance of them (besides the philosophy of informatics) turns out to be, must be left for future occasions to decide.

References

1. Albertazzi, L. 2001a. Back to the origins, in [3, pp. 1–27].
2. Albertazzi, L. 2001b. The primitives of representation, in [3, pp. 29–60].
3. Albertazzi, L. (ed.), 2001c. *The Dawn of Cognitive Science: Early European Contributors*, Dordrecht: Kluwer.
4. de Alfaro, L. 2003. Game models for open systems, *Proceedings of the International Symposium on the Verification Theory in Practice*, Springer.
5. Churchland, P.M. 1993. *Matter and Consciousness*, Cambridge, MA: MIT Press.
6. Dienes, Z., Perner, J. 1999. A theory of implicit and explicit knowledge, *Behavioural and Brain Sciences* 22, 735–755.
7. Dreyfus, H.L. (ed.), 1982. *Husserl, Intentionality and Cognitive Science*, Cambridge, MA: MIT Press.
8. Fagin, R., Halpern, J.Y. 1988. Belief, awareness, and limited reasoning, *Artificial Intelligence* 34, 39–76.
9. Gabbay, D., Johnson, R.H., Ohlbach, H.J., Woods, J. (eds), 2002. *Handbook of the Logic of Argument and Inference*, Amsterdam: North Holland.
10. Hintikka, J. 1962. *Knowledge and Belief*, Itchaca: Cornell University Press.
11. Hintikka, J. 1989. The Cartesian cogito, epistemic logic and neuroscience: some surprising interrelations, *Synthese* 83, 133–157.
12. Husserl, E. 1900–1/1970, *Logical Investigations*, London: Routledge.
13. Husserl, E. 1913/1982. *Ideas Pertaining to Pure Phenomenology and to a Phenomenological Philosophy* 1, Dordrecht: Kluwer.
14. Lakoff, G., Johnson, M. 1999. *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*, New York: Basic Books.
15. Manna, Z., Pnueli, A. 1993. Models for reactivity, *Acta Informatica* 30, 609–678.
16. Pietarinen, A.-V. 2002. Knowledge constructions for AI, Hacid, M.-S., et al., (eds), *Foundations of Intelligent Systems*, Springer, 303–311.
17. Pietarinen, A.-V. 2003. What do epistemic logic and cognitive science have to do with each other?, *Cognitive Systems Research* 4, 169–190.
18. Pietarinen, A.-V. 2004a. Diagrammatic logic and game-playing, Malcolm, G. (ed.), *Multidisciplinary Studies of Visual Representations and Interpretations*, Elsevier.
19. Pietarinen, A.-V. 2004b. Semantic games in logic and epistemology, Gabbay, D. et al., (eds), *Logic, Epistemology and the Unity of Science*, Kluwer.
20. Pietarinen, A.-V. 2004c. Towards cognitive informatics: awareness in the brain, logic and cognitive neuroscience, Wang, Y. (ed.), *Cognitive Informatics: Exploring the Natural Intelligence*, World Scientific.

21. Pietarinen, A.-V. 2004d. Peirce's magic lantern: moving pictures of thought, *Transactions of the Charles S. Peirce Society: A Quarterly Journal of American Philosophy*.
22. Pietarinen, A.-V. 2004e. Multi-agent systems and game theory: a Peircean manifesto, *International Journal of General Systems*.
23. Pietarinen, A.-V. 2004f. The evolution of semantics and language games for meaning, *Interaction Studies*.
24. Rantala, V. 1982. Impossible worlds and logical omniscience, *Acta Philosophica Fennica* 35, 106–115.
25. Varela, F.J., Thompson, E., Rosch, E. 1991. *The Embodied Mind*, Cambridge, MA: MIT Press.
26. Wansing, H. 1990. A general possible worlds framework for reasoning about knowledge and belief, *Studia Logica* 49, 523–539.
27. Weiskrantz, L. 1997. *Consciousness Lost and Found: A Neuropsychological Exploration*, Oxford: Oxford University Press.
28. Young, A.W. 1994. Neuropsychology of awareness, in Revonsuo, A., Kamppinen, M. (eds), *Consciousness in Philosophy and Cognitive Neuroscience*, Hillsdale: Lawrence Erlbaum, 173–203.