for representing the intelligible order (since for us, nonconceptual sensory impressions are caught up in the inferential web of language). The second and more important point is that the causal interaction of the automaton with the environment, as a protosemantic navigation of the world, exhibits a noncategorial orderliness that our concepts in their inferential relationships reflect and illustrate, disambiguate, and make explicit. In other words, the causal-heuristic interaction of the automaton instantiates precisely the structures from which our semantic structures have (in part) evolved.

For these reasons, if handled correctly, this analogical circle is a virtuous rather than a vicious circle. It describes a species of 'as-if' (*als ob*) arguments that play the role of *regulative* theoretical and practical judgements. We cannot nonlinguistically see how the automaton sees the world nonlinguistically, but will treat the automaton's view *as if*—under appropriate approximations—we could view it linguistically. Thus, through a series of controlled analogies, the circle enables us to construct, step by step, from an intelligence that is neither discursive nor apperceptive, an intelligence that is no longer analogically posited because it has the form of a full-blooded discursive apperceptive intelligence. What it takes to analogically posit a nonconceptual awareness is exactly what it takes to elaborate this nonconceptual awareness into a conceptual awareness; and what must be added to the analogically posited awareness in order for it to be no longer analogical—to move from nonconceptual awareness to conceptual awareness—is exactly the same as what is needed to develop the non-apperceptive intelligence into an apperceptive intelligence, consciousness into self-consciousness.

Taking into account these considerations, this is the purpose of our story: to cautiously exploit the conceptual resources of our natural language to outline the nonconceptual awareness of this automaton, without imputing to it something like an awareness of the content of its experience, an awareness of its awareness as its own, an ability to use concepts, or any mastery of language; to locate this inchoate awareness or preconceptual form of intelligence as the minimum condition for the realization of a temporally discursive apperceptive intelligence; and to specify what is required in order for this analogically posited heuristic intelligence to be elaborated into a general intelligence endowed with a generative cognitive complexity that is no longer analogically posited, whose abilities are not merely heuristic, but whose heuristic abilities are fundamentally caught up in its non-heuristic linguistically-enabled cognitive capacities. By pursuing these objectives, we will outline what we might call a Kantian-Hegelian program for the construction of general intelligence.

Rather than reducing the discursive apperceptive subject to an awareness at the level of causal structure, we attempt to reconstruct rational agency from a naturalized 'causally reducible' agent, or, in other words, to reconstruct sapient general intelligence from sentient intelligence. In contrast to a reductionist approach, this reconstruction does not overextend the naturalized account of sentience to sapience by virtue of the global reducibility of the latter to the former. The former is necessary but not sufficient for the realization of the latter and its cognitive-practical *abilities*. Naturalization is supposed to be a two-way street between causes and reasons, but all too often it presents itself as a policed one-way traffic from reasons to causes, with no time, intellectual budget for, or interest in anything that moves in the other direction. Here I emphasize 'abilities' because, following the argument of the previous chapter, the self-consciousness of discursive apperceptive intelligence should be understood as a generative framework of theoretical-practical abilities, and not simply as an introspective or amplified form of consciousness.

Knowledge is not awareness or consciousness, and self-consciousness is neither awareness of the phenomenal self nor knowledge of the empirical self, but a consciousness that issues from and is licensed by the powers of understanding and reason—concepts and judgements—which are enmeshed in intersubjective, formal, and inferential linguistic activities. The self of self-consciousness is not the self of phenomenal reality, but a self whose selfhood is a transcendental dimension necessary for judging and being judged; it is essentially like a program constructible under its logical autonomy. This formally and socially instantiated self is the locus of claiming rational authority and responsibility for what is being said and done. Put differently, the self of self-consciousness is a being of the concept not only in the sense that it falls under the concept but also in the sense that its acts issue from and exhibit the concept. For this reason, it is necessary to differentiate the concept of self-consciousness as used by Kant and Hegel from the ordinary intuitive concept of self-consciousness, which can signify either phenomenal introspection or empirical awareness.

One objection to the virtuous circle I have described is given by a figure I will call the greedy sceptic—one who wants to have the cake of semantic apocalypse, but who also eats it in order to fuel an all-out assault on the conceptual structure of thinking.[106] Broadly speaking, the greedy sceptic is someone akin to an exaggerated version of Plato's Meno, introduced as a student of the sophist Gorgias, the master of eristic arguments on the nonexistent. The greedy sceptic assertively claims that we do not know anything and we will never know anything, while at the same time confidently laying out a lavish theory of what he takes to be the case. In short, the greedy sceptic is someone who casually slips in and out of his complicated relationship with knowledge as he pleases. The greedy sceptic may valorize neuroscience to remind us that what we call knowledge is just a figment of our blind brain, or that our talk of the a priori and semantic content refers only to neurological hallucinations or pedestrian metaphysical fantasies. He may believe in science as that which discredits not only our concepts but conceptualization in general, but will at the same time dismiss mathematics, its epistemological status, and the a priori application of mathematical concepts in the sciences as another pseudo-semantic absurdity. In our case, the greedy sceptic of mind is the one who believes there is not and will never be a knowledge of human mind or intelligence, but who nonetheless goes on to put forward a theory of superintelligence or disconnected posthuman intelligence, or a panpsychist account of mind. In other words, the greedy sceptic is not committed to the labour of the virtuous circle of analogy in so far as he already *knows* what intelligence is,

106 For a defence of the Blind Brain theory and the Semantic Apocalypse thesis, see S. Bakker, *The Last Magic Show: A Blind Brain Theory of the Appearance of Consciousness* (2012), <https://www.academia.edu/1502945/The_Last_Magic_Show_A_Blind_Brain_Theory_of_the_Appearance_of_Consciousness>.

despite the fact that he has invalidated the conceptions of knowledge and mind as the guarantor for the intelligibility of intelligence.

The strategy of the greedy sceptic is to characterize discursive self-consciousness as merely a reflexive or epiphenomenal model of phenomenal consciousness. Once a naturalized account of phenomenal consciousness is provided—and such a naturalistic account can indeed be provided, for mere consciousness is nothing but a natural phenomenon—then by virtue of the above characterization of self-consciousness as belonging to the same genus as that of phenomenal consciousness, the greedy sceptic can conveniently characterize the normative-intentional vocabularies of the former as mere by-products of a-rational processes that conditioned the latter. It is not that rational consciousness is not conditioned by nonrational processes (a naturalistic thesis to which this book is fully committed), but being conditioned by nonrational processes is not the same as being constituted or piloted by them.

'If the trick of consciousness can be performed by a vast system of unconscious modular processes,' the greedy sceptic claims, 'then by virtue of being a genus of phenomenal consciousness, discursive apperceptive awareness can also be executed by implementing the same processes.' 'And to the extent that phenomenal consciousness is blind to the activity of these modular processes that occasion it,' the sceptic continues, 'self-consciousness is ultimately but a special—that is, more illusory—kind of blindness, another parochial heuristic device among others, but one with the illusion of being somewhat special.' It is important to note that the greedy sceptic is not really a thoroughgoing reductionist per se, but someone who is bent on the trivialization of the semantic, formal, or a priori dimensions of cognition. But the greedy sceptic describes the blind brain or unconsciousness qua subpersonal information-processing modules in terms of vocabularies and formal relations—for how can one describe anything other than by resorting to formal and semantically-loaded vocabularies of thought? There is no description—whether scientific or philosophical—that can be provided without formal and semantic aspects of thought or theory. The greedy sceptic is analogous to someone who eats his own cake of semantic apocalypse while—to the ridicule of others—not realizing that the cake is actually made of semantic ingredients. Thus the greedy sceptic advocate of the blind brain ends up exposing himself as being doubly blind according to his own standards. And even if he admits to his performative contradiction—'I, the emperor am naked'—this does not mean that there is no real contradiction, that he, as a matter of fact, is not really naked. The telltale acrobatics of the sophist begin with the admission that 'by the way, I am a sophist, so by virtue of this admission, I am not a sophist'. Plato's strike on the Eleatic stranger can be revived here: the admission of sophistry or performative contradiction does not let the sophist off the hook, for the sophist does not really know what makes him a sophist, and thus, by virtue of his ignorance—his semantic naivety—he is still a sophist. The learned sophism of the greedy sceptic makes him akin to those Lorenz Puntel describes as 'hikers who follow their paths while maintaining that all of us, themselves included, are blind, and that there are no paths'.[107]

Contrary to this greedy sceptical approach, the claim here is that self-consciousness is not merely causally structured consciousness. What makes it apperceptive in the last instance is not to be found within the causal structure of nonconceptual awareness, or, for that matter, in the unconscious modular processes that condition it. Discursive apperceptive awareness involves a specific type of cognition: it operates via normative judgements that are both *linguistically oriented* and *linguistically driven*. These judgements supervene upon heuristic abilities, which in turn are caught up in and conceptually utilized and refined by them. The point is that normative judgements are not heuristics; nor do they essentially need to be dependent on some heuristic core, behavioural regularity, or even causal connection with the world in order to count as judgements. A glance at today's logic and theoretical computer science—if not the ever suspect organon of philosophy—should set one straight with regard to the issue of mistaking judgement for the heuristic.

In globally reducing sapience to the empirical facets of phenomenal consciousness and the modular processes that condition it, the greedy sceptic elides the distinctions between conceptual awareness and phenomenal

107 Puntel, *Structure and Being*, 64.

awareness, judgments and heuristics, the structure of discursive intentionality and the structure of proto-intentionality, the space of reasons and the space of causes. This greedy figure will always tell us that the overlap between the space of reasons and causes is too fuzzy to be considered as a criterion of sapience, but such a critique is predicated upon what the fuzziness actually is. Without a determination of the degree of fuzziness, the greedy sceptic merely resorts to the path of least resistance, propping up the fuzziness as an unintelligible principle—saying that the state of affairs is 'more complex than you think', without ever elaborating on what is meant by complexity—in order to perpetually advance his uncritical view. The sceptic's global deflationary approach to discursive self-consciousness is a side effect of his inflationary account of empirical-phenomenological consciousness, within which lies a colossal stack of muddled, dusty, and unchecked metaphysical assumptions.

This illicit merger can be traced back to a principal theoretical lacuna in understanding the linguistic sociality of discursive apperceptive intelligence, how this linguistic sociality determines semantic content, and how the semantic contentfulness of apperceptive intentionality differs qualitatively from the proto-intentionality of the merely phenomenal consciousness whose description at both empirical and phenomenological levels it makes possible. The global reduction of discursive apperceptive awareness can ultimately be traced back to a misconception of what language really is—what it does, and how it does it. Linguistic interaction is the fabric of thought, rather than a means for its transmission, a mere medium of communication.

The emphasis on the essentially linguistic character of sapience is nothing new, of course. It is a familiar topic that is perhaps even deemed well worn and antiquated in some philosophical circles. But what has been for the most part absent in pro-linguistic arguments is a concern with the fundamental interactive structure of language itself: not merely an interaction steeped *in* language, but an interaction which itself *is* language. This is a picture of language not as a symbolic totality with an ineffable essence—a medium of communication, a tool for labelling items in the world, or a system of social discourse in which language is the facilitator of mutual recognition—but as a sui generis framework of interaction-as-computation furnished with different classes of complexity and the different cognitive abilities associated with them.

It is in this sense that the reconstruction of apperceptive intelligence from heuristic intelligence becomes an inquiry into what is required in order for this sui generis interactive computational framework to be constructed—a process of construction that brings about an intelligence of a different type. But to embark on this journey, we must first begin the story of our hypothetical automaton as a preconceptual pre-apperceptive intelligence.

THE AUTOMATON'S STORY, A THOUGHT EXPERIMENT

> STORY 1: The automaton is sentient of the environment: A heap of black accompanied by a low monotonous noise. Suddenly, a faint rustling noise perturbs the scene. A short while later, as the sound becomes louder, a mass of protruding fuzzy grey appears at the opposite end of the heap of black and moves toward it. As the mass of grey reaches the heap of black, it stops, makes contact with the heap of black, moves away, makes contact with it and then moves away, again and again. The perturbing sound has shifted to a shrill and high-pitched sound. Later, the mass of grey recedes from the heap of black. The sound becomes a rustling noise and continues for a while after the mass of grey completely disappears. All that is left on the scene is a rectangular heap of black accompanied by a low monotonous noise.

> STORY 2: The automaton has just seen a reenactment of the monolith scene in *2001: A Space Odyssey*. A featureless black monolith is erected in a landscape whose serenity is broken by the steps of an approaching monkey. Once the monkey finds the monolith, it starts examining it, screams, panics, runs away, and comes back to touch it again. The monkey continues this game for a while until it finally gets bored and goes away.

The world of the automaton is uncannily analogous to the prelinguistic world-picture of an infant. The first narration is a rough analogical linguistic story about how the automaton registers the environment, the second a non-analogical linguistic narration—essentially a story of and within language—in which such registers are inferentially caught up. More specifically, the items in the second story are linguistic items that possess semantic features by virtue of the inferential roles they play in our collective linguistic practices, whereas the items in the first story are nonconceptual representations of items in the environment, i.e., registers at the level of internal variable states of the automaton's wiring structure which causally mediates between sensory inputs and behavioural outputs. The automaton is aware *of* the items in the environment (a heap of black, a rustling), but in the second story we are aware of the items *as* something (*as* a monolith, *as* the steps of an approaching monkey). We perceive or *take* things *as* such and such, whereas the nonconceptual content of the automaton's experience is neither seeing or hearing *as*... nor seeing or hearing *that*..., but seeing and hearing *of*.... What it sees *of* the black monolith is the colour on the facing part of its surface and its shape on the facing side (let's use the term seeing$_1$ to designate this seeing-of). But what we see of the monolith is not merely the impoverished seeing$_1$, for what we *see*$_2$ *of* the object is its colour through and through and its shape all around; we also see it *as* a black rectangular cuboid, a black monolith. Kant would describe the former as an image, qua function of the imagination, whereas the latter proceeds via the subsumption of this image, qua singular representation, under a concept through schemata which serve as universal procedures for constructing a model of that concept (monolith) out of multiple and associable, synchronically and diachronically synthesized perspectival aspects of an item in the world.

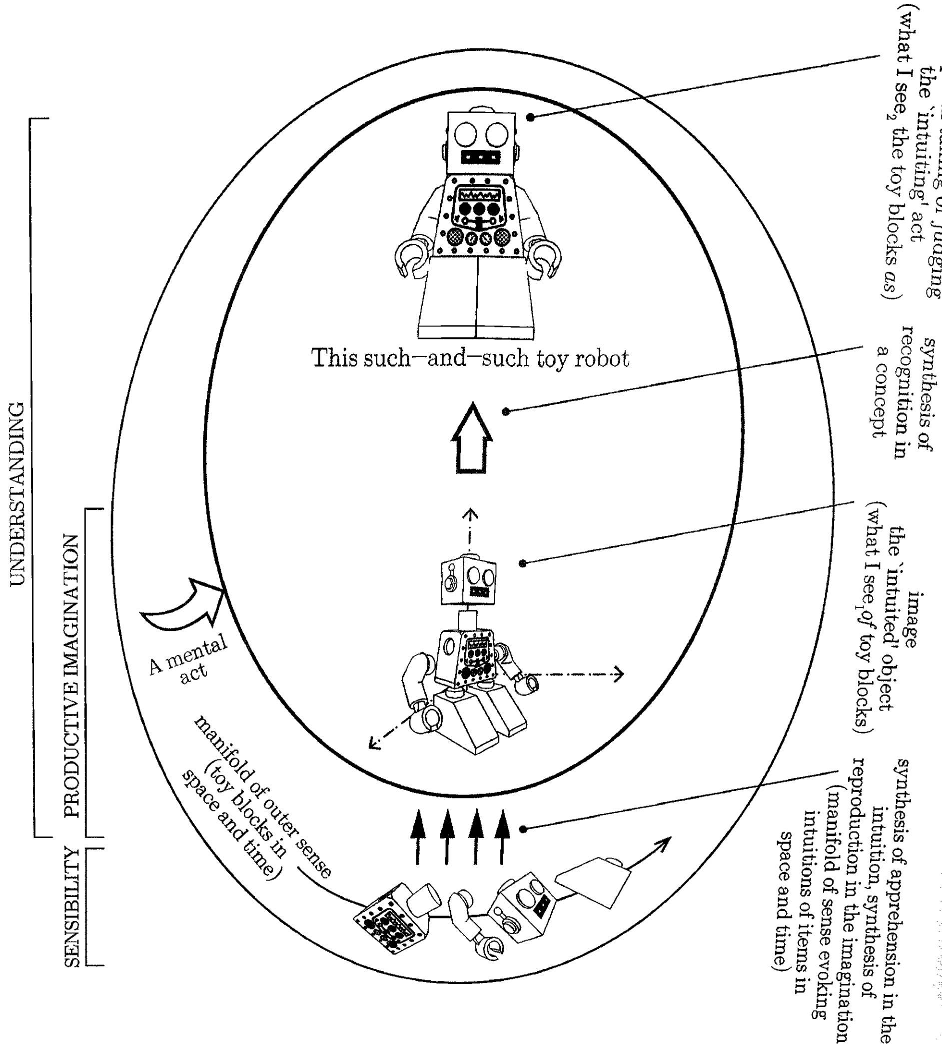
We can think of 'seeing$_1$ *of*' in the first story and 'seeing$_2$ *as*' in the second story in terms of making a Lego model—say, a toy robot—using Lego blocks of different shapes and colours. From a shifting *perspectival* point of view, the blocks with their various shapes and colours correspond to the diverse yet rudimentary images of our Lego model (the point-of-viewish *intuiteds*) which are obtained as our automaton confronts the Lego blocks in such and such perspectival imagings (e.g., this Lego block or heap of black is facing the automaton edgewise). Their role in our model construction is particular and contingent. The shapes and colours of the blocks are the raw 'sense-given matter' of the intuited items qua images. The pictorial motif of our Lego model-building corresponds to the conceptual representation of these intuited items in acts of judgements (the intuitings). The function of the pictorial motif is to determine the colours and shapes of the blocks in such a way that it becomes possible to put them together so as to construct the specific Lego model in question. In other words, the pictorial motif encapsulates the function of the concept of *a robot* that *determines* images (the right blocks) as different aspects of only that object. It is only because the colours and shapes of these blocks hang together in the right way—perspectivally in space and time, synchronically and diachronically—that we are able to synthesize the pictorial motif of our Lego model-building. And respectively, it is only because the blocks (the images) can be put together in the right way—in accordance with a rule, i.e., the concept of a robot—that we are able to *conceive* them *as* associable and multiple aspects of *one* such-and-such robot.

In seeing$_1$, we are dealing with local variations and rudimentary—point-of-viewish—invariant aspects of particular items. Whereas in seeing$_2$, we have the function of productive imagination by means of which categories or pure concepts of understanding—general and universal invariances—are applied to the intuited items, like blueprints or instructions necessary for the construction of a toy object. But in seeing$_2$ we are also capable of bringing and constructing images under specific concepts (e.g., this such-and-such toy robot). The combination of these abilities turns seeing$_2$ into a complex act of imagination and understanding. In more contemporary terms, seeing$_2$ involves object individuation and object simulation, general and specific forms of classification, and reidentification of local invariants across different contexts. Units of experience are perceptual-takings or the class of -ings-as (e.g., grasping or conceiving x as ...), not the class of -ings-of. Local variations (manifolds) of items (shape, colour, etc.) supplied by the latter (sensory and perspectival) class are not by themselves sufficient for object-construction since object (*gegenstand*) individuation, classification, and recognition involve the construction of types of invariances which form a network of associations and implications. Such invariances cannot be obtained without a complex interplay between the sensory intuitions given in imagination and the pure concepts of understanding—overseen by what Kant would call the faculty of judgement.[108] Categories are logical functions. It is only when they are brought into conformity with sensory intuition through schematism and synthetic a priori principles that they can become rules for generating perceptual-takings—that is, rules for constructing something as an object of representations. Schemata can be said to be rules of construction pertaining to the process of providing a concept with its singular representation qua image (*Bild*), e.g., triangularity and a triangle. On the other hand, synthetic a priori principles can be understood as *general* rules of unity in the integration or synthesis of appearances.

CONSTRUCTING A LEGO MODEL IN THE STYLE OF KANT'S THREEFOLD SYNTHESIS

The Lego model building process—which is necessary for the transition from seeing$_1$ to seeing$_2$—as a whole corresponds to Kant's threefold synthesis: namely, synthesis of apprehension in the intuition, synthesis of reproduction in the imagination, and synthesis of recognition in the concept. The synthesis of apprehension delineates the first constructive role of imagination in pulling together a synchronic manifold of sensations by antecedently taking up the sense impressions into its activity, apprehension.

108 'In short, we do not perceive of the object what might be called "categorial" features. For the image construct does not have categorial features. It has an empirical structure which we can specify by using words which stand for perceptible qualities and relations. But it does not have logical structure; notness, or-ness, allness, some-ness are not features of the image-model. They are features of judgment. More generally we can say that the image-model does not have grammatical structure. (It will be remembered that we are construing mental judgments as analogous to sentences. A judgment, we said, is, as it were, a Mentalese sentence episode.) And, of course, Kant's categories are grammatical classifications. They classify the grammatical structures and functions of Mentalese.' W. Sellars, *In the Space of Reasons: Selected Essays of Wilfrid Sellars* (Cambridge, MA: Harvard University Press, 2007), 463.



Kant's threefold synthesis: What sets apart Kant's account of experience and mindedness from that of Hume is the emphasis on a *multilevel construction* rather than *basic empirical associations*. From a computational perspective, it is the mode of integration or synthesis of different processes or algorithms that distinguishes mental acts and critical faculties from mere quantitative algorithmically realizable abilities. Indeed, Kant's account of the threefold synthesis is not merely a psychological argument whereby the summary list of mental abilities are given. It is, more fundamentally, an epistemological argument concerning the analysis of specific modes of cognition required for objective knowledge and critical judgment.

It introduces order into the confusion of simultaneous impressions by giving them temporal and spatial locations, and thus differentiating them. In doing so, the synthesis of apprehension brings about the condition of the intelligibility of impressions as *distinct* (spatiotemporally structured) impressions qua appearances available for further construction and structuring. The second synthesis, the synthesis of reproduction, signifies the second constructive role of imagination in combining and reproducing the sensory manifold diachronically, carrying over its earlier elements in order to construct a *stable* image qua *singular* representation of an item in the world. It establishes temporal associations between appearances that the synthesis of apprehension has located in space and time in a certain way, rudimentarily structured out of the undifferentiated homogeneity of simultaneous impressions.

These two syntheses are the figurative part of the building process associated with imagination as a constructive-simulating capacity whose function is to 'represent an object even without its presence in intuition'[109] and which is unavailable to pure sensibility. They are, accordingly, what we might call figurative syntheses. The third synthesis, the synthesis of recognition, strictly designates the role of apperceptive consciousness in perception—that which must be 'added to pure imagination in order to make its function intellectual', since 'in itself the synthesis of imagination, although exercised a priori, is nevertheless always sensible, for it combines the manifold only as it *appears* in intuition.'[110] The synthesis of recognition requires both the act of recognizing a past representation as related to the present one and the act of recognizing past and present representations as belonging to one object via the function of a concept. The third synthesis then involves different a priori acts of cognition (*erkenntnis*) which produce full-blooded judgements.

To avoid terminological confusion, it is best to provide brief definitions for these recurring elements of the Kantian vocabulary: *Sensations* are the 'immediate' results of the mind being causally affected by objects. In other

---

109 Kant, *Critique of Pure Reason*, 256.

110 Ibid., 240.

words, a sensation is the effect of an object on our representational capacities to the extent that we are affected by the sensible object or *gegenstand*. *Sensibility* refers to the capacity or receptivity to acquire representations through being affected by objects qua particular or individual items. The ability to have thoughts of individual items is *intuition*. And that through which cognition relates to objects (individual items) through 'sensations' is *empirical intuition*. The generic and indeterminate object of empirical intuition is an *appearance*. The 'matter' of empirical intuitions is sensation, and its form is sensibility, which allows the manifold of appearances to be initially structured. Appearances are the elements that are picked up and built into singular representations of items in space and time qua intuitions. Finally, intuitions differ from sensations since they are types of cognition while sensations are not. The raw content of intuitions is what is *intuited* from sense-given materials (sense impressions of an item located in space and time). The form of intuitions is intuition as an act of *intuiting* which is a singular act of *conceptual* representing.

The syntheses essential for the transition from seeing$_1$ to seeing$_2$ are dynamic acts of integration and need not to be thought in either purely top-down or purely bottom-up fashion, but rather in the mesoscopic or mid-level fashion introduced in the previous chapter in the context of big toy models. As briefly introduced in our discussion of Chu spaces and toy models, syntheses can instead be understood as matrices of interactions between the integrating-organizing acts belonging to causal-empirical and logical-inferential domains and the various rules of integrative transition (moving from one level of unity to another) that must necessarily obtain between them in order to have anything like seeing as, hearing as, smelling as, i.e., conceptual awareness as knowledge or experiences endowed with (nonprivate) epistemic status. These rules of integrative transition can indeed be defined as necessary computations—typed or untyped, statistical or logico-linguistic, context-sensitive or context-free, centralized or distributed—which can enter into interaction with one another asynchronously or synchronously.

In this sense, transcendental logic, as that which supplies concepts with sensory intuition and applies classificatory concepts to intuitions, can be understood as a computational search space. As such, it is perhaps more fruitful to elaborate transcendental logic by way of information theory and computational complexity theory where each level of necessary conditions for the possibility of having mind can be construed in terms of the increase in information processing abilities and the new types of computational problems (Turing-completeness, NP-completeness and NP-hardness, etc.) that can be solved at that specific level. Moreover, this computational view of transcendental logic coincides with the paradigm of deep functional analysis presented in the first chapter, or what the late Hilary Putnam dubbed 'liberal functionalism'—a functionalist view of the mind as a collection of world-structuring abilities which require an anti-individualistic picture. A organism is only a system insofar as it is in realtime transactions with the environment. The functionalist view of such a system cannot be greedily reductionist, because its information processing abilities 'seek their own level of interpretation'.[111]

A DIGRESSION ON MODELLING FIGURATIVE SYNTHESES

Equipping our automaton with objective—in the sense of *Gegenständlichkeit* rather than *Objektität*—figurative syntheses, then, would require the application of a mixture of mid-level models similar to the generative models utilized in the predictive processing (PP) paradigm.[112] Such models are based on prior probability and statistical estimates which function as representations employed to predict current and future sensory input as well as the source of such input. Source detection is possible in so far as the estimates are hierarchically organized in order to track features at

111 H. Putnam and L. Peruzzo, *Mind, Body and World in the Philosophy of Hilary Putnam: Léo Peruzzo in conversation with Putnam* (2015), *Trans/Form/Ação* 38:2 (2015), <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0101-31732015000200211>.

112 On schematism and predictive processing see L.R. Swanson, 'The Predictive Processing Paradigm Has Roots in Kant', *Frontiers Systems Neuroscience* 10:79 (2016), <https://dx.doi.org/10.3389%2Ffnsys.2016.00079>.

different temporal and spatial scales, and this hierarchical distribution in turn enables estimates at different levels to be predictive of one another.

These estimates, however, must be probabilistically constrained, otherwise predictions would be impossible for any sensory-neural condition. Without constraints as the ground for predictions and likelihood estimates, our automaton or intelligent system would not be able to winnow the sensory-neural possibilities and converge on a set of predictive hypotheses. Such probabilistic constraints are inbuilt inductive biases which are necessary for any form of predictive processing system and are defined in terms of probability priors. The hierarchical system of priors—moving from more fundamental or abstract to less—enables the development of advanced representational systems which not only effectively single out hypotheses from a set of possible hypotheses but also handle different levels or types of hypotheses in order to explain the data. This hierarchical architecture permits differentiation between basic representations of, for example, a round Lego block and a cubic one, since the prior probability that these two blocks are colocalized in space and time is negligibly small.[113]

Deep entrenchment of more fundamental priors or constraints canalize and guide upper-level estimates and less abstract priors. Such *priors upon priors* are called hyperpriors. Andy Clark and Link Swanson have identified hyperpriors with the brute constraints imposed by space and time, explicitly as Kantian forms of intuition and appearance. These constraints can range from hard restrictions on spatiotemporal bilocalization or colocalization of sensible items to limitations of bodily actions (e.g., either turning left or right). In Swanson's words:

> Abstract internal knowledge of space and time—spatial and temporal hyperpriors—are thought to narrow and restrict large swaths of possible hypothesis spaces, thereby aiding the formation of decisive perceptual predictions regarding the external objects causing incoming stimuli. This narrowing of possible hypotheses is critical to the entire probabilistic

113 See A. Clark, 'Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science', *Behavioral and Brain Sciences* 36 (2013), 181–253.

> inference process—without it the required Bayesian computations become intractable. Spatial and temporal hyperpriors can thus be usefully conceived of as necessary conditions on the possibility of probabilistic perceptions of external objects. [...] Kant's proposal that space and time are features of cognition that form, constrain and restrict possible perceptions of outer objects is echoed in explanations of the role of hyperpriors in PP accounts of perception. Without spatial and temporal hyperpriors, the objects of perception that putatively result from PP would be impossible.[114]

In predictive processing systems, incoming sensory inputs are not passively received but are contrasted with the existing representational repertoire, i.e., they inform representational or image updates. It is precisely this update-function rather than sensory input that results in prediction error minimization in such a way that updates adhere to the norms of Bayesian inference. Moreover, the applications of such predictive models can be made more exact with the help of adequate formalizations provided by Category theory and Topos theory, whose geometric and topological richness have been studied in relation to neural modelling, constructive memory, Kantian schematism, and figurative syntheses (see the works of Ehresmann, Gómez-Ramirez, and Healy).[115]

For instance, the commutative diagrams used to define the concept of colimit can provide a formalized map of the neural paths, compositions, transformations, and categories necessary for the construction of complex image-models out of simpler ones. For example, we can think of a colimit diagram for the process of construction of an obelisk as an image-model

---

114 Swanson, 'The Predictive Processing Paradigm Has Roots in Kant'.

115 A.C. Ehresmann, J.-P. Vanbremeersch, *Memory Evolutive Systems: Hierarchy, Emergence, Cognition* (Amsterdam: Elsevier, 2007); J. Gómez-Ramirez, *A New Foundation for Representation in Cognitive and Brain Science* (Dordrecht: Springer, 2013); and M.J. Healy, 'Colimits in Memory: Category Theory and Neural Systems', in *Proceedings of the International Joint Conference on Neural Networks*, IJCNN '99, vol. 1 (1999), 492–6.

qua mental object embodied in the neural organization using simpler mental images—a square and an oblong rectangle—which are at different (lower) levels of the neural organization. A colimit then can be said to be a category-theoretical diagram for the construction-cum-compression of complex mental images from simpler ones using commutative diagrams and universal properties.[116] Details and associations between the square and the oblong rectangle are compressed or streamlined into the image of a cuboid.

A colimit diagram captures how the neural category *Neur*, which is composed of neurons and the synaptic paths—represented by morphisms—between them at the microscopic $level_1$ can be constructively raised up to the macroscopic $level_n$ of the mental objects category *Ment*, which is composed of mental images or image-models and the morphisms or structural relationships between them. Then $N-1$ levels signify different scales or levels of organization which lie between the microscopic category *Neur* and the macroscopic category *Ment*, moving from the fine-grained scale of local networks of neurons to the coarse-grained scale of neural events as for example recordable by functional MRI.

---

116 In mathematics—particularly category theory—a commutative diagram expresses the generalization of a system of equations and its internal symmetries. Take for instance, $x * y = y * x$ where x and y are operators and $*$ is the operator multiplication. This equation states that x and y operators commute with respect to $*$. The commutative diagram can be seen as the category theoretical generalized equivalent of such a statement. As for the universal property, it roughly means that the property of construction on a particular mathematical object can be expressed in terms of its relations to all other objects. One of the main motivations behind the concept of universal property is to forgo the concrete details regarding the construction of a particular object (e.g., a proof) and to instead use an effective account of the construction that is not simply limited to a particular object but rather concerns the construction of the object in terms of its relations (i.e., associated morphisms) with other objects. In this respect, the universal property is analogous—with some caveats regarding the distinction between sensible objects of experience and mathematical objects—to Kantian categories or pure concepts of understanding, which are not derived from a particular object or encounter with an item in the world, but instead are a priori universal and necessary concepts by which objects (*Gegenstände*) are structured and ordered.

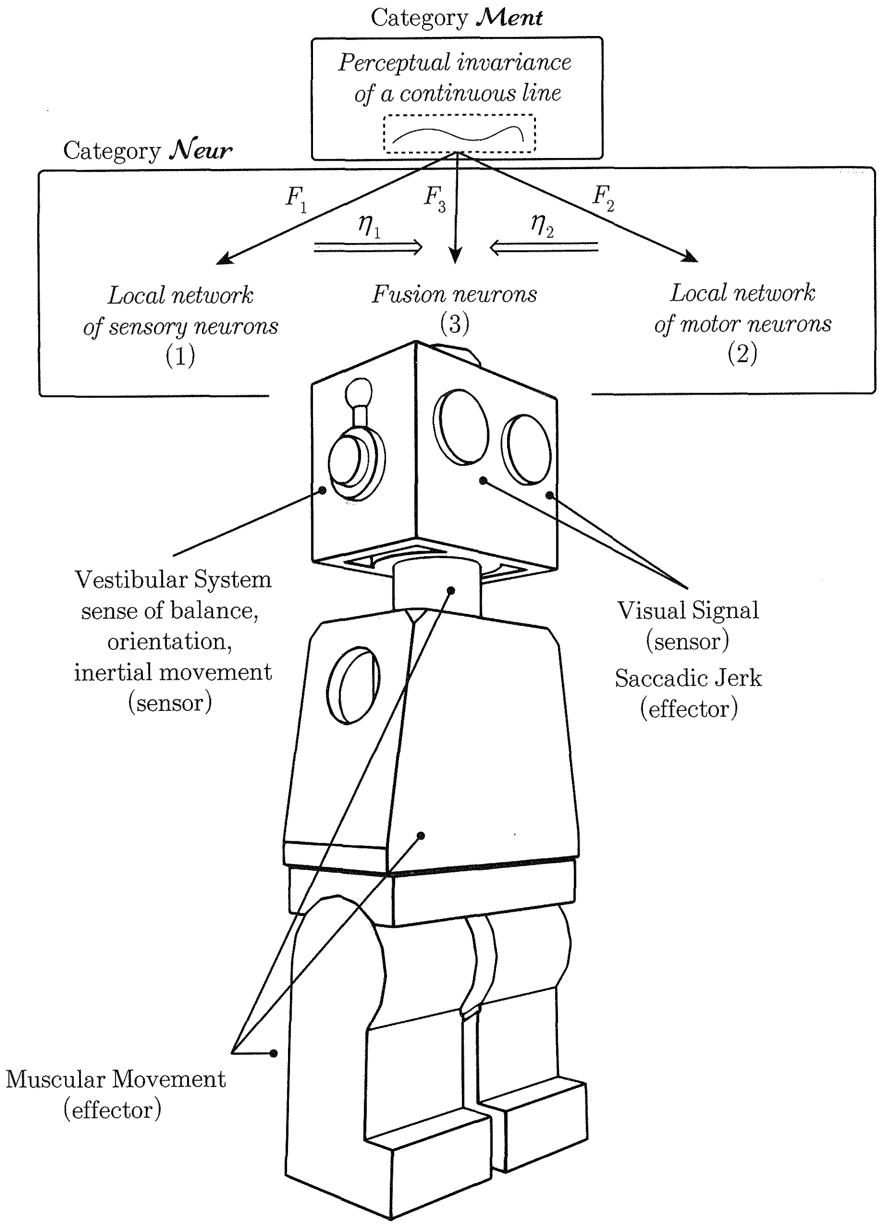
In this respect, in the context of the neural organization and mental images, the colimit can be understood as a map from a commutative diagram in the category *Ment* to a commutative diagram in the category *Neur* via the functor F: *Ment* → *Neur*, while the hierarchy of *Ment*—levels of image-models shifting from simple to complex—can be represented via natural transformations η_n between given functors (η_1: $F_1 \Rightarrow F_2$, η_1: $F_2 \Rightarrow F_3$).[117] Here, F_1 and F_2 map *Ment* to two local networks of afferent neurons in the category *Neur* which are responsible for processing two distinct sensory inputs (for example, one visual and the other haptic). F_3, on the other hand, maps from *Ment* to a local network of neurons which is a sensor fusion of the aforementioned afferent neurons.

In this configuration, other objects can be introduced to the neural category, such as motor neurons associated with the agent's effectors. A seemingly simple but in fact highly complicated image-model such as the abstract image of a continuous line as the root of the concept of line can be modelled in this manner via a colimit diagram in which natural transformations obtain between functors from *Ment* to *Neur* whose objects are both local sensory neurons (visual signals from the ocular system, sense of gravity, balance and spatial orientation from the vestibular system, etc.) and local networks of motor neurons connected with bodily actions, specifically those involving embodied spatial gestures of direction and orientation (e.g., the saccade of the eyes and muscular movements). The fusion between sensors and effectors, afferent and efferent neurons, would then account for the complex image of a continuous line as rooted in a perceptual invariant generated by the stabilization and integration of the senses of inertial movement, direction, orientation and certain locomotory actions.[118]

---

117 Given categories C and D, and functors $F, G : C \to D$, the natural transformation η from F to G is a family of morphisms. It assigns to every object x in C of a morphism $\eta_x : F(x) \to G(x)$ in D such that for every morphism $f : x \to y$ in C, the condition or structural diagram $\eta_y \circ F(f) = G(f) \circ \eta_x$ is satisfied or commutes in D.

118 On the links between perceptual invariances and the concept of continuous line in mathematics see F. Bailly and G. Longo, *Mathematics and the Natural Sciences:*



Modelling a sensor-effector system—a robot—capable of forming a stable perceptual invariance.