As with all models and paradigms, the theoretical scope and practical applications of the above models should be treated with modesty and critical vigilance. While such models can be regarded as models of choice with a certain theoretical strength and range of application within the big tinkertoy model of AGI or of our own representational systems, they should not by any means be viewed as complete or global and comprehensive. There are in fact significant flaws in both their theoretical underpinnings and their range of application, whether in the context of artificial intelligence or cognitive science. The strong modular constructivism implied by the category-theoretical formalization of neural organization—particularly the colimit diagrams—does not leave any room for modelling the necessary pruning mechanisms which are responsible for repairing or destroying unneeded or deleterious neural patterns. As a result, the complex microscopic view of the brain turns into yet another variation of the well worn metaphor of the brain as a receptacle of whatever is thrown into it without ever having a chance of filtering out or destroying constructed and entrenched patterns in the neural structure. When the role of neural unlearning is removed from the equation or downplayed, the relevance of neural learning to concept learning progressively diminishes.

The rise of category-theoretical models in the study of complex systems, particularly in neural architecture, calls for the development of a pertinent critique of both the conceptual underpinning and the range of application of category theory or any other mathematical formalization. While the generality of category theory makes it a powerful tool for the study of mathematical structures, it is well-suited neither for the study of all types of mathematical structures nor for the modelling of all forms of physical systems and phenomena. The question of structure as that which addresses the relation of the mind to the world is chiefly the question of constructing theories within which *different* models can be compared and tested. Context-independent appraisal and application of mathematical

---

The Physical Singularity of Life (London: Imperial College Press, 2011), 66–8; and A. Berthoz, *The Brain's Sense of Movement* (Cambridge, MA: Harvard University Press, 2006).

models, or any model for that matter, is an absurdity. The context is firstly given by the theoretical system within which the model operates, and then by the analysis of the correspondence between, on the one hand, the type and scale of structures a theory aims to investigate, and, on the other, the specific mathematical formalism indexed by the model for the study of the target structure. Absent this analysis, not only will the application of the mathematical model be too arbitrary, it will also result in a complete distortion of the phenomenon under study by the misapplication of the mathematical structure presupposed in the model. Here category-theoretical models of neural structures are no exception.

Category theory's strong reliance on *commutativity*—as a mathematical structural constraint—or what Ion Baianu has described as 'the stricture specific to Abelian theories, categories or ontologies'—makes category-theoretical models ill-suited for application to a broad range of complex physical systems and phenomena at scales where the distinction between objects and processes begins to fade away.[119] Physical systems or behaviours involving symmetry-breaking, genuine asynchronous processing, and irreversible dynamics cannot be adequately encoded by models enmeshed in a mathematical framework that places strong strictures on mathematical structures such as commutativity or internal symmetry.[120] For example, myriad forms of neural tasks involve true asynchronous processing of events or states which are spatiotemporally separated and have no symmetry

119 I.C. Baianu, R. Brown et al., 'A Category Theory and Higher Dimensional Algebra Approach to Complex Systems Biology, Meta-Systems and Ontological Theory of Levels', *Acta Universitatis Apulensis* 52 (2011), 11–144.

120 In the context of category theory, a mathematical internal symmetry can be understood as a commutativity of a concatenation of morphisms. This property allows any object of an abstract category—which has a unique identity 1_x—to be replaced by its identity morphisms. Internal symmetry as the condition for the concatenation of morphisms can be expressed by a basic example:

$$f : x \to y,\ g : y \to z \Rightarrow h : x \to z,$$

that is, the concatenation of the morphisms f and g is replaced by the unique morphism $h=g \circ f$.

between events and states. It is far from obvious that category-theoretical models can sufficiently encode such aspects of neural organization.

In his work *Simulation and Similarity*,[121] Michael Weisberg offers a detailed specification of models as structuring systems. Regardless of whether the model is descriptive, explanatory, or predictive, or whether it is concrete (e.g., an orrery), mathematical, logical, or computational, the model is comprised of a structure formulated in a specific theoretical framework. The structure has two poles, the model description and the model construals. The description is a set of formulas or equations which describe the range of model applications based on its structure. In short, the model description provides us with information regarding what the model is (its structure) and how it can be applied (its dynamic). But the description is not the structuring kernel of the model itself since the structure contains a diverse range of information regarding the scope, assignment, and fidelity criteria (i.e., constraints which specify the representational, dynamic and resolution or scale-sensitive specifications of the model) which are implicitly encapsulated in the model description. All such information has not only theoretical assumptions, but also metatheoretical (metalogical, metamathematical, ...) assumptions which should be carefully analysed.

Therefore, consideration of this often-ignored dimension of modelling-metatheoretical assumptions is absolutely indispensable not only for the correct application of a model to the appropriate type of data and the scale at which a specific phenomenon is studied, but also for the appropriate choice of the mathematical (or logical and computational) framework and the type of constraints imposed by that framework on the mathematical structures that are supposed to encode a physical phenomenon. Absent this analysis, we can never be sure whether our empirical models of a target system such as the brain distort the data, whether they are sub-optimal models for the study of the target system or are being applied to the wrong structural level of the phenomenon or system in question.

As such, the scope and application of the predictive processing paradigm cannot be overextended. Clark's claim that the predictive processing

121 M. Weisberg, *Simulation and Similarity* (Oxford: Oxford University Press, 2013).

paradigm not only accounts for rudimentary representational systems but also the scientific method is based on a crude view of scientific theories.[122] No predictive-inductive method or model can by itself represent scientific theories, since the construction of scientific theories requires both a ranked plurality of methods and a complex semantic dimension without which no inductive method can converge on a single hypothesis or a set of hypotheses that can be said to be true.[123] Without the semantic dimension of theories, the predictive processing paradigm of scientific theory construction is no more than an extension of our inductive biases. A scientific theory that is solely based on a form of predictive-inductive processing can only offer a version of objective reality that is in full conformity with the local and contingently constituted inductive biases of the subject and its deep-seated intuitions. But the fact that science progressively breaks itself from our common intuitions should be counted as evidence that scientific theories cannot be construed as an overblown version of the predictive brain.

Finally, inductive biases and hyperpriors—particularly as the constraints of space and time—should be understood as constraints on the order of appearances—that is to say, they are features of the experiencing subject or the conscious organism. Identifying them as features of objective reality is an unwarranted metaphysical commitment whose implicit assumption is the existence of a *given* or preestablished harmony or one-to-one correspondence between reality and the perceptual mechanisms of the brain. The defender of such a view would have to explain a number of glaring problems with it, including explaining why modern physics continues to shatter our most cherished inductively-enabled intuitions regarding the universe, or why it took modern *homo sapiens* almost two hundred thousand years to develop Euclidean geometry, and after that almost two thousand years to discover non-Euclidean geometry. If the structure and dynamics of our scientific theories and natural languages could easily be analysed or explained away in terms of inductive biases, then how could we even recognize or describe such inbuilt biases? It would mean we can

122 See Clark, 'Whatever Next?'.

123 For a detailed discussion of this point, see the Appendix.

only resort to our inbuilt inductive biases to justify or explain inductive biases—which is viciously circular, a petitio principii.

A more modest claim would be that inbuilt inductive biases are enabling constraints which have become entrenched through the course of natural evolution. Strictly speaking, inductive biases are local and contingent features of the conscious organism. Some may even be purely psychological features. These inbuilt inductive biases have enabled the development of otherwise more complex representational systems which are not merely inductive, probabilistic or statistical. In this sense, the enabling role of inductive biases does not mean that they have constituted or play a central role in our theoretical activities. In other words, enablement neither means constitution nor calls for continued reliance on the enabling conditions.

Yet even this modest claim should be treated with caution. Following the previous discussions regarding the critique of the transcendental structure, inbuilt inductive biases should be taken as features of the transcendental structure of the experiencing subject and, as such, should be challenged and progressively suspended in a Hegelian manner. First we ought to distinguish psychological and nonpsychological features, then attempt to model variations of nonpsychological features which are not only alternatives to our own, but also have a more expansive traction on objective reality in so far as they enlarge the field of experience and therefore detect features of objective reality which our experience either distorts or cannot come to grips with. This means that, even if we admit that inbuilt inductive biases—such as the enabling spatiotemporal constraints—can be both psychological features and nonpsychological features *specific* to the local and contingently constituted subject of experience, we can never be sure of the extent to which our particular experience distorts objective reality. Here, by objective reality, I also have in mind the objective descriptions of the brain qua representational system.

Even when we strive to be attentive to the distinction between the particular characteristics and features of our experience and the characteristics of objective reality, we might indeed smuggle in bits of the former to the latter. This is a topic that we shall explore in depth in the excursus on Boltzmann and time in chapter 4. To conclude, short of a theoretical and practical critique of the transcendental structures, the necessary link between the intelligible and intelligence cannot be renewed: the intelligible remains within the confines of what intelligence perceives to be an objective description of reality but is in fact a psychological representation, or an account which misrepresents the transcendentally ideal features of experience, or the local characteristics of the subject, as global and necessary characteristics of reality. Consequently, our objective description of ourselves in the world becomes yet another manifest self-portrait and our speculations about a future artificial general intelligence invariably reiterate—like the portrait of Dorian Gray—ever more distorted pictures of ourselves.

SEEING$_1$ WITH THE AUTOMATON

Having summarily examined the experience of the discursive apperceptive intelligence of the items in the world of Story 2, the syntheses it requires, and the candidate models for its figurative syntheses, we can now return to the first story, that of the automaton's nonconceptual encounter with the same items.

In Story 1, the encounter of the automaton with an item in the environment is an analogue of an Aristotelian this-such (*tode ti*: this-something-or-other), namely, the impression of an object as a 'materiate individual substance'. The automaton's this-suches, however, are best linguistically expressed by the application of dummy quantifiers that designate form ('this much of', 'a quantum of', 'a heap of') and mass-terms that designate matter ('stuff', 'blob') to occurrent sensible properties (a shaped and coloured facing side of the object)—as in 'a heap of black', 'a protruding mass of fuzzy grey'. In more straightforward terms, the automaton sees$_1$ an expanse of grey as an object rather than as an object whose surface is grey. For the automaton, grey and black are not 'categorial' kinds of experience or manners of experiencing$_2$ (an object looking *greyly* or *blackly*), but the objects of its de facto experience *tout court*. The automaton is conscious of *a heap of black* as *the object* of its awareness—although at this time the possessive pronoun 'its' is not yet part of the automaton's thoroughly nonconceptual awareness, i.e., the automaton is not yet able to ascribe this awareness to itself as its own. This is another difference between the first and the second story: what is not given in the world of the as yet thoroughly nonconceptual awareness of the automaton is precisely the categorial structure. At this stage, trafficking this categorial structure into the story about the automaton's analogically-posited experience$_1$ (or awareness$_1$) of items in the world would be precisely a function of what Sellars calls 'the myth of the given': 'the idea that the categorial structure of the world—if it has a categorial structure—imposes itself on the mind as a seal imposes an image on melted wax'.[124]

The automaton's sensing is acategorial, since categoriality belongs to the domain of the conceptual in which sensing is caught up. To ascribe categorial structure of any kind to the mere sensing of the automaton would be to relapse into a sceptical empiricism for which the intelligibility of what is sensed is given. But also, by analogically developing the sensing automaton into a thinking automaton, we avoid the dogmatic rationalism for which the relation between thinking and being is *given* not in the world but in the deductive rules of reason.

In order to attempt to faithfully represent the logico-phenomenological *form* of the automaton's 'way of experiencing' and, correspondingly, the features relevant to the automaton's mode of awareness, in a fashion that it is 'consistently and from the beginning identified with [the automaton's] modes of representation',[125] we therefore have not only to avoid using

124 Ibid., 237.

125 'It is "phenomenology" in that its aim is not to characterize a world but rather "ways of experiencing a world," that is, elements, aspects, or features of modes of awareness. It is, however, "logical" phenomenology in that such modes of awareness are consistently and from the beginning identified with modes of representation. What separates such "logical phenomenology" from "pure" (Husserlian) phenomenology, in other words, is the crucial acknowledgement that our only possible conceptual access to the structures constitutive of such modes of awareness (= modes of representation) is through a grasp of the logico-semantic structures instantiated in the inferential interrelationships of elements of our own system(s) of linguistic representations—and that this is so whether the awarenesses themselves are thought of as instantiating an "internalized" natural language or as the

'grey' and 'black' as adjectives ('a grey item') or as adverbs ('an item that is experienced greyly'), but also the kind of mass terms (such as 'rock', 'water', 'flesh') that smuggle in a categorial commitment to causal powers and propensities constitutive of the nature of such stuffs. We should instead use phenomenological this-suches that have a logical form appropriate to the automaton's mode of awareness and somehow convey the suspension of categorial commitments, particularly those pertaining to the causal propensities of items in the environment. This is the reason behind the peculiar language of the first story wherein the 'protruding fuzzy grey' as a whole conveys some stuff that plays the analogical functional role of a 'countable noun' (monkey) relevant to the general form of the automaton's nonconceptual awareness$_1$ (seeing$_1$, hearing$_1$). But in its 'stuffiness', the 'protruding fuzzy grey' is also distant from ordinary mass terms such as 'fur' or 'protoplasm', use of which would imply the automaton's categorial commitment to the nature (causal powers and propensities) of such stuffs. And finally, to demarcate the individuated form of such stuffs, we can use mock-up quantity-specifying terms that are functionally analogous to but removed from the measurement specifications of the terms we ordinarily use to quantify mass terms, such as a 'bucket' of water, a 'piece' of meat, etc. In this fashion, a 'heap of black' is really this-much (expanse) of black (stuff), individuating form and matter.

On top of these distinctions between the first and second stories, there is another less noticeable difference: the inferential connections between

---

thoroughgoing nonconceptual pure positional awareness of pre- or nonlinguistic experiences. Whereas a "pure" (Husserlian) phenomenologist thinks of himself as using his own language "directly," to describe (nonlinguistic) modes of awareness ("ways of experiencing a world") with which he is (somehow directly) acquainted, a "logical phenomenologist" consistently sees himself as using his own language "at one remove," to illustrate (through the inferential relationships) normatively obtaining among its elements) systematic and orderly relationships among aspects or features of modes of representation which may variously be imputed (actually or hypothetically) to beings, including himself, as their modes of awareness or "ways of experiencing a world."' J. Rosenberg, *The Thinking Self* (Philadelphia, PA: Temple University Press, 1986), 100.

sense impressions from different sensory modalities, visual and auditory. In the second story, these inferential connections are relayed by implicit genitives that associate an impression specific to one sensory modality with another impression specific to a different sensory modality, the sound (steps) *of* an approaching monkey (visualized), the screams *of* an excited monkey, etc. The first story, on the other hand, does not have such explicit associations between sense impressions of different modalities. It only contains rudimentary protocol-like transitions or transformation rules rather than proper inferential or objective-conceptual rules of transition, e.g.:

mass of fuzzy grey, contact with the heap of black →
screeching noise from the direction of mass of fuzzy grey

but not

excited monkey touching the monolith ↛ *monkey screams*

We cannot assume that, for the automaton, the shrieking sound is associated with the mass of protruding fuzzy grey as opposed to the heap of black or, for that matter, another item in the environment, or that it says anything (about any such association) at all.[126]

The goal of this comparative excursion was to highlight the features implicit in the automaton's nonconceptual encounter with the world, by highlighting its striking differences from the awareness peculiar to discursive apperceptive intelligence (Story 2). Now that we have a better idea what features and abilities the automaton is lacking, it is easier to figure out what must be added to the automaton in order for it to have categorially and inferentially structured experience and to bring its inchoate encounter with the world under the power of judgment.

However, despite these profound differences between Stories 1 and 2 as two *distinct* types of narratives, there is also a connection between them: the experiential content of the automaton's fully nonconceptual awareness still

126 For more details about such rudimentary protocol-like transitions, see chapter 5.

exhibits a non-categorial orderly structure that the inferential relationships among the concepts of the second story reflect and illustrate. Not only can this rudimentary orderliness be instantiated without language; it is a structuring that constitutes the minimum condition for the possibility of discursive apperceptive intelligence. But what is this minimum and necessary orderliness? It is the perspectival orderliness of space and time as forms of intuition, an encounter with the world from a nonconceptual point of view.

GETTING STUFF IN PERSPECTIVE

In Story 1, the automaton is aware of the presence of an expanse of grey stuff; it is aware of it—in analogical terms—as approaching or receding. It is also aware of an expanse of black stuff. But it is not just aware of the grey stuff approaching or receding, it is in fact nonconceptually aware of its moving toward or away *from* the expanse of black stuff, and similarly stopping in *front* or moving *behind* the black stuff, disappearing and appearing. In a nutshell, the automaton has a sense of movement, and with that, a rudimentary sense of space and of the presence and perspectival spatial relations between items *in* space. This is because the automaton is *situated* in space as a privileged endocentric frame of reference. But we ought to be cautious here. The automaton has no *concept* of these spatial relations, it is not aware of the conceptual contrast between something that simply disappears (goes behind something else) and something that ceases to exist (appearances and reality). Nor does it possess a *self*-concept. Nevertheless, the wired and programmed (structural and behavioural) system of the automaton can exhibit a spatial orderliness in its encounter with items in the environment that our linguistic terms such as 'approaching', 'receding from', 'moving in front of' and 'moving behind' (as opposed to 'ceasing to exist') inferentially instantiate and articulate. This rudimentary structured spatial encounter with the world can of course be detailed in terms of the complexity of the nervous system, but it also has a necessary form—not in the sense that it could not be otherwise, but in the sense that it is applied to all of the automaton's encounters with particular items in the world.

Coming back to the predator-prey example, both predator and prey behave in a way that exhibits a spatial orderliness in their goal-oriented activities (chasing and escaping). As the prey moves away, the predator follows its movement and orientation in space by means of saccadic eye movement—one of the fastest and most ancient predatory gestures—before orienting itself in space to chase the object of its hunt. When the prey hides behind a rock, it does not give up as if the prey had ceased to exist but instead follows the shortest path to catch the prey behind the visual obstacle. The predator has a sense of space only by virtue of its capacity to differentiate itself from its surroundings in order to effectively engage in its goal-oriented activity (catching the prey), a sense that is profoundly different from our own wherein items stand in inferentially articulated spatial and topological relations to one another. It is this functional differentiation from the surrounding space that serves as an egocentric (endocentric) frame of reference for the predator, a perspectivally situated position in space.

But the predator neither reflects on itself as being differentiated from the surrounding space, nor is it aware of having a privileged position within it—a perspectival stance. It just *has* a perspectival position in space by virtue of its structural capacity to behaviourally differentiate itself so as to successfully engage in *its* goal-oriented activity, and, by virtue of this, it *occupies* (without conceptually representing to itself) a frame of reference. It is this structurally and behaviourally posited endocentric frame of reference that permits the predator to entertain not only spatial relations between items and itself, but also a limited range of spatial relations between one item and another—that is, to treat not only itself, but also items in the environment, as frames of reference. This *endocentric* frame of reference—itself an index of the predator's situatedness in space and its spatial perspectival encounter with the world—is what enables the predator to have a constrained *allocentric* view of items as situated in space, and thus as spatially related to one another.

The origin of the nervous system, as René Thom and Alain Berthoz articulate, can be traced back to the solving of a fundamental problem for the self-preservation and successful execution of the organism's teleological activities:[127] *How can an organism differentiate itself from its food and from the predator?* Absent a contrasting index for differentiating itself from its food, the organism risks autophagy, and without any robust differentiating cue, the organism can neither successfully secure its food nor evade predators. To solve this problem, the organism must first differentiate itself from space via successful responsiveness to various forms of stimuli (electrochemical signal, physical pressure, light, etc.); but in order for it to effectively secure its food and elude predators, such parochial stimulus-driven contrast is not enough. The organism must form a 'mobile' perspectival frame of reference to respond appropriately to changes in the parameters of spatial relationships between itself and items in its environment. This is one of the oldest roles of the nervous system as a basic 'organ of alienation':[128] to mobilize and develop ur-alienation (minimal contrast with the environment through mere responsiveness to stimuli) into an alienation (a *designated* and *designating* discontinuity in space simulated by the nervous system as a self-model)[129] that enables the organism to structure its surrounding space and, in doing so, to derive an awareness-structuring orderliness from the rudimentary spatial relationships between items situated in that space. This enabling alienation is the capacity for spatial differentiation between items in the environment through the mobilization of the organism's successful self-differentiation from its surroundings, a 'perspectival pure positional awareness of items-in-relation-to-one-another'.[130]

In order for the automaton to have this perspectival pure positional awareness, it must have not only a sufficient structure for reliable differential responsiveness, but also a behavioural goal-oriented architecture—and that is exactly what our hypothetical agent has been supplied with.

---

127 See Berthoz, *The Brain's Sense of Movement*; and R. Thom, *Structural Stability and Morphogenesis: An Outline of a General Theory of Models* (Reading, MA: W.A. Benjamin, 1975).

128 Thom, *Structural Stability and Morphogenesis*, 299.

129 See Metzinger, *Being No One*.

130 Rosenberg, *The Thinking Self*, 111.

We can now say that the world of the automaton contains a privileged position, a position that *we* conceptualize and inferentially model on *our* spatial concepts but which *it* occupies and upon which it acts. The world of the automaton, then, contains a privileged and structure-conferring position only in the sense that its awareness$_1$ *of* the world is an orderly and coherent spatially perspectival awareness of items. But what is more interesting is that this spatially perspectival awareness$_1$ can be laid out precisely—albeit still analogically—in terms of our subjective spatial awareness$_2$ of items *as* moving toward and moving away, coming in front of and going behind. In other words, the automaton has a coherent spatial awareness of the items without necessarily possessing the conceptual resources to be able to tell the difference between appearances and reality, between, for example, an object whose shape and size remains constant as it retreats and an object that appears to be shrinking or morphing into something else as it moves away—as in the predator-prey example, where the sentient predator does not need the inferential relations between spatial concepts in order not to mistake a prey that moves behind a rock for an object that has now ceased to exist simply because the rock has occluded its shape and colour. Similarly, when the prey runs away, the predator's spatial awareness need not be armed with perspectival concepts in order not to take the running prey as an object that *seemingly* shrinks or morphs into something else. The predator continues the chase and captures the prey with pinpoint accuracy.

This is all to say that, even though our automaton is still a thoroughly nonconceptual non-apperceptive intelligence, its spatial perspectival awareness$_1$ of items is precisely already a spatial awareness of items *as* (*as* moving behind, moving away and not *of*$_1$ items ceasing to exist or shrinking and morphing into something else). For this spatial perspectival 'awareness-of the-items-as…' to be instantiated, the automaton does not need the conceptual distinction between what merely seems to be the case from an occupied pure perspective and what veridically appears to be the case, mastery of language, inferential relationships between perspectival concepts or even self-awareness. All it needs is to have (i.e., to occupy) a privileged point of view in space by virtue of a *sufficient structure coupled with a functional system* *of goal-oriented activities* that allow it to adopt one or multiple awareness-structuring perspectival frame(s) of reference in the world.

Looking at the ordinary spatial prepositions of our natural languages, especially those concerning localization, physical accessibility, and contact between objects (from, toward, behind, on, under...), we can see them as exhibiting what Claude Vandeloise calls a form of 'naive physics' of space, 'differing from scientific physics as radically as natural languages differ from formal ones'.[131] Although the automaton's naive physics of space is limited to simple spatial relationships that, for example, involve access to the field of sensory differentiation (grey stuff approaching or retreating), nonprepositional contact between objects (the chunk of red touches the mass of black, but not a red thick book *on* the table) and simple orientations (screechy noise from the mass of fuzzy grey, but not the scream *of* a monkey). More hybrid spatial relations and propositions of this naive physics of space such as bearer-burden relations, container-contained relations and complex localization (like the uses of *at* or near/far distinctions) are not yet on the menu for our automaton.

Serving as an armature for our higher order awareness-structuring spatial concepts, at its core this naive physics is built on a necessary family of perspectival spatial representations which can be instantiated without inferential relationships obtained from spatial concepts. This family of spatial representations demarcates a noncategorial yet orderly spatial perspectival encounter with the world that is a necessary condition of a conceptual and thus categorially-structured encounter with the world, and hence for an automaton with a full-blooded apperceptive awareness (awareness$_2$ as). In displaying exactly this spatial perspectival awareness, our hypothetical automaton has thus fulfilled the first necessary condition of possibility for the realization of general intelligence.

131 C. Vandeloise, *Spatial Prepositions* (Chicago: University of Chicago Press, 1991), 14.

THE AUTOMATON'S TEMPORAL PERSPECTIVAL AWARENESS

Now that the first minimum condition necessary for the realization of discursive apperceptive intelligence has been outlined, we can move on to the second minimum condition: a temporal perspectival awareness—that is, a structured and structuring temporal encounter with the world. For the automaton this means having the ability to locate itself in time, and thereby achieving a time-consciousness of both the world-history and its own history. Here self-locating in time means a rudimentary capacity to be aware of successive sensory affections produced by objects in its environment and to actively—but nonconceptually—respond to such affections. Having this rudimentary capacity is not as simple as it may seem. On one level, it requires a perspectival spatial orderliness of the kind we have been looking at. But on a different level, which, as we shall see, is even more crucial, it requires a temporal orderliness that relies on more fundamental *successively interrelated* capacities:

(1) The capacity to synthesize *compresent sensations* (the simultaneity of the manifold of senses). The role of synthesis here is an operation of combining, binding, and gluing partial maps of objects that come as compresent sensations (registered local invariances of shapes, colours, smells, etc.) into spatiotemporal global maps of objects.

(2) The capacity to synthesize the *simultaneous states* of objects into sequences of occurrences. Think of this synthesis as what is required for the automaton to capture 'the movement' of a moving car.

(3) The capacity to be aware of representations of objects and sequences of occurrences *as* representations and sequences. In a Kantian sense, this is tantamount to having the capacity to report representations of items and occurrences in the world to the mind. At this level, the capacity to report 'being affected' qua mind-state—'being aware of being thus (successively) affected'—does not by any means imply a form of knowledge. In other words, for the automaton, having this capacity does not imply

a conceptual awareness of the mind-state; this is something we should reserve only for an apperceptive agency. This awareness is merely an attentional mobilization of the global workspace that was introduced earlier. And similarly, the (conceptual) knowledge of the mind-state should not be conflated with this rudimentary awareness of mind-states, which is necessary but not sufficient for the realization of a conceptual awareness.

I said above that these capacities are successively interrelated because they are built on top of one another. (2) shares the components of (1), and (3) shares the components of both (1) and (2). But the reason they should be treated as distinct capacities is that each successive capacity adds a new component that is not available in the previous one(s). Our particular focus in this section will be on capacity (3): What needs to be added to a sequence of representations in order for the automaton to be aware of this sequence *as* a sequence of representations, to have a capacity by virtue of which it can be aware of being successively affected so as to be able respond to the manner in which it is so affected?

TIME AND MEMORY

Alongside the spatial elements in Story 1, we also find temporal elements—not only in explicit terms such as 'a short while later', but also, more predominantly, as implicit elements of motions. Our task now is to examine the automaton's temporal perspectival awareness, the structure of this awareness, and the consequences of its having this minimal temporal orderliness. But in order to move forward we have to bring into play a necessary structural component of the automaton that we have thus far not fully utilized, namely its constructive memory.

Let us call the registers of items that the automaton encounters in its environment—events which occasion alterations in the internal states of the automaton and which are manifested at the level of the global workspace—*impressions*. Now that the memory system is fully in place, the automaton is able to retain these impressions and access them. But, as discussed above, the retrieval of a retained impression is tantamount to the reconstruction of that impression. In other words, each time a retained impression is accessed, it is not merely called up but rather reconstructed, and each of these reconstructions is then added to the memory—i.e., becomes part of the situation that guides the further construction of memories.

According to this memory schema, the automaton's impression of an item in the environment—for example the monkey as a mass of fuzzy grey (*tode ti*)—can then take two general forms: a retained impression (formed by the *presence* of the mass of fuzzy grey), and a reproduced version of the retained impression of the mass of fuzzy grey (reconstructed in the absence of the actual item in the environment). Both forms mark the presence of the nonconceptual content of the *tode-ti* impression, but in one form this presence implies synchronicity with the actual item, while in the other form it implies diachronicity.

Analogically approximating the nonconceptual and tenseless awareness of the automaton, the impression of the object (I_0) in its retained and reproduced forms can be represented in the following manner:

I_o be_1 (impression)

which denotes the retained impression in which the 'presence of the nonconceptual content' is derived synchronously from the presence of the actual item.

I_o be_2 (reproduction)

which denotes the 'presence of the nonconceptual content' reproduced in the absence of the actual item, which is to say, derived (and reconstructed) diachronically from the retained impression I_0 be_1.

What is common to both forms is the presence of the content of this-such—represented by the tenseless verb *be*. It is tempting to interpret the $_1$ and $_2$ in be_1 and be_2 as indices of times, and hence to conclude that be_2 is the past tense of be_1 (*the automaton saw*...), in so doing attributing to the automaton a temporal awareness, or even tensed thought. But at this point $_1$ and $_2$ merely represent two distinct *types* of impression-content: one that occurs in the presence of the actual item and another that occurs in its absence. These two general types signify respectively the impression (be_1), and the reproduction of that impression bearing the mark of both being retained and being reconstructed (be_2). More succinctly, be_1 and be_2 present I_0 respectively as the content of an impression and as the content of a reproduction. From our perspective, the automaton's *this–such*$_1$ and *this–such*$_2$ clearly signify present and past tenses of impressions of the items (the automaton sees a monkey; the automaton saw a monkey). But we should avoid ventriloquizing the machine, and instead equip it in such a way that it will become capable of drawing such a conclusion all by itself. In other words, we should allow our toy automaton to have a genuine temporal perspectival awareness that *legitimately* admits our analogical observations regarding what it is aware of. And even then this temporal awareness is not by any means close to any form of time-consciousness, even though it may be the incipient germ of such consciousness.

The automaton, at this stage, is aware$_1$ *of* this-such$_1$ and this-such$_2$ but it is not aware$_2$ of them *as* present and past. Nor in fact is it aware of them as temporally related. In order for the automaton to have a pure situational temporal awareness, it must first be able to temporally relate these impressions to one another. This temporal relation is already available at the level of the automaton's structure, in the shape of the causal ordering of its registers, and therefore as a sequence of alterations in the internal states of the automaton's wiring. Hypothetically speaking, and following Rosenberg's version of this thought experiment,[132] if we could attach a monitor to the automaton and treat it as a computer with a visual display unit, we might see its visual registers displayed in a certain hue with a high degree of saturation. For the purposes of this example, we can think of the actual item in the environment as a key being pressed on the keyboard attached to this computer, and the saturated hue as its corresponding register. Moreover, these registers do not disappear from the screen, since they are stored somewhere in the automaton's memory.

132 Rosenberg, *The Thinking Self.*

Now, as more keystrokes are made, something happens on the monitor: as new saturated registers appear on the monitor, the saturation of the stored registers decreases. The high-saturated range of the hue corresponds to impressions synchronous with the presence of an item in the environment, i.e., registers that appear on the monitor as the keystrokes are being made; while the low-saturated range corresponds to reproductions, i.e., registers of the keystrokes that have 'previously' been made, and which are now retained in the memory.

What is important to note in this example is that the difference in saturation is a difference made by causal correspondences or mappings between displayed registers and keystrokes qua independently existing items in the environment, and that these causal correspondences stand in de facto temporal relationships to one another. But the automaton is not aware of this temporal ordering of causal correspondences (keystrokes made *before* and keystrokes made *after*, registered as stored and registered as being displayed); nor can it infer from the TV in its head—the monitor displaying registers in various degrees of saturation—any temporal relationship. It is we, not the automaton, who infer a temporal order from the changes in the saturation in correspondence with the sequences of keystrokes. But could we say that, if we sufficiently train the automaton via the application of statistical inference, then the automaton will eventually learn to recognize the ordering or sequence of impressions, and thus legitimize our attribution of temporal awareness to it? Not at all—temporal awareness precisely cannot be reduced to such a trained response: the representation of a sequence, or ordering, is not the same as a sequence of representations. The ability to differentiate a sequence of representations does not amount to the ability to represent a sequence. In the same way, a sequence of awarenesses of items in the world does not bring about an awareness of a sequence qua temporal ordering. Devising a trick to achieve the former does not generate the latter capacity.

Indeed, this is part of Kant's attack on Hume's sceptical reductionist system, specifically his Separability Principle, the idea that 'every thing, that is different is distinguishable, and every thing that is distinguishable