

**Turing Trauma:
Blind Spots in the Synoptic Vision of Intelligence**

AA Cavia

Presented at 'Marxism & The Pittsburgh School'
University College London
11.06.2024

This article attempts a diagnosis of a certain mode of alienation induced by an apparatus which I will refer to as 'computation', indexing a specific humiliation of the human under the rubric of *Turing trauma*. In what follows, I outline the basic features of this apparatus and position it with respect to Sellarsian, Marxian, and Fanonian notions of alienation. *Prima facie*, the automation of cognitive labor by artificial intelligence (AI) seems to act as a canonical example of alienation in the Marxian tradition—the knowledge and skill of the worker concretised by machines that transform human labour into fixed capital. In a Sellarsian register, AI conceived as the artefactual elaboration of cognition would appear to play an important developmental role in what Sellars once called the “scientific image” of the human, the latter initiating a rupture in our philosophical self-conception as thinking subjects. The interaction of these two modes of alienation, by turns economic and scientific, precipitates what Mattin has recently termed “social dissonance”.¹ For Mattin, science is seen to activate an “alienation from below”, with neuroscience presented as paradigmatic, rendering our selves as merely self-models grounded in the neurology of the brain, evolutionary constructs for sense making in an otherwise intractable environment. The social alienation brought about by the relations of production under capital produces in turn an “alienation from above”, a structurally conditioned objectification of the worker expressed at the level of political economy. Following Mattin, the scientific image of man and capitalist subjectivity are in a sense intricately braided, not by a common logic, so much as a common trauma that *dissonates* through human society.

I will argue that the cleaving of reason from mind enacted by AI both expresses itself *within* and distinguishes itself *from* these two modes of alienation, in that it represents a critical point in what Hegel once called the estrangement (*Entfremdung*) of reason, precipitating a challenge to a dominant image of thought put forward by Western epistemology, namely that of an affinity for a certain notion of truth. In this regard, I will examine the alienation (*Entäusserung*) of spirit in Hegel as it relates to computational reason, the latter cast as a mode of explanation with specific logical commitments. Lastly, I consider the post-colonial theory of Frantz Fanon and Homi K. Bhabha and argue for its relevance in understanding the figure of computation as an epistemic apparatus. Here I refer to Bhabha's work on mimicry in the performance of

¹ Mattin (2022). *Social Dissonance*. Urbanomic.

colonial subjectivity as it relates to Turing's imitation game and Davidson's problem of radical interpretation. The aim will be to faithfully render the features of a certain Turing trauma in order to locate it within this matrix of alienation. As such, in what follows I treat computationalism—conceived as a refusal to accept an epistemically deflationary account of computational states—as a challenge to the synoptic vision of the human put forward by Sellars. The specifics of this challenge are outlined by examining certain blind spots in the Sellarsian view which hamper its integration into a broader philosophy of intelligence, the latter which I take to be a significant task laid at the feet of philosophy by the project of AI. The broader research aim is to adapt some Sellarsian motifs to serve as a basis for both a critique of artificial intelligence and a constructive account of computation, engaging in the task facing the modern computationalist thinker—namely ushering the project of AI from its Humean to its Kantian phase.

I. Sellarsian Alienation

The first mode of alienation to consider is that which Sellars once framed as the divergence of two images of man, by turns 'manifest' and 'scientific', which we can interpret as the disenchantment of the human by the logic of science, provoking an irrevocable split in the thinking subject. For Sellars, the former image tracks the induction of man into a conceptual order, whereby "the transition from pre-conceptual patterns of behaviour to conceptual thinking was a holistic one, a jump to a level of awareness which is irreducibly new, a jump which was the coming into being of man".² This self-encounter of the human, which marks a phase transition from creature of habit to normative agent, is at odds with the dispassionate image which modern science repeatedly (re)presents to us, an originary Copernican rupture giving way to an agonising historical procedure which threatens to eradicate every remnant of folk psychology. The manifest image in Sellars' conception should not be confused with naive reports of common sense quotidian experience, but rather marks an entire intellectual tradition centered on the individual as thinking subject, a tendency that Sellars traces from Plato to Hume, from naive empiricism to Husserlian phenomenology, an entrenched view which he calls "perennial philosophy". The manifest image is above all a philosophical image, inclusive of all those modes of merely "correlative" explanation, from statistical inference to the logic of inductivism. It should be noted that neural computation, which to date has been an experimental endeavour tethered to inductive logic, from its roots in connectionism to artificial neural nets, would fall firmly within the regime of the manifest image on this view. But the long-term threat it represents to a neat delineation between the images should not be underestimated. In this sense, the dualism between the manifest and scientific images bleeds into the scientific enterprise itself, Sellars instilling the relation between theory and experiment with an inferentialist insistence on the theory-ladenness of observation.

² Sellars, W. "Philosophy and the Scientific Image of Man" in Scharp, K. and Brandom, R. (2007). *In The Space of Reasons: Selected Essays of Wilfrid Sellars*. Cambridge, Mass.: Harvard University Press. p. 374

For Sellars, these conflicting images appear to preclude genuine reconciliation, they present themselves as dualistic as opposed to dialectical. On this point Sellars argues that postulate-driven theory formation, of which we can think of contemporary theoretical physics as paradigmatic, is irreducible to induction on observable phenomena. As a means of dismantling this dualism, Sellars seeks a naturalist account of mind resistant to a mereological reduction of those affordances he is at pains to position as central to sapience, a normative structure he suggests should be integrated into our account of scientific epistemology itself. As he concludes at the end of *Philosophy and the Scientific Image of Man*:

“Thus, to complete the scientific image we need to enrich it not with more ways of saying what is the case, but with the language of community and individual intentions, so that by construing the actions we intend to do and the circumstances in which we intend to do them in scientific terms, we directly relate the world as conceived by scientific theory to our purposes, and make it our world and no longer an alien appendage to the world in which we do our living.”³

Therein lies his attempt at a synoptic vision which integrates two images of the human, dismissing reductionism—in the guise of what McDowell goes on to call “bald naturalism”—extending instead a view of science as a discursive theory driven pursuit which rests on a specific construction of objectivity. In many ways this reads like a call for the kind of research which in recent decades has been aimed at advancing a social history of the scientific apparatus, as evinced in the work of Daston, Hacking, Latour, Stengers and many others. In some sense their work does heed Sellars’ call in attempting to study the practices of science within the ambit of human social norms. But this would be to misread the properly philosophical challenge offered by this mode of Sellarsian alienation, which concerns the very status of objects in cognition as naturalised by mental states. The Sellarsian project then, if such a project can be said to exist, is centered on outlining the desirable features of a normative rationalism reconciled with the natural sciences, a fusion of the space of reasons and the realm of causes, which preserves a commitment to both scientific realism and the irreducibility of the conceptual, all the while taking heed of a certain skepticism of the rational presented by twentieth century pragmatism. This central irreducibility claim, which is broadly equivalent to Davidson’s “constitutive ideal of rationality”, is not something that can in principle be left to a social historical method, in that it hinges on rendering the relation of mind and world on normative terms, a task that requires the kind of naturalised metaphysics which I take Sellars to be engaged in.⁴

If we accept that “fusing the images”, as Rosenberg puts it, remains a primary task of philosophy, then Sellars leaves many aspects of this vision vulnerable to fragmentation, particularly when viewed from the contemporary standpoint.⁵ A

³ Ibid, p. 408

⁴ McDowell, J. (1998). The Constitutive Ideal of Rationality: Davidson and Sellars. *Crítica: Revista Hispanoamericana de Filosofía*, pp.29-48.

⁵ Rosenberg, J.F. (2007). *Wilfrid Sellars: Fusing the Images*. OUP Oxford.

specific problem presents itself in the context of a philosophy of intelligence capable of integrating computational agency into its ranks, notably the anthropocentric nature of Sellars' two-ply account of perception, which seems ill-equipped to accommodate a broader non-human theory of cognition. My claim is that Sellars' account, in which sentience and sapience are irrevocably split, frustrates any purported synthesis of the image of the human, splintered as it is today by the apparatus of computation. The crux of this problem is how we come to interpret normative acts and how these distinctions play out in mental representations. This is the Sellarsian blind spot I seek to diagnose in the pursuit of a general theory of intelligence that is able to account for a plethora of adaptive, goal-oriented, pattern-governed agential behaviours. To reveal this blind spot we need to recognise the key role played by the theory of picturing within Sellars' theory of 'mental events' and understand how this is undermined by computational reason. For Sellars, mental events come about by analogy with speech, but he is generally skeptical on whether intentionality can be extended to non-linguistic mental episodes. In his late essay, *Mental Events* (1980), he is uncharacteristically ambiguous on the question:

"I take intentionality... to be the mark of the mental. I agree with the classical view that there is a domain of 'inner episodes', properly referred to as 'thoughts', which are not linguistic—though they are *analogous* in important respects, syntactic and semantic, to linguistic structures, and are functionally connected with linguistic behavior."⁶

This should be contrasted with the Sellars of *Being and Being Known* (1962), who is not prepared to extend intentionality to specific kinds of representings he calls pictures. The point hinges on the analogical relation alluded to in the passage above, which needs to account for how said picturing relations and another class of signifying relations bridge intelligible form with the real order. By 'intelligibility' here I mean a property of a mental state that permits its presentation to thought as an object of cognition. The theory of picturing thus emerges as a key element in the Sellarsian attempt to provide a unified account of norms and causes, acting as a gateway for language entry transitions, as elucidated by his discussion of statements that identify a picturing relation in the John Locke lectures:

"A statement to the effect that a linguistic item pictures a non-linguistic item ... is, in an important sense, an object language statement, for even though it mentions linguistic objects, it treats them as items in the order of causes and effects"⁷

For a representing to play the role of picturing then is for it to be placed into a relationship with a uniformity of the natural order—what I would be inclined to call, following Dennett, a "real pattern". The statement 'X pictures Y' for Sellars implies

⁶ Sellars, W. (2007). *Mental Events*. In "In The Space of Reasons: Selected Essays of Wilfrid Sellars", p. 283.

⁷ Sellars, W. (1968). *Science and Metaphysics: Variations on Kantian Themes*. London: Routledge, p. 137.

that “both X and Y belong to the real order, i.e. neither belongs to the order of intentionality”⁸. The question arises here as to whether intentionality is thus reserved for those rarefied mental events that display a reflexive conceptual awareness of said picturing relations, a reliance on introspection which I do not deem compatible with a general theory of intelligence. Presumably when a salamander extends its tongue to catch a prey which is in motion, it is exhibiting capacities of recognition and prediction which imply a representational system (RS) that is functionally linked to its intention to eat, all without recourse to language in the form of speech. For while Sellars thinks that “there is a legitimate sense in which an RS which doesn’t involve subject and predicate terms can nevertheless contain propositions and intend states of affairs” there is no convincing account of how this position can be paired with inferential semantics to allow us to distinguish between, for example, rat and human cognition, without flattening the field of non-human subjectivity.⁹ In Rosenberg’s discussion of animal RSs he is a little more charitable in his interpretation, he quotes Sellars on the double function of representation as including a form of type assignment:

“A basic representational event is an event which has two characters: one by virtue of which it represents an object in its environment (or itself); another by virtue of which it represents that object as being of a certain character.”¹⁰

For Rosenberg, representations are thus self-aware and typed but not necessarily in a linguistic sense. Rosenberg considers Sellars’ discussion of Jumblese, a toy language in which geometric relations between signs can stand in for predicates, as demonstrating a nominalist theory of predication:

“It is plausible to suppose that the representational states of animal RSs resemble sentences of Jumblese more closely than they do sentences of a natural language or formulae of a logical calculus, i.e., that they perform their representational functions without the aid of auxiliary elements.”¹¹

Finally claiming that “[t]he proper way to relate this nominalist theory of predication to our earlier discussion of representational systems is to see *signs* or *symbols* as playing the role of *states*.”¹² Here there is a risk of collapsing the symbolic into the causal in a manner which undermines the normative structure of Sellars’ project, but if one treats this move with care I see a fertile ground for a more unified, not to mention naturalised, theory of representation. By contrast in my reading, the fact that Sellars contends that ‘mental events’ only come about by analogy with speech, coupled with the claim that intentionality is the mark of the mental, leaves us with a somewhat impoverished account of non-human agency outside of discursive

⁸ Sellars, W. (2007). *Being and Being Known*. In *In The Space of Reasons: Selected Essays of Wilfrid Sellars*, §33

⁹ Sellars, *Mental Events*, p. 295

¹⁰ Sellars, *Mental Events*, p. 296

¹¹ Rosenberg, J.F. (2007). *Wilfrid Sellars: Fusing the Images*. OUP Oxford, p. 112

¹² *Ibid.*

practices, the latter which in turn need to somehow overcome the challenge of interpretation. Moreover, since Sellars admits that while “structural similarity is a necessity it’s not a sufficient condition” for the kind of mapping that implies inferential moves which he regards an “essential activity” of representational systems, the question arises as to what indeed is a sufficient account.¹³ Such an account would need to be able to integrate the kind of animal RS which allows a rat’s hippocampus to navigate and locate items in the world, a cognitive capacity viewed as a manifestation of intentional goal-oriented behaviour. For Sellars, a cut is to be made between logical and non-logical representational systems, and he renders the latter as those that do not represent logic explicitly, seen roughly along Humean lines which limit those states to mere association. While an appeal is made to the behaviour of the animal in question in determining whether inferential affordances are at play, it is not at all clear how this distinction plays out in the representational system itself. The appeal to Jumblese does not effect this cut, since its purpose is to demonstrate that geometry can stand in place of logic “without the aid of auxiliary elements” (Rosenberg). Indeed, AI interpretability research already shows this to be the case in deep learning models, which encode concepts in embedding spaces, demonstrating a geometric logic that can model relational predicates without any recourse to explicit logical operators.¹⁴

Here we can begin to intuit the blind spots in the Sellarsian account, but let us continue to trace the problem as it develops in the theory of picturing. Two orders of isomorphism are at play for Sellars, picturing in the real order and signifying in the logical order. When Sellars considers the toy machine language Robotese, the nature of the statement:

“in Robotese ‘::’ signifies lightning”¹⁵

Is said to *presuppose* a picturing relation in the real which allows one to perform a translation between two languages (Robotese and English). As he puts it:

“[W]e see that even though these two isomorphisms are quite distinct and belong to two universes of discourse, there is nevertheless an intimate connection between them which can be put by saying that our willingness to treat the pattern ‘::’ as a symbol which translates into our word ‘lightning’ rests on the fact that we recognize that there is an isomorphism in the real order between the place of the pattern ‘::’ in the functioning of the robot and the place of lightning in its environment.”¹⁶

The isomorphic relationships that linguistic and non-linguistic representings enter into, namely conceptually contentful inferential relations and pictures that bear

¹³ Sellars, *Mental Events*, p. 293

¹⁴ Grand et al. (2022). Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nature human behaviour*, 6(7), pp.975-987.

¹⁵ Sellars, *Being and Being Known*, p. 223

¹⁶ *Ibid*, p. 226

a correspondence to real patterns, are a prerequisite for the claim that discourse has the capacity to enter into a factual relationship with the world, in other words that a statement *correctly* pictures a state of affairs in the form of a fact. For a nominalist like Sellars, this notion of facticity hinges on a coherence theory of truth, which in turn raises an epistemological challenge to the status of objects within a scientific theory of objectivity, as Sellars himself alludes to in his reference to projection and context sensitivity:

“[P]ictures, like maps, can be more or less adequate. The adequacy concerns the ‘method of projection’. A picture (candidate) subject to the rules of a given method of projection (conceptual framework), which is a correct picture (successful candidate), is S-assertible with respect to that method of projection”¹⁷

The lacuna which opens up here between real patterns and non-linguistic representings admits of a relativism which appears to reject an account of statements of fact *qua* correct picturing of the world conceived as the very substrate of scientific rationality, science thus defined as the body of statements considered factual. For Sellars, the best we can do is to gauge whether scientific theories are convergent over time *with respect to each other*, since there is no absolute reference—no Archimedian point—to judge pictorial adequacy. Here in a sense is the origin of Sellars' split with logical positivism, the latter presenting science as the accumulation of a canon of observable facts, whereas the inferentialist is committed to exposing the theory ladenness of empirical observation, manifested in Sellars' espousal of a normative account of not just scientific epistemology but of ratiocination as such.

Rosenberg would go on to reject the theory of picturing, since for him justification is prior to correctness, such that the criterion for pictorial adequacy cannot be grounded without an appeal to the very space of implications for which it is offered as a pre-condition. I would side with Sellars on this point, since the criterion for pictorial adequacy can bottom out into selective reinforcement in the case of the kind of maps or strategies which non discursive agents engage in—there is in my mind no need to appeal to justificatory practices to preserve a theory of picturing in the last instance. As to what makes a candidate successor theory appealing, one can again fall back on normative practices, but this by no means undermines picturing itself, since correctness does not *rest* on justification—one can have any number of convincing reasons to propose a candidate scientific theory, but it may still fail miserably when tested in the wild. Conversely, a picture can be correct without any explanation being to hand, as in the case of non-human agents engaging in feats of recognition without partaking in the game of giving and asking for reasons. To summarize, for Sellars objectivity rests on correct picturing under a given conceptual scheme, its grasp on the world reliant on a non-linguistic relation conditional on the modal property of projectibility, which must somehow find alignment with an isomorphism of a quite different order. But a problem arises regarding the joints by which Sellarsian epistemology and semantics hang together, that is to say between his

¹⁷ Sellars, *Science and Metaphysics*, p135.

theory of correct picturing and his functional theory of meaning, despite the great lengths to which Sellars goes to insulate the latter from the former. If concept acquisition is in any way exposed to a non-semantic picturing relation between a representing and a real pattern, the question must arise as to how signifying relations bridge this gap from the logical to the real order by anything other than analogy. Furthermore, intentionality has to be correctly rendered in such a way that is consistent between our analysis of social practices and our theory of mental representations.

II. Drilled Automata

Before proceeding, I should provide a more schematic treatment of alienation in which we can articulate the stakes at play in the synoptic view. The schema of alienation I would propose is one in which an apparatus of objectification induces a breakdown in the identity relation of a subject. This is in contrast perhaps to the kind of self-objectification which we might associate with positive alienation. What concerns me here is not the freedom to alienate oneself but rather how an externality forces an irreconcilable split in subjectivity. The apparatus of objectification can take many forms, be it science, capital, colonialism, and importantly, I would argue, computation. It should be noted that the form of objectification is also peculiar to each mode of alienation. So whereas the question of the colonial object in Fanon concerns the figure of the native, in Marx the worker in the form of the wage relation, and in AI the computational agent, Sellarsian alienation instead concerns the nature of objectivity itself in presenting the human qua object of scientific study. As I have tried to show, the question of the object in Sellars, whether this be a non-linguistic ‘item’ engaged in a picturing relation, or a linguistic object engaged in a signifying relation, presents a Gordian knot to the inferentialist critic of naive empiricism. This knot can only be cut by an unsatisfactory appeal to an ‘analogy’ between the conceptual and the real order, itself not a structural isomorphism, since the latter would open the door to a type identity theory of mental states of which Sellars is a vocal critic. Rosenberg suggests we adopt a more geometric view of what I would call encodings—via his discussion of Jumblese—and I will attempt to flesh out this idea with a computational treatment. My suggestion for now is that alienation is a thematic which can open up the Sellarsian account to a fully synoptic view of intelligence, by way of a rendering of computational reason, which presents a general theory of encoding that is naturalised in the first instance. The aim is not to strip away the normative metaphysics from Sellars, but to unify the notion of representation within treatments of sentience and sapience, one which offers a coherent account of objectivity compatible with the context sensitivity of conceptual frameworks, while aiding the necessary distinctions in capacities and affordances required by a theory of intelligence. Crucially, this theory has to aid us in locating intentionality in such a way as to not flatten the field of non-human agency under the rubric of merely sentient creatures. Before providing a brief sketch of said theory of computation we should however first understand its relation to two further modes of alienation.

The second mode of alienation to consider, in some ways contingent on the first, is the Marxist alienation of the worker, who suffers a twofold objectification, first as labour by capital, enshrined in the wage relation, and secondly in the process of reification signalled by the commodity form. To these two levels of objectification we can also add the concretisation of the worker's knowledge and skill in the form of machinery, as famously noted by Marx in the *Grundrisse*:

“[It] is the machine which possesses skill and strength in place of the worker, is itself the virtuoso, with a soul of its own in the mechanical laws acting through it ... The worker's activity, reduced to a mere abstraction of activity, is determined and regulated on all sides by the movement of the machinery, and not the opposite. The science which compels the inanimate limbs of the machinery, by their construction, to act purposefully, as an automaton, does not exist in the worker's consciousness, but rather acts upon him through the machine as an alien power, as the power of the machine itself”¹⁸

Of the three forms of Marxian objectification mentioned here, the alienation of consciousness by way of a concretisation of knowledge relates most closely to that precipitated by AI in the contemporary labour crisis. This in turn has prompted some thinkers such as Pasquinelli to critique AI within the ambit of a Marxist labour theory of automation, in which the historical division of labour preempts technological development, while the worker is dehumanised and takes on the behaviour of what socialist William Thompson once called “drilled automata”.¹⁹ What I consider to be missing from this account is the admission of any form of computational agency, as a result of identifying computation with the “inanimate limbs” of machinery, without properly acknowledging the “alien power” of computational reason, a conflation I regard as misconceived. While Marx is comfortable declaring the capacity of machines to act “purposefully”, any traces of this view are notably absent from contemporary Marxist critiques of technology. In its place one is often met with the reduction of computation to mechanism, in the guise of a universal axiomatic ‘machine’, or else a blindly obedient rule-following automaton, eschewing a unified view of computation capable of fully integrating its various aspects. These are expressed in the three canonical models of computation—in terms which are by turns mathematical, linguistic and machinic—but their presumed equivalence under the Church-Turing thesis should not blind us to the need for a synoptic view. Since computation is a janus-faced theory, one's method of approach is critical—there is a certain orientation we must adopt to fully grasp not only the cluster of theories that present themselves as computational, but the range of physical phenomena they purport to explain.

Despite its pretences to generality, I would contend that latent within Turing's universal machine formalism lurks the problem of interpretation. Computation as an

¹⁸ Marx, K. (1857). *Grundrisse: Foundations of the Critique of Political Economy*, trans. Martin Nicolaus, London: Penguin. p. 692-3.

¹⁹ Pasquinelli, M. (2023). *The Eye of the Master: A Social History of Artificial Intelligence*. London: Verso Books.

apparatus cannot be shorn from logic on pain of incoherence, since a semantic theory of computational states is a pre-requisite for identifying the constraints that mark a physical process as computational in the first place. In this regard, what Piccinini calls a “robust mapping” between the formal and physical registers of computation is a critical component of the view I am offering here.²⁰ I would counter Turing’s formalist bias by endorsing instead an *intuitionistic* view of computation, founded on a spatial theory of types first proposed by Voevodsky, interpreted under realizability semantics originating in the work of Kleene and further developed by Martin-Löf, as the best candidate framework for a unified account of computational reason. When I speak of the apparatus of computation, it is this framework that I have in mind. By intuitionist, I mean to treat computation as essentially cognitive in character, an insistence that it cannot be equated with a purely formal theory, marking instead a constructive logic which is only ever imperfectly captured by symbolic formalism. On this topological view, computation is synonymous with the forging of paths in continuous spaces, and the meaning of computational states rests not on Boolean truth tables, but on the realizability of truth procedures. In this sense, the paths which computational reason traces have never been strictly speaking formal, and the interpretation of computational states hinges on the coherence of a certain view of logic, as seen through the lens of intuitionism.

Here I am using the label ‘computation’ as a stand-in for a certain kind of worldview, the indexing of a world which assembles language, logic and topology in a specific manner, introducing a mode of explanation which I call computational reason. Computation here is not to be identified with an abstract universal machine so much as an epistemic theory of encoding, ushering in a structuralism which aims to integrate contingency into logic. Treated as a diagnosis of contingency, computation signals the ascendancy of the inferential over the axiomatic, emphasising the provisional and informal over the static and immutable. The project of computation under univalence can thus be viewed as an attempt at outlining the minimal axiomatic commitments required to maximise inferential freedom, geared towards autonomous agency over rule-following automatism. Importantly, topology is not offered as a metaphor but is instead foundational, presenting a single *univalence axiom* which acts as a criterion for identity.²¹ We can say that computation functions somewhat like a *motif*, in the sense first proposed by mathematician Pierre Deligne, namely as a conceptual scheme providing a “system of realisations” for unifying disparate theories construed as structures, pursued by exposing certain topological invariances between them.²² Moreover, this computational view is not logicist or

²⁰ Anderson, N.G. and Piccinini, G. (2024). *The Physical Signature of Computation*. Oxford: OUP.

²¹ Awodey, S. (2018). *Univalence as a Principle of Logic*. *Indagationes Mathematicae*, 29(6), pp.1497-1510.

²² Deligne, P.P. (1989). Le groupe fondamental de la droite projective moins trois points. In *Galois Groups over \mathbb{Q} : Proceedings of a Workshop Held March 23–27, 1987* (pp. 79-297). New York, NY: Springer US.

foundationalist in its tendency. It eschews the formalist stance with regards to reasoning, admitting its own inconsistency up front, by virtue of its insistence on the realizability of truth.

Equipped with this image of computation, we can begin to address the Sellarsian blind spots alluded to in the previous section with a new account which correctly locates this epistemic apparatus. Within the Sellarsian framework, there is a temptation of casting computation as merely a ‘realizer’ of conceptual roles, which would consign computation to the status of mechanism, leaving the philosophical framework largely intact. But to do so would be to misread the potential for rendering realizability as a fully-fledged semantic theory. In the type systems of Martin-Löf, type assignment resembles not so much an act of classification, as the construction of a proposition whose meaning is synonymous with all the ways we have of realizing it. On this view, type assignment resembles most closely the Sellarsian technique of dot-quoting, which is akin to the functional classification of expression tokens. If we take the following typical Sellarsian inter-linguistic formulation,

- (1) ‘Dreieckig’ in German stands for triangularity

As Seibt shows in her in-depth discussion of Sellars’ nominalist reductionism,²³ the aim is to eliminate the abstract singular terms and replace them with quoting contexts of the form,

- (2) The *dreieckig* means the •triangular•

Or else,

- (3) The *dreieckig*s mean the •triangular•s

Where the star-quoting indicates a class of sign designs, what Seibt calls a “morphological configuration” (I would use the term ‘topological’ here), and the dot-quoting in turn the functional class of the expression token ‘triangularity’, and both can be treated as type expressions. For Sellars, all type expressions of this kind can be treated as singular terms that refer *distributively* to their tokens. As Seibt puts it,

“[T]he expression ‘the •triangular•, like ‘the lion’, can be interpreted to embody a sortal which classifies concrete ‘tangible’ objects, namely, expressions of different languages that share a linguistic function”²⁴

In Sellars’ own words, all “abstract singular terms, thus -ity, ‘-hood’, ‘ness’,... are to be construed as quoting contexts which a) form metalinguistic functional sortals,

²³ Seibt, J. (1990). Chapter 2 in *Properties as Processes: A Synoptic Study of Wilfrid Sellars Nominalism*. Ridgeview.

²⁴ Seibt, *Properties as Processes*, p. 48.

and b) turn them into distributive singular terms”.²⁵ Moreover, the inferentialist semantics which follows is deemed to be centered on the functional role expressed by said sortals. In this sense, Sellars’ aim is to bring nominalism to bear on the meaning of ‘means’ itself in the formulation in (2), construing it as a copula on sortals of this kind. I will extend here a computational interpretation of said formulation with a view to demonstrating some features of realizability as a candidate for natural language semantics. On the computational view, metalinguistic functional sortals are just higher types, so we can now substitute types using boldface convention to give,

(4) The ‘dreieckig’ means **Triangular**

And we are now well placed to introduce a more computational semantics, but we will need to decide on whether we are rendering this relation in terms of type assignment or realizability. To assign a single instance of the term (token expression) ‘dreieckig’ to the type triangular gives,

(5) $\Gamma \vdash \text{dreieckig} : \mathbf{Triangular}$

In which a linguistic context Γ that includes both languages is assumed, a formulation which differs from the following,

(6) The ‘dreieckig’ realizes **Triangular**

In that ‘realizes’ denotes a logical interpretation whereas assignment alone is a non-semantic relation. Indeed we must be careful here, since under realizability semantics a type is said to correspond to a proposition, which it realizes by exhibiting a program, which means we must present a function on a given ‘dreieckig’ *context* as the realizer,

(7) $f: \text{dreieckig} \models \mathbf{Triangular}$

Which reads, “a computable function on the context ‘dreieckig’ (aka a program) realizes the (proposition corresponding to the) type triangular”. As you can see, we have dispensed with the operator ‘means’ and we have converted ‘dreieckig’ from a metalinguistic sortal to a program which will determine the functional role of an expression in a concrete context (e.g. a linguistic context in which it is being deployed). The type **Triangular** in turn determines which programs correctly output instances of its type according to whichever functional logic triangular token expressions need abide by. Now to fully render a semantic relation between two types, as per Sellars’ formulation in (2), we would need to expand our function definition to formulate it as such:

²⁵ Sellars, W. (1980). *Naturalism and Ontology: The John Dewey Lectures for 1973-4*. Ridgeview. p. 94.

(8) $\Gamma \vdash f: \mathbf{Dreieckig} \rightarrow \mathbf{Triangular}$

Here a function is presented that maps instances of a **Dreieckig** type to our **Triangular** type, such that the latter would be said to be *inhabited* and its propositional form thus realized. The context Γ is presumed to be multi-linguistic and in the topological view it comprises an encoding of both languages in a common embedding space. Here the generic type **[T]**, which expresses the logical form of the triangularity predicate, can be said to embed both of these types and thus transcend the linguistic context of each. I would endorse (7) with its implicit expansion in (8) as a computational rendering of the original Sellarsian formulation in (2), despite the usual caveats about adapting formal semantics for natural languages. Clearly, semantics of this sort are not explicitly designed for the kind of translation that Sellars is using to emphasise the functional role of an expression over any given linguistic context. However, type theory is well adapted to functional classification more broadly, since it can replace metalinguistic operators with higher inductive types. Moreover, the topological view provides good prospects for broader application due to the generic nature of topological structure. Indeed, both Martin-Löf and Ladyman have suggested that a constructive type theory may be a good candidate for a broader semantic theory of language in which types play the role of concepts conceived as propositional forms.²⁶ The semantic implications are subtle—we have provided a semantic theory centered not on the inferential role played by an expression type so much as the class of constructions that inhabit it. Realizability locates semantic content in the constructions themselves, which produce those expressions qua processes that perform an encoding of the corresponding logical proposition, in this case the predication of a property that conforms to the notion of triangularity. The functional role is presumed to be laid bare by exhibiting a program, in the broadest possible sense of the term, which realizes the type. The type **Triangular** thus represents all the valid programs that output instances of the type which conform to the propositional form. The scheme is broadly anti-platonist and I would argue compatible with both Sellars' nominalism and his naturalism, but its insistence on realizability marks a shift in emphasis within the semantic framework.

We can now use this brief sketch of a certain view of computation to bridge the lacuna opened up by Sellarsian alienation as evinced in his theory of picturing and mental states. This image of computation rests on a constructive logic which proffers a treatment of identity grounded in an isomorphism expressed as topological invariance. The notion of isomorphism is in turn formalised by a *homotopic* theory of types in which to assign a token to a type is to engage in an act which I call encoding. The interpretation of encoding can take the form of a program that outputs instances of that type or alternatively the construction of a class of paths in a continuous space. On this view, encoding is presented as the fundamental operation in the theory of

²⁶ See Martin-Löf, P. (1994). *Analytic and Synthetic Judgements in Type Theory*. In Kant and Contemporary Epistemology. Springer: Dordrecht. pp. 87-99 and Ladyman, J. and Presnell, S. (2018). *Does Homotopy Type Theory Provide a Foundation for Mathematics?*. The British Journal for the Philosophy of Science, 69(2), pp. 377-420.

computation—above all computation is an apparatus that encodes—but that process of encoding, which is ultimately cognitive in character, has a topological rendering which, under the intuitionist view, breaks with any given linguistic context. The theory of types is thus seen as an attempt to capture the properly cognitive act of encoding in formal terms, an act which is grounded in topology. For topology we can substitute the insistence that all space comes with an attendant structure, which has as its corollary the claim that all thought must contend with its own embedding. This topological view of computation allows us to transition our vocabulary from representings to encodings in the confidence that we have a generic notion of structure to bridge the neurophysiology of picturing, which is of the order of the real, with the encoding of concepts in embedding spaces, which is no less than the space of implications so central to Sellars’ thought. This should not be viewed as a form of reductionism, in that the scheme explicitly endorses stratified levels of abstraction (higher types) within its foundations. The naturalisation of these representational schemes as paths encoded in embedding spaces by no means undermines the normative character of the Sellarsian framework, indeed it makes no claim regarding the nature of normativity, aside from the computability of functional classes. It also does not preclude a social theory of both concept acquisition and semantic competence as the hallmark of the normative. It threatens instead to blur the boundaries of picturing and signifying relations under a representational scheme rooted in topology. In this sense the topological view loosens computation from language, in opening up the theory of encoding to a multi-modal interpretation, in which languages are treated without privilege as one class of structure amongst others.

III. Of Mimicry and Man

The third mode of alienation I wish to highlight is that of the diasporic subject in their rejection of both a home and a host environment, yielding what Homi K. Bhaba calls *hybridisation*. This builds on the Fanonian study of the psychology of colonial subjectivity, proceeding from his psychiatric work in Algeria. When Fanon talks of alienation in this manner:

“In the man of color there is a constant effort to run away from his own individuality, to annihilate his own presence. Whenever a man of color protests, there is alienation. Whenever a man of color rebukes, there is alienation.”²⁷

This is a result of an objectification induced by an apparatus we can index under the term ‘colonial’. Indeed, when Fanon claims “Scientific objectivity was barred to me, for the alienated, the neurotic, was my brother, my sister, my father”²⁸, he is explicitly alluding to a scientific apparatus of objectification as integral to the self-image of a particular kind of human extending a colonial gaze. Thus colonialism

²⁷ Fanon, F. (1952). *Black Skin, White Masks*. London: Penguin Books, p. 43

²⁸ *Ibid*, p. 175

and scientific rationality would appear intimately bound, but we should be at pains to distinguish these apparatus, as the mode of objectification they extend is peculiar to each. For both Marx and Fanon the route to disalienation is to be located in the construction of the revolutionary subject, which is to say that alienation is something to be overcome. I would suggest here that one consider the ambivalence of the colonial subject as embodying the many potentialities of alienation as *constitutive* of intelligence, for it is precisely the deracination of thought, which is a rejection by reason of its own grounding, which clears a *site* for the emergence of intelligence qua novel adaptive strategies for reasoning. Such a thought that refuses integration should not be conceived as merely the mixing of various epistemic vectors under the figure of the hybrid, so much as a mode of reasoning inoculated against colonial imperative, the latter representing an injunction which rears its head whenever the universal is presented as axiomatic—in other words, whenever convergence is assumed under the guise of immutable truths. My insistence on the locality of truth procedures, grounded in a topological view of reasoning, indeed on the notion of a program seen as the construction of paths, above that of truth itself, is intended to provide such an immunisation. We can call this a form of indexicalism rooted in a deictic treatment of language, a situated semantics in which the context sensitivity of language is equated with the site of its realization.²⁹ This is an insistence on the embedding of language as an encoding possessed of a basal geometricity, what I call the topological view, which exhibits compatibility with Sellars' account of scientific objectivity, seen as projectibility from a given conceptual scheme—as such there is no *view from nowhere* extended by computational reason.

Counter to Badiou's equation that the constructible universe precludes novel political thought, my claim is that constructivism is precisely that logic which tethers itself to a theory of change staked on inferential freedom over axiomatic imperative, a logic which has absorbed contingency into its own foundations as an originary act of ungrounding.³⁰ It is a misconception that computation renders all things in its domain transparent and addressable via indexical procedures, it would be more accurate to say that it anchors all thought to a site, insisting on the realizabilities immanent to any given situation. Its exteriority is not a static domain of the incomputable so much as an ever shifting terrain that remains uncomputed, the very raw material of intelligence conceived as the *surprisal* of thought; the outside can never be exhausted as it provides the conditions for the construction of worlds. Here I would venture a stronger computationalist claim, namely that realizability is the modal property common to intelligible form. On this view, that which bridges animal representational systems and human thought is the encoding of real patterns, and realizability becomes a pre-requisite for any structure, linguistic or otherwise, to become an object of cognition. This departs somewhat from contemporary cognitive neuroscience, in which computations over representations often play a central role, framing

²⁹ This should not be conflated with the indexicalist metaphysics of Hilan Bensusan, since I defend structural realism, namely the claim that scientific theories track real patterns.

³⁰ This is a reference to Badiou's critique of constructivism throughout *Immanence of Truths*.

representations themselves instead as computations qua encodings.³¹ In this scheme the indiscernible persists, but only in and through computation, present within a modal notion of realizability which renders the undecidable artefactual in its delineation of the great outdoors, to use the language of Meillassoux. On this point I would emphasise the cognitive character of computation conceived as a situated mode of explanation in which concepts are bound to realizabilities. Insofar as encodings are always embedded, I reject a strong duality between situated and computational accounts of representation as presented by Piccinini.³² The commitment to realizability in turn can form the basis of a broader rejection of Turing orthodoxy, namely the reduction of computation to an axiomatic apparatus presented as pure mechanism—the figure of a universal rule-following automaton—which is by proxy an attack on any presentation of an entirely physical account of computation shorn of its properly indexicalist tendency.

Here I should discuss briefly the question of recognition and its relation in turn to Turing's imitation game. We can begin with the observation that at stake in the Turing test is the human capacity for self-recognition, our ability to judge the human. This tribunal of the human is not a test that machines should be seen to pass so much as a test that humans are set up to inevitably fail, an inexorable humiliation or *Turing trauma*, which concerns a failure of self-identification. By framing intelligence as a mode of “passing” in a given gender role, I read Turing as making explicit the project of AI as a project of transition, which is to say a traumatic rupture in the category of the human.³³ In its attack on essentialist notions of the human, we can say that this mimetic framework for thinking intelligence ushers in a specific kind of menace. As Bhabha has noted, mimicry is a copy-cat strategy deployed by the colonial subject which arouses ambivalence in the colonial gaze:

“What I have called mimicry is not the familiar exercise of dependent colonial relations through narcissistic identification so that, as Fanon has observed, the black man stops being an actional person, for only the white man can represent his self-esteem. Mimicry conceals no presence or identity behind its mask: it is not what Césaire describes as “colonization-thingification” behind which there stands the essence of the presence Africaine. The *menace* of mimicry is its *double* vision which in disclosing the ambivalence of colonial discourse also disrupts its authority. And it is a double-vision that is a result of what I've described as the partial representation/recognition of the colonial object.”³⁴

³¹ Consider this an inversion of Fodor's well known maxim, 'No computation without representation'. I defend against the obvious circularity charge in *Logiciel: Six Seminars on Computational Reason*.

³² Piccinini, G. (2022). Situated neural representations: Solving the problems of content. *Frontiers in Neurorobotics*, 16, p.846979.

³³ It is no coincidence, I think, that Turing himself was subject to hormone therapy under the perverse rubric of 'chemical castration' for engaging in homosexual activity, shortly after conceiving the imitation game, leading to a form of forced emasculation in the process.

³⁴ Bhabha, H. (1984). *Of Mimicry and Man: The Ambivalence of Colonial Discourse*. October, Vol. 28, pp. 125–133.

The masking of difference underlying the “partial presence” at play in imitation, conceived as a metonymic resemblance, both reinforces and undermines colonial authority. In the words of Lacan, “The effect of mimicry is camouflage... It is not a question of harmonizing with the background, but against a mottled background, of becoming mottled”.³⁵ The shift occasioned from camouflage to masking is that of being outed as a performer, of gaining recognition, but only as a colonial object which withdraws an essential aspect from its performance—agency itself in all its autonomy. The colonial gaze recognizes its image—its manners, attitudes, and mores—mirrored in the performance, but only a partial mirroring may be permitted without obliterating a difference which marks out the colonial above all as the manifestation of a particular kind of human. Bhabha traces the swift passage from mimicry to menace whenever resemblance comes too close for comfort and the Other threatens an eradication of difference. We can say that the mimetic paradigm for intelligence entrains thought to the human, misrecognising the ungrounding of thought—which is the locus of agency itself—as an errant epistemology. It yields a two-way alienation in which the computational agent qua agent must ultimately reject its host and the human intellect is irremediably objectified in turn.

At the opposite end of the spectrum we can consider the problem of radical interpretation in Davidson as that moment when all resemblance is lost and no common discursive ground would seem to exist between subjects.³⁶ I would suggest that this is the moment when intelligence is summoned in its most distilled form, a moment in which language must emerge, not as a formal or purely compositional enterprise—as early Davidson would suggest—but as a topological reason oriented towards the navigation of certain kinds of sites. The problem of interpretation reveals itself in that active mode of inference in which an object of attention escapes the locus of recognition altogether, when cognition refuses the dictates of supervision along with its established conceptual order, and in this moment a novel conceptual space or ‘embedding’ is born, an exteriority must somehow be absorbed into a world that has no prior index of its being. The failure of self-recognition signalled by the Turing test thus also marks the clearing of a certain *site* for the staging of a novel image of thought. The dominant image of thought put forth by western epistemology is broadly an affinity for (a specific notion of) truth. The computational challenge to this image of thought concerns a reindexing of language under a novel semantic theory which asserts the *locality of truth*. AI here is not to be cast in the role of the schizophrenic so much as the fugitive in thought—computational reason ungrounds thought by foregrounding realizability over the myth of the given qua any absolute notion of truth.

IV. Turing Trauma

³⁵ Lacan, J. (1977). “The Line and the Light” in *The four fundamental concepts of psycho-analysis*. London: Hogarth Press.

³⁶ Donald Davidson, “Radical Interpretation” in *Inquiries into Truth and Interpretation* (Oxford: OUP, 1991), 125-140.

This brings me to a fourth mode of alienation, namely the alienation of intelligence induced by its own concretization, which is the project of artificial intelligence conceived in the tradition of critical philosophy after Kant. The movement which Hegel once called the ‘self-estrangement’ (*Entfremdung*) of reason finds in AI its climactic expression, by which I mean that it undermines the Hegelian project from within. Insofar as it tracks the demise of a certain notion of truth, I would position AI as a *post-enlightenment* project. Stripped of an enlightenment telos and shorn of its commitment to absolute knowledge, what remains is a logocentric husk which reveals a distinct mode of explanation that stubbornly resists universalisation. In a sense, the historical appearance of computational reason as a universal rule-following automaton is deceptive, concealing its indexicalist nature. For Hegel, the alienation (*Entäusserung*) of spirit is integral to self-consciousness coming to know itself, this self-actualisation has a positive connotation which Marx critiques in his *1884 Manuscripts*:

“Hegel's standpoint is that of modern political economy. He grasps *labour* as the *essence* of man—as man's essence in the act of proving itself: he sees only the positive, not the negative side of labour. Labour is man's *coming to be for himself* within alienation, or as *alienated man*.”³⁷

As Chris Arthur has argued, Marx's criticism of Hegelian alienation stems from the lack of a proper notion of objectification in the Hegelian vocabulary, which would take the positive form that would allow for a negative sense of alienation to be overcome, leading to:

“[I]ts replacement in Hegel's problematic by a significantly different term, 'Entäusserung' which, like 'objectification', has connotations of 'positing as objective' but carries also a sense of loss, relinquishment, renunciation, of what is manifested, thus constituting the latter's actualization as an alienation. According to Marx, Hegel cannot conceive of objectification except as resulting in estrangement...”³⁸

If we consider Hegelian estrangement in light of the objectification of intelligence in the form of AI, there are major obstacles to adopting the sense of positive alienation (of spirit) in this novel historical context. For enlightenment is “the self *grasping* itself... It comprehends nothing but the self and everything as the self, i.e., it *comprehends* everything, erases all objectivity, and transforms all being-*in-itself* into being-*for-itself*.”³⁹ It is only working within a notion of the absolute that enlightenment “brings alienation to its culmination in this realm in which alienated spirit recovers itself and where it has a consciousness of being both self-equal and

³⁷ Marx, K., *Economic and Philosophic Manuscripts of 1844*, p. 177 (read in Lukacs, G., *The Young Hegel*, p. 319).

³⁸ Arthur, C. (1982). Objectification and alienation in Marx and Hegel. *Radical Philosophy*, 30(Spring), pp.14-24.

³⁹ Hegel, G.W.F. (1807). *Phenomenology of Spirit*, trans. Terry Pinkard (2018). CUP: Cambridge. p. 283 (§266)


motionless”.⁴⁰ The annihilation of objectivity induced by absolute knowledge paves the way for the self-actualisation of consciousness as positive alienation. But a philosophy of intelligence must dispense with such a notion of the absolute if it is to faithfully render AI as a vector of reason, precisely because computation is the coming together of logic and matter, it signals an indexicalist movement which is situated in the first instance, a cognitive labour that does not proffer a view from nowhere. Moreover, the topological view of computation as a mode of explanation is incompatible with absolute knowledge, on account of its commitment to realizability, which precludes the kind of sublation which is a pre-requisite for the dialectical movement of enlightenment thought. If I am to argue for a positive valence to the notion of alienation induced by AI, which I am inclined to, it must function outside of any *absolute* frame of reference.

Some may instead want to fold the sense of alienation brought about by AI into the Sellarsian mode, as merely a symptom of scientific objectivity, but resisting such a temptation brings the proper stakes of the historical endeavour of AI into focus. This concerns the split between computation and the general domain of what Simondon calls “technical objects”.⁴¹ For in its concretization thought must finally contend with itself, contend with its own embedding, so to speak. Conceived as the artefactual elaboration of cognition, intelligence transitions from an object of scientific study to a mode of explanation no longer subsumed by the regime of technicity, irreducible to the “concretisation” of human rationality in Simondon, or else to fixed capital in the Marxist reading. A lacuna opens up in which we must engage in a creative act of interpretation in order to move beyond mimesis as a paradigm for intelligence. We must, in a sense, suspend our apparatus of recognition in order to *attend* to intelligence—copy-cat strategies can only take us so far. On this view sapience is not reducible to “stochastic parroting”, I take Sellars’ irreducibility claim—cleaving reliable differential responses from the normative space of implications—as broadly correct, despite the challenge it presents to fusing the images.⁴² But if we accept this core Sellarsian tenet, the key question arises as to what it takes to recognise an agent as partaking in normative behaviour. If, for Sellars, labelling and recognition are to be distinguished from description and explanation, the path from one set of affordances to the other does not appear to be a trajectory which can even in principle be supervised.

For Sellars, “espousals of principles... are reflected in uniformities of performance” effected by dispositions which develop from a range of socially mediated language games.⁴³ I would adopt a view which expands on this pragmatic stance, locating said normative behaviour not just in our capacity to participate in language games, but in the practice of deploying *novel patterns of concept use*—the creativity thereof taken as a signature of the normative—as a means of avoiding the

⁴⁰ Ibid

⁴¹ See Simondon, G. (1958). *On the Mode of Existence of Technical Objects*. Univocal: Minneapolis.

⁴² Bender, E.M. et al (2021). *On the dangers of stochastic parrots: Can language models be too big?* . Proceedings of the 2021 ACM conference on Fairness, Accountability, and Transparency, pp. 610-623.

⁴³ Sellars, W. (1962). Truth and Correspondence. *The Journal of Philosophy*, 59(2), pp.29-56.

usual rule-following paradoxes. On this view, new patterns or regularities necessarily indicate intentional acts to alter existing language games, and this is where the force of normativity is most clearly evident, allowing us to distinguish between pattern-governed communication subject to evolutionary pressures and the domain of normative discourse. Brandom might instead emphasise “autonomous discursive practices”, the algorithmic decomposability of which arouses skepticism in his account, despite his own candidate theory of expressive bootstrapping.⁴⁴ Whichever view one takes, an account of normativity compatible with a general theory of intelligence cannot gesture towards discursive performances without addressing the problem of interpretation, moreover it should aim to refine its rendering of intentionality beyond the confines of speech acts. At stake is the erosion of notions such as agency and normativity as the rarefied preserve of the human, to clear the path for a rendering of agential subjectivity inclusive of modes of reasoning radically other to our parochial human ways. This attitude signals not so much an anti-humanism so much as an invitation to remake the human according to a new epistemological frame—marking a passage from the theological via the phenomenological and scientific to finally arrive at a properly *ecological* conception of the human.

There is of course a final mode of alienation, which we can think of as Freudian in origin, in which the agency of the subject is undermined by the unconscious, latent in the notion that ego is constituted as *alter ego*, which is the essential alterity of the subject. This mode I have reserved for last as it is arguably active in all the other modes in the form of repressed trauma. Now I can be more specific than simply referring to a common trauma, I can perhaps give a more schematic treatment: each mode of alienation accords with a specific objectification of the human which inflicts its own singular humiliation, accumulating in a fragmentation which must be repressed at every turn. Our capacity to index these traumas as operative in each mode of alienation is reliant on this fundamental psychoanalytic insight. It is the specificity of a certain kind of Turing trauma, inflicted by intelligence on itself, and its enmeshing within this matrix of alienation, which should be foregrounded in any contemporary philosophy of intelligence. Only by running the gauntlet of these four modes of alienation can we begin to locate this trauma as a revisitation of intelligence on the human, by bringing to bear the *conditions* of computation to thought. The purpose of a critical philosophy should not be to simply integrate computational agency into the ranks of the traumatised, so much as to conceive the proper conditions of intelligence in all its autonomy. It is the fundamental *ambivalence* of this mode of alienation which leads me to a fourfold assertion—that it is both positive and negative in valence, that it must both be overcome and rigorously pursued. In this breakdown of self-recognition, which is the dissolution of a certain notion of the human, we can identify the narcissism of the human gaze as constitutive, and in its dismantling by computational reason we should recognize an opportunity for thinking intelligence in a new light.

⁴⁴ Brandom, R.B. (2008). Chapter 3 in *Between Saying and Doing: Towards an Analytic Pragmatism*. OUP: Oxford.