URBANOMIC
sequence PRESS

How to Make Yourself an AGI

Reza Negarestani expands upon the major themes of his new book *Intelligence and Spirit* in this edited and expanded version of his conversation with Robin Mackay at the launch of the book in New York in November 2018.

Robin Mackay: Before I invite Reza to discuss some of the themes and ambitions of *Intelligence and Spirit*, I'd like to give a short, somewhat personal introduction to the book by relaying my impressions of *Intelligence and Spirit* as an editor and as a reader.

To first give a very broad outline, *Intelligence and Spirit* is a painstaking interrogation of the notions of intelligence and artificiality. It begins with the question: What would it mean to speak, philosophically, of an *artificial general intelligence*, an AGI, whose capacities would, at least, equal our own? Is the human the correct or the only model to start with in trying to conceptualise such an intelligence? And, looking at the various proposals and programmes of research into artificial intelligence, and increasingly into the broader notion of artificial *general* intelligence, are we able to clarify what exactly we mean by intelligence? The book then culminates in a vision of philosophy itself as a program for the artificialization of intelligence, or a program for artificialization *as* intelligence.

The central chapters of the book use a 'toy model' to determine the conditions of possibility for the construction of such an intelligence. What would a simple automaton with sensors need to be equipped with in order to attain what we would recognise as intelligence? This is a kind of Kantian thought-experiment in transcendental philosophy, but one inflected, crucially, by two of Hegel's crucial insights. Firstly, that intelligence, or *Geist*, can only be defined *functionally*—in terms of what it *does*; and secondly, that it does not belong to the individual but always implies a community, a *collective mind* and, in particular, a shared *language*. And there is a strong emphasis on functionalism, on collectivity, and on language, although none of these are treated in an entirely familiar sense, and with their convergence there emerges and entirely novel philosophical approach to the intertwined notions of intelligence, artificiality, and the practice of philosophy itself.

In order to determine further this conception of a *community* of automata, the toy model is further specified using concepts drawn from computer science, research into artificial speech, and interactive logic, to mention just a few; in short, by selecting those currently available resources best able to satisfy the conditions for the emergence, within this community, of artificial general intelligence.

Answering to the profound and apparently rather diffuse philosophical questions which the book opens up, then, we have this simplified model that is functionally specified, experimented with, with parts taken from various places being added, swapped out, and refined. Returning from this functionalist staging of the problem, the final chapter returns to a philosophical vision in the

grandest sense, a vision of philosophy as a project at once transcendental and functionalist, as *having always been a program for the construction of artificial intelligence*, in the form of a striving to understand what we are as intelligences, how we can live up to our capacities, and how we can make ourselves better—because intelligence cannot be separated from its tendency to upgrade itself, which Reza interprets in terms of Plato's concept of the Good.

What's startling about the use of philosophy here is not just that it is articulated with many other disciplines, with other practices, but also that, rather than this being a philosophy *of* artificial intelligence—as if it involved some relation of application to an independent object, it becomes evident that philosophy itself, as a programme of constructing *an outside view of ourselves*, of understanding how changes in our self-conception necessarily lead to the transformation of our collective modes of acting (intelligence *is* what it *does*), is already a programme for artificialization, for artificializing ourselves. And in particular how Kant, and Hegel, and the program of German Idealism in general, unknowingly furnish the functional blueprints for a future artificial intelligence.

This is obviously then not straightforwardly a book about AI, about the prospects of AI, about its technical possibilities as seen from the present, in a kind of pop-science mode. Instead it's using the problematic of AI as a way to develop a more general *philosophy of intelligence*. And this in turn involves us in some acute political, ethical, and existential questions about ourselves: about the capacity of we humans to merit the name of intelligence, about the process of allowing ourselves to be invaded by other intelligent processes; about allowing intelligence to *grow out of us*, or, inversely, our ability to *make ourselves* an AGI.

Part of the importance of this book lies in the way it addresses an amorphous murk of uncertainties, fears and hopes that we are surrounded by in the present hysterical moment. I've always loved Deleuze's suggestion that every true philosophy begins with a cry of anguish; that this is what it means to need a concept: to have something to scream. And we have plenty to scream about. The piecemeal absorption of the human, and human sociality, into some other form of unfathomable machine; uncertainty and proliferating doomsday/singularity scenarios; the apparent erosion of any viable concept of human agency; a total dependency on technology paired with a chronic difficulty in upholding any liberatory idea of technology-as-progress; and—perhaps most upsettingly—we seem to also be surrounded by a wilful squandering of what little human intelligence is left…. Not to mention, for those of us who have invested our time in philosophy, the question of whether its resources can possibly address this set of apparently ineluctable processes which seem refractory to any sort of philosophical judgment.

What is extraordinary in *Intelligence & Spirit* is Reza's diligence and conscientiousness in addressing these questions, his resolute refusal of both hysteria and the illusion that we can achieve a sort of instant relief through philosophy. Throughout the book he insists on tracking these unsettling questions so tenaciously that they are often transformed into something unrecognizable. Sometimes we're led, from the 'big' questions we want to ask, into dauntingly technical expositions of recent fields of research. I think there are few people who will read the

book without learning something about some field of which they were previously unaware. Whether it's Hegelian spirit or nonmonotonic logic, there will be some new encounter here.

I see the story arc of the book like this: a set of forays out from this ambient amorphous anguish and the huge questions it poses, into increasingly determinate and actionable *engineering questions*, via a series of meticulously plotted navigational paths or zooms between different layers or scales of the same questions. It becomes a kind of slow-motion scream in which the acute emotional urgency is attenuated, and all kinds of unexpected details emerge. One result of this approach is that it is difficult to separate the book's methodology from the task it sets out to achieve, because in a sense it is also a book about methodology, about *how*—and how not—to do philosophy today, how to make something constructive of the scream without betraying it.

The kind of thought that is going on in this book marks a decisive departure from the expectations of a certain type of philosophy or theory with which most of us are familiar, and in particular, to say it somewhat schematically, from a certain coupling of the ontological and the political which, in the context of this ambient political, existential, technological distress, has become something like a salve, a philosophical arnica for the afflicted. Because an ontology—whether it's mathematical, object-oriented, new materialist, or whatever…an ontology always allows us to say, firstly, *everything is* x. This overextension of a single conceptual articulation affords us a clear and stabilised field for thought and action so that, from there, we can move forwards within the bounds of this unified vision of the world to say, secondly *if only we could do* y—that is, the idea that some great shift in our conception of the structure of things could potentially relieve us of the burden of a tangled, layered, enmeshed situation now seen to be, ultimately, simple and radically subvertible.

Intelligence and Spirit absolutely departs from these cartoonish ontologies whose emancipatory promises are as brittle as their concepts are overextended. In this book you won't find any *everything is*, nor any *if only we could*. In fact *Intelligence and Spirit* even deprives us of the stability of any recognizable *we*, because it demands that we rethink intelligence as a virtual collectivity, one that is yet to be fulfilled as a concrete project. The short-circuit of the ontological and the political is effectively replaced by a *functionalist* approach to problems inherited more from systems engineering than from philosophy—in this sense, the book sees Reza going back to his former life as a systems engineer, bringing the sensibilities and the responsibilities of the engineer into philosophy.

The wager here is that, if there is any optimism to be had in this situation, it begins with a commitment to the complexity and multilevel nature of the problems, and in particular an acknowledgement that, in order to map out the question of intelligence in all of its aspects and its multiple scales, we will have to assemble and articulate different modes of knowledge, understanding that intelligence is neither simple nor homogeneous.

Rather than a grand architecture that would afford us the classical satisfactions of philosophy, what comes out of the book is a tentative patchwork of specifications for interlocking functional modules that could be swapped out later in light of new advances in multiple fields—or what

Reza has called a kind of *philosophical Lego*, big toy models. With this approach there begins a long-term labour, both arduous and playful, the labour of the continual conception and transformation of ourselves. Or rather, we are invited to recognise the historical achievements of this labour (the 'labour of the inhuman' as defined in a previous work), and to explicitly address the continuing of this vector forward as a concrete task. Reza even speaks of our *responsibility* as intelligent beings to continue to artificialize ourselves since, deprived of its movement of continual expansion, trapped in a box, whether it's a human frame or a computational gadget, intelligence ceases to be. This in a sense is the ultimate extrapolation of Brandomian pragmatism: once we conceive correctly of the nature of our intelligence, this *conception* immediately presents us with a concrete task of self-*transformation*, failing which, we consign ourselves to admitting that we are fundamentally disinterested in thinking at all.

In presenting us with this stringent relation between conception and transformation, the ultimate questions the book asks us are the following: Is our scream more than a plea for comfort, are we prepared to follow its consequences, and above all *are we ready to start building*? Are we any longer willing or capable of taking on a philosophical challenge, or are we happy languishing within our cognitive niches, taking up positions that are vulgarisations of historical ideas left unrevised by contemporary contributions to knowledge?

This slow-motion scream born of anguish and frustration at the shortcomings of actually-existing-intelligence ends up giving us a minimal but robust form of philosophical optimism that results from ruling out well-worn philosophical consolations, quick-fix tropes to which we may have become all too accustomed. To characterise this subtle *optimism* I could cite a recent tweet by the writer and technology consultant Ventakesh Rao: 'Systems engineering is the art of turning moral crises into architectural opportunities.'

Intelligence and Spirit demands that we revise our conception of time, placing ourselves within a historical sequence that far surpasses our individual lifespan, but in which we are nevertheless agents. It examines the problem of intelligence with the cheerful pragmatism of an engineer who, unable to simply pop the hood of his own braincase to peep inside, gets out his Lego and begins to build a model.

*

Robin Mackay: Unless you have any response to my mischaracterizations, Reza…

Reza Negarestani: No, at this point I'm sure you're more well-versed in *Intelligence and Spirit* than I am!

RM: …then perhaps we should begin with the basic terms in which the book is framed. Now, philosophy of mind is a recognised specialism within philosophy, and we could cite various philosophical projects that propose a theory of knowledge or of reason. But is there something more specific that you understand by 'philosophy of intelligence'?

RN: Intelligence as philosophically understood is a higher-order domain than the mind. We have intelligent behaviours (which are prevalent in nature), then mind (the organ or dimension of structure, or a set of faculties necessary for structuring the world or rendering it intelligible) and then intelligence. Intelligence is in the business of determining what to think and what to do with the capacities of the mind, using the faculties of mind to constantly renegotiate its place in the world (all possible sectors of the intelligible). So in a sense, philosophy of intelligence takes as its premise the mind as the organ of structure, in the vein of the transcendental turn. Yet what philosophy of intelligence arrives at is the question of *what can be done with the mind*, all things considered. Throughout the course of the book, then, we see that this problem cannot be coherently answered if we take the paradigm of mindedness as something stable, if we regard the list of faculties and transcendental structures as immutable. So philosophy of intelligence not only renegotiates the very concept of mind, but also investigates the prospects of what can be done—theoretically and practically—with a concept of mind not as a thing but as an ongoing project. To this extent, intelligence is what engineers its reality by enriching the very reality of which it is a part. It becomes a concrete movement that graduates like a child from the transcendentally passive paradigm of settling on what the mind is, what reality is, and how they are related to one another, to the domain of transcendental proactivity where new forms of intuitions are put forward, the given list of faculties is renegotiated, and the limits of theoretical and practical cognitions are revised by refashioning how the mind and its correlative reality appear to us.

Philosophy of mind, of course, is one of the most ancient strains of philosophy. It begins explicitly with Parmenides and Plato, going through various metamorphoses throughout the Middle Ages, then Descartes, then Hume, then Kant, Hegel and so on and so forth. I think we should also talk a bit about German Idealism at this point, precisely because the book purports to hijack certain resources from German Idealism without necessarily strictly abiding by its theses.

I think that, within the history of philosophy, German Idealism can be said to be a form of critical project built at the intersection between philosophy of action, philosophy of knowledge, and philosophy of mind. As a program, it can be defined in terms of how the relations between mind, knowledge and action can be elaborated into a global or comprehensive system within which these interconnections be methodologically investigated using different tools, but also modified so that the system as a whole can evolve. So when I say the book is about German Idealism, I mean it is a work that strives to elaborate various interconnections between knowledge, action and mind. And from this perspective, philosophy of intelligence is about how to modify these links rather than regarding them as set or permanent; how to imagine a different course of evolution for a system that primarily addresses the perennial questions of philosophy: What to think and what to do (with all the epistemological and methodological intricacies that naturally come with those big questions).

Now let me a take a step back and unpack your question, ‘Why should we talk about philosophy of intelligence as a contrast to philosophy of mind?’ I think, at least in the book, there is a hierarchy of what you might call philosophical ‘systems’, with corresponding methodologies at

each level. At first we are in the domain of intelligent behaviours: problem solving, task-orientated attention systems, and so on. Today, cognitive sciences—but also computer science—have shown us that many things that we thought were unique to us can in fact be captured by certain kinds of algorithms, which you might call 'artificial realisations' of certain processes—pattern recognitive processes—that are prevalent in nature. To appropriate Nick Szabo's claim, the real competition might be between the qualitative picture of the mind and the view that every aspect of mindedness can be realized algorithm-by-algorithm. The argument laid out in *Intelligence and Spirit* is that intelligence is not to be equated with these intelligent behaviours. They form the bottom level of the hierarchy.

And then we have philosophy of mind. Philosophy of mind is introduced in terms of the transcendental turn, in terms of how critical philosophy emerged from precritical philosophy. You can think of precritical philosophy as a philosophy in which nature, or a supposed reality, gives us a structure. It reveals its own secrets to us spontaneously like a tell-tale heart. But with the transcendental turn in philosophy, the situation switches to a different paradigm. In the previous paradigm, the precritical one, the mind was a *tabula rasa*, a blank slate, and the world was data. With the transcendental turn, the mind becomes the structuring factor or the data, and the world becomes the blank slate.

So philosophy of mind is essentially what you might call a philosophy that deals with the broadly understood conditions of possibility of structure, with the understanding that structure can also be understood as synonymous with intelligibility in the broadest possible sense: the intelligibility of the world, but also the intelligibility of our thoughts, practices, and values with regard to ourselves in the world. What might strike some as controversial here is that this paradigm of mindedness in terms of the question of structure is qualitative, not quantitative. Mere accumulation of formal learning algorithms cannot yield something like epistemic and objective criteria, systematic modes of appraisal and revision. It is how these natural or quantitative processes are qualitatively integrated that is important, not the simple fact that we are just a bundle of such processes.

And then on the next level, philosophy of intelligence takes the question of the intelligible and, correspondingly, philosophy of mind one step further by asking what can be done with the intelligible and toward what ends—turning the intelligible into a concrete labour in which the historical illusions of totalized and completed ideas of the mind and the world are dissolved, in favour of reworking the boundaries and enriching reality.

RM: In the book your reappropriation of the history of philosophy is enmeshed with the thoroughly contemporary question of artificial intelligence. So can AI research really learn something by going back to the transcendental turn, to German Idealism, or even to Plato? What do the various projects of AI have to gain from going back to what would appear to be centuries-old, perhaps even obsolete, philosophical conceptions of mind?

RN: I think, to answer this question, it would perhaps be beneficial to say a little bit about the original ambitions of AI (the artificial realization of the mind) and how it later turned into hard AI

(task specific algorithms) and how, in response to the abandonment of the original ambitions, AGI research was proposed.

AI started to emerge in the early twentieth century, and became quite evident as a field in the mid-twentieth century. When we read the unpublished essay by Alan Turing, 'Intelligent Machinery', where he introduces the idea of a child AI, even within his own classical Church-Turing paradigm of computation, he is already thinking not about specific algorithms—task-orientated, problem-solving algorithms, or what you might call intelligent behaviours—but about the paradigm of mindedness. He actually thinks of AI as a philosophical problem—and yes, I absolutely do think that AI *is* a philosophical field, ultimately—but the thing is that, here, a problem arises after this initial ambition in the sense that, just like philosophy, just like any field, when we have some sort of ideal, when we have some sort of general theme, the concretisation of this idea is beholden to the available methodologies and models. And of course, throughout the next few decades, the kinds of models of mindedness that were proposed all failed to satisfy this initial idea of AI, which is *absolutely* the idea of AGI: a human-level AI, an intelligence capable of doing *anything* we can do, if not more.

Nevertheless, over the past few decades, there has been a divergence between what you might call 'narrow AI' and AGI, narrow AI being the idea of task-orientated and mere problem-solving algorithms like navigation in a maze, or even making a coffee that tastes really good to us humans. But why on earth would an AGI want to make a coffee that tastes good to the established transcendental structure of human experience? That's just design from the point of view of human experiential biases. It's as if AI hasn't even reached the realm of the philosophical controversies between Descartes, Hume, and Kant. It has not yet exited the kingdom of precritical medieval hubris.

However, AGI is the grandchild of the initial idea of artificial intelligence qua mindedness, making an agent or a multi-agent system where the agent goes through various stages of development, infant, child, etc. The agent becomes capable not just of navigating its immediate world, but also of postulating a different and far more expansive world by arriving at 'new facts of experience'—a task which, as Boltzmann remarked in his lectures on gas theory, requires a thoroughgoing critique of ourselves as observers and agents that tend to cherish their entrenched biases and evolutionarily given methods. To entertain the possibility of new forms of experience through which the scope of intelligible reality can be expanded—this is the task by which both the human and AGI are defined. Settling on the existing structure and facts of human experience as the benchmark for what is human or what is AGI only attests to how badly we as humans have failed to reimagine ourselves as humans. Conservative ideas and tests for AGI are nothing but conformist counterparts of the parochial concepts of the human. AGI and the human become almost undistinguishable in their task to overcome such instances of givenness and parochialism.

So yes, why should computer science or, for that matter, research in AI, go back to these antiquated (*so-called* antiquated) philosophies? Because asking *What is mind?* is already a philosophical question, and it is a badly framed one, precisely because the idea of mind is as

vague as the idea of body in the Cartesian system. We just don't know what this question is about! I think that, starting with Descartes, even though he endorses a metaphysical dualism between mind and body, philosophy begins to address a series of questions which put us in a better or more coherent orientation that enables us to answer this question in a more refined way. With Kant, it becomes even more refined, in that, for Kant, 'mind' is no longer a *thing* but a system, a multi-level system. At each level, we are dealing with different kinds of constraints, different kinds of rules. It is not a *unified thing*—it is, simply, *not a thing*. It is a functional hierarchy whose different levels need to be approached by different models, methods, and descriptive vocabularies. This multi-level picture is far more in keeping with the new research on AGI and even the concept of computation in theoretical computer science than with classical forms of research in AI which in one way or another take the mind as a flattened thing or structure: If you come up with that master key, you will surely unlock the powers of the mind. There has never been, never is, and never will be a master key. Mind is the domain of many gates, each requiring different keys.

RM: But does that justify the notion of 'mindedness'? How can you justify positing it without it being a kind of spiritual supplement, an extra magical ingredient which goes into the mix alongside the algorithmic reproductions of a whole raft of intelligent behaviours?

RN: The problem that arises here is that that no single specialized algorithm can actually do the job of conceptualisation, be it in the context of the ordinary natural language or in the domain of scientific theory construction. In fact, when an AI researcher talks about mindedness or intelligent behaviour, he or she is modelling that intelligent behaviour, the function of that intelligent behaviour, however implicitly, *analogically with respect to the human mind, how we reason theoretically and practically*, and with regard to the conceptual activities that *we* do, but which sometimes we are unaware of doing.

Philosophy of mind, or cognitive science in general, tries to make these conceptual activities explicit so as to show that, when we are talking about any kind of index of intelligence in the world, we are actually talking about it *in analogy* with human conceptual behaviours which are holistic. Of course, the methodological problem of artificially realizing these conceptual behaviours is still an open question. Can we capture them statistically, computationally? My answer would be yes, but that requires different levels of statistical description, different models of computation in tandem with the idea that when we are talking about mind, we are actually talking about a range of complex behaviours which are generated by a large number of very different processes and mechanisms.

RM: In what sense are human conceptual behaviours 'holistic'?

RN: In the sense that, when we talk about conceptual activities, we're talking about the content of our inferences and judgements; and so we are already presupposing—even though we might not be fully conscious of this fact—that every judgement we make about X is inferentially linked with judgements about not-X. To say *this is blue* means that *this is coloured*—another judgement—but also that *this is not red*, *this is not white*, *this is not black*, *this is not an intrinsic*

property of a metal element. So, you see, we are working in a holistic network, a web of conceptualisation and inferences. This web requires something more than the task of a specific task-oriented algorithm. Whatever we do is always implicated in this web, whether we are conscious of it or not. The critical task is to become conscious of what we say when we say X, what premises or presuppositions and what consequences or implications does it hold? To assert *X is* in this sense we must determine two things: what X is *not* in the web of inferential connections, as in the rudimentary example of the judgement 'This is blue.' And what is followed by and what follows from asserting X?

Of course, the idea that we speak in terms of analogy with our own conceptual behaviour raises a further question: If every behaviour we recognize in the world is recognized in analogy to our own paradigms of theoretical and practical reasoning, then does this mean that we are infinitely projecting our image into the universe—particularly if the way in which we reason is tethered to the particular transcendental structures of our experience (neurological diversity, language, culture, etc)? This is a sceptical question that should be sufficiently investigated. And yes, I think this sceptical question is what is missing in the current AGI research. To frame these questions about AI and to respond to them, we have to once again get back to the works of critical philosophy, that is, to *understand* problems before trying to *solve* them.

RM: In the model proposed by Kant's critical works, we have a first example of what the movement from the relative simplicity of precritical metaphysics to what even the most admiring of Kant scholars would admit is a more messy model. The question being: given that we have this kind of apparently consistent experience, what conditions must have to be in place in order for this to be the case? In asking that question, he develops an incredibly circuitous, complex, multilevel model, which in some respects looks far less satisfying than the metaphysics it replaces.

In that sense, he is the first to treat mind as a system, a system which, in a certain sense, he seeks to reverse-engineer. How does this approach help us avoid simplistic or overextended concepts of mindedness that we might otherwise be tempted to use? I've already referred to the fact that in this book you have finally outed yourself as an engineer in philosopher's clothing….

RN: Basically, even though my background is *actually* engineering, this whole relevance of engineering with regard to philosophy only came extremely late to me—I was already thinking about it, but in quite a naive sense. It came to me by way of the late work of Rudolf Carnap—who is, by the way, a hidden figure throughout the entire book. Carnap began as a logical empiricist. He was a positivist—his 1928 book *The Logical Structure of the World* is a monumental work of logical positivism. However, by 1934, Carnap had fundamentally betrayed the original theses of the Vienna Circle and its vision of positivism. He had moved towards a fundamentally different vision which he continued to refine towards the end of his life. Essentially, Carnap's main idea, what bothers him—what you called 'ontological cartoons'—is that philosophy, in the traditional sense of big ontological and metaphysical questions, is always, as he puts it, 'the opiate of the educated'.

So we say: *What is life? What is intelligence? What is mind? What is justice? What is the good?* Big questions that don't just excite graduate students, but also humans in general! And then, under the auspices of such questions, we come up with such tantalising answers: *If we could only do this, we will have answered all of our questions!* Big ideas beget practical illusions of grandeur. But Carnap instead thinks about what he calls 'conceptual engineering'. Carnap's 'conceptual engineering' is the idea that these overarching upper concepts are all *vague concepts*—what he calls *explicandum*, in the sense that these concepts mean different things in different contexts for different language-using agents. In fact, many incommensurable questions can be posed under the apparently unifying facade of these concepts. So he proposes a process, or an ideal of an engineering process, that he calls *explication*, and which would turn these vague concepts into more refined concepts which are called the *explicata*.

RM: Yet Carnap thinks that there are, even so, real philosophical problems, correct? This is not the same programme as Wittgenstein's dissolution of 'pseudo-problems' by treating them as mere problems of language.

RN: The Vienna Circle Carnap comes from a Wittgensteinian position, but after *The Logical Syntax of Language*, after a fever that leads him to sever his last links with the Wittgensteinian strings, Carnap turns against both his earlier logical empiricist commitments and his Wittgensteinian influences. Why? In a sense, for both thinkers, the question of language is the question of structure. And if one posits something outside of the dimension of structure, one ends up peddling those pseudo-problematic nasties about immediate knowledge of the world that are the hallmarks of precritical philosophy. In fact, even logical empiricism, Carnap's earlier position, is fundamentally against the kind of naive empiricism that Sellars admonishes as an instance of the myth of the categorial given. But then why is the later Carnap the epitome of anti-Wittgensteinianism? The answer lies in how they understand language. For Wittgenstein, the limits of language are the limits of the world and existence. For Carnap, after attending the Gödel seminars, language is the ocean of meta-languages and meta-logics, the unbound ocean of possibilities. To see the limits of language, one does not step into the extra-linguistic, but must adopt a new meta-logical position with regard to language, inhabit a new more expansive language. Even though Wittgenstein is apparently anti-Kantian, he endorses the main thesis of Kant regarding the subordination of language to the intuitive, especially in his picturing theory of language. Carnap unshackles the vision of language from this Kantian metaphysical cage. For Carnap, language is not about the world, or anything extra-linguistic. Anyone who says otherwise will have to pay a metaphysical high price. One can say that Carnap in fact breaks away from the metaphysical dualities of mind or language and the world by reinventing their classical problems on the level of theory as a system of object constitution.

RM: This is something that is covered in detail in the book, where language plays an important role—indeed, it is hailed as the *Dasein of Geist*!— but where you anticipate artificial languages that could outstrip natural language both syntactically and semantically: the domain of artificial general language.

RN: Yes, language generally understood—i.e. beyond the existing scope of the ordinary natural languages—is the dimension of structure, and structure is a matter of worldbuilding. We can only expand our representations by expanding the structure, our resources for world-representation are indebted to our resources of world-building, toying around with the possibilities of structure, diversifying it and enriching it. Of course, this raises the question of how such free play can be related back to our rudimentary experiences so as to avert the

danger of naïve idealism or logicism in the vein of *Aufbau*. This is the question of the so-called 'protocol sentences' and the possibilities of a true physicalistic scientific language which I don't have time to elaborate here. Essentially, the mature Carnap thinks like Poincare: How can we stop being the 'victims of our particular habituations'? How can we diverge from our so-called habituated facts of experience by constructing new experience-constituting languages or paradigms such that we can habituate our experiences to new objective ways? This is the revival of the Enlightenment question. Carnap admitted that the Enlightenment paradigm has been corrupted, it has become a recipe for conformity to the order of *is*. But the real ambition of the Enlightenment as he understood it is to move from the order of *is* to the order of what *should be* or what *might be*. The principle of tolerance and the idea of diversifying the formal artificial languages are instances of moving from what is the case (in the order of appearances) to what might actually be the case.

We usually think of Carnap as that positivistic or radical conventionalist guy, someone who thinks rules alone are sufficient. But no, to the extent that he thinks rules alone are not sufficient, he is not conventionalist; and to the extent that he revises his thoughts about how elementary experiences relate to linguistic sentences (of a broader linguistic-logical domain), he is not that logical empiricist guy. In any case, the late Carnap is against Wittgenstein. His position is something between Hegel's vision of language as the Dasein of Geist, Poincare's idea that new structures make new experiences just as Riemannian geometry opens up new spaces for the observer, and Helmholtz's emphasis on perception and sensory processing, without ever eliding the distinctions between such views or expanding the conclusions reached by one insight to another.

RM: So, for Carnap, the refining of 'big' philosophical questions into engineering problems is quite different to a dismissal of them as 'pseudo-problems'.

RN: Yes: for instance, Erwin Schrödinger asks, in his celebrated book, *What is Life?* But what does that mean? Life means different things at different scales and contexts, just as the concept of hardness for an engineer means different things at different scales of a metal beam. At the level of macroscopic elasticity, at the level of the crystallographic structure of the metal, at the level of the nanometric scale-length, there are many different concepts of hardness, which are not commensurable with each other. So an engineer always wants to know the exact context and scale of the question that is being posed. When you're talking about mind, it is not—as Kant would have said, I'm sure—it is not about a *thing*, about a unified thing, a uniform thing to which we have full access. No, in fact, we just don't know what this question means, unless and until we try to refine the concept at different scales, each according to its own constraints. *May each explanation and description of the mechanisms and functions find its own scale*—that's what the late Putnam endorsed as a sort of recipe against both greedy reductionist and anti-reductionist approaches, prevalent in analytic and continental philosophy respectively.

RM: This multilevel view involves leaving behind what you call the 'flat picture' of intelligence for a more nuanced view of the complex task ahead. And this risks bringing into play more psychological reflexes: anyone who proposes such an approach is fated to be regarded as a philosophical spoilsport by those for whom the 'bigness' of those concepts is precisely what is compelling and bracing, perhaps satisfying an emotional rather than a cognitive need.

RN: Yes. There is a historical lesson to be learned here. In a recent essay on Carnap, André Carus makes a distinction between what he calls ‘drifters’ and ‘engineers’. Given this idea that we should always refine concepts at different scales, then when it comes to questions like What is mind? What is life? What is justice? What is good?, do we really have any chance of convergence? Can we actually ask ‘What is life?’ without drifting toward an ever greater fragmentation where the concept of life at this or that scale becomes fundamentally incompatible with the question as we pose it in the ordinary sense? Does this mean that we can never get any fundamental answer and in fact there are no fundamental questions? Is it just going to be fragmentation all the way down, so that the whole ideal of the global concept shatters into pieces? This is the drifter paradigm: going all the way down, drifting into nowhere. But engineers are not just drifters. Engineers always have a global concept *as well as* these local ramifications at multiple scales and contexts. The drift paradigm is espoused only as a methodological way of working around and refining global or universal concepts. A good example of this in the history of science is the concept of gravitation, which has changed and completely fragmented with regard to local theoretical frameworks after Newton. But then we also see there is a global Copernican paradigm here that allows us to, at some point, reintegrate these back without just drifting into ever more fragmentation. Everything that we do with the concept of gravitation should respond to the stringencies set by Newton’s theory as a special case. Einstein’s theory of gravitation is an expansion on Newton’s but also arises in response to the constraints established by Newton’s theory. The real universal or global concept can only be extracted by a historical analysis of how these two theories are connected. Another allegory would be the idea of a machine code as a metaphor for the global concept. Machine code is what you might call the unificatory paradigm. However, when we are working with a computer we always use interfaces and apps. We don’t know what is actually happening at the level of the machine code—but if we knew what was happening at the level of the machine language we could change the paradigm of how to redesign these interfaces, these fragmentations, glue them back together, refine them, more fragmentation, reintegration, more fragmentation, so that we can come up with entirely new interfaces. And this is the idea of refinement that Carnap associates with the ideal of engineering as the very ideal of Enlightenment: the back-and-forth movement between the drifter and the engineer, fragmentation and integration, local navigations and global orientations.

But there is also another thing about engineering and how it is connected with philosophy, in fact the most important thing. Think of science as a discourse about the order of *is*—What *is* the case? Essentially, we are in the realm of theoretical intelligibilities. Whereas the question of rationality, as early as Hume, is about the difference between *is* and *ought*, *fact* and *value*. The question of engineering goes further: it is the difference between the order of *is* and the question of *what might be*, namely, the space of possibilities. Essentially, this is the very vision of an engineer, which comes with some sort of balance between the messiness of reality, the constraints of reality, and the space, or the unbound ocean, of possibilities. And for the history of philosophy, particularly for Carnap, that is the expression of human autonomy: No matter how the order of *is* appears to us, we must gamble in favour of *possibilities,* because only the latter category can actually lead us out of the status quo, i.e. reality shackled to the established actuality. Only to the extent that we can imagine new worlds, possible and counterfactual, can

we understand the logic of our current actual world. And only to the extent that we can imagine and actualize what might be in contrast to what is the case, are we endowed with autonomy.

Of course, Carnap's engineering ideal is usually taken to be an ideal of the scientific enlightenment. But as many commentators have argued—for example, Carus, Haslanger and Novaes—these ambitions can be expanded into the realm of social change and in fact can be reconciled with the ambitions of the Frankfurt School.

RM: Along with the drifter/engineer distinction, we might consider another distinction you make very early in the book, between Dionysian and Apollonian modes of philosophizing. Part of what I was hoping to get across in my introduction was the fact that there's a fascinating and sometimes infuriating mix of the two in this book. That is to say, you really do play the Carnapian philosophical spoilsport at various points…

RN: Yes, definitely, and I'm proud of it!

RM: …but there are also movement of grand philosophical ascent and huge creative ambition, which you refer to as the Dionysian mode. How do you see your position in relation to those two?

RN: Every philosopher, unfortunately, whether she admits it or not, works in the context of the history of philosophy. I really do agree with Brandom that philosophy is that which has a history, rather than just a nature or a mere past. Philosophy has a movement, a historical movement, a positive sequence of conceptions and transformations.

So what does it mean to engage with the history of philosophy; to *do* philosophy? Well, there are two ways of doing this. One is what you might call this Dionysian adventurism. I would say the scarecrow of Hegel is a good example of it, where he says that 'philosophy is thought apprehended in its own time'. We can think of this slogan quite heretically. When I write about Hegel, I don't care about his theological vision, his teleological vision. I just try to, heretically, extract some of the critical aspects of Hegel, mutate them, reengineer them in my own contemporary context, without being faithful to the canon of Hegel as a sacred text. Philosophy is an organ of judging the ancestors, not a medium of being impeded by the tradition and the canon. But to judge the ancestors, the previous judges, we have to recognize them, even though recognition of the past should not be equated with being hindered by the past. In this sense, modernism is yet to be understood as the meaning of philosophical thought, insurrection against all manifestations of experience as established or given to us in advance. It is only the time of thought that can rebel against the temporality of experience, whether it be that of the human or of something else.

RM: To pick up on the Hegel question, because I have to ask you about the unforgivable sin of bringing back Hegel again: you don't think it's your duty, in building on his work, to explicitly distance yourself from all of the other, possibly objectionable elements of Hegel's thought?

RN: Yes, definitely, but even more so I would say that, for me, the important thing is not just that I don't want to hear, 'Oh, Reza just became Hegelian, theological….'

RM: That's what they're saying!

RN: Yes, I know!

RM: To paraphrase Foucault, Hegel is like Michael Myers in *Halloween*: whenever you think you've killed him, he's always just waiting in a bedroom closet waiting to jump out at you.

RN: But the whole point is that the task of philosophy is to reinvent thought according to the contemporary moment. But to go back to the two ways: philosophy is not all about Dionysian adventurism and full exploration, it is also about putting a halt to this adventure so that you can look back into the deep history of philosophy, compare it not just with the Hegelian canon but also Kant, Descartes, Hume and so on and so forth. This is the Apollonian side of philosophy, digging through what the tradition has left us…. The Dionysian is one way to do it, and of course it has a price. But for either of these two ways, there is a price to be paid. Sorry to say this, but to me a philosopher is a great hedge-funder between the Dionysian and Apollonian approaches, not because one of these paradigms might actually pay off at the end of the day but because we are just premature when it comes to seeing the history of philosophy as the unfolding of intelligence. We either want to hold to the tradition or become absolute insurgents against it. To me these tactics by themselves are not sufficiently equipped to understand what the history of philosophy is; they are just bipolar reactions which distort the task of philosophy, what it is and what it can become.

Usually, continental philosophers think that Apollonian ones are basically just scholars who are philosopher-wannabes, whereas Apollonians think that Continental philosophers, Dionysian adventurists, are just trying to find some sort of facade of novelty. But I think that these are, in fact, pathologies and something quite recent. Analytic and Continental are brands, and philosophy should never buy into brands. If we look at the beginning of the twentieth century with Carnap—he came from a line of inquiry opened up by Cassirer, who never abided by such distinctions and Frege's revolt against psychologism—these distinctions never held. My friend Adam Berg said, it's like Lamborghini and Ferrari. Who actually drives Lamborghinis and Ferraris in Italy?! These are just brands. Philosophers should never buy into brands. Luxury brands prevent us from understanding what people actually drive. They pigeonhole us into not seeing all those innovations and detailed works that have gone into the production of a generic automobile. Sorry, I'm getting extra mileage from this metaphor.

RM: But before we talk about how you then get from Hegel to Plato, I think it's an interesting point at which to talk about another malign influence: it's notable that there is a parallelism, up to a certain crucial juncture, between what is coupled together in *Intelligence and Spirit* as a philosophy of intelligence and its pragmatic consequences, and the central trope in Nick Land's work, right from his earliest writings, where the trope of Kantian philosophy as a program for artificial intelligence is already present, the idea of a transcendental reverse-engineering. For

Land, Kantian critique is a philosophical diagram of the dismantling and artificializing processes characteristic of modernity, and the uneasy and unstable compromise between this unleashing of exploratory intelligence and the attempt to rein it back in to the subject, the nation state, and so on. It's on this basis that Land can speak of 'being on the side of intelligence', being on the side of this critical process of disassembly and artificialization operated from the future—and as we know this process that is assembling intelligence out of the human is ultimately piloted by, or even synonymous with, the 'templexical' processes of capitalism. And arguably, a normative choice is then proposed as to whether we act so as to promote our absorption, our being harnessed by this process of intelligence or hopelessly react against it.

What happens when you bring Hegel back into this? It seems that the scenario is similar, but the future intelligence that is harnessing us becomes one that operates through language, through sociality, and through norms, rather than, precisely, against all of those: they become enabling conditions of intelligence rather than drag factors.

RN: There are many different reasons for the introduction of Hegel. I think one of the reasons—and we might disagree on this—is that, although Kant initiated the transcendental turn, at the end of the day he is in fact a conservative thinker, to the extent that, for him, the idea of an 'agent' is not only confused but also, as Hegel would have said, 'parochial'. Not only parochial with regard to the constitution of rational agenthood, but also in terms of how this rational agent renegotiates its position within a reality in which neither the rational self nor the world of which it is a part can be taken for granted.

For Hegel, the idea of the enrichment of the rational self or agent goes hand-in-hand with the enrichment of reality, and that's the Hegelian odyssey, whereas in Kant you have a sort of transcendental passivity. This question arises as early as the transcendental aesthetic in *The Critique of Pure Reason*, in which the forms of space and time, as forms of organising for intuitions, are pre-established—and that's it. All we get is this kind of sterile phenomenalism which gives us some sort of distorted lens onto a yet more distorted reality, and we can only refine them up to a point, but what if this is just a vicious circle? What if the phenomenal reality we use and refine in order to gain further access to what you might call the Real, is just simply a projection of certain kinds of agential biases onto the world, like Dorian Gray's picture—What if the world in all its infinity is just an infinitely distorted, ugly picture of our own selves?

I was once an adamant Kantian, but at this point the poison of rational scepticism has crept under my skin. While I see no way of undoing the transcendental turn without resurrecting precritical philosophy, I also don't see how exactly we can adopt orthodox Kantian methods without sounding like we are dabbling in scholastic amusements with regard to sensory processes, logic, language, etc.

So to answer your question, Hegel adopts a certain kind of scepticism. I think he is a sceptical thinker, but not in a passive Pyrrhonic sense. It is in fact the project of Hegel to reactivate the organon of a scepticism, namely the labour of investigation (*skeptikos*) within the order of

reason; to question the givenness of those experience-constituting categories which, for Kant, are just there and are never questioned with regard to the agent.

So that is one of the reasons for introducing Hegel; but there are many, many other reasons, one of which is the question of history. Intelligence is essentially a self-conceiving agent —a self-conceiving agent that has a history precisely because it has a conception of itself, according to which it conceptually transforms itself. Of course, everything can go wrong and pathological at this point. Your conception of yourself or what you appear to yourself can widely diverge from what you really are. But nevertheless this is what, for Hegel, constitutes the history of *Geist*, this correlation between conception and transformation. What I take myself to be is a ground for how I *ought* to transform myself, and how I transform myself is, again, the ground for a new conception of myself.

These are all reasons why I introduced Hegel, not just as a corrective to Kant's project but as something more; something that Kant couldn't imagine. For Kant, the idea of philosophy is what you might call a passive understanding of the world, whereas for Hegel it is not just a passive understanding of the world but also involves intervening and enriching the reality of the world. Hegel, like Plato, is an unrepentant Promethean and of course an arch-idealist, but without falling in the traps of naive idealism. To enrich the reality of the world one has to be sceptical of the broad structure of agential understanding. To push through the facade of what appears to us, to break that supposedly neat link between categories of understanding and the intuited items in the world. Hegel's idea of Reason not as faculty of general understanding but as the organon of critique is precisely what initiates the critique of Kant without reversing the basic gesture of transcendental turn, which must be upheld.

RM: That was a good way of sidestepping the question of the influence of Land, but I will persevere because your simultaneous fidelity and betrayal (has Nick been *aufheben*?!) interests me: If one thing is clear it's that you share with his brand of 'accelerationism' the idea of being 'on the side of intelligence' as opposed to being committed to salvaging any particular image of the human. But for you, what does it mean to be on the side of intelligence?

RN: To be on the side of intelligence is to be on the side of thinking or, broadly, cognitions, with the understanding that the consequences of thinking might diverge vastly from psychological or opinion-based premises. Here of course, the difference with Land can be highlighted in the sense that intelligence is always beholden to at least three forms of intelligibility—theoretical, practical, and axiological—and there is no way to flatten the difference between them without falling back on some sort of the myth of the given (theoretical, practical, or axiological givens). Each requires its specific norms of inquiry. No matter how much intelligence accumulates facts about the world, it wouldn't be able to function without a fact-value distinction, which requires a sense of self-conception (its revisable position with regard to the world in which neither the self nor the world are treated as pre-established). Self-conception as the hallmark of intelligence is all about agenthood (the axiological gap between the self of which there is a conception and the possibility to transform it in accordance with a concept of the world which is expandable). Nick mistakes rational agenthood with a perspectival phenomenal self (much like eliminativists) that

can be explained away by some physical law, but the minded self is a-perspectival. It is only an agent by virtue of a deprivatized space (geist) which enables it in the first place not only to be multi-perspectival (virtually adopting a collective perspective) but also to be a-perspectival, to see the force of concept as a revisable entity of thinking that is not determined by any particular agent, whether an individual experiencing subject or a particular historical agent or judge.

In disconnecting the concept of intelligence from rational agenthood, we run the risk of rationalizing nature as intelligence, or even as an intelligent design. I mean, if Land truly believes that capitalism as intelligence is non-agentic and is also a natural teleological drive toward complexification of intelligence, then why does it matter that there are some troublemaking Muslims, that there is no neoreactionary island, etc? Surely Capitalism can take care of itself with or without adversarial and/or favourable agents? Then ultimately, why do we even need to talk about a neoreactionary future which is entirely based on agents' contribution to the accelerating paradigm of capitalism? Why do we need to define intelligence as adversarial competition or through the slogan 'war is god'? If acceleration is real, then there is no need for any sort of agent-based society (neoreactionary or Marxist), and if it is not real, then agents will bring about the acceleration of capital toward its telos. These are pure contradictions.

RM: Now I want to ask you a somewhat related question that asked myself over and over while reading *Intelligence and Spirit*—and maybe this is also coming from the part of me that wants to stubbornly resist the philosophical spoilsport. Your work in *Cyclonopedia* was obviously indebted to a delirious tradition of philosophy which engaged with psychoanalysis, myth, fiction, and the unconscious. And it repeatedly struck me, contemplating this idea that intelligence is tied to collectivity, and therefore is tied to language and to social norms—even though intelligence operates so as to revise and update social norms continually: What happened to the unconscious? What happened to that part of 'mindedness' that always acts obliquely to the social program and to collective norms, that acts underneath language, the thing that programs us from elsewhere? Where is the unconscious in this? Or else, this collective notion of intelligence, of the essentially deprivatized mind, is that itself the unconscious, since it implies that your mind was never 'yours' to begin with?

RN: If I had talked about the unconscious, you'd have to add five hundred more pages…! Coming back to the overarching concepts which I oppose, the very question of the unconscious, even in the tradition of psychoanalysis, is quite vague. It is what you might call an excessively big concept…

RM: It's an *explicandum*?

RN: Yes, exactly. Essentially, what do we mean by the unconscious? Do we mean what neuroscience calls the mechanisms and computational processes that act beneath the threshold of an attentional system or global workspace? What you might call 'consciousness' is the idea of the theatre that is carried out onstage, and the unconscious is all the plots, the props that are being carried away backstage, and you never see them…. So, is this what we mean by

the unconscious? Or do we mean something that is more insidious? In the later theory of Freud, the unconscious is something like in Nietzsche or Marx, a kind of reality that is suspected of having hijacked the entire edifice of consciousness. Again, the idea of consciousness is, of course, a big thing, and the whole book tries to show that there are different grades of consciousness and thus we cannot posit a single notion of the unconscious for them.

So, in that first sense of 'unconscious' I would say that yes, the question is already in fact answered in the book, in terms of the computational infrastructure of thinking. But if we mean by unconsciousness something more like the later Freud's theory, as a kind of reality that works *against* consciousness, and which we cannot really access directly because it only manifests itself in the incongruities of our actions and beliefs, like hysteria, if this is what we mean, then how can we in fact talk about the unconscious without resorting to the resources of the consciousness? But we assumed that the latter is subsumed by the former. So then a vicious circle of diagnosis comes to the foreground: How can we coherently talk about that which potentially commandeers the criteria of our conscious coherency? Isn't this like Kant's idea of the phenomenon as a gateway to the noumenon as that which cannot be known but can be thought? But Kant, in opposition to his dictum that we should never confound an 'as if' judgement with a constitutive judgement, goes on to treat the noumenon as a postulate of thought as something that actually conditions thinking. This is too incoherent. How can we think X about Y if Y has subsumed X? We can see a similar example of this metaphysical totality in the real subsumption thesis, where Capitalism has hijacked not just all substantive social relations but also formal social practices which enable us to actually coherently talk about Capitalism.

There is another thing: I think that we have talked about this in terms of the difference between Sándor Ferenczi and Freud himself. For Ferenczi, and also for Otto Rank, the unconscious is not a reality that is inaccessible, because if we say that it is a reality, it's like the noumenon for Kant, but it also hijacks the phenomena, then we have no access to it, so then why are we talking about it and why are we talking about psychoanalysis? But for Ferenczi it is a different thing. It is a duty of consciousness, of self-consciousness...

[Loud and intense sirens]

RN: They heard us talking about Nick Land….

RN: It's the Epistemology Police!

…so, for Ferenczi, the idea of the unconscious is not the idea of a noumenon, of a given reality, or a real subsumption of capital, but it is a duty of self-consciousness to posit as a postulate, as a hypothesis, the reality of the unconscious, because thought's duty is to always to mark its own constraints. Here, the idea of the unconscious does not become a contradiction or an opposition to the idea of intelligence or consciousness or those epistemological requirements that the book talks about, but in fact becomes the maturation of a thought that is already armed with the epistemological instruments to highlight its constraints, because it finds itself in a confrontation

with a reality or nature that does not share its ambitions. When consciousness becomes mature, in order for it to increase its maturation, it has to mark its own limitations by showing that it might, as a matter of fact, be under the influence of external causal factors. But such influence can never be taken as wholesale (a metaphysical totality) because it will fundamentally undermine whatever we can say either in defence of consciousness or in recognition of unconsciousness as a real factor. Every claim we make might be justified in an arbitrary manner.

RM: There is another way of asking this question: In the toy model, you start with an automaton with rudimentary sensory capacities, and then you have to give it the ability to orient itself in space, the ability to communicate with its fellow automata, the ability to orient itself in time…. But you also speak about a *nisus* or a striving, defining intelligence in terms of its striving to create something better than itself. So for you this is a striving for the Good—in a Platonic sense which you elaborate on in the final chapter. For others, of course, it is the desiring machine of capitalism that provides a *nisus* that assembles intelligence. And I've certainly read some people writing on AI who say, for example, you would never be able to produce an environment where real intelligence would emerge without endowing its agents with sexuality, or some similar 'irrational drive'. Or, we could also ask, would such an intelligence need to dream in order to be intelligent? Or would it dream *because* it's intelligent? In other words, is the thirst for the Good or for self-cultivation enough of a driver to define and to bring about intelligence in a full enough sense?

RN: I think that the frame of this question is wrong. It's a kind of pseudo-problem. It's not really the task of AI to talk about this. It is the first task of AI to determine, exactly like Kant, a non-arbitrary or critical list of conditions of possibility. But, of course, to create an actual AI means that, from a set of abstract problems, which are theoretical, undergirding our project, we have to move towards the concrete. And that's not something that we can just talk about—'Oh, well, it requires a sexuality, it requires emotion', and so on and so forth. Such questions need to go in conjunction and in parallel with the maturation of our science and human self-consciousness, and that's why AGI is not something *out there*, that we could make these kinds of statements about. It's just what you might call an encapsulation of how we axiologically, theoretically and practically think about ourselves. The pathologies of AGI are the pathologies of the human mind, of how we think about ourselves. How can we demand that this relatively young research program to incorporate problems of emotion and sexuality when these concepts are still so vague to us humans? It's like to fault AGI from the perspective of human ignorance. First of all, who said that an AGI cannot have such concepts just because it is realized by computational and statistical processes? I just think we are projecting our poverty of concepts such as emotion and sexuality to machines. My answer is: Grow up! How do you expect a machine to show emotion if you don't know what the emotion consists of. This is just reactionary humanism. We are positing our own ignorance in the form of vague demands upon the machine.

Now with regard to the question of emotion, again, like the unconscious, it's a cluttered concept. When we talk about our emotions, are we talking about cognitive or experiential mental episodes, like hopings and wishings (cognitive) or the feeling of sadness, the experience of

being cozy (the experiential)? If we are talking about the latter, are we regarding them as implicit inferential knowledge or some sort of immediate and private knowledge, the non-inferential knowledge of an episode? If the answer to the latter is positive, then how can we thwart the threat of the myth of given? As early as the Stoics, emotions are considered to have belief-contents, albeit implicit ones. To actually know what an emotion is, one has to step into the domain of the inferential and justification of the content.

Surely, from an evolutionary perspective, emotions have been useful heuristic tools, ones which have conditioned our *ordinary* sense of reasoning to be this way rather than another way. Many discoveries have been achieved by the use of emotional apparatus. But really that does not actually explain what emotions are, nor does it justify that the emotional ways of discovery are epistemologically justified. To use emotions or affects to explain or justify their epistemological rights is hardly anything more than *petitio principii*. It is only in tandem with the maturation of the scientific rational method that we can recognize the import of these emotional ways of discovery, but also recognize that emotions by themselves cannot gives us anything with regard to epistemological claims about the objective realm. And therefore, we must not only use conceptual reasoning to distinguish and determine the contents of such tools and their respective discoveries, but also to change what is given to us by evolution. To devise new emotions and affective ways of navigation is to go beyond what is evolutionarily given to us, to rise into the domain of conceptual rationality where not only the content of emotions can be distinguished, justified or modified, but new emotions can be put forward. This, I'm afraid, is the Stoic lesson 101 which we have long forgotten.

RM: In order to be intelligent, an agent has to have this *nisus* or striving, a striving to be a better version of itself. But what we know about ourselves is that our intelligence developed from very different, blind strivings.

RN: Yes, absolutely, but the whole point is that being conditioned by natural processes doesn't mean being constituted or piloted by them. We in fact see the blind processes to the extent that we are in the possession of theories, concepts and inferences.

RM: So this is where we come back to taking the functional rather than the genetic point of view on intelligence.

RN: Yes, and this comes back again to Freud's idea of the unconscious. The idea is that, okay, the unconscious, whether it is a Freudian unconscious, whether it is a noumenal reality in a Landian sense, or any other kind of stuff…. First of all, we should say whether actions can actually be explained by such indexes of reality or not, by these causal factors. And if they can be explained, then according to what scientific methods? What kind of robust methods? It wouldn't be a free association or a free, exegetical hermeneutics! That would be just like saying a demon possessed this witch, and because she has these symptoms, we should burn her and burn the demon with her; when it comes to the idea of free association of symptom with cause, we are dealing with pure occultism.

RM: Let's finally come back to a high-level view of *Intelligence and Spirit*. If the book essentially stands against both nihilistic resignation and the idea of a magical revolutionary emancipation, and configures the task of emancipation as one that extends way beyond our individual life spans, then what part can any of us hope to play in that? And what part do you see yourself, as a philosopher, playing in that emancipation of intelligence from its cage, and from the shortcomings of actually existing human intelligence?

RN: This is a very difficult question, and not only because it is something that inevitably leads to vague answers, but precisely because you are asking an individual—me—this question. Any individual who thinks he or she can answer this question is a psychopath rather than someone who is truly faithful to the multifaceted aspects of the question itself. So I can only talk about myself, as someone who thinks I am a philosopher—I might not be, but nevertheless let's pretend that I am. And to that extent, I would say that, for me, the question becomes extremely important.

Why is it that we have arrived at this particular historical moment in which theoretical and practical cognitions—the augmentation of our cognitions, the augmentation of intelligence with the understanding that intelligence is also collective through and through—is now, instead of being promoted, being debased, by both the Left and the Right? This is something to think about.

Intelligence is nothing but that which makes a world and a history for itself, but the thing is that—exactly like the scenario that we have gone through—we claimed autonomy from the clutches of our parents, for better or worse. We recognise them, we commemorate them, but that commemoration does not mean that we allow our parents to impede what we can become. The same thing applies for future intelligence, whatever or whoever it might be. This is a kind of a general paradigm: What does it mean to live in the age of general intelligence where we are merely particular forms of intelligence, special cases so to speak? What does it mean to live in the prehistory of intelligence? That is the ultimate question any agent that considers itself intelligent should entertain. Absent that, we fall into the trap of thinking ourselves as fundaments, as centres of the universe and as completed totalities beyond which nothing can be imagined. I see both posthumanism and conservative humanism as unwilling to methodologically deal with this question. If the human is this fixed X, then let's just abide by its vision, for everything else is just a supernatural vagary (conservative humanism). Or if the human is this established X and we want to get beyond its narrow scope, then let's talk about all kinds of fanciful stuff like sentient lava lamps, intelligent spuds or unbounded intelligences, and maybe even god as the register of absolute contingency (the posthuman flight of fancy). These are both parasitic upon a notion of a fixed idea of the human, they both feed on a notion of humanity that is precritical and prescientific.

I don't want to sound as if I am endorsing practical resignation. There are in fact certain kinds of contributions that we can make. For me, as a former engineer, I don't in fact want to think about these super-revolutionary ambitions; I want to be fundamentally practically modest, even boringly so. And for me as a philosopher, the most important thing is the idea of education.

Education is always and all the time connected to the philosophy of intelligence and the philosophy of mind. And by that I do not mean higher education, I mean the broad spectrum of education, from nurturing to developmental psychology, and so on. If we don't take this idea of education seriously as the basis of what we can do here and now, then any kind of future anticipation is going to fail.

We get overexcited by our revolutionary paradigms, by what we have achieved, but then we see, two days later, twenty years later, that we are back in to square one, if not worse. Education is absolutely, for me, the most concrete contribution I can make. And the idea of education, right now, not only in Western countries but across the globe, is fundamentally pathological. Why are the so-called revolutionaries not talking about education anymore, as something that is deeply, fundamentally tied to the history of intelligence and to concrete political change?

Now, it is not just a Western pathology that education is market-driven. Education is market-driven everywhere today. But here we see something far more insidious than the marketization of education. I talk about this a little bit in the book, but not directly in relation to education. We are witnessing a kind of historical bipolarisation as to what education consists of, between the Left and the Right. On the left, we see education as being about the virtues of intersubjectivity, with minimal regard to the purview of scientific facts. But when you go too deep into your subjectivity without the scientific facts, it becomes something akin to methodological individualism where different individual preferences and choices—even though they might be purely psychological—are taken as facts. Whereas on the right we see a different kind of pathology: the minimisation of intersubjectivity and the hyperinflation of facts. But as early as Hume, and in fact even from Plato, there is such a thing as a fact-value distinction. You just cannot conflate them with one another. You always need to triangulate them with regard to one another, and that is a labour of intelligibility. You cannot just have intersubjectivity without scientific facts, nor can you think you can simply derive social values, political values, political paradigms, from mere scientific fact accumulation. These are both pathological.

For me, all of this is just a first step, and I'm just trying to actually work on the details of what would be a system of education, an education in which we can determine the good life of an intelligence which has not yet fully determined what it is, where it is in the world and what it should do; an intelligence which is still in the process of developing its methods of inquiry with regard to its position in the world, so as to cultivate itself by enriching the universe it inhabits.