Running After Emotional AI. Assessing Virtual Experiences of Smart Cooperators

Dr. Alex Card. (DD.. MPhil., PhD.)

Postgraduate Dissertation Proposal (2016)
National Institute for the Studies of Sciences & Technology (ECyT) — University of Salamanca (SP).

Abstract *. This writing's purpose is to expose a perspective landscape over the advancements in Artificial Intelligence (AI), oriented to assess the attainment of Strong AI. Rooted on epistemic scepticism, the main consideration points out a dilemma upon certain synthetic systems's genesis, where a definition of the problem is proposed through contemporary Philosophy of Emotions, Technology and Neurophilosophy. A prediction threshold is discussed, balancing autonomy of technology vs. genetic inference, the 'forms of being' of 'Intelligence', while conclusions are derived, maintaining emotional communities as critical token in the emergence of evaluable traits.*

Artificial Intelligence, Forms of Being, Technical Autonomy, Scientific Conditions.

THE HEAD issue of discussion in this article deals with how scientists and technologists could assess whether certain artificial systems (AS)[1] are thinking intelligently. As a metaphor of the whole text, let this question be its summary: 'If a "black box" could think… would we think in turn the strange box thinks of box-things?' The hypothesis notes that we (human like thinkers) would start to work with a bunch of assumptions projected to ASs without knowing in depth which questions are we able to really answer and how; hence the 'blackened' or occult feature of artificial thinking. The aim of the writing is to point out a dilemma on the creation of such artificial intelligences (AI)[2], and to sketch a possible solution for a definition of what to assess and how. The article will revise a satisfactorily wide panorama of literature regarding contemporary philosophy of life, of technology, of neuroscience and of emotions, for making these texts to relate with those of the positive fields in experimental limits biology, neurotechnia and virtuality, rendering an epistemological study of their grounding processes for AI universally accepted base traits: artificial life (A-Life), cognition, memory and emotional experiences. The three parts of the text covering these interests will come by *I*, accounting the aforementioned epistemic access complication, extracting some problems its creation exposes; by *II*, introducing the 'forms of being' and 'being forms' duality, as some troubles with simulation given this; and by *III*, sketching the 'ecological integrity' problem and its derivations, for finally considering emotional experience of communal intelligence, drawing some conclusive remarks about its possible solvency, at least in its definition, at the end; albeit regarding the fact the text supports its points of view from epistemic scepticism.

Hereinafter I offer the main argument proposed. Beginning with the epistemological access to AI, any assessor is said to face a sibling concept, the axiological

—

I feel thankful to anatomist dr. Ángel Peña-Melián, at UCM.

1 'AS' will be used as universal term for a general conception of any artificially electron-based system with information integrative dynamics.

2 'AI' is a very universal term in positive fields. This text will specialise in just one of its forms, usually called 'strong AI' and its derivatives, through an epistemic perspective.

validation of the evaluated affairs as intelligent, cognitive, alive (and the related sort). This perspective engages them with a genetic dilemma, or the dilemma of generating such intelligence or such intelligent beings by deciding us (human-like thinkers) the base of their (non-human) conditions of genesis. The dilemma can be unfolded as a result of hesitating to choose between two approaches: *autonomy of technology* for participating of its own creation vs. *simulating natural conditions* of pre-existent organic intelligences; finding that at any case we will encounter grounding problems.

The dilemma can be decomposed in such problems, as for how to assess something that could be autonomous enough to be manifested "as-intelligent", albeit not being this obvious, evident nor easily assessable by our human devising —problems with 'forms of being' and 'being forms', the anthropic principle—; and for how to define its genetic conditions in the case we used simulation at some extent —problems with simulating something we are not utterly familiar with its architecture nor activity—. In general terms this writing posits that any problem will round about or lay on a bigger one: how to access the "integrity" (*infra*) of artificial events (studying how they are ecologically imbricated as artificiality). These integrity traits of AIs I confer here offer us an interesting clue about how to make evident resemblances of shared integrity traits of humans's and even of experimenters's ecology towards AIs. What matters is it epistemologically situates them for the main concern, an axiology of decisions, between which forms of AI are we looking for, if looking for anything at all, sketching an able epistemic definition of the problem. Such ecology raises not less than the idea that AIs don't grow in vacuum, but from communities, and introducing a premise, that *responsibility* is a higher condition for the mere possibility of a choral, communal genesis of intelligence, axiological experimentation could finally end up suggesting emotional experiences as well in AI come with that virtual community the ecology entails. The conclusive landscape of all the argument is a devitation (a final dual manifestation of the dilemma impossible to discern properly). I called this devitation the *Reliant's Regress*: we need to rely on one of those perspectives, rather autonomy or simulation, though as it will be seen, even if we introduce just one of them, when assessing harder and harder forms of AI and A-Life, and even for creating them, we will need the other one recurrently, leading us to decide and define which autonomy and which simulated communal conditions, and why those and no others. Thus, epistemic scepticism is released as a preventive measure in socially, historically oriented programmes of the token.

I

Expert managers and expert systems —using Markov models, chains of probable facts and Bayesian predicting nets— act by incorporating new serial events into the step list of programmed functions to operate by, and dealing with very specialised and partialised instructional habits, they attain resolution of given problems. Combinatorial systems get generated then and reproduce an 'incorporatory tactic' of progressive options-building operations, decision-making routines and error-solving panoramas; the kind of processes that describe *incorporation* itself. As a consequence of the rearrangement of information, the new order of issues to be observed by certain prediction corrector gets a binary value: correct or incorrect. The correct-like ones are reinforced, the other lost. These incorporatory reinforcements have been called Hebbian memory tactics[3], and are commonly used by simulationists and neuroscientists for tracking neural models of mnesicoception and location in tests. The complica-

3 Hebb (1964) places the definition of electronic organisation of *learning processes* as memory tactics in this way. Some experimental research carried this sense of learning in simulated models of visual perception (Pritchard, Heron & Hebb 1960), and it was transported to neurocognitive simulators very soon: one of the most successful was the *NETtalk* experiment (Sejnowski & Rosenberg 1987), which used probabilistic adaptation for speaking-reading written text.

tion with simulation is that biological labour does not actually work this way: we find that neural nets come to be ecologically, historically and genetically varied and reproduced by fitting the needs of specific range conditioners (relevance is no longer a trait to be suitable of obtaining due just to a better-and-better progressive biological evolution). The idea of perfection in biology got transmorphed into plasticity by the end of 20th century with astonishing facility.

Now, for us to consider an expert manager like these as an AI it shall be intelligent in such a way we assess its actions in *that* proper manner. The relevant thing is how could we introduce problems that have to be solved in *that "intelligent" way* by previously non-programmed incorporatory tactics. If we finally are able to evaluate "the manager has intelligently managed the situation", it could be assessed as intelligent[4]. But, it would be nothing more than a simulation, not just what the AI has done, but its very process of making itself: it would be an emulation, using the correct term, of an organic final solution, though mechanically proceeded following an artificial tactic of reinforcement. The AI is not actually deciding how to do something, nor why, just computing probable solutions within a variable range of tactical programs, featuring the solvency of a conditioned by the experimenters situation, and "acting" *as organic specimens would do*[5]. But we shall not be infantile here, this acting is just a reflection of something that has actually never been satisfactorily manifested: intelligence *as such*[6]. The rising complexity of the item offers a scale problem by which to divide soft, bland or weak forms of intelligence incorporated in virtuality, and harder, stronger ones[7]. Technologist Raymond Kurzweil proposes as well a clear distinction between strong AI (examples in 2012, 330ss) and weak ones (*ibid*., 319-330) in the base of usefully instrumental AI that reaches to fix problems so far only humans have been able to solve. Deeper definitions (introducing limits of agents) would be studied afterwards, so I will just note this strong AI as *ƒ*AI (from *forte*), and specify that I will focus just on the evaluation of attaining this kind of intelligence by artificial means hereafter. Now I would like to introduce the dilemma considering how to expose the genesis of *ƒ*AI and why it is complicated to solve with unimodal perspectives in the field.

. *The genetic grounding dilemma*

The conflict is noticeable as between two principles, the *principle of autonomy of technology*, and like its parity, the *principle of scientific authority*. The first one generally argues the drift of irreversible operations in the advancement of technology would cause a sort of autonomy in the way decisions about how to produce, create and industrialise certain technical objects are made regardless of social needs and collective control[8]. Bimber Bruce (1996, 95-116) analyses this autonomy as well as a form of determinism, depending on basic

4 Different experiments have been actualising this idea from the old times of AI to modern simulation, though this concept entails just a definition of what human-like conceptions of intelligence can afford. It doesn't fit but an anthropic principle of what is to be intelligent as-we-are. Further readings will be exposed.

5 The actual implications in philosophy of virtuality come from the assumption that Wittgenstein was right in his idea of a thinking machine (2009, §360), that behaving "as something similar to men-thinking" would be said it actually thinks. This concept of simulation has been critical in anthropology and in philosophy of technology for its instrumental validation.

6 This definition of intelligence as such is that of an order, an arrangement, a structure (Goertzel 1993). More than 'the being that is intelligent', the actual 'being intelligent'. This difference of 'forms of being' and 'being forms' will be treated in more depth in *II*. The problem of conceptualising intelligence as an object instead of a highly complex order can be seen as what Morton (2013) calls a problem of *hyperobjects* on contemporary concerns in science. Objects as climate that overdimensionate the regular scale of daily ones.

7 As Johnson-Laird (1988, 148) would come to expose, a robot or an AI that follows blindly the programmer's programs, shall not be considered more than a mechanic.

8 This principle comes from Mumford's (1964, 5-8; 1970) idea of a *Technological Imperative* by which the technic's development would imply medium's alterations: whether accepting them or not is decided in virtue of an imperative need of technology. This technology, then, could follow a drift of advance seemingly autonomous, Winner (1979, 66) exposes.

explanations that will suffice for evaluating the current state of socially conceived technics, where the same society has less and less power upon its technology[9]. The problem considered here in the genetic isolation of autonomously created AI is that its scientific practice could lose the axiological control upon how it did happen, why, and which ascriptions of intelligence does it actually recognises as itself bearing them.

The second principle gets to scientific decision, standing for a regulation of that same AI genesis practice (in this case), by which to control the definition of the conditions the affair that is being created will be conformed by. This control rules not just the ground of the creation but as well its final result, the specific traits-space to be fitted by it[10]. The dilemma comes as this: if accepting the 1st principle, Necessity Conditions are put (conditions for technical genesis are decided, as for example a basic energetic supply) albeit the genetic process itself is unknown for it is the machine which autonomously enhances the process. If accepting the 2nd principle, technological meddlings and human interventions appear, for a model of an organic trait is wanted to be simulated (a characteristic habit, a morphology, the very process of genesis, and so on).

What is left of our experiments as a result? If simulation takes place, a technical *Rule-Following protocol* will break forth, exposing rules that are artificial and not equal to the organism's ones, but that resemble, in the expression and manifestation of a behaviour or organic activity, to biological ones. It would be as written before, a reflection of an "acting process" that, as in a theatre, would have its stage, where observers would see a biology-like guise, but that in backstage is performing a computing protocol of retro-engineered patterns that actualise the final result. It would be not an artificially intelligent behaviour, but a protocol of actions that are just apparently biological. The problem: rules are not willingly followed nor decided, the machine does not accept whether to follow the rule assigned or not. This *acceptance* will be crucial later (*infra*).

On the other hand, if autonomy takes place, a *Plural Conditioning process* would break forth, where artificial protocols generate activity patterns (in a system of actions) that have been created after the event of humanly, scientifically decided inferences (those would be the Necessity Conditions catering as, for example, supplying the experiment with electricity or the like). That artificial "black box", using the previous metaphoric expression, could be considered primitively autonomous if the protocol has been produced without intervention. The problem: experimenters wouldn't know if the machine is utterly responsible of its system of actions, nor whether it has grown sufficiently enough their complexity scale to demand 'feeling the need of being responsible' of its own actions —as put in the upcoming section, here the problem is about the 'forms of being' and 'being forms' duality—.

In both cases, we could see, experiments are socially situated, oriented to succeed at the planned results: creating intelligence as-defined-by-us, and definable-by-us. Assessing results is a matter of assessing if those AIs are actually *ƒ*AIs… or a fraud.

II

Facing evaluation of scientific justifications, experiments and results; maybe achieved due to technological prowess; previous queries are needed to be solved: who are we to assess that affair; how are we able to get an epistemic access to the affair for assessing it;

9 This Ilkka Niiniluoto (1984, 239) treats as a «politically desirable» social answer to a needed goal, given the case of such need: i.e., building or not Space Stations is desirably decided in terms of scientific knowledge, though rising up a Nuclear PS is desirable in terms of economic energetic resources, both dependent of social and ecological reasons.

10 Latour (2001, 310) at some extent following Weinberg's laws, explains how science is precisely the reaching study of non-human ontologies that gradually get socialised and exposed to human ones, getting injected their human features. What specially scientists do want, he says, is to be sure enough of knowing what amount of autonomous ontologies of the studied non-human affair their observations bear, and which proportion of humanised, creative injected ontologies they are handling. In a similar concern cf. Cromby (2007).

and maybe a much more delicate one, what is that affair consisting of, from our definitions of its traits by us attributed to it. I will deal with those queries upon *ƒ*AI in two points by this section: appealing to the difference of forms, I will present a duality that has been concerning researchers in A-Life for a quite long time (carrying the arguments of autonomy's "black box uncertainty" close to the anthropic principle and rule-following), and then attending to the second part of the dilemma, simulationism's problems in artificial nervous systems (ANS) and nets, that cause interventionism's regress.

. Introducing 'forms of being' & 'being forms'

Kurzweil puts a rather incipient argument to assess the attainment of *ƒ*AI, focused on listening simulating-based research programmes from which to extract the main features of that intelligence; in this case, anthropomorphic traits of brain-like interactions that with the discoveries of retro-engineering will reinforce the knowledge of how machines can be thought to be as-intelligent-as-humans are (Kurzweil 2012, 335-338). This anthropic principle[11] can be sorely discredited if taking into account the rigmarole of 'forms of life' and the 'life forms' that will be studied in a moment. Using an analogy from the traditional duality which conflates 'forms of life' and 'life forms', I wrote 'forms of being' and 'being forms' extracting the meaning of the precise attribution scientific practices wanted to place their focus of research (on forms of cognition, on forms of memory, of learning, of intelligence…; and then forms of being cognitive, being mnesic, being intelligent…). Again the question is why "being like human" is the thing to expect when talking about intelligent, alive, cognitive, and the sort. Why *like us*? Why an *us*, an item that we know, is the last thing to be disputed, regarding the fact that we don't even understand the basic processes that have us standing? Will retro-engineering map the final answers that easily, when almost not even handling the proper questions to make? This simple, trivial silly note, that there could be different 'forms of being', turns out to be the most prevailing simulationists's nuisance: fully focused simulation-based research not only cancels the autonomy of technics's advance, genesis and growth (for we shall doubt of a machine's fate dominated by humans's decision making approaches —anthropic principle applied—), but it also suppresses the possibility of assisting to the emergence of new 'forms of being' that the methodology immediately neglects —against a multiplicity of forms perspective—. Below this, the idea of genetic conditions appears. Carr & Rees spot an important feature to take into account: «It is conceivable that some form of intelligence could exist without all of these features [water, helium dominance, certain astro-planetary distribution] but thermodynamic disequilibrium is perhaps the only prerequisite that one can demand with real conviction.» (Carr & Raes 1979, 612). This interesting domain of necessity in systemic dynamism is a deep concern: why those patterns? Where do they affect and how? Is it that there could be Necessity Conditions that are immune to human-like reasoning's inference, as for example, the need of energy for any life-like performance, being that as it could be?

The duality of "forms" problem is observable in Carr's & Raes's account of 'forms of life' too, as in two systems intertwined. The first system is a complex compensation of forms of being *x*, while this *x* is just *as it is* for the *x*-bearers. Their observations transmit *x*-traits towards their landscape, but it is a transmitting practice just in the sense of the bearers's having *x*-traits as a fundamental integrator practice attachable to their existence and observations. Take that *x* to be 'Life', for instance. A 'form of being' of 'Life' is an actualisation of a live being of transmitting Life-traits to a landscape of affairs of what could *be alive as such* (life in other planets, A-Life, and so on). Nonetheless, that *being alive*

11 Other pro and counter notes about the anthropic principle problem with A-Life, AI and others (cosmology, exoplanets and the like) have been set out: cf. Smith (1985); Kurzweil & Grossmann (2004); Moravec (1988; 1989).

as such of landscape's affairs is just a mere consideration of 'Life' utterly imbricated in the way that a specific live being realises 'forms of Life' as an integrator practice of *that* feature towards others: i.e., a rock is not a live affair of the landscape just in the sense 'Life'-bearers's observations and beliefs assess that affair as not bearing *that* trait *as such*, where *such* means "like the bearers" (as in the case of "like us" assumptions… why *like us*? Just because we find a definition that satisfies our realm of considering liveliness as a known trait? Aren't we neglecting the 'other forms of being' of 'Life'?). That is the idea behind the first system. As it can be seen it consumes itself as a circular process: just live beings as bearers of '*that* form of being' of 'Life' attain to integrate other 'forms of Life' along their landscape as the way assessors are alive[12]. The second system is easier to explain: it engages the various possibilities of 'being forms' as if what previously has been treated as the quality of a 'form of being' was now the very 'being'. If 'being *x*' is being *as such* (attention: not being the adjective form of *x*; that would be 'being a form of being *x*'), then *x* would fit a scaled observation that turns out to be the observer itself. In the case of ontological theories, panpsychism and hard forms of holism are examples; in the case of AI, it could be very common in the future to think too that *f*AI is not a sci-fi cyborg, but a program dimensioned to the scale of complexity of an ambiance. The AI *would then be* its ambiance (*infra*). Closing the duality, it affects the experimental variables of how to assess those *exes*, because it is quite possible that as far as the duality remains, following the first system we would be unable to attain a proper axiology of intelligence or life but of our forms of being intelligent and alive, while in the same way following the second system, it would be also possible we could stay open to discover that *that* intelligence is not but an act, instead of an actualisation of a 'form of being' of 'Intelligence'. This implies that, searching for *f*AI, we are both launched to theorise and assess artificial intelligence as 'a feature of an AS', *and* 'as the very AS' itself. This means scientists will be 1, able to ask in which sense could we say 'this AS is intelligent', as it bearing *that* property for it is a trait of its highly complex order and arrangement of electronic dynamics; and 2, able to ask about how to create intelligence, life or cognition *as such* as a program (the very system, and not an attribute of the system: to create the ability in isolation *as* the very system). Dealing with morphism comes with a complicated duality at a first encounter[13].

As Stefan Helmreich puts it, those terms seem to figure in logical operations of conceptualisation in particular theoretical backgrounds that liminary reconstitute epoch by epoch the appliance of such terms in different fields: deductive, inductive and abductive formalisations of the words, he says, expose classically and unconventionally adapted uses of the concepts of 'Life', 'live-like organisms' and 'live-like behaving systems'. Since its Romantic definition coming from Raunkiær,

12 The complex problem of 'Life as form' has been an extensive focus of attention in the artificial rediscovery of the classic theme of life, the *bios*, and the variations of biota: Evelyn Fox Keller (1995; 2003a) has been one of the most influential explorers of this topic in social epistemology of biology, starting with rhetoric voices of the terms (also Doyle 1997; and Kauffman 1995 explore that perspective). Chemical exploration with philosophical inspection upon 'forms of life' have been proposed as well: the famous marxist geneticist JBS Haldane (1947), Prigogine (1993) and Prigogine & Stengers (1997), Oyama (2000) or Luisi (2006) attend this other line of research.

13 Helmreich & Roosth (2016) have explored this paradox in depth. Appearing with the German term *Lebensformen* (forms of life) as a habit of diffusion and reproduction conducts in the Romantic 19th century's botanic usage, the expression liminary transformed into 'the forms of life' as the haptic study of the morphology nature exhibits and is 'embodied within'. '*Life-forms*', as in botany hyphonated recovering the German root, was then understood as the varieties of life, and '*life forms*' as in the formation of living things, objects of study that appear to be alive or lively to certain theoretical accounts. It is true that Wittgenstein (2009, §19ss & §241ss) famously used the term, as a particular manner of living, as a pattern of acts that is publicly modulated by norms and accepted by an epistemic community of life-beings that live that way. Though at any case, its possibilities don't get exhausted in that conception when applied to the technical practice (cf. the work of Fischer 2003 for plurality of voices). What is remarkable from Wittgenstein's account is the public aspect of *accepting* rules linked with the 'forms of life' idea.

to Goethe and Humboldt, artificially-involved scientific practice didn't do but to increase the complexity of how to assess *what* is life, as *being a possibility of it*, scaling this as a mereology of limits problem. We could read this dialogue between 'life forms' and 'forms of life' (or any other topic: AI was for it the 'pretty little baby question' of previous A-Life interests) as an example of the effectiveness this idea actually projects to the other notions implied in *ƒ*AI: cognition, intelligence, memory, emotional experiences and communities, intentionality, and so forth. For the sake of simplicity and to generalise this collection of used terms by scientists and performers of researches on the artificial, I employed the expression 'forms of being', as 'forms of being the ascription (normally the adjective form of the previous notions) scientists attribute to ASs' ('forms of being intelligent', 'forms of being cognitive'…). These 'forms of being' are cast out of the same paradoxical scepticism about *what* they actually are, how to define what is the ascription of "being intelligent", to finally be attributable or not to an agent evaluated.

One generally spread use of term 'ascription' would concede a *ƒ*AI ascribes *as intelligent* for it is empirically intelligent as for itself; while a generally spread use of term 'attribution' would concede that same *ƒ*AI is intelligent just in the case it filters inside our dimension of definitions for 'intelligence', causing this *our recognition of that ƒAI as intelligent* (cf. Goldie's approach to 'personal perspective swifts' in narrative accounts of traits attribution in 2000, *ch*. VI, 151ss; 2002b, *p*. II, IV; 2003, 203-205). One is more privative, the other more perspective. In any case, the use of 'ascription' will need an axiological conviction, a discernment belief, for which to exercise it in 3rd person, for if the very *ƒ*AI ascribes its own form of "being intelligent", this ascription —not specially a self-belief, we could interpret it is intelligent without using beliefs 'as we do'— would be invisible to us (a "black box" problem), and therefore we could never, at this case, attribute intelligence to an already *intelligent-in-its-way ƒ*AI, returning to the rigmarole of 'forms of being' applied to artificiality. This is, all our assessments depend upon how we want (or we can afford) our experiments to be proved, provoked, defined, orchestrated and calibrated; and not just the experiment as such, but the instruments the experiments consist of along with their act of experimenting. This points towards orientation of scientific goals, technical asymmetries that blur exploration, interests and other ecological and economic conditions, that again situate the process of scientific practices into epistemic scepticism and social, naturalised epistemology[14].

It seems the crucial token is (Doyle 1997, 120-123) to assure life scientists get to experimentally observe reproducible 'lifelike behaviour', since this reproduction of a conduct will manifest the idea that 'life as fact', in its coherent definition for a computationally based economy of its behaving, is an effect and not a core issue, but a responsive declaration of ordering matter *that* way, and not of matter itself. Life then will not be a thing *per se*, in their definition (not a 'being form' program), not causing things to be alive, but a clarified effect of matter. Indeed if life is an effect and a property of ordering

14 The epistemic discussion of a final theory (Galison 2004; 2008) and of fragmented theories due to the power of implicated agents of scientific explicative production has dealt with this item for decades (cf. studies in models, Keller 2003b; in political interests and impediments, Harding 1991; 1993; in social choral construing of scientific validation, Longino 1990; 2000; in scientific decision by pluralism, Cartgright 1999; in scientific practice and plural accommodation, Kitcher 1993; 2003; 2004; or historical scientific variations through perspectives, Giere 1999). The interest of experimental justification is one of the most intriguing problems social, situated and comparative epistemology can study: the contingency between *Instrumental Realism* and *Scientific Instrumentalism*, for justifying via experiments scientific knowledge (also its acceptance), can be tremendously important to the axiological practice of a *ƒ*AI, for we shall consider what technologists need for assessing if what they are observing is, in fact, actually there or invented though, not by artificial genesis, but by instrumental blurring (cf. studies in experimental validation and comparative epistemology, Woodwards 1997; 2003; in constructivism-instrumentalism theses, van Fraassen 1976; in scepticism upon experimental justification, Harré 2003; in less critic positions with experimental explanation, Franklin 1986; 1999; Franklin & Howson 1988; and in construction of instrumental ontologies of experimental observation, Hacking 1981; 1983).

matter[15], the theories upon it will tend to deal with the conditions for life, and then the conduct is what a simulationist is looking for to see. But this falls into the trap of the 'forms of being' duality: in which way could we accept this transformation of the conditions for life as an implementation of the scientific goal to 'what shall life be' instead of 'what life could and could not be' —and in it a difference between strong and bland A-Life and AI not as we could say so, but as it could be—?[16] In which sense could we say they are observing an autonomous phenomenon that looks like what has been defined before by social and historical grounds as alive or intelligent, and not a marked, mocked up, rendered by previous conditions that must follow this or that schemata… and that are, then, no more than operative response to what has been ordered, and not an "accepted" response of what the thing we try to assess whether alive, cognitive, mnesic or not need to react to? Why do experimenters and assessors think alternative 'forms of being' and 'being forms' possibilities of manifestation are so easily neglectable? Why is the anthropic principle that indulgent with experimental justification, specially over the contents of what are humans able to find in their experiments to instrumentally provoke the oriented result? And why do we think it is so imminent that stronger appreciations of life, but more implicitly intelligence, have just one monolithic 'form of being' (*ours*, a human-like behaving one) and no other 'being forms' are possible to take into account? Let's explore the operative problem of evaluative attempts.

. *Rule-following*

Rules are operations that can be followed, actualised in the sense of giving birth some order or command as an act. While inspecting "the insides" of an AS, as for example the cases of ANS in neurotechnia or artificial models of neural connectivity, biological pathways, and the like, certain problems appear for assessing those rules are being followed appropriately enough for saying the follower system is doing it cognitively, emotionally, mnesically or intelligently. Now 1, applying Wittgenstein's road to the notions, this is differentiating between 'following the rule' and 'accepting to follow the rule' as a game. The first one is an operative act ('do *x*', 'then *y*', 'then *z*'…), completely apart of decision, away of interaction; the later an actualisation of the act of playing the game (by which the rule gets actualised), remodelled by intimate contact and being able to be decided. If the AI or A-Life has no means of deciding, it shall not be considered for a scaling strong definition of intelligent nor alive. And 2, this also fits the notion of 1st personal perspectives (1stP) and 3rd ones. The 1stP grasps being an interoception of a habit or practice (as when playing chess… accepting the rules of chess), the 3rdP a mere observation that has no private confidence or intimate cognisance of that habit (it would just consist in following the behaviour of that who ordered the values in the way moving chessmen from black to white spaces holds. For example, a 1stP understands pain as '*my* pain', which is not transferable to others, while 3rdP exemplifies the conventions on shouting, retracting the body and communicating pain, to be simulated by the audience of someone's "pain-like utterances" as a similar situation they have lived before albeit upon the base of their own personal pain. Following rules is accepting them, not just believing oneself is following them.

Wittgenstein apparently uses this idea in (2009, §68), as if rules were not utterly fixed nor limited to a pure definition of the game they could be actualising: i.e. Tennis has a nice lot of rules, though no rule expresses how high must players tee off balls, though they will be playing a strange 'form of tennis' if touching the clouds every time. While developing his notion of 'family re-

15 This need of some theoreticians of artificiality comes from CG Langton, ideas were life and intelligence are proposed as a property of ordering matter, not as matter itself (the ordering-scale problem): Langton (1988, 74).

16 This difference is substantially accepted: cf. Helmreich (2007) 'life-as-we-know-it', 'life-as-it-could-be' and 'life-as-it-will-be-becoming'.

semblances' with the same argument[17], those forms of following the rule are "familiarly related" among all of them in the sense their degree of acceptance comes to be more and more or less and less similar among their practices. *Acceptance* would be, then, something that actualises intelligent rule-following programs, and in this sense, using the idea of a family resemblance between the different 'forms of being alive or cognitive or intelligent', we could say those strong forms are scale-relatives of others "that do not accept", but in the very situation of playing the game with the same rules, though without same conditions of rule-actualisation.

In (*ibid.*, §201-202) Wittgenstein manifests the interest of true prosecutions of rules as games of following rules, albeit not just believing oneself is following it his own way. Regarding my interpretation of what he wanted to show, *by accepting to follow a rule one accepts the public value of a community of followers of that rule in which one includes himself*. It is quite important to remark the triviality of following a rule without accepting it, for it will imply a vague or bland understanding and use of the term, as put in his example (*ibid.,* §158) of a reading machine that "reads" just by following the rule of *transcribing* 'letters as by graphics manifested' *into* 'letters as by sound manifested', with their precise linkages and forms of expression.

Including his previous concepts, "reading" just as a translation is an operation of being following a rule of transformation, but not a game of accepting to follow the rule. The machine is not accepting it because our attributions of 'it reading' are made adopting its own 'mechanical form of being', at least, the one we think it exposes, since 'reading' is something just cognitive systems could do.

. *Simulating… with?*

Cognitive systems like ANSs, learning programs or artificial brain memory topics are taking advantage of 21st century early research lines in neurotechnia and artificial modelling of virtual organisms. Neural networking has been widely studied as a gate-controlling dynamic system of electronic open and closed interactions that, in a gathering step, contribute to amplifying or minimising the electric voltage (action potential) of certain cells, specially neurons, causing a particular metabolic stress in partialised regions of cortical, subcortical and peripheral organic matterials that, as pivoting around complex actions, define a plurally orchestrated behaviour, attributable to a general organism (the problem of whether organs, organisms or ecosystems are proper agents, sub-agents or any other category will be briefly observed later on).

With these lines, plenty processes have been theorised as valid, accepted scientific content, due to their experimental nature: since at experimental conditions a notion appears to be found useful, the standard opinion is to include it into systemic knowledge of justified beliefs (scientifically justified by technological means), by public discussion, triggering a modern boom in fields of simulation[18]; not only studies in cognition, life and intelligence were credited, as well new genetics and genomic artificial understanding bloomed along… and along these, significant advances have been achieved too, specially in modelling human neural cortices and basic organic activity, mnesic and cognitive processes. New mathematical views won great acceptance in stochastic approaches to neuron recounting, or precision of instrumental calibration, and due to it clinical impli-

17 He exposes family resemblances (and their "parental" diffusion) in (Wittgenstein 2009, §67-79). Cf. Rosch & Mervis (1975, 579ss) for a critical discussion of terms.

18 Anthropological voices and epistemological discussion of the phenomenon concerned some authors, considering the fast evolution the field is seeing (cf. studies in observation critical with scientific simulation, Hyles 1995; in historic debates of functionalism and construction in neuroscience, Reddy 2010; or in our relation as already artificial organisms embedded within the network of virtually social interactions, as the concept of wetwares extends the notion of biosynthetics in Doyle 2003). In A-Life the boom was highly prominent (cf. the progressive evolution of the scientific concepts from Moravec 1989; to Emmeche 1991; to Riskin 2007; to Geraci 2010).

cation reinforced its derivative studies[19]: if experimental treatments carry on significant applications, then the core of scientific knowledge must get justified. I feel tempted to argue, nonetheless, they would follow an informal fallacy of circular implication if taking this lines as valid: 'suppose that if neural simulation and virtual nets get justified if they are proved to cure' (that is their main argument); it has been proved the model can be used to cure some particular conditions or guiding surgeries, improving technical advance in pharmacology; 'then the model must be justified because it works'. But as it might be seen, it falls into a *petitio principii* fallacy. Instrumental Justification is the case of this misunderstanding: however much something works or is used for curing, this does not imply that what is used to heal is valid to explain the nature of the activity of that affair, because while instrumental assumptions need not to answer to why-like inconveniences, explanatory ones do need to get, at least, a causal probative relationship[20].

One of these inconveniences is why neurons; which usually are the main item to have in mind for the most artificial neural network systems; are as they are and where are their natural conditions of existence lying on? A big critique to the previous circular vision comes with cell kinds multiplicity in the Central Nervous System (CNS), discovered in recent years, at the early beginnings of the present century. Neurons actually correspond to the ~10% of the global matter in CNS, while as far as contemporary neurotechnia can afford with exploratory technology, the collaborating 90% rest would be performed by glial bodies, with completely renewed organic activities[21], if not other cells yet unknown. If neuroglial-dependent processes occur to be the basis (nowadays) for almost every basic organic actualisation of biologically followed cognitive protocols, why the vast majority of ANSs and neuronetworks simulate just nuclei of neuronal activity (also glial-dependant for even the mere metabolisation and transformation of glutamate to *gamma*-aminobutyric acid, two utterly basic CNS's transmitters) of grey matter and neglect the Necessity Conditions of those stabilised points of interest they look at? Why shall not every other cell be simulated instead of instrumentally calibrated as a bland conditioner, that is suitable to be proposed as a waver, an bioalterator of voltage easily computable with simple predictive adjustments into the important thing: neuronal activity? Why this neuro-centrism, in this case, is allowed and accepted, as if

19 Rall (1964; 1970) was one of the first major neural networking simulationists in applying probabilistic neurotechnical knowledge in favour of ANSs, inventing a very commonly used neurotopology about the 60s-80s. Computational neurosimulation gave birth to new forms of scientific proceeding, specially in the field of perception and cognitive analysis (cf. studies in neural electrodynamism, Rosenblatt 1962; Abbott & LeMasson 1993; in neuropsychology, Eeckman 1993; identification models, Zipser 1992; connectivity and modulism, Among-Snir & Segev 1996; Dayan & Abbott 2001), and in neuropsychiatry (cf. studies in modelling personality, Tomkins & Messick 1963; on electronic implants, Abbott 2002; in computer simulation of bionic reactions to mental conditions's processes, Moser et al 1970; Colby 1976; as well as the *MedTronic Projects*).

20 Alternative theorising, emergentism, and integrativism started with ideas of Sherrington (1947) in neurotechnology, and were not deluded in time: following this approach, Hopfield (1982) famously combined the computational with probably emergent features of systems (as ANS could be thought of emergent ones). Recent panoramic studies also give an abundant idea of emergentism in the positive field (cf. Damoiseaux & Greicius 2009). *Scientific Scepticism* navigates this alternative process too, specially in language-cognition and semantic memory projects in ASs (cf. studies of comparative data, Gustafsson & Balkenius 2004; and natural vs. synthetic system's perception of grammars, Friederici et al 2006).

21 From the most vital to the most trivial activity of innervated organisms, they have been observed to be more than brain's glue (Pfrieger & Soltys 2006; Sherwood, Stimpson & Raghanti 2006; Allen & Barres 2009; Kettenmann et al 2013). They intervene from metabolisation of the commonest neurotransmitters to fagocitation, from memory and self-location perception (Nishiyama et al 2002; Claudel 2006; Slezak; Kim et al 2011: works regarding the advances of the Japanese discovery of glial protein S100beta's effects on mnesicoception) to psycho- and neuropsychiatric diseases where some glial-based failure happens to occur (cf. Shapiro, Bialowas-McGoey & Whitaker-Azmitia 2010 for S100beta's effects in Down syndrome and Alzheimer's disease). Those findings were enabled because staining processes have covered this focus of research anew: myeline, astrocytes, oligodendrocytes and more glial external biopaths's crucial activity were markable by newer techniques that actualised them.

no fundamental change in CNS's theory making happened? This is an important issue to be discussed and decided in future research. 'What shall be incorporated or eliminated in scientifically achievable simulations and virtual construing, and why?' If complexity scales soar, mustn't any of their definitional parameters be coherently incorporated (accommodated in the case of theory making) as well for fitting the rhythm of freshly adaptive scientific final theories?[22] Any epistemic community, and more importantly scientific ones, shall be able to decide properly their axiologies of 'affairs to be experimentally justified and justifiable': computerisation could indeed be highly beneficial for healing or assistance, instrumentally speaking, though highly weak when explanation comes. Could simplicity be the engine moving simulation, deciding those axiologies?[23]

Present resources are just for arguing in favour of a perspective approach, handling not only the rejectionist direction saying conditions for intelligence are no more accessible than those we actually profess, but as well for having in mind the so-called 'further alternative origination' (goethean *Ursprung*) of higher scale orders that could fit a for the nonce unknown actualisation of intelligence, life or cognition. An inceptive order that inhabits its own 'intelligent' directions and conditions of existence, in the way Evelyn Fox Keller (2005, 1068-70) deals with the idea of a 'Kantian organism', or following the dissertation of Whewell (1855, 256), as an order "of another world", with different adaptive concessions. Against the ideas of epistemic relativism and scepticism, even Laudan & Leplin (1991, 451-452) still feel useful some extent of doubt in theoretic acceptance and experimental justification.

It is visible in their 'three familiar theses' on 1, the variability of the range of the observable, 2, the need of auxiliaries in prediction and 3, the instabilities of auxiliary assumptions. It's this third one the most tantalising feature supporting difficulty of *ƒ*AI's attainment, if applied. It reads 'auxiliary information providing premises for the derivation of observational consequences from theory is unstable in two aspects: it is defeasible and it is augmentable'. And we observe that was true with the case: the mere idea that the body of premises scientists adopt to ground their coherent definition of alive, cognitive or intelligent *is, at any scale*, *dependent* of historically, socially, politically and exploratory oriented contributions —that is the point specially Laudan resists to— managing what is life, cognition and intelligence to be attributed to a particular affair, in this case, an artificial creature. Those partial definitions are defeasible in the sense they could appear to be unconsciously fraudulent or incongruent with new collected data, but as well augmentable via the same reason; that exploratory means of accessing to the immediacy are changing harshly and adopting severe perspectives of partialisation and specialisation, consisting of a quite deep effort to nail differently specialised sources of information for enjoying of a wide range panorama. This completes the role of scale-frame problems in epistemic landscapes about the growth of scientific knowledge, and the role of complexity problems as consequences of these previous developments, specially regarding 'forms of being' and simulation (failing the access to variations of the conventional in both cases: the handicap of the anthropic principle). This leads to a *suitable scepticism in its epistemology,* as I pointed out beforehand in a

22 Maybe Galison (1996) and Haraway (1991) were right, and no final general theory can be settle down to be fully hierarchically accepted without intervention and instability. Maybe there is a limit of simulation (the idea of computer as the world, and the world as computer), as Galison exposes. Maybe there is also a counter-move in recent decades, of a certain virtualisation of theory contents, as Haraway argues.

23 Weinberg (2001, 50) exposes this idea of a final theory as a final possibility of an axiology: the final theory is, in its nature, also the limit of scientific explanation through justification. Previously Weinberg (1993, 148-149) concerned about how to decide upon theoretical adceptance, accommodation, coherency… where some times, he points out, *simplicity* is used as an instrumental axiology for theory decision (the so written, he exemplifies, 'elegance of a physical theory'). Instead of an axiology we have generated in turn an aesthetic of experimental justification, and elected it as a proxy for internationally oriented interests.

distinct way: «whenever we perform an experiment, whether by computer or some other means, we are dealing with situations in which the initial conditions are given with a finite precision and lead, for chaotic systems, to a break of time symmetry.» (Prigogine & Stengers 1997, 105). If virtuality as an intelligence-ambiance or as a form of intelligence is to be thought a chaotic system, any experimental depletion of its basic architecture (its virtual time, space, resources, energetic dispositions…) will mean predictive conditions that are initial conditions for chaos-resulting bases of ƒAI: following electrochemical lines of what is shared between 'our human integrated world' and 'the virtual world of ƒAI', that overall condition needs to be energy equilibrium (as in Carr & Raes 1979, *supra*). Exhaustions, variations or mutations on this condition are precisely a third possibility to take into account for emergent phases of artificial genesis (starring now a genetic trilemma): 1, they could be programmed (intervention), 2, rendered by the very ƒAI (autonomy) but as well 3, accidentally given (perturbation); all of these cases being different strategies to be faced as varieties of the experimental performance —and maybe even needed for success, if treated as intentional, autonomous or accidental calibration (cf. note 14) of the experiment's instruments—. According to Prigogine's chemical dissertations upon time-life relations in other texts (2012, 95), those ruptures of time symmetries are precisely what lively organic molecules express, in a very precise manner regarding the irreversible operations of chemical connectivity between matter and states of matter — as Langton's claim concluded likewise, life, intelligence, cognition, shall then be not a property of matter itself, but of the order of matter—.

A final extrapolation of Prigogine's thought relates this affair of time and life with complexity. He puts the example of biologic alteration and irreversible operations in evolutionary processes: as brain's strategic complexities of growth soar, possibilities of chemical instability increase as well within the dynamic system of autonomy and biological independence of the medium (*ibid*., 94), which is exactly what modern conceptions of CNS and any biologically adapted system propose: organs are not born from teleological routines of perfection nor relevance, but because of need, and their development is full of conception errors, sketches, "spandrels" (as Gould noted). In this sense, *complexity and instability are parallel*, among other factors due to irreversibility requirements in chaotic dynamisms (cf. Prigogine 1993, 67-69; and Skarda & Freeman 1990 for early approach to chaotic systems in neurosimulation).

III

Concerning these AS's Necessity Conditions, this final part of the text will study which problem is being defined as the characteristic one; how sections *I* and *II* finally have an application to technical axiology, and how emotional experiences, coming from emotional communities, can be denoted as the grounding integrity trait below those conditions of need for assessing plausible ƒAI's attainment (at least instrumentally).

. The 'Integrity Problem' and derivatives

This problem rests with how humans integrate information as in contrast with any other 'forms of being' of 'Intelligence' and the rest tokens could do. *Integrity* would understand experience actualisation as some form of evidenciation of what an experimenter manages as "the given facts of an scenario of affairs", regarding how integration processes compose the basis for those affairs to 'be of his scenario'. This so-called integrity process works as an inclusion operation not just regarding perception, but also towards emotion, movement, agency, dispositions, intentions… a whole ecology of those affairs that an experimenter feels as 'concerning to his immediacy', with independence of their actual being real or fantastic. What is generally a raw fact; that different organisms do feel immediacy in behalf of their ecology of integrating events into their scenario of affairs (vision in horses vs. rats; ultrasound hearing vs. its deaf, etc.); turns a very interesting point if facing it with *autonomy of technology*: cancelling the

anthropic principle, what is left of human axiology if artificial forms of integrity differentiate their ecology in such a scale that would seem unable to be epistemically accessible by human assessors? How would A-Life, cognition or intelligence (even extraterrestrial life as in Whewell's plurality of worlds) in an autonomous strong artificial form be assessed to be attained? Emergentism in neurotechnia, biochemistry and neuropsychiatry, as the spontaneous activity of brain and organic tissue, are bland examples of this ecology of acts and 'given facts-response habituation', where definitional boundaries are transcended, sometimes stepping in the line and others stepping out. Unexplainability of the nature of thermococci, tardigrades (and other extremo- and mesophiles), viruses, higher manifestations of criptobiosis, randomness in specific pathologies (latests research in oncogenetics), and the like are also known examples transcending the definition of several 'forms of being' of 'Something that humans hitherto didn't really have a true epistemic access to'. For those organisms, any transformation of any 'given facts-explanation' is then responsive in behalf of an habituation that still works ecologically: in such way, moreover, that if the perceivers lack some of their own organic integrity's nuances, they variate the explanatory ability they have and need —i.e.: blind-born people need not light-like explanation, for it doesn't fit any epistemic channel or access to what serves for them to attain a coherent explanation of the 'given facts' before them; they need and actually manage another kind—. Again, how to evidence what is given to an experimenter must be considered (concerning this approach) a problem of integrity, a problem of the limits of the experimenter's integrity as defined above. It doesn't have to irrevocably mean the experimenter is a fully responsive epistemic agent[24], though at least an ecologically situated epistemic integrator (as if actively thought… albeit we could play with the use of 'integrant', more passive, or 'integer', more independent) of an epistemic community, where humans could be included or not.

Integrity is the thing to resolve if provoking serious questions upon what an artificial creature must 'feel' as evidences given to it, and as well as the proper world, the sense of things to come, to explain, the evidenciation habits it would or could develop and so forth. Is this limit a boundary we shall create, a condition creators of it shall posit, or is it something the artificiality must autonomously make? Which nuances of this habit shall it show or even bear, rather conditioned by creators or expose autonomously? How could we know about it if we do not share, if the case comes, the same, similar, or analogous integrity habits, the same evidenciation of 'given facts' routine? This all applies to an artificial 'form of being' of 'intelligence', but as well for a particular 'being form' of 'Intelligence' (not just for intelligent artificial creatures, but as well for intelligence as the very creature). It is an important query that would go this way: 1, they could integrate affairs we do not, for if they developed a different tendency of recognition, and even of cognition (so far, it could be possible their forms of being alive, intelligent, cognitive or any other trait were different from ours, and to our definitions of them), *that* they perceived could appear incomprehensible to us, or non-existent (it is a trouble concerning our lack of epistemic access to the same immediacy they 'feel', a lack of scenario actualisation); or this other: 2, they could develop not just a habit, but themselves, in a way that results apparently incomprehensible for us, in such a way that their intelligence conducts, their cognitive manifestations or their core ecology basis appears to us unable to be instrumentally measured or evaluated, for they would

24 An empiricist definition of 'autonomous agents' can be extracted from Castelfranchi (1995) where independent action-control of the system's agency is thought questionless. Against ecologically conflicting action-controlling, Galliers (1988, 162; 165) considers an axiological need, that agents do not have fighting intentionalities (applying non-rational schematas into chaotic systems, this precision is not actually verisimilar, though this will clash his *rationality assumption*, applied as well in Cohen & Lavesque 1990 agents logic). On the other hand, Bates et al (1992 and other texts) integrate emotional roads into goal-oriented agents as fundamental trait of AI. Cf. as well Maes (1994, 71) for AI cooperating responses with human users at same 'work environments'.

come to fail our red line of what shall or shall not be considered an intelligent, alive, cognitive... artificiality (it would be a trouble of axiology and the definition of what those forms of life, cognition or intelligence 'must be for us'). The heart of the dilemma still consists of integrity. There is no coherent methodology for assuring ƒAI's integrity is, must be, shall be or could be like the already known.

Conceiving another perspective, a virtual intelligent ambiance, the main access to that world comes inevitably anew: it is an access without epistemic certainty, as anthropologist Roger Bartra (2004, 41-42) exposes in a critical analysis of a similar case with Kant's voyage to his so-called pneumatic world, *mundus intelligibilis*. A yet more abstract conception is that of ƒAI as an artificial cosmos itself, a 'reality' in certain manner for the objects it determines. As Deutsch (1997, 139) points out, for 'reality' to be definable as such by the 'real' observers, it shall fit a very low premise: it being a self-contained organisation from which nothing from a hypothetical outside can be constitutive of things or orders of things from the inside and, if this happened, that will not be reality as such, but a formal manifestation of reality that has family resemblances (using Wittgenstein's idea) with other manifestations of reality — like a Russian doll, the greater ones are closer to a better and better definition of reality—[25]. David Chalmers in (1996, 305) studies a raw form of panpsychism with a double perspective approach inside-outside, where experience and physics come to be the way the mental matter modulates; precisely that change in an ƒAI would be a cosmos of beliefs constituted by interpersonal manifestations of actualised different intelligent ambiances. This idea comes to artificiality then in the sense of an intelligence-ambiance instead of an intelligent artefactual creature, and if this gets scaled into an ANS, wouldn't be the virtual world or cosmos of the ƒAI created in virtue of a pseudo-panpsychism of the similar kind? All the "real-like" of the virtual would be a saturation of the artificial, a hyper-technologic soap certain ƒAIs instantiated. "Their" world would be in itself they themselves instantiating it by an autonomous program. We would be assessing "intelligence" as a virtual system that, with the proper crane, could be put into the order of another AS and turn it intelligent[26].

Nevertheless, those questions are formulable just in the case of hard manners of artificiality, ƒAI, ƒA-Life or ƒANS, because if those 'different without analogy habits or actualisation of events' are not definable by us (humans), and scientists cannot introduce them inside the Necessity Conditions, are thus emergent traits of those systems; if occurring autonomously, independently, away from our human capability of action; showing for us conspicuously their 'forms of being' if noticeable. Another way of conceiving a virtual world is by exploring the current thoughts on organic division of perceptive labour and the modalities of perception. In cases of complex perception, how much of a virtual fact has that we use to think is actually there? Away of representationalistic approaches, how much of a "legitimate" creation has the way we smell, we touch, we taste? How much of virtual is our understanding of painful and painless events? How much of virtual is a phantom limb or spectrum pain sensation for a patient suffering of this kind of torment? How real is the dimension our standards of real allow us to think of it? [27]

25 Cf. Morton (2000, 49-53) for an analysis of an artificial cosmos in Deutsch's concepts. Cf. Weihel (1999, 220) for a critical consideration of virtuality as worlds, like Galison's 'the world as computer' example. Cf. an early conception in logic and mind philosophy as well in Brunner (1986).

26 As with embryogenetic mobility of rombencephalon's upper areas, that moved elsewhere could come to generate a cerebellum.

27 For a neurophilosophical interpretation of olfaction, cf. Le Gérer (2002); Shepherd (2012) and Gottfried & Dolan (2003) in constructivism and virtual cognition. In spectrum extremities as delirium, cf. a new experimental revision in Gardner-Thorpe & Pearn (2006); in virtual pain cf. a recent vision from clinical structuralism in Subedi & Grossberg (2011); and in virtual self-location in spectra, cf. an early work of Goldberg & Bloom (1990) related with Ramachandran's late results.

. Discussing emotion and communities for ƒAI

Whether an ambience as intelligence itself being the AS generated, or an intelligent form of an AS, tensions between artificiality and virtuality don't consume the questions about applying anthropic principles for defining if intelligence is a virtual feature of an artefactual creation, or an artefact by itself. Is it conceivable as an instrument for the sake of usage-need processes, a mental tool included in what we define as 'organic forms of intelligence' —and is it needed to be transmittable to ASs in the same way—? At any step, it seems specially important to assess scaling complexities of these needing processes as both, quality-focused phenomena of study and quantity-focused ones: if we define the token as a trait-bearing process, it shall be pointed out too how many of those systems are needed for that trait to actually appear bearable, to be actualised for the renewed order of matter (accepting Langton).

Dawkins (1983, 113) gives an interesting perspective through ecosystems: presenting recombinant operations at certain scale, some of them irreversible levels, the systemic units will interact «following laws appropriate to that level, laws which are not conveniently reducible to laws at lower levels». Those units, that go from genes to lowered chemical chains to organs and biota are specially susceptible, then, to those scaling laws of complexity, by following and adapting them to historical contexts of their evolution (cf. Diéguez 2012, ch. VII, for a critical analysis of teleology and functionalism in evolution). But as even Dawkins (1983, 117) exposes and admits, those units do not work in isolation, but in pools of alliances, *succeeding one in the presence of others, which in turn,* as he says*, succeed in the presence of it*, conforming a discrete transformation (a vehicle he then relates as the extended phenotype, which is his thesis for adaptive mechanisms). Communal or choral conceptions of atomistic relations are affordable to 'forms of being' of *ƒ*AI theorising as well, but rarely applicable to 'being forms' of 'Intelligence' (in the sense that creating "the tool" before its proper user, loads the experiment with plenty uncertainty). This all directs discussion towards how in community the instrumental operation of intelligence is created, deciding first that what scientists and technologists would actually be insisting in is over particular forms of ASs that proceed in community as what we openly, widely define alive, intelligent, cognitive and the rest. This communal perspective wires the previous problem of integrity and its derivatives, about how to assess what we don't have epistemic access to, or affairs by instruments whereof epistemic channels hitherto present unknown for us.

The idea, in any case, is not to develop a monolithic solution about the first or the second principle of the genetic dilemma presented in *I*, but to achieve an equilibrium due to the impossibility of decision in its axiology, between autonomy and conditioning, via our human definitions of the unknown events (open protocols of prediction). The community thesis gets close to that balanced perspective: on the *autonomy of technology principle* hand, the choral solution exposes not *how* actualisation of *ƒ*AI's integrity must be, nor it confers *which* mereology must it proceed with (virtual limits of artificiality can work with highly scaled chaotic operations that could fit, in a few decades, unimaginable degrees of computation). On the *scientific authority principle* one, chorality relates a global foundation: nothing stays in vacuum. Its ecology of conditions is the one tip to accept, for then including rule-following protocols oriented to achieve one special characteristic of the order to be present, *experiences of accepting the rules*. Clearly, those rules are not necessarily conditioned by creators: it is conceivable certain synthetic emergence of rule-following and rule genesis without interventionism. Grounding briefly the arguments of experiments this way, it could be installed in a plausible equilibrium, though falling into some fundamental problems of each side, to which epistemology would had to give up to some extent: hence the devitation.

Maintenance of community as such demands that "the common" stays as it is, *publicly* obvious for all integrators. The core item remains the same because of this: the important thing is not to follow the common rule, but accepting to follow it. The inceptive direction

is that of a choral gregariousness that accepts commonly to follow the rule, making that the law to respect. This difference between respecting a law and following a rule is the key concept for emotional communities[28], to discern, to be able to differentiate between law to respect, by acceptance of common rules, and the mere addressing an operation in a certain way. The core to understanding the intimate actualisation of emotional communities is *the manner a rule turns into law*, and how this law persists as a legal manifestation of a particular habit or customary act. Judith Butler (2006, 201-203) examines this manner as a concept of philosophy of violence, and exposes that the founding act of the law is precisely the same act of violence by which this law gets reaffirmed as valid[29]. Rule-following by acceptance and not by vague apposition of acts is what constitutes the community that, if translated into *ƒ*ASs, would precisely comprehend a trait of common recognition of publicly integrated forms of intelligence.

To accept the chorality condition by an AI, and its manners of producing what is needed for the community, is also an adaptive constituent of 'being in that community' (i.e.: it has been regarded without being in community, no acceptance-like traits are seen to be epistemologically provable as what we could define as 'intelligent' to be assessable. If assessors need acceptance-like traits for assessing strong forms of AI, those could have an opportunity of generating autonomously such an emergent feature in a condition of choral ground, following the equilibrium perspective). It is easily assimilable that a form of violence could happen to be applied (even autonomously) as well in virtual scenarios. This ability for structuring emotional events within the very ecological integrity of *ƒ*AI would create interesting observed repetitions of patterns, from the biological to the artificial; patterns that could be tracked and assessed (hypothetically) as emotional experiences of acceptance. Violence as automata, spontaneous arrangement of processes, and as ecological integrator of AS's integrity must be studied in further works regarding the technological development of Necessity Conditions for assessable *ƒ*AI. An argument in favour of violence as a founding trait of intelligence can be as elastic as the concept gets managed, though reasons against violence as an integrative token of intelligence[30] can be highly considerable if taking into account the weight of the anthropic principle in the definition of the assessable forms of intelligence we are dealing with: maybe those who belief violence is thoroughly indispensable are focusing just in their own acknowledged 'forms of being' of 'Intelligence *as such*' (where *as such* actually means 'as violence-based'), neglecting other 'forms of being' of 'Intelligence' just on the basis of simplicity, confounding aesthetics instead of axiology in complexity-grounded experimental justification[31]. Again, when complementing the dilemma with the problem of defining affairs as-we-know-them the results do not come out impossible, indeed, but instrumentally justified, in-

28 The term 'emotional community' or 'space' is of regular use nowadays in philosophy of emotions. It was deeply studied by Barbara Rosenwein (2006) and historian of pain Joanna Bourke (2003; 2014) as an aggregative social protocol of maintenance established in rough circumstances of well-being (historical spiritual crises, war terror and suffering, loss and grief…) favouring the stability of biographical location (self-narratives, self-description, self-beliefs) incorporated in new theorising (cf. the famous writings of 'emotional refuge' in ethnography and history, Reddy 2001; Leys 2007; Dixon 2003; in politics, Massumi 2002; in clinical practice, Greenberg 2008; in genomics and emotional communities, TallBear 2013; Abu Al-Haj 2012; and in neuropsychiatry, Craig 2008).

29 Notion Butler extracts from reading Benjamin's (1999) studies on the pressure of violence for law-based processes's genesis.

30 Cf. Whitwell's research with birds raised in predator-free insularity conditions in New Zeeland, Massey University (2008). Other visions of limits in artificial adaptation are also explorable through what R Beer (1990; 1996) calls 'minimally cognitive agents'.

31 As with violence arguments, those of pain are similar to the ecological integrity of emotional experiences resembling responsibility and discernment: pain-events are to be thought as the biography-centre motor for perceptive learning and intelligent memory-based biological precaution. In their outline of the raising experiments with Scottish terriers, Thompson & Melzack (1956) showed an alternative empiricist perspective upon self-recognition of pain as a complex emotion, inhibited in adults if at growth pain-events are eliminated, deleting the generability of that experience.

stead of explained. Assessors will be neutral adopting certain axiological scepticism: shall we base our willingly oriented artificial intelligence in pain, violence, loss, mourning… just for the sake of achieving what we roughly understand as conditioners of intelligent intentional behaviour? Where are the moral implications?

My claim is about an equilibrium, and my concern over what actually causes the balancing epistemic access, profitable for the positive fields to carry along with. Regarding other voices on the nature of emotions, consider the notion of emotional experiences as practices (cf. Scheer 2012), or styles (cf. Gammerl 2012) or as instrumental kinds, instead of natural kinds. With this account, a *ƒ*AI would develop certain structures of order as habits of conduct by the needs of its own proper dynamism, its supposed autonomously generated ecological integrity. Rule-following acceptance, then, instead of based in pressure of community, would respond to their own scaled virtual grounding rhythms, and no other (in the sense we cannot conveniently reduce laws from some level to whatever others: that's the sense of scaling complexity). My final point is that one of the most intrinsically explorable traits any intelligent ecologically installed integrator must expose (there in biology exposed, but meta-theoretically given, as something beyond its scientific interests) is responsibility. The emotion of responsibility, as naturalised as Jonas's (1995) 'must-being' action or Ben Ze'ev (2003; 260) proposals, touches its definition as the feeling within a dynamic system by which to response to what is needed to be dealt with in a global space of action. The emotional experience of feeling oneself responsible for something, a form of a 'feeling towards' in Goldie's terminology (cf. 2002a), is not a bad translation into narrative perspectives of biography of 'the proper feeling' (Goldie 2003) to be felt. This condition applied to acceptance in rule-following permits to the *ƒ*AI a licit understanding of what is the ecological necessity of an scenario of affairs concerning a rule, and how to assess if accepting or not the systemic actions it depicts, tracing a respectful axiology of independence. Rethinking if this independence is allowed as a high scale manifestation of complexity through decision will depend on that specific domain of complexity, the frame through which the *ƒ*AI is systemically integrated. If it shared with us those traits, it would be assessable in certain equilibrium. Maybe they come bearing a different way of 'being responsible', but how many ways have responsibility that are in biosystems incomprehensible to us? How many forms of feeling responsible for something could we consider? The epistemic access comes, at any case, with what is to be assessable too, not with what is to be assessed. That is the dilemma's complication I have been exposing hitherto, placing artificiality fields with a *Reliant's Regress*: we need to give certain independence to a considerable *ƒ*AI to be generated, though in the same sense for us to assessing it as independent it needs to satisfy grounding conditions established by the experimenter's definition of the searched traits, cancelling autonomy. The whiting biting its tail.

. *Results & final remarks*

Attainment of *ƒ*AI as an instrument of itself, and not just as a machine, would include an intimate perspective that could serve to be a prominent axiological trait for assessing if such AI has autonomously developed intelligent features, evident for itself. Making an analogy out of this, it would be like a violin or a piano, easily conceivable as 'machines that produce' (in contrast to 'apparatus that prepare' —from *apparare*—) environmental structures… not just sounds but an order of experiences. This is to say, that order produced with the help of the instrument is, *instrumentally actualising* certain meaning to certain people, organisms and so on. Applying this metaphor, if a *ƒ*AI is an instrument of itself, it means something with its playing itself (Mumford 1952, 63; 1951). This 'meaning something with' invites us to think, I argue, if autonomously provided by the machine, that it exposes attempts at licit scale forms of intelligence (intentionality, culture provoking acts, collective interpretation of facts, sense of intangible immediacy, and so forth) in correlation with the *responsibility principle* management of decision making

in rule-following, achieving some extent of equilibrium between the parts of the dilemma: epistemic grounding conditioning (autonomy/simulation) and axiological definability (anthropocentrism and responsibility). As a basic result: among other factors, a plausible axiology for assessing scientific and technological attainment of *ƒ*AI shall study the balancing points of those problems, considering three underpinning tokens, concerning 1, the ecology of affairs in AS's landscape of usage and integrity; 2, its range of independence in accepting rule-following; and finally 3, the responsibility dynamism in emotional communities of *ƒ*AIs, defining their system of actions and autonomy in licit conditioning.

. Conclusive remarks

An account on the epistemic access and the axiological viability in AI has been offered, pointing out a genetic dilemma. A revision of the thematic implications in the current fields of neurotechnology, simulation, theory making and philosophy of emotions sorted the different problems the text wanted to make visible. As a general conclusion about the scientific and technological practice for assessing attainment of strong scales of artificial intelligence ended up with a devitation, the *Reliant's Regress,* in regard of the problem of 'forms of being' of 'Intelligence' and the variability of the limits that determine human definitions. Dealing with epistemic scepticism, an instrumental equilibrium between the several parts of that dilemma has been proposed, reforming a critical perspective towards contemporary experimentation in artificial and virtual fields. These concluding factors, open to further evaluation, are presented in the text to be taken into account for a renewed application of neurophilosophy and philosophy of technology to current aims, and the ones to come…

—

Abbott, A (2002) «Brain implants show promise against obsessive disorder» in *Nature 419*: 658.

Abbott, LF & LeMasson, G (1993) «Analysis of neuron models with dynamically regulated conductances» in *Neur Comput 5*: 823-842.

Abu Al-Haj, N (2012) *The Genealogical Science: The Search for Jewish Origins and the Politics of Epistemology*, Chicago UP: Chicago.

Agmon-Snir, H & Segev, I (1996) «The concept of decision points as a tool in analyzing dendritic computation» in Bower, JM (ed.) *Computational Neuroscience*, 41-46, Academy P: NY.

Allen, NJ & Barres, BA (2009) «Neuroscience: glia — more than just brain glue» in *Nature 457*: 675-677.

Bartra, R (2004) *El Duelo de los Ángeles. Locura Sublime, Tedio y Melancolía en el Pensamiento Moderno*, Pre-Textos: Valencia.

Bates, J, Bryan Loyall, a & Scott Reilly, W (1992) «Integrating reactivity, goals and emotion in a broad agent» in *Carnegie-Mellon U Computer Science Report*: 92-142.

Beer, RD (1990) *Intelligence as Adaptive Behavior: An Experiment in Computational Neuroethology*, Academic Press.

— (1996) «Toward the evolution of dynamical neural networks for minimally cognitive behavior» in Maes, P et al. (eds.) *From Animals to Animats 4: Proceedings of the Fourth International Conference on adaptive Behavior*, 421-429, MIT P.

Bourke, J (2014) *The Story of Pain*, Oxford UP: NY.

— (2003) «Fear and anxiety: Writing about emotion in modern history» in *History Workshop J*: 111-133.

Butler, J (2006) «Critique, coercion, and sacred life in Benjamin's "Critique of Violence"» in de Vries, H & Sullivan, LE (eds.) *Political Theologies*, 201-219, Fordham UP: NY

Bruce, B (1996) «Tres caras del determinismo tecnológico» in Smith, M & Marx, L (eds.) *Historia y Determinismo Tecnológico*, Alianza: Madrid.

Bruner, J (1986) *Actual Minds, Possible Worlds*, Harvard UP: Cambridge.

Carr, BJ & Rees, MJ (1979) «The Anthropic Principle and the Structure of the Physical World» in *Nature 278*: 605-612.

Cartwright, N (1999) *The Dappled World*, Cambridge UP: Cambridge.

Castelfranchi, C (1995) «Social power» in *Decentralized AI—Proceeding 1st MAAMAW 89*: 49-62.

Chalmers, D (1996) *The Conscious Mind*, Oxford UP: NY.

Claudel, RM (2006) «Memory in astrocytes. A hypothesis» en *Theor Biol Med Model 3*: 2 (rev.).

Cohen, PR & Lavesque, HJ (1990) «Intention is choice with commitment» in *AI 42*: 213-261.

Colby, KM (1976) «Clinical implications of a simulation model of paranoid process» in *Arc Gen Psychiatry 33:* 854-857.

Craig, A D (2008) «Interoception of Emotion: a Neuroanatomical Perspective» in Lewis, M (ed.) *Handbook of Emotions*, 272-289, Guilford: NY

Cromby, J (2007) «Integrating social science with neuroscience: Potentials and problems» in *BioSocieties 2*: 149-169.

Damoiseaux, JS & Greicius, MD (2009) «Greater than the sum of its parts: a review of studies combining structural connectivity and resting-state functional connectivity» in *Brain Struct Funct 213:* 525-533.

Dayan, P & Abbott, L (2001) *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, MIT P: Cambridge.

Dawkins, R (1983) *The Extended Phenotype*, Oxford UP: NY.

Deutsch, D (1997) *The Fabric of Reality*, Penguin: London.

Diéguez, A (2012) *La Vida Bajo Escrutinio*, Buridán: BCN.

Dixon, T (2003) *From Passions to Emotions*, Cambridge UP: Cambridge.

Doyle, R (1997) *On Beyond Living: Rhetorical Transformations of the Life Sciences*, Stanford UP: Stanford.

— (2003) *Wetwares: Experiments in Postvital Living*, Minnesota UP: Minneapolis.

Eeckman, F (ed.) (1993) *Computation and the Neural Systems*, Kluwer: Norwell.

Emmeche, C (1991) *The Garden in the Machine: The Emerging Sci-*

ence of Artificial Life, Princeton UP: Princeton.
Fischer, MJ (2003) *Emergent Forms of Life and the Anthropological Voice*, Duke UP: Durham.
Friederici, AD et al (2006) «The brain differentiates human and non-human grammars: Functional localization and structural connectivity» in *Proc Natl Acad Sci USA 103*: 2548-63.
Franklin, A (1986) *The Neglect of Experiment*, Cambridge UP: Cambridge.
— (1999) *Can that be Right? Essays on experiment, evidence, and Science*. Boston Studies in the Philosophy of Science, Springer.
Franklin, A & Howson, C (1988) «It probably is a valid experimental result: a Bayesian approach to the epistemology of experiment», in *Studies in the History and Philosophy of Science 19*: 419–427.
Galison, P (1996) «Computer simulation and the trading zone» in Galison, P & Stump, DJ (eds.) *The Disunity of Science: Boundaries, Contexts, and Power*, 118-157, Stanford UP: Stanford.
— (2004) «Specific theory» in *Critical Enquiry 30, 2:* 379-383.
— (2008) «Ten problems in history and philosophy of science» in *Isis 99*: 111-124.
Galliers, JR (1988) *A Theoretical Framework for Computer Models of Cooperative Dialogue, Acknowledging Multi-Agent Conflict*, Phd Thesis: Open University.
Gammerl, B (2012) «Emotional styles: Concepts and challenges» in *Theory Practice 16:* 161-175.
Gardner,-Thorpe, C & Pearn, J (2006) «The Cotard syndrome. Report of two patients: with a review of the extended spectrum of "délire des négations"» in *Eu J Neurology 11*: 563-6.
Geraci, R (2010) *Apocalyptic AI: Visions of Heaven in Robotics, Artificial Intelligence, and Virtual Reality*, Oxford UP: Oxford.
Giere, R (1999) *Science without Laws*, Chicago UP: Chicago.
Goertzel, B (1993) *The Structure of Intelligence*, Spring-Verlag: NY.
Goldberg, G & Bloom, KK (1990) «The alien hand sign: localization, lateralization and recovery» in *Am J Phys Med Rehab 69*: 228-38.
Goldie, PL (2000) *The Emotions*, Oxford UP: Oxford.
— (2002a) «Emotions, feelings and intentionality» in *Phenom Cogn Sci 1*: 235-254.
— (2002b) «Emotion, personality and simulation» in Goldie, PL (ed.) *Understanding Emotions: Mind and Morals*, 97-109, Ashgate P: Aldershot.
— (2003) «Narrative and perspective; values and appropriate emotions» in Hatzimoysis, A (ed.) *Philosophy and the emotions*, 201-220, Cambridge UP: Cambridge.
Gottfried, JA & Dolan, RJ (2003) «The nose smells what the eye sees: Crossmodal visual facilitation of human olfactory perception» in *Neuron 39:* 375-386.
Greenberg, L (2008) «The clinical application of emotion to psychotherapy» in Lewis, M et al., *Handbook of Emotions*, 88-101, Guilford: NY.
Gustafsson, M & Balkenius, C (2004) «Using semantic web techniques for validation of cognitive models against neuroscientific data» in *Swedish Artificial Intelligence Society-Swedish Society for Learning Systems Workshop* (On Line) -www.lucs.lu.se/People/Christian.Balkenius/PDF/Gustafsson.Balkenius.2004.pdf.-[consulted 8.1.2005].
Hacking, I (1981) «Do we see through a microscope?» in *Pacific Philosophical Quarterly 63*: 305-322.
— (1983) *Representing and Intervening*, Cambridge: Cambridge.
Haldane, JBS (1947) *What is Life?*, B&G: NY.
Haraway, D (1991) *Simians, Cyborgs, and Women*, Routledge: NY.
Harding, S (1991) *Whose Science? Whose Knowledge?* Cornell UP: NY.
— (1993) «Rethinking standpoint epistemology: what is strong objectivity?» in Alcoff, L & Potter, E (eds.) *Feminist Epistemologies*, Routledge: London.
Harré, R (2003) «The materiality of instruments in a metaphysics for experiments» in Radder, H (ed.) (2003) *The Philosophy of Scientific Experimentation*, pg. 19-38, Pittsburg UP: Pittsburg.
Hayles, KN (1995) «Simulated nature and natural simulations: rethinking the relation between the beholder and the world» in Cronon, W (ed.) *Uncommon Ground*, 409-425, Norton: NY.
Hebb, DO (1964) *Organization of Behaviour*, Wiley: NY.
Helmreich, S (2007) «'Life is a Verb': Inflections of Artificial Life in Cultural Context» in *Artificial Life 13, 2*: 189-201.
Helmreich, S & Roosth, S (2016) «Life forms: a keyword entry» in Helmreich, S (2016) *Sounding the Limits of Life*, 19-34, Princeton UP: Princeton.
Hopfield, JJ (1982) «Neural networks and physical systems with emergent collective computational abilities» in *Proc Natl Acad Sci USA 79*: 2554-2558.
Johnson-Laird, P (1988) *The Computer and the Mind*, Harvard UP: Cambridge.
Jonas, H (1995) *El Principio de Responsabilidad*, Herder: BCN
Kauffman, S (1995) *At Home in the Universe: The Search for the Laws of Self-Orgnaization and Complexity*, Oxford UP: Oxford.
Keller, EF (1995) *Refiguring Life: Metaphors of Twentieth-Century Biology*, Columbia UP: NY.
— (2003a) *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines*, Harvard: Cambridge.
— (2003b) «Models, simulation, and "computer experiments"» in Radder, H (ed.) *The Philosophy of Scientific Experimentation*, 198-215, Pittsburg UP: Pittsburg.
— (2005) «Ecosystems, organisms, and machines» in *Bioscience 55, 12*: 1069-1074.
Kettenmann, H & Kirchhoff, F & Verkhratsky, A (2013) «Microglia: New roles for the synaptic stripper» en *Neuron 77*: 10-18.
Kim, HS, Seto-Ohshima, A, Nishiyama, H & Itohara, S (2011) «Normal delay eyeblink in conditioning in mice devoid of astrocytic S100B» in *Neurosci Lett 489*: 148-153.
Kitcher, P (1993) *The Advancement of Science*, Oxford U P: Oxford.
— (2003) «What kind of science should be done?» in Lightman, A, Sarewitz, D & dresser, C (eds.) *Living with the Genie*, 201-224, Island P: Washington.
— (2004) «Evolutionary theory and the social uses of biology» in *Biology Philosophy 19*: 1-15.
Kurzweil, R (2005) *The Singularity is Near: When Humans Transcend Biology*, quoted through (2012) *La Singularidad está Cerca*, LolaBooks: Berlin.
— & Grossman, T (2004) *Fantastic Voyage: Live Long Enough to Live Forever*, Rodale: NY.
Langton, CG (1988) «Towards Artificial Life» in *Whole Earth 58*: 74- 79.
Latour, B (2011) *La Esperanza de Pandora*, Gedisa: BCN.
Laudan, L & Leplin, J (1991) «Empirical equivalence and underdetermination» in *J Phil 88, 9*: 449-472.
Le Gérer, A (ed.) (2002) «Olfaction and cognition: A philosophical and psychological view» in *Olfaction, Taste, Cognition*, Cambridge UP: NY.
Leys, R (2007) *From Guilty to Shame,* Princeton UP: Woodstock.
Longino, H (1990) *Science as Social Knowledge,* Princeton UP: Princeton.
— (2000) «Reply to P Kitcher» in *Philosophy of Science 69*, 573-577.
Luisi, PL (2006) *The Emergence of Life: From Chemical Origins to Synthetic Biology*, Cambridge UP: Cambridge.
Maes, P (1994) «Social interface agents: acquiring competence by learning from users and other agents» in *Software Agents Report*: 71-78.
Massey University, University of the New Zealand (2008) (On line) «Native birds feel no fear when facing foes» in http://www.mas-

sey.ac.nz/massey/about-massey/news/article.cfm?mnarticle_uuid=422a8039-96BF-57Fe-a306-47F73aCa91BC [consulted in 16.2.2016].
Massumi, B (2002) *Parables of the Virtual*, NC: Durham.
Moravec, H (1988) *Mind Children: The Future of Robot and Human Intelligence*, Harvard UP: Cambridge.
— (1989) «Human culture: a genetic takeover underway» in Langton, CG (ed.) *Artificial Life*, 167-199, Addison-Wesley: Redwood.
Morton, T (2013) *Hyperobjects: Philosophy and Ecology after the End of The World*, Minnesota UP: Minneapolis.
Morton, O (2000) «The computable cosmos of David Deutsch» in *American Scholar 69, 3:* 51-67.
Moser, U et al. (1970) «Computer simulation of a model of neurotic defence process» in *Int J Psychoanalisis 50*: 53-64.
Mumford, L (1951) *The Conduct of Life*, Harcourt: NY.
— (1952) *Art and Technics*, Oxford UP: London.
— (1964) «Authoritarian and democratic technics» in *Tech, Culture 1*: 1-8.
— (1970) *The Myth of the Machine II*, Harcourt: NY.
Niiniluoto, I (1984) *Is Science Progressive?*, Riedel: Dordrecht.
Nishiyama, H et al (2002) «Glial protein S100B modultaes long-term neuronal synaptic plasticity» en *Proc Natl Acad Sci USA 99*: 4037-42.
Oyama, S (2000) *The Ontogeny of Information, Developmental Systems and Evolution*, Duke UP: Durham.
Prigogine, I (1993) «Time, dynamics and chaos» in Holte, J (ed.) *Nobel Conference XXVI. Chaos: The New Science*, 55-88, UPA: Lanham.
— (2012) *El Nacimiento del Tiempo*, Tusquets: BCN.
Prigogine, I & Stengers, I (collaborator) (1997) *Time, Chaos and the New Laws of Nature*, Free Press: London.
Pritchard, RM, Heron, W & Hebb, DO (1960) «Visual perception approached by the method of stabilized images» in *Canad J Psychol 14*: 67-77.
Rall, W (1964) «Theoretical significance of dendritic trees for neuronal imput-output relations» in *Neural Theory and Modeling*: 73-97.
— (1970) «Cable properties of dendrites and effect on synaptic location» in Andersen, P & Jansen, JKS (eds.) *Excitatory Synap- tic Mechanisms*, 175-187, Universitetsforlaget: Oslo.
Reddy, WM (2001) *The Navigation of Feeling*, Cambridge UP: Cambridge.
— (2010) «Neuroscience and the fallacies of functionalism» in *History Theory 49*: 412-425.
Richards, RJ (2002) *The Romantic Conception of Life: Science and Philosophy in the Age of Goethe*, Chicago UP: Chicago.
Riskin, J (ed.) (2007) *Genesis Redux: Essays in the History and Philosophy of Artificial Life*, Chicago UP: Chicago.
Rosenblatt, F (1962) *Principles of Neurodynamics*, Spartan: NY.
Rosenwein, B (2006) *Emotional Communities in the Early Middle Ages*, Cornell UP: Ithaca-NY.
Sejnowski, TJ & Rosenberg, CR (1987) «Parallel networks that learn to pronounce English text» in *Complex Systems 1*: 145-168.
Scheer, M (2012) «Are emotions a kind of practice (and is that what makes them have history)?» in *History Theory 51*: 193-220.
Shapiro, LA, Bialowas-McGoey, LA & Whitaker-Azmitia, PM (2010) «Effects of S100B on serotonergic plasticity and neuroinflammation in the hippocampus in Down syndrome and Alzheimer's disease» in *Cardiovasc Psychiatry Neurol 2010, pii*: 153657.
Sherwood, CC, Stimpson, CD & Raghanti, MA et al (2006) «Evolution of increased glia-neuron ratios in frontal cortex» in *Proc Natl Acad Sci USA 103*: 13606-13611
Sherrington, CS (1947) *The Integrative Action of the Nervous System*, Yale: New Haven.
Shepherd, G (2012) *Neurogastronomy. How the Brain Creates Flavor and Why it Matters*, Columbia UP: NY.
Skarda, CA & Freeman, WJ (1990) «Chaos and the new science of the brain» in *Concepts in Neuroscience 1, 2*: 275-85.
Slezak, M & Pfrieger, FW & Soltys, Z (2006) «Synaptic plasticity, astrocytes and morphological homeostasis» in *J Physiology 99*: 84-91.
Smith, Q (1985) «The anthropic principle and many-worlds cosmologies» in *Australasian J Phil 63, 3*: 336-348.
Subedi, B & Grossberg, GT (2011) «Phantom limb pain: mechanisms and treatment approaches» in *Pain Res Treat (Ebub)*: 864605.
TallBear, K (2013) *Native American DNA: Tribal Belonging and the False Promise of Genetic Science*, Minnesota UP: Minneapolis.
Thompson, WR & Melzack, R (1956) «Early environment» in *Sci Am 194*: 38-42.
Tomkins, SS & Messick, S (eds.) (1963) *Computer Simulation of Personality*, Wiley: NY.
van Fraassen, BC (1980) *The Scientific Image*, Clarendon: Oxford.
— (1976) «To save the phenomena» in Boyd, P; Gasper, P & Trout, JD (eds.) (1992) *The Philosophy of Science*, 187-194, MIT: Cambridge.
Weibel, P (1999) «Virtual worlds: the emperor's new bodies» in Druckery, T (ed.) *Ars Electronica*, 207-223, MIT P: Cambridge.
Weinberg, S (1993) *Dreams of a Final Theory*, Pantheon: NY.
— (2001) «Can science explain everything? Anything?» in *NY Rev Books*: 50.
Whewell, W (1855) *The Plurality of Worlds*, GL: Boston.
Winner, L (1979) *Tecnología Autónoma*, GG: BCN.
Wittgenstein, L (2009) «Investigaciones Filosóficas» in *Wittgenstein, v. II*, Gredos: BCN.
Woodwards, J (1997) «Explanation, invariance and intervention» in *Philosophy of Science 64 (Supp.):* 26-41.
— (2003) «Experimentation, causal inference and instrumental realism» in Radder, H (ed.) *The Philosophy of Scientific Experimentation*, 87-118, Pittsburg UP: Pittsburg.
Zipser, D (1992) «Identification models of the nervous system» in *Neurosc 47*: 853-862.

.