COMPLEXITY AND UNCERTAINTY

A PRUDENTIAL APPROACH TO NANOTECHNOLOGY[1]

Jean-Pierre Dupuy[2]
March 2004

[1] A contribution to the work in progress of the "Foresighting the New Technology Wave" High-Level Expert Group, European Commission, Brussels.

[2] Ecole Polytechnique, Paris, and Stanford University. Email: jpdupuy@stanford.edu.

SUMMARY

We have become capable of tampering with, and triggering off, *complex* phenomena. As a consequence we have to confront a new kind of *uncertainty*. The "Precautionary Principle" is of little help in that task. *Anticipating the consequences* of our technological choices is at the same time more important and more difficult than ever. What is desperately required is a novel *science of the future*.

TABLE OF CONTENTS

1. The debate about molecular manufacturing

Eric Drexler, the inventor of the notion of nanotechnology, and Christine Peterson, the President of the Foresight Institute, are notoriously keen to make the distinction between "near-term nanotechnology" and "advanced nanotechnology". The former refers to any technology smaller than microtechnology, e.g. nanoparticles; the latter to " complete control of the physical structure of matter, all the way down to the atomic level.[3]" It is of course advanced nanotechnology, also known as molecular manufacturing, that will have major societal impact and possibly entail major risks, provided that ... it will see the light of day.

As is well known, controversy is still raging about the physical, technical, industrial, economical feasibility of molecular manufacturing. As Peterson puts it, "Until this issue has been put to rest, neither a funded molecular manufacturing R&D project nor effective study of societal implications can be carried out. [...] We urgently need a basic feasibility review in which molecular manufacturing's proponents and critics can present their technical cases to a group of unbiased physicists for analysis[4]."

In July 2003, the UK Economic and Social Research Council published a report entitled "The Social and Economic Challenges of Nanotechnology". It pointed to the current debate "about whether the radical view of nanotechnology, leading to molecular manufacturing, is feasible or practical, whether by the route sketched out by Drexler or some other means. Those who consider this radical view of nanotechnology to be feasible are divided as to whether it will lead to a positive or negative outcome for society. This debate takes for granted that nanotechnology will have a revolutionary effect on society, and the contrasting visions are correspondingly utopian or dystopian."

On 18 November 2003, the US Senate passed the 21st Century Nanotechnology Research and Development Act, "to authorize appropriations for nanoscience, nanoengineering and nanotechnology research, and for other purposes". It called for a one-time study on the responsible development of nanotechnology "including, but not limited to, *self-replicating nanoscale machines or devices; the release of such machines in natural environments*; encryption; the development of defensive technologies; the use of nanotechnology in the enhancement of

[3] Christine Peterson, a testimony given before the U.S. House of Representatives Committee on Science, April 9, 2003.
[4] Ibid.

human intelligence; and the use of nanotechnology in developing AI.[5]" Many have interpreted this as an opportunity and a challenge to those who support Drexler's vision of molecular manufacturing to make their case, or even as an endorsement of the feasibility of that program. In contrast, the studies performed by the UK's Royal Society/Royal Academy of Engineering are still wondering what nanotechnology is all about, without the least mention of molecular manufacturing.

Richard Smalley, the Nobel laureate in chemistry who was one of the discoverers of the fullerene (C 60), has been challenging Eric Drexler on the possibility of molecular manufacturing. Recently the former accused the latter of scaring children with stories of self-replicating nanobots going haywire, and the latter replied by saying, "U.S. progress in molecular manufacturing has been impeded by the dangerous illusion that it is infeasible. [...]Building with atomic precision will dramatically extend the range of potential products and decrease environmental impact as well. The resulting abilities will be so powerful that, in a competitive world, failure to develop molecular manufacturing would be equivalent to unilateral disarmament.[6]"

The debate between the two men has also been quite technical, and it is all about the limitations of chemistry. Smalley asserts that atoms cannot simply be pushed together to make them react as desired, in the manner fancied by Drexler, but that their chemical environment must be controlled in great detail, through a many-dimensional hyperspace, and that this cannot be achieved with simple robotics. Drexler rejoins that such components of cells as enzymes or ribosomes are able to do precise and reliable chemistry. Smalley agrees but adds that this can occur only under water. Drexler replies that his proposal does assert that chemistry in dry surfaces and a vacuum ("machine-phase chemistry") can be quite flexible and efficient, since holding a molecule in one place can have a strong catalytic effect. Drexler ends his statements by calling for further research, beginning with an independent scientific review of molecular manufacturing concepts.

An advocate of Drexler's program recently wrote:

---

[5] My emphasis.

[6] "Nanotechnology. Drexler and Smalley make the case for and against 'molecular assemblers'", *Chemical & Engineering News*, December 1, 2003. See http://pubs.acs.org/cen/coverstory/8148/8148counterpoint.html.

> *Failure to anticipate the development of molecular manufacturing could have serious consequences*. Simple physics theories, conservatively applied, predict that the technology will be dangerously powerful. A working molecular nanotechnology will likely require the design and enforcement of policies to control the use of compact advanced manufacturing systems and their products. But panicked last-minute policy will be bad policy—simultaneously oppressive and ineffective. The military implications are even more perilous. Molecular manufacturing systems are expected to be able to produce weapons as powerful as nuclear bombs, but much more selective, easier to manufacture, and easier to use. If a powerful nation suddenly realizes that molecular manufacturing is possible, and discovers that rival nations are already making material progress, they may react violently, or may enter into an arms race that will probably be unstable and thus may result in war with weapons of unprecedented power.
>
> On the positive side, molecular manufacturing may be able to mitigate many of the world's humanitarian and environmental crises. Advancing its development by even a year or two could alleviate untold suffering, raising standards of living worldwide while sharply reducing our environmental footprint. However, rapid and effective humanitarian use may also depend on sound policy developed well in advance[7].

My opinion on this is the following. The Smalley – Drexler debate is a red herring, and we should refrain from taking a position about it, even if we had the scientific and technological expertise to do so. There is no doubt that molecular manufacturing is feasible once we regard molecular biology itself as a form of it. *The issue is not one of essence but of point of view*. As soon as we construe the cell as natural machinery, the possibility of tampering with it becomes a forgone conclusion. If the feasibility of molecular self-assembly is beyond question, it is because we have developed a view of nature and the living system that is akin to our own artifacts.

---

[7] Chris Phoenix, "Of Chemistry, Nanobots, and Policy", Center for Responsible Nanotechnology, December 2003. http://crnano.org/Debate.htm.

2. Complexity and Self-organization

It is often asserted that the starting point of nanotechnology was the classic talk given by Feynman in 1959[8], in which he said: "The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. [...] It would be, in principle, possible (I think) for a physicist to synthesize any chemical substance that the chemist writes down. Give the orders and the physicist synthesizes it. How? Put the atoms down where the chemist says, and so you make the substance." Today's champions of nanotech add: "We need to apply at the molecular scale the concept that has demonstrated its effectiveness at the macroscopic scale: making parts go where we want by *putting* them where we want![9]"

I tend to disagree. If such were the essence of (advanced) nanotechnology, the worries that it raises would rest on sheer ignorance. As *Nature* science writer Philip Ball puts it in his excellent essay, "2003: nanotechnology in the firing line"[10]:

> In March [2003], the Royal Institution (RI) in London hosted a day-long seminar on nanotech called "Atom by atom", which I personally found useful for hearing a broad cross-section of opinions on what has become known as *nanoethics*. [...] First, the worry was raised that what is qualitatively new about nanotech is that it allows, for the first time, the manipulation of matter at the atomic scale. This may be a common view, and it must force us to ask: how can it be that we live in a society where it is not generally appreciated that this is what chemistry has done in a rational and informed way for the past two centuries and more? How have we let that happen? It is becoming increasingly clear that the debate about the ultimate scope and possibilities of nanotech revolve around questions of basic chemistry [...]. The knowledge vacuum in which much public debate of nanotech is taking place exists because we have little public understanding of chemistry: what it is, what it does, and what it can do.

---

[8] "There's Plenty of Room At the Bottom".

[9] See http://www.zyvex.com/nano/.

[10] Nanotechweb.org, 23 December 2003. http://www.nanotechweb.org/articles/society/2/12/1/1

Writing about nanoethics, Ball goes on to say:

> Questions about safety, equity, military involvement and openness are ones that pertain to many other areas of science and technology. It would be a grave and possibly dangerous distortion if nanotechnology were to come to be seen as a discipline that raises unprecedented ethical and moral issues. In this respect, I think it genuinely does differ from some aspects of biotechnological research, which broach entirely new moral questions. Yet it is perhaps the first major field of science, applied science or technology - call it what you will - to have emerged in a social climate that is sensitized in advance to the need for ethical debate in emerging technologies.
>
> [...]Yet the pragmatic truth is that if nanotechnology does not acknowledge some kind of ethical dimension, it will be forced upon it in any case. Those working in the field know that nanotech is not really a discipline at all, that it has no coherent aims and is not the sole concern of any one industrial sector. But even funding agencies speak of it as though this were not so. To the public mind, organizations such as the US National Nanotechnology Initiative surely suggest by their very existence that nanotech has some unity, and it is therefore quite proper that people will want to be reassured that its ethical aspects are being considered.

Here I cannot follow Philip Ball. I believe him to be wrong on two major accounts. I believe there is indeed some kind of unity behind the nanotech enterprise and even behind the NBIC convergence; but that this unity lies at the level of the metaphysical research program that underpins such convergence. I also believe that the ethical issues raised by it are to a large extent novel and that they find their source in the very *ideas* that govern the field.

In order to substantiate those two claims, I submit that the origin of the new field is to be sought in another classic conference, the one John von Neumann gave at Caltech in 1948 on complexity and self-reproducing automata.

Turing's and Church's theses were very influential at the time, and they had been supplemented by cyberneticians Warren McCulloch and Walter Pitts' major finding on the properties of neural networks. Cybernetics'

Credo was then: every behavior that is unambiguously describable in a finite number of words is computable by a network of formal neurons---a remarkable statement, as John von Neumann recognized. However, he put forward the following objection: is it reasonable to assume as a practical matter that our most complex behaviors are describable in their totality, without ambiguity, using a finite number of words? In specific cases it is always possible: our capacity, for example, to recognize the same triangular form in two empirical triangles displaying differences in line, size, and position can be so described. But would this be possible if it were a matter of globally characterizing our capacity for establishing "visual analogies"? In that case, von Neumann conjectured, *it may be that the simplest way to describe a behavior is to describe the structure that generates it*. It is meaningless, under these circumstances, to "discover" that such a behavior can be embodied in a neural network since it is not possible to define the behavior other than by describing the network itself.

Von Neumann thus posed the question of complexity, foreseeing that it would become the great question for science in the future. Complexity implied for him in this case the futility of the constructive approach of McCulloch and Pitts, which reduced a function to a structure---leaving unanswered the question of what a complex structure is capable[11].

It was of course in the course of his work on automata theory that von Neumann was to refine this notion of complexity. Assuming a magnitude of a thermodynamical type, he conjectured that below a certain threshold it would be degenerative, meaning that the degree of organization could only decrease, but that above this threshold an increase in complexity became possible. Now this threshold of complexity, he supposed, is also the point at which the structure of an object becomes simpler than the description of its properties. Soon, JVN prophesied, the builder of automata would find himself as helpless before his creation as we feel ourselves to be in the presence of complex natural phenomena.[12]

At any rate, JVN was thus founding the so-called *bottom-up approach* aka reverse engineering. In keeping with that philosophy, the engineers of the future will not be any more the ones who devise and design a structure

---

[11] Here as elsewhere, the irony of intellectual history is great. Marvin Minsky, who wrote his doctoral thesis under von Neumann, regarded his teacher's attack on McCulloch's approach as an aberration, an admission of weakness, a lack of faith in what he himself, John von Neumann, had managed to accomplish. Now, as is well known, Eric Drexler wrote his dissertation on nanotech under Minsky's supervision!

[12] On all that, see my *The Mechanization of the Mind*, Princeton University Press, 2000.

capable of fulfilling a function that has been assigned to them. The engineers of the future will be the ones who know they are successful when they are surprised by their own creations. If one of your goals is to reproduce life, to fabricate life, you have to be able to simulate one of its most essential properties, namely the capacity to complexify itself always more.

Admittedly, not all of nanotech falls under the category of complexity. However, the scope covered by it, especially in the case of the NBIC convergence, is much wider and relevant than the implications of a possible Drexler-type molecular manufacturing. Even more importantly, the novel kind of uncertainty that is brought about by those new technologies is intimately linked with their being able to set off complex phenomena in the Neumannian sense.

3. Unchaining Complexity

"The unleashed power of the atom has changed everything
save our modes of thinking, and we thus drift toward unparalleled catastrophe."
Albert Einstein

It would be a mistake to think that, although novel, our current situation before the consequences of our technological choices is not the outcome of a long historical process. Discontinuities and ruptures must always be analyzed against the background of continuous dynamics. In her masterly study of the frailties of human action, *Human Condition*[13], Hannah Arendt brought out the fundamental paradox of our time: as human powers increase through technological progress, we are less and less equipped to control the consequences of our actions. A long excerpt is worth quoting here, as its relevance for our topic cannot be overstated – and we should keep in mind that this was written in 1958:

> [...] the attempt to eliminate action because of its uncertainty and to save human affairs from their frailty by dealing with them as though they were or could become the planned products of human making has first of all resulted in channeling the human capacity for action, for beginning new and spontaneous processes which without men never would come into existence, into an attitude toward nature which up to the latest stage of the modern age had been one of exploring natural laws and fabricating objects out of

[13] The University of Chicago Press, 1958.

> natural material. To what extent we have begun to *act into nature*, in the literal sense of the word, is perhaps best illustrated by a recent casual remark of a scientist who quite seriously suggested that "*basic research is when I am doing what I don't know what I am doing*." [Wernher von Braun, December 1957].
>
> This started harmlessly enough with the experiment in which men were no longer content to observe, to register, and contemplate whatever nature was willing to yield in her own appearance, but began to prescribe conditions and to provoke natural processes. What then developed into an ever-increasing skill in *unchaining elemental processes*, which, without the interference of men, would have lain dormant and perhaps never have come to pass, has finally ended in a veritable art of '*making*' *nature*, that is, of creating 'natural' processes which without men would never exist and which earthly nature by herself seems incapable of accomplishing [...].
>
> The very fact that natural sciences have become exclusively sciences of process and, in their last stage, *sciences of potentially irreversible, irremediable 'processes of no return'* is a clear indication that, whatever the brain power necessary to start them, *the actual underlying human capacity which alone could bring about this development is no 'theoretical' capacity, neither contemplation nor reason, but the human ability to act* – to start new unprecedented processes whose outcome remains uncertain and unpredictable whether they are let loose in the human or the natural realm.
>
> In this aspect of action [...] processes are started whose outcome is unpredictable, so that *uncertainty rather than frailty becomes the decisive character of human affairs*[14].

No doubt that with an incredible prescience this analysis applies perfectly well to the NBIC convergence, in particular on two scores. Firstly, the ambition to (re-) make nature is an important dimension of what I called the metaphysical underpinnings of the field. If the NBIC converging technologies purport to take over Nature's and Life's job and become the

[14] P. 230-232. My emphasis.

engineers of evolution, it is because they have redefined Nature and Life in terms that belong to the realm of artifacts. See how one of their most vocal champions, Damien Broderick, rewrites the history of life, or, as he puts it, of "living replicators":

> *Genetic algorithms* in planetary numbers lurched about on the surface of the earth and under the sea, and indeed as we now know deep within it, for billions of years, replicating and mutating and being winnowed via the success of their expressions – that is, the bodies they *manufactured*, competing for survival in the macro world. At last, the entire living ecology of the planet has *accumulated, and represents a colossal quantity of compressed, schematic information*.[15]

Once life has thus been transmogrified into an artifact, the next step is to ask oneself whether the human mind couldn't do better. The same author asks rhetorically, "Is it likely that nanosystems, designed by human minds, will bypass all this ¨Darwinian wandering, and leap straight to design success?[16]"

Secondly, as explained before, it will be an inevitable temptation, not to say a task or a duty, for the nanotechnologists of the future to set off processes upon which they have no control. The sorcerer's apprentice myth must be updated: it is neither by error nor by terror that Man will be dispossessed of his own creations but by design.

There is no need for Drexlerian self-assemblers to come into existence for this to happen. The paradigm of *complex, self-organizing systems* envisioned by von Neumann is stepping ahead at an accelerated pace, both in science and in technology. It is in the process of shoving away and replacing the old metaphors inherited from the cybernetic paradigm, like the ones that treat the mind or the genome as computer programs. In science, the central dogmas of molecular biology received a severe blow on two occasions recently. First, with the discovery that the genome of an adult, differentiated cell can be "reprogrammed" with the cooperation of maternal cytoplasm – hence the technologies of nucleus transfer, including therapeutic and reproductive cloning. Secondly, with the discovery of prions, which showed that self-replication does not require

---

[15] Damien Broderick, *The Spike*, Forge, New York, 2001, p. 116. My emphasis.
[16] Ibid., p. 118.

DNA. As a result, the sequencing of the human genome appears to be not the end of the road but its timid beginning. Proteinomics and Complexity are becoming the catchwords in biology, relegating Genomics to the realm of passé ideas.

In technology, new feats are being flaunted every passing week. Again, the time has not come – and may never come – when we manufacture self-replicating machinery that mimics the self-replication of living materials. However, we are taking more and more control of living materials and their capacity for self-organization and we use them to mimic smart machinery or perform mechanical functions.

Examples are plenty. To name just a few: in December 2003, IBM managed to create silicon memory chips using a template provided by *a plastic polymer that organizes itself naturally*. One application of the technology could be to design flash memory chips with cells roughly 1/100th the size of the cells currently required to store a piece of data. More broadly, IBM said, "the successful research suggests that *polymer-based self-assembly techniques* could be used to build other kinds of microchips in the future, when more features shrink to such small scales that current production techniques become impractical[17]". On the same month, scientists from DuPont, the University of Illinois at Urbana-Champaign and the MIT used the self-assembly of DNA to sort carbon nanotubes according to their diameter and electronic properties. DuPont said, "spontaneous self-assembly of nucleic acid bases occurs on a variety of inorganic surfaces. This phenomenon, considered as an important prebiotic process relevant to the origin of life, has led us to seek a new function for nucleic acids in the manipulation of inorganic nanomaterials, where interfacial interactions dominate." The feat will have momentous applications, since "the separation of carbon nanotubes is the single greatest impediment to their technological application.[18]" Last November, scientists in Israel built transistors out of carbon nanotubes using DNA as a template. A Technion-Israel scientist said, "What we've done is to bring biology to *self-assemble an electronic device* in a test tube [...] The DNA serves as a scaffold, a template that will determine where the carbon

[17] See Barnaby Feder, "I.B.M. set to unveil chip-making advance", *New York Times*, December 8, 2003: http://www.siliconinvestor.com/stocktalk/msg.gsp?msgid=19572729.
[18] Liz Kalaugher, "DNA sorts out nanotubes", *Nanotechweb.org*,, 3 December 2003: http://www.nanotechweb.org/articles/news/2/12/1/1.

nanotubes will sit. That's the beauty of using biology.[19]" And so on and so forth.

4. A new kind of uncertainty and the irrelevance of the Precautionary Principle

Our tampering with, and setting off complex processes, in the technical, Neumannian sense of the word "complex", brings about a kind of uncertainty that is radically novel. In particular, it is completely alien to the distinctions upon which rests the Precautionary Principle.

The precautionary principle introduces what initially appears to be an interesting distinction between two types of risks: "known" risks and "potential" risks. It is on this distinction that the difference between prevention and precaution is made to rest: precaution would be to potential risks what prevention is to known risks.

A closer look reveals 1) that the expression "potential risk" is poorly chosen, and that what it designates is not a risk waiting to be realized, but a hypothetical risk, one that is only a matter of conjecture; 2) that the distinction between known risks and hypothetical risks (the term I will adopt here) corresponds to an old standby of economic thought, the distinction that John Maynard Keynes and Frank Knight independently proposed in 1921 between *risk* and *uncertainty*. A risk can in principle be quantified in terms of objective probabilities based on observable frequencies; when such quantification is not possible, one enters the realm of uncertainty.

The problem is that economic thought and the decision theory underlying it were destined to abandon this distinction as of the 1950s in the wake of the exploit successfully performed by Leonard Savage with the introduction of the concept of subjective probability and the corresponding philosophy of choice under conditions of uncertainty: Bayesianism. In Savage's axiomatics, probabilities no longer correspond to any sort of regularity found in nature, but simply to the coherence displayed by a given agent's choices. In philosophical language, every uncertainty is treated as an *epistemic* uncertainty, meaning an uncertainty associated with the agent's state of knowledge. It is easy to see that the

---

[19] Kenneth Chang, " Smaller Computer Chips Built Using DNA as Template", *New York Times*, November 21, 2003: http://www.nytimes.com/2003/11/21/science/21DNA.html?ex=1075525200&en=67948bd27029a142&ei=5070.

introduction of subjective probabilities erases the distinction between uncertainty and risk, between risk and the risk of risk, between precaution and prevention. If a probability is unknown, a probability distribution is assigned to it "subjectively". Then the probabilities are composed following the computation rules of the same name. No difference remains compared to the case where objective probabilities are available from the outset. Uncertainty owing to lack of knowledge is brought down to the same plane as intrinsic uncertainty due to the random nature of the event under consideration. A risk economist and an insurance theorist do not see and cannot see any essential difference between prevention and precaution and, indeed, reduce the latter to the former. In truth, one observes that applications of the "precautionary principle" generally boil down to little more than a glorified version of "cost-benefit" analysis.

Against the prevailing economism, I believe it is urgent to safeguard the idea that *all is not epistemic uncertainty*. One could however argue from a philosophical standpoint that such is really the case. The fall of a die is what supplied most of our languages with the words for chance or accident. Now, the fall of a die is a physical phenomenon that is viewed today as a low-stability deterministic system, sensitive to initial conditions, and therefore unpredictable — a "*deterministic chaos*," in current parlance. But an omniscient being — the God of whom Laplace did not judge it necessary to postulate the existence — would be able to predict on which side the die is going to fall. Could one not then say that what is uncertain for us, but not for this mathematician-God, is uncertain only because of lack of knowledge on our part? And therefore that this uncertainty, too, is epistemic and subjective?

The correct conclusion is a different one. If a random occurrence is unpredictable for us, this is not because of a lack of knowledge that could be overcome by more extensive research; it is because only an infinite calculator could predict a future which, given our finiteness, we will forever be unable to anticipate. Our finiteness obviously cannot be placed on the same level as the state of our knowledge. The former is an unalterable aspect of the human condition; the latter, a contingent fact, which could at any moment be different from what it is. We are therefore right to treat the random event's uncertainty *for us* as an objective uncertainty, even though this uncertainty would vanish for an infinite observer.

Now, our situation with respect to the complex phenomena we are about to unleash is also one of objective, and not epistemic, uncertainty. The novel feature this time is that we are not dealing with a random

occurrence either. Neither random, nor epistemically uncertain, the type of "risk" that we are confronting is a monster from the standpoint of classic distinctions. Indeed, it merits a special treatment, which the precautionary principle is incapable of giving it.

We know today that what makes a complex system, (e. g. a network of molecules connected by chemical reactions or a trophic system) robust is exactly what makes it exceedingly vulnerable if and when certain circumstances are met. As Albert-László Barabási puts it, this "coexistence of robustness and vulnerability plays a key role in understanding the behavior of most complex systems. [...] topology, robustness, and vulnerability cannot be fully separated from one another. All complex systems have their Achilles' heel.[20]" Complexity gives those systems an extraordinary stability and a no less remarkable resilience. They can hold their own against all sorts of aggressions and find ways of adapting to maintain their stability. This is only true up to a certain point, however. Beyond certain *tipping points*, they veer over abruptly into something different, in the fashion of phase changes of matter, collapsing completely or else forming other types of systems that can have properties highly undesirable for people. In mathematics, such discontinuities are called *catastrophes*. This sudden loss of resilience gives complex systems a particularity which no engineer could transpose into an artificial system without being immediately fired from his job: the alarm signals go off only when it is too late. And in most cases we do not even know where these tipping points are located. Our uncertainty regarding the behavior of complex systems has thus nothing to do with a temporary insufficiency of our knowledge, it has everything to do with objective, structural properties of complex systems.

On the other hand, this uncertainty is not of the kind that is attached to random events and it is not amenable to the concept of probability. The key notion here is that of informational *incompressibility*, which is a form of *essential unpredictability*. In keeping with von Neumann's intuitions on complexity, a complex process is defined today as one for which the simplest model is the process itself. The only way to determine the future of the system is to run it: there are no shortcuts. This is a radical uncertainty: in contrast with a deterministic chaos – the source of randomness –, perfect knowledge of the initial conditions would not be enough to predict the future states of the system. Its unpredictability is irremediable.

---

[20] *Linked. The New Science of Networks*, Perseus Publishing, Cambridge (Mass.), 2002, p. 118 and 121-122.

When the precautionary principle states that the "absence of certainties, given the current state of scientific and technical knowledge, must not delay the adoption of effective and proportionate preventive measures aimed at forestalling a risk of grave and irreversible damage to the environment at an economically acceptable cost", it is clear that it places itself from the outset within the framework of epistemic uncertainty. The presupposition is that *we know we are in a situation of uncertainty*. It is an axiom of epistemic logic that if I do not know p, then I know that I do not know p. Yet, as soon as we depart from this framework, we must entertain the possibility that we do not know that we do not know something. An analogous situation obtains in the realm of perception with the blind spot, that area of the retina unserved by the optic nerve. At the very center of our field of vision, we do not see, but our brain behaves in such a way that we do not see that we do not see. In cases where the uncertainty is such that it entails that the uncertainty itself is uncertain, it is impossible to know whether or not the conditions for the application of the precautionary principle have been met. If we apply the principle to itself, it will invalidate itself before our eyes.

Moreover, "given the current state of scientific and technical knowledge" implies that a scientific research effort could overcome the uncertainty in question, whose existence is viewed as purely contingent. It is a safe bet that a "precautionary policy" will inevitably include the edict that research efforts must be pursued — as if the gap between what is known and what needs to be known could be filled by a supplementary effort on the part of the knowing subject. But it is not uncommon to encounter cases in which *the progress of knowledge comports an increase in uncertainty* for the decision-maker, something that is inconceivable within the framework of epistemic uncertainty. Sometimes, to learn more is to discover hidden complexities that make us realize that the mastery we thought we had over phenomena was in part illusory.

5. Toward a new science of the future

"We have met the Enemy and He is Us."
Pogo Possum

5.1. *In Search of an Ethics of the Future*

German philosopher Hans Jonas's fundamental work, *The Imperative of Responsibility*[21], cogently explains why we need a radically new ethics to rule our relation to the future in the "technological age". This "Ethics of the Future" [*Ethik **für** die Zukunft*] - meaning not a future ethics, but an ethics *for* the future, for the sake of the future, i.e. the future must become the major object of our concern – starts from a philosophical aporia. Given the magnitude of the possible consequences of our technological choices, it is an absolute obligation for us to try and anticipate those consequences, assess them, and ground our choices on this assessment. Couched in philosophical parlance, this is tantamount to saying that when the stakes are high, we cannot afford not to choose consequentialism[22], rather than a form of deontology[23], as our guiding moral doctrine. However, the very same reasons that make consequentialism compelling, and therefore oblige us to anticipate the future, make it impossible for us to do so. Unleashing complex processes is a very perilous activity that both demands foreknowledge and prohibits it. To take just an illustration:

> The unpredictable behaviour of nanoscale objects means that engineers will not know how to make nanomachines until they actually start building them[24].

Now, one of the very few unassailably universal ethical principles is that *ought* implies *can*. There is no obligation to do that which one can not do. However, in the technological age, we do have an ardent obligation that we cannot fulfill: anticipating the future. That is the ethical aporia.

Is there a way out? Jonas's credo, which I share, is that there is no ethics without metaphysics. Only a radical change in metaphysics can allow us

---

[21] Hans Jonas, *The Imperative of Responsibility. In Search of an Ethics for the Technological Age*, University of Chicago Press, 1985.
[22] Consequentialism as a moral doctrine has it that what counts in evaluating an action is its consequences for all individuals concerned.
[23] A deontological doctrine evaluates the rightness of an action in terms of its conformity to a norm or a rule, such as the Kantian categorical imperative.
[24] *The Economist*, March 2003.

to escape from the ethical aporia. The major stumbling block of our current, implicit metaphysics of temporality turns out to be our conception of the *future as indeterminate*. From our belief in free will – we might act otherwise – we derive the conclusion that the future is *not real*, in the philosophical sense: "future contingents", i.e. propositions about actions taken by a free agent in the future, e.g. "John will pay back his debt tomorrow", are held to have no truth value. They are neither true nor false. If the future is not real, it is not something that we can have cognizance of. If the future is not real, it is not something that projects its shadow onto the present. Even when we know that a catastrophe is about to happen, we do not believe it: we do not believe what we know. If the future is not real, there is nothing in it that we should fear, or hope for.

The derivation from free will to the unreality of the future is a sheer logical fallacy, although it would require some hard philosophical work to prove it[25]. Here I will content myself with exhibiting the sketch of an alternative metaphysics in which free will combines with a particularly hard version of the reality of the future.

Before I broach the metaphysical and final part of this discussion, I should like to add a further ethical reflection that compounds the need we are in to bestow some measure of reality onto the future.

I am referring to the concept of "moral luck" in moral philosophy. I will introduce it with the help of two contrasting thought experiments. In the first, one must reach into an urn containing an infinite number of balls and pull one out at random. Two thirds of the balls are black and only one third are white. The idea is to bet on the color of the ball before seeing it. Obviously, one should bet on black. And if one pulls out another ball, one should bet on black again. In fact, one should *always* bet on black, even though one foresees that one out of three times on average this will be an incorrect guess. Suppose that a white ball comes out, so that one discovers that the guess was incorrect. Does this *a posteriori* discovery justify a retrospective change of mind about the rationality of the bet that one

---

[25] See my *Pour un catastrophisme éclairé*, Paris, Seuil, 2002. See also Jean-Pierre Dupuy, "Philosophical Foundations of a New Concept of Equilibrium in the Social Sciences: Projected Equilibrium", *Philosophical Studies*, **100**, 2000, p. 323-345; Jean-Pierre Dupuy, "Two temporalities, two rationalities: a new look at Newcomb's paradox", in P. Bourgine et B. Walliser (eds.), *Economics and Cognitive Science*, Pergamon, 1992, p. 191-220; Jean-Pierre Dupuy, «Common knowledge, common sense», *Theory and Decision*, 27, 1989, p. 37-62. Jean-Pierre Dupuy (ed.), *Self-deception and Paradoxes of Rationality*, C.S.L.I. Publications, Stanford University, 1998.

made? No, of course not; one was right to choose black, even if the next ball to come out happened to be white. Where probabilities are concerned, the information as it becomes available can have no conceivable retroactive impact on one's judgment regarding the rationality of a past decision made in the face of an uncertain or risky future. This is a limitation of probabilistic judgment that has no equivalent in the case of moral judgment. Here we touch upon a second serious deficiency of the precautionary principle. As it is unable to depart from the normativity proper to the calculus of probabilities and the cost-benefit approach, it fails to capture what constitutes the essence of ethical normativity concerning choice in a situation of uncertainty.

A man spends the evening at a cocktail party. Fully aware that he has drunk more than is wise, he nevertheless decides to drive his car home. It is raining, the road is wet, the light turns red, and he slams on the brakes, but a little too late: after briefly skidding, the car comes to a halt *just past* the pedestrian crosswalk. Two scenarios are possible: Either there was nobody in the crosswalk, and the man has escaped with no more than a retrospective fright. Or else the man ran over and killed a child. The judgment of the law, of course, but above all that of morality, will not be the same in both cases. Here is a variant: The man was sober when he drove his car. He has nothing for which to reproach himself. But there is a child whom he runs over and kills, or else there is not. Once more, the unpredictable outcome will have a retroactive impact on the way the man's conduct is judged by others and also by the man himself.

Here is a more complex example devised by the British philosopher Bernard Williams,[26] which I will simplify considerably. A painter — we'll call him "Gauguin" for the sake of convenience — decides to leave his wife and children and take off for Tahiti in order to live a different life which, he hopes, will allow him to paint the masterpieces that it is his ambition to create. Is he right to do so? Is it moral to do so? Williams defends with great subtlety the thesis that any possible justification of his action can only be retrospective. Only the success or failure of his venture will make it possible for us — and him — to cast judgment. Yet whether Gauguin becomes a painter of genius or not is in part a matter of luck — the luck of being able to become what one hopes to be. When Gauguin makes his painful decision, he cannot know what, as the saying goes, the future holds in store for him. To say that he is making a bet would be incredibly reductive. With its appearance of paradox, the concept of "moral luck" provides just what was missing in the means at our disposal

[26]Bernard Williams, *Moral Luck,* Cambridge, Cambridge University Press, 1981.

for describing what is at stake in this type of decision made under conditions of uncertainty.

Like Bernard Williams' Gauguin, but on an entirely different scale, humanity taken as a collective subject has made a choice in the development of its potential capabilities which brings it under the jurisdiction of moral luck. It may be that its choice will lead to great and irreversible catastrophes; it may be that it will find the means to avert them, to get around them, or to get past them. No one can tell which way it will go. The judgment can only be retrospective. However*, it is possible to anticipate, not the judgment itself, but the fact that it must depend on what will be known once the "veil of ignorance" cloaking the future is lifted*. Thus, there is still time to insure that our descendants will never be able to say "*too late!*" — a too late that would mean that they find themselves in a situation where no human life worthy of the name is possible.

Hence the bold metaphysical move advocated by Hans Jonas. The idea is to *project oneself* into the future and look back at our present and evaluate it from there. This temporal *loop* between future and past I call the metaphysics of *projected time*. As we are going to see, it makes sense only if one accepts that the future is not only real but also fixed.

5.2. *A Critique of the Scenario Approach*

For the last half century, futurology has been equated with the scenario approach. If some credit is granted the foregoing, it appears that this method is no longer appropriate to tackle the kind of radical uncertainty that we are confronting.

Ever since its beginnings the scenario approach has gone to great lengths to distinguish itself from mere forecast or foresight, held to be an extension into the future of trends observed in the past. We can forecast the future state of a physical system, it is said, but not what we shall decide to do. It all started in the 50s when a Frenchman, Gaston Berger, coined the term "*Prospective*" – a substantive formed after "Retrospective" – to designate a new way to relate to the future. That this new way had nothing to do with the project or the ambition of anticipating, that is, *knowing* the future, was clearly expressed in the following excerpt from a lecture given by another Frenchman, Bertrand de Jouvenel, in 1964:

> The purpose is to generate a habit, the habit of forward-looking. We feel that as this grows into a habit, we, or our successors, shall develop in this exercise greater skill, thanks to self-criticism and mutual criticism. At the outset we encountered in the authors we solicited a great reluctance to embark upon such speculation. They said it was unscholarly, which of course it is, but it happens to be necessary. It is unscholarly perforce because there are no facts on the future. Cicero quite rightly contrasted past occurrences and occurrences to come with the contrasted expressions *facta* and *futura*: *facta*, what is accomplished and can be taken as solid; *futura*, what shall come into being, and is as yet 'undone,' or fluid. This contrast leads me to assert vigorously: '*there can be no science of the future*.' *The future is not the realm of the 'true or false' but the realm of 'possibles*.'

Another term coined by Jouvenel that was promised to a bright ... future was "*Futuribles*"[27], meaning precisely the open diversity of *possible futures*. The exploration of that diversity was to become the scenario approach.

Again, the premises on which the whole enterprise rests are at best arbitrary metaphysical presuppositions and ones, to repeat, that we can no longer afford to entertain. If we do not bring ourselves to believe in the reality of the future, we'll never be able to measure up to the challenges that lie ahead[28]. And those who claim that those presuppositions derive from the freedom of the will are just committing a serious philosophical blunder. Thus Michel Godet, one of the foremost among today's "prospectivists", could write

> All who claim to foretell or forecast the future are inevitably liars, for the future is not written anywhere – it is still to be built. This is fortunate, for without this uncertainty, human activity would lose its degree of freedom and its meaning – the hope of a desired future. If the future were totally

[27] The tradition launched by Bertrand de Jouvenel continues today in a journal called *Futuribles*, edited by his own son Hugues.

[28] Another early proponent of "Prospective" was Robert Jungk. In 1960, as he was interviewing victims of the Hiroshima atomic bomb, he met a man under 50 who looked 80. That man said to him in a sedate manner: "How could all those intelligent people have dropped this bomb without thinking of the consequences?" That encounter was what prompted Jungk to devote his life to futures studies.

> foreseeable and certain, the present would become unlivable. Certainty is death. Because the future has to be built, it also cannot be conceived as a simple continuation of the past[29].

This passage is also typical of a confusion that spoils much of what is being offered as a justification of the scenario approach. On the one hand, the alleged irreducible multiplicity of the "futuribles" is explained by the *ontological indeterminacy* of the future: since we "build", "invent" the future, there is nothing to know about it. On the other hand, the same multiplicity is interpreted as the inevitable reflection of our inability to know the future *with certainty*. The confusion of ontological indeterminacy with *epistemic uncertainty* is a very serious one, as explained above[30].

To underline the weaknesses of the philosophical foundations of the scenario method is not to deny its many virtues. There is no question that it has helped individuals, groups, and nations to find new ways to coordinate through a jointly worked-out image of the future shared by all. However that has been achieved in a paradoxical way. The method aimed at emphasizing the importance of the future while it denied its reality. Hence the essential question, is there a way to protect the democratic virtues of the scenario approach while jettisoning its flawed metaphysics?

5.3. *From Occurring Time to Projected Time*

If the future is ontologically indeterminate shouldn't we say the same about the past? After all, there was a time when our past was the future of

---

[29] Michel Godet and Fabrice Roubelat, "Creating the future: the use and misuse of scenarios", *Long Range Planning*, 29, 2, 1996.

[30] The World Business Council for Sustainable Development (WBCSD) "Global Scenarios 2000-2050" Summary Brochure provides another illustration of this glaring confusion. On the one hand, we read that "Unlike forecasts, which impose patterns extrapolated from the past onto the future, scenarios are plausible, pertinent, alternative stories that are concerned more with strategic thinking than with strategic planning." We also read that "scenarios recognis(e) that possibilities are influenced by a wide range of people." Here we are clearly on the side of the indeterminacy of the future due to people's faculty to make (strategic) choices. On the other hand, we are also told that a crucial step in the making of scenarios is "to identify and analyse driving forces that will shape the environment. What will persist and can be forecast (for example, demography in many exercises), and what may change and is unknown? Following the identification of the driving forces, we can now contemplate a set of plausible storylines." The uncertainty is here clearly epistemic.

its own past. French biographer André Maurois once went so far as to write:

> There is no privileged past (...) There is an infinitude of Pasts, all equally valid (...) At each and every instant of Time, however brief you suppose it, the line of events forks like the stem of a tree putting forth twin branches[31].

Dutch historian Johan Huizinga had already paved the way by writing:

> The historian must (...) constantly put himself at a point in the past at which the known factors will seem to permit different outcomes. If he speaks of Salamis, then it must be as if the Persians might still win; if he speaks of the coup d'Etat of Brumaire, then it must remain to be seen if Bonaparte will be ignominiously repulsed[32].

The few historians who take this line of thought seriously are those who do not shy away from writing what goes today by the name of "Counterfactual History" or "Virtual History". Those "What if?" historians try and put forward more or less convincing answers to such questions as, What if there had been no French Revolution? What if Hitler had invaded Britain? What if the Soviets had won the Cold War? And, of course, the Pascalian one, What if Cleopatra's nose had been different?

Among professional historians, though, widespread is the opinion that this kind of exercise is a mere "parlour game" or a "red herring"[33]. From Marxists and other materialists this opinion doesn't come as a surprise but it is much more widely shared than that. It is worth quoting British idealist philosopher Michael Oakeshott on this:

> It is possible that had St Paul been captured and killed when his friends lowered him from the walls of Damascus, the Christian religion might never have become the centre of our civilisation. And on that account, the spread of Christianity might be attributed to St Paul's escape ... But when events are treated in this manner, they cease at once to be historical events. The result is not merely bad or doubtful

[31] Quoted by Niall Ferguson in his *Virtual History*, Picador, London, 1997, p. 1.
[32] Ibid.
[33] These dismissive phrases are from E. H. Carr. Quoted by Niall Ferguson, p. 4.

> history, but the complete rejection of history (...) The distinction (...) between essential and incidental events does not belong to historical thought at all[34].

The opposition between historians who see only historical necessity and those who are sensitive to the metaphysical postulation that things might be different from what they turned out to be, can and must be transcended. The metaphysical tools exist that allow us to carry out this *Aufhebung*. We owe them to French philosopher Henri Bergson and his brilliant student Jean-Paul Sartre. The idea is that as long as human beings live, they are absolutely free, and their freedom resides entirely in their capacity to choose, that is, to invent their lives. Future-oriented counterfactual propositions such as, "If I were to do this, the consequences would or might be that, and I am entirely responsible for them, whatever they turn out to be", make full sense. However, as soon as "death has turned life into destiny", to quote another famous existentialist, backward-looking counterfactual propositions such as, "Had I had more time to devote to my work, I would have written the novel of the century", are completely devoid of meaning and serve as mere alibis or cheap excuses – the stuff "bad faith" is made of[35].

In that kind of metaphysics, counterfactual propositions are admissible only when they are future-oriented. When we look back at the past, we see only necessity. There is nothing else than that which has happened, no possibility that never came to actuality. When history unfolds, then, possibilities become actual, but something strange happens to the branches that were not selected. It is not that they have become impossible: it turns out that they were never possible! As history proceeds in its course, it interjects necessity back into the past. Necessity is only retrospective.

In the framework of this metaphysics the parties to the debate about the meaning of virtual history appear to suffer from symmetrical blind spots. The "What if?" historians argue as if the possibilities that did not become actual kept existing forever, in a kind of eternal limbo. The mainstream historians who refuse to ascribe any meaning to counterfactuals reason as

---

[34] Michael Oakeshott, *Experience and its Modes*, Cambridge, 1933; quoted by Niall Ferguson, p. 6-7.

[35] In Sartre's plays, the dead keep talking to each other and even make definitive philosophical claims such as, "Hell is other people"! The only thing they wish to do, but can *no longer* do, is "choose their past". The latter has become inert, sentenced to be forever part of the "In itself".

if agents endowed with free will didn't make any difference in the way events occur.

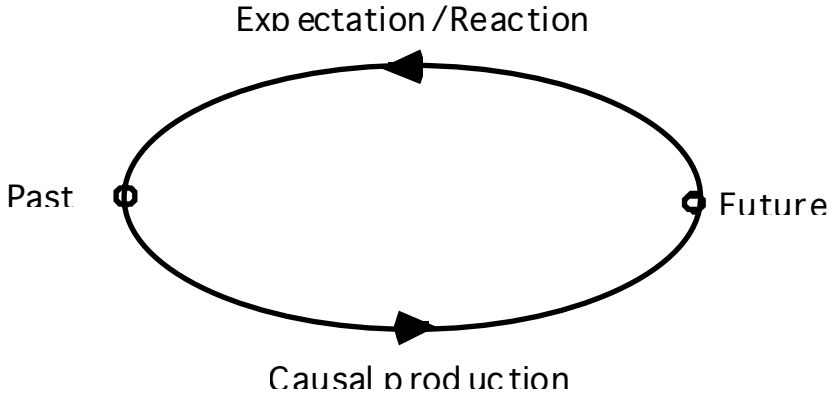
Back to the future. Following Hans Jonas, as explained before, my task has been to reestablish the future in its ontological status of a *real* entity. Bergsonian – Sartrean metaphysics permits exactly that: project yourself into the future and look back from there at the present. Seen from the present the future was open, but seen from the vantage point of the future, the path that led to it appears to have been necessary. We were free to choose, to be sure, but what we chose appears to have been our destiny[36].

At this stage non-philosophers are probably thinking that this is all speculative bla-bla-bla that has no bearing whatsoever on the real world. One couldn't be more plainly wrong.

The temporal experience I am trying to describe – and which, again, I call "projected time" -, is ours on a daily basis. It is facilitated, encouraged, organized, not to say imposed by numerous features of our social institutions. All around us, more or less authoritative voices are heard that proclaim what the more or less near future will be: the next day's traffic on the freeway, the result of the upcoming elections, the rates of inflation and growth for the coming year, the changing levels of greenhouse gases, etc. The *futurists* and sundry other prognosticators, whose appellation lacks the grandeur of the prophet's, know full well, as do we, that this future they announce to us as if it were written in the stars is a future of our own making. We do not rebel against what could pass for a metaphysical scandal (except, on occasion, in the voting booth). It is the coherence of this mode of coordination with regard to the future that I have endeavored to bring out.

A sine qua non must be respected for that coherence to be the case: a *closure condition*, as shown in the following graph. Projected time takes the form of a loop, in which past and future reciprocally determine each other.

[36] This is a famous Heideggerian *philosopheme*.



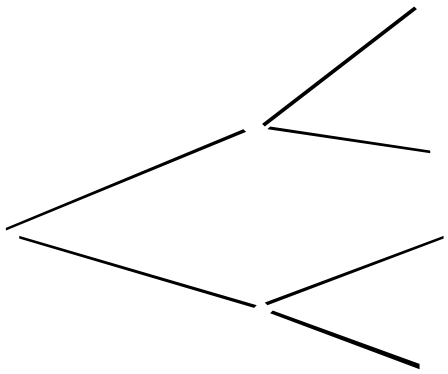
Projected time

To foretell the future in projected time, it is necessary to seek the loop's *fixed point*, where an expectation (on the part of the past with regard to the future) and a causal production (of the future by the past) coincide. The predictor, *knowing that his prediction is going to produce causal effects in the world*, must take account of this fact if he wants the future to confirm what he foretold. Traditionally, which is to say in a world dominated by religion, this is the role of the prophet, and especially that of the biblical prophet.[37] He is an extraordinary individual, often excentric, who does not go unnoticed. His prophecies have an effect on the world and the course of events for these purely human and social reasons, but also because those who listen to them believe that the word of the prophet is the word of Yahveh and that this word, which cannot be heard directly, has the power of making the very thing it announces come to pass. We would say today that the prophet's word has a *performative* power: by saying things, it brings them into existence. Now, the prophet knows that. One might be tempted to conclude that the prophet has the power of a revolutionary: he speaks so that things will change in the direction he intends to give them. This would be to forget the fatalist aspect of prophecy: it describes the events to come as they are written on the great scroll of history, immutable and ineluctable. Revolutionary prophecy has preserved this highly paradoxical mix of fatalism and voluntarism that characterizes biblical prophecy. Marxism is the most striking illustration of this.

[37]To his misfortune and above all that of his compatriots, the ancient prophet (such as the Trojans Laocoon and Cassandra) was not heeded; his words were scattered by the wind.

However, I am speaking of prophecy, here, in a purely secular and technical sense. The prophet is the one who, more prosaically, seeks out the *fixed point* of the problem, the point where voluntarism achieves the very thing that fatality dictates. The prophecy includes itself in its own discourse; it sees itself realizing what it announces as destiny. In this sense, as I said before, prophets are legion in our modern democratic societies, founded on science and technology. What is missing is the realization that this way of relating to the future, which is neither building, inventing or creating it, nor abiding by its necessity, requires a special metaphysics.

Perhaps the best way to bring out the specificity of the metaphysics of projected time is to ponder the fact that there is no such closure or looping condition as regards our "ordinary" metaphysics, in which time bifurcates into a series of successive branches, the actual world constituting one path among these. I have dubbed this metaphysics of temporality "occurring time"; it is structured like a decision tree:



Occurring time

Obviously the scenario approach presupposes the metaphysics of occurring time. But that is also the case of the metaphysical structure of prevention. Prevention consists in taking action to insure that an unwanted possibility is relegated to the ontological realm of non-actualized possibilities. The catastrophe, even though it does not take place, retains the status of a possibility, not in the sense that it would still be possible for it to take place, but in the sense that it will forever remain true that it could have taken place. When one announces, *in order to avert it,* that a catastrophe is coming, this announcement does not possess the status of a *prediction,* in the strict sense of the term: it does not claim to say what the future will be, but only what it would have been had one failed to take

preventive measures. *There is no need for any loop to close here*: the announced future does not have to coincide with the actual future, the forecast does not have to come true, for the announced or forecast "future" is not in fact the future at all, but a possible world that is and will remain not actual.[38]

By contrast, in projected time, the future is held to be fixed, which means that any event that is not part of the present or the future is an impossible event. It immediately follows that in projected time, prudence can never take the form of prevention. Once again, prevention assumes that the undesirable event that one prevents is an unrealized possibility. The event must be possible for us to have a reason to act; but if our action is effective, it will not take place. This is unthinkable within the framework of projected time.

Such notions as "anticipatory self-defense", "preemptive attack", or "preventive war" do not make any sense in projected time. They correspond to a paradox exemplified by a classic figure from literature and philosophy, the killer judge. The killer judge "neutralizes" (murders) the criminals of whom it is "written" that they will commit a crime, but the consequence of the neutralization in question is precisely that the crime will not be committed![39] The paradox derives from the failure of the past prediction and the future event to come together in a closed loop. But, I repeat, the very idea of such a loop makes no sense in our ordinary metaphysics.

5.4. *Conclusion. Exploring the set of projected equilibria as a substitute for the scenario approach*

We should take very seriously the idea that there is a "co-evolution of technology and society" (Arie Rip). The dynamics of technological development is embedded in society. The consequences of the development of nanotechnology will concern society as well as technology itself. Technology and society shape one another.

The future of nanotechnology, therefore, depends on the way society is going to react to the anticipations that are being made of this future. If

[38]For an illustration, one may think of those traffic warnings whose purpose is precisely to steer motorists away from routes that are otherwise expected to be clogged with too many motorists.

[39]Here I am thinking of Voltaire's *Zadig*. The American science fiction writer Philip K. Dick produced a subtle variation on the theme in his story "Minority Report." Spielberg's movie is not up to the same standard, alas.

those anticipations are produced through the scenario method, they will be of no help in the resolution of the ethical problem. They won't restore the future in its status of a real entity of which our knowledge must be as precise as possible. I have argued that the most effective way to ascribe reality to the future is to reason in the framework of projected time. But, then, we are confronted with a problem of reflexivity. This "we" refers to all groups, lobbies, expert groups, administrations, institutions that purport to shape the future through its anticipation, anticipation made public. We are the "prophets" of today, in the technical sense explained above. We have to explore the fixed points of the temporal loop that links the future to the past and then to the future again. Those fixed points I have called "projected equilibria".

Alexei Grinbaum and I have called "*ongoing normative assessment*" the methodology that corresponds to the determination of these projected equilibria. One can succinctly capture the spirit of this approach with the following words: it is a matter of obtaining through research, public deliberation, and all other means, an image of the future sufficiently optimistic to be desirable and sufficiently credible to trigger the actions that will bring about its own realization. It is easy to see that this definition can make sense only within the metaphysics of projected time, whose characteristic loop between past and future it describes precisely. Here coordination is achieved on the basis of an *image* of the future capable of insuring a closed loop between the causal production of the future and the self-fulfilling expectation of it.

I have said before that prevention made no sense in projected time. What can take its place then? Are there projected equilibria that may protect us against a major disaster, if such a denouement is in the offing? The search for an answer to that question I have called "enlightened doomsaying".

From the outset it appears that this search is bound to run into an irremediable paradox. It is a matter of achieving coordination on the basis of a negative project taking the form of a fixed future that *one does not want*. One might try to transpose the above characterization of the methodology of ongoing normative assessment into the following terms: "to obtain through scientific futurology and a meditation on human goals an image of the future sufficiently catastrophic to be repulsive and sufficiently credible to trigger the actions that will block its realization" — but this formulation would fail to take account of an essential element. Such an enterprise would seem to be hobbled from the outset by a prohibitive defect: self-contradiction. *If one succeeds in avoiding the*

undesirable future, how can one say that coordination was achieved by fixing one's sights on that same future? The paradox is unresolved.

In order to spell out what my solution to this paradox is, it would be necessary to enter into the technical details of a metaphysical development, and this is not the place to do so.[40] I will content myself with conveying a fleeting idea of the schema on which my solution is based. Everything turns on a random occurrence — but one whose nature and structure defy the traditional categories that I discussed in the first sections of this work.

The problem is to see what type of fixed point is capable of insuring the closure of the loop that links the future to the past in projected time. We know that the catastrophe cannot be this fixed point: the signals it would send back toward the past would trigger actions that would keep the catastrophic future from being realized. If the deterrent effect of the catastrophe worked perfectly, it would be self-obliterating. For the signals from the future to reach the past without triggering the very thing that would obliterate their source, there must subsist, inscribed in the future, an *imperfection in the closure of the loop*. I proposed above a transposition of our definition of ongoing normative assessment, in order to suggest what could serve as a maxim for a rational form of doomsaying. I added that as soon as it was enunciated, this maxim collapsed into self-refutation. Now we can see how it could be amended so as to save it from this undesirable fate. The new formulation would be: "to obtain… an image of the future sufficiently catastrophic to be repulsive and sufficiently credible to trigger the actions that would block its realization, *barring an accident*."

One may want to quantify the probability of this accident. Let us say that it is an epsilon, **e**, by definition weak or very weak. The foregoing explanation can then be summed up very concisely: it is because there is a probability **e** that the deterrence will not work that it works with a probability 1-**e**. What might look like a tautology (it would obviously be one in the metaphysics of occurring time) is absolutely not one here, since the preceding proposition is not true for $\mathbf{e} = 0$.[41] The fact that the

[40]I will take the liberty of referring the interested reader to my *Pour un catastrophisme éclairé*.

[41]The discontinuity at $\mathbf{e} = 0$ suggests that something like an uncertainty principle is at work here, or rather an indeterminacy [*Unbestimmtheit*] principle. The probabilities **e** and 1-**e** behave like probabilities in quantum mechanics. The fixed point must be conceived here as the *superposition* of two states, one being the accidental *and*

deterrence will not work with a strictly positive probability **e** is what allows for the inscription of the catastrophe in the future, and it is this inscription that makes the deterrence effective, *with a margin of error **e***. Note that it would be quite incorrect to say that it is the *possibility* of the error, with the probability **e**, that saves the effectiveness of the deterrence — as if the error and the absence of error constituted two paths branching out from a fork in the road. There are no branching paths in projected time. The error is not merely possible, it is actual: it is inscribed in time, rather like a slip of the pen. In other words, the very thing that threatens us may be our salvation.

---

preordained occurrence of the catastrophe, the other its non-occurrence. I cannot pursue this line of reasoning any further here.