Luciano Floridi

There were many reasons why I was initially drawn to informational issues. Let me try to summarise them under four headings.

The first reason was anti-metaphysical. I was drawn to what I later defined as the philosophy of information because, in the late eighties, I was looking for a conceptual framework in which psychologism, introspection, armchair speculations and all those linguistic (or perhaps one should say, Anglo-Saxon, or Indo-European) intuitions could be monitored, tamed and kept under tight control. I shared with Popper a desire for an "epistemology without the knowing subject", as the title of one of his papers declared. The sort of philosophy popular at the time smacked too much of bad metaphysics, a sort of betrayal of the purer and cleaner approach defended by Analytic as well as Neopositivist philosophy, which I admired so much (since then, I have somewhat repented and now I consider myself an ex-analytic philosopher). Since I was interested in epistemology and logic, the move from knowledge to information and from inferential to computational processes was almost natural.

The second reason was related to what I like to describe as *methodological minimalism*. I was looking for a more "impoverished" approach, which could allow me to work on more elementary concepts, less worn down by centuries of speculation, and more easily manageable. It seemed that, if one could have any hope of answering difficult questions about complicated issues concerning knowledge, meaning, the mind, the nature of reality or morality, it made sense to try to tackle them at the lowest and less committed level at which one could possibly work. Informational and computational ideas provided such a minimalist approach. To give a concrete example, my interest in artificial agents was motivated by the classic idea that less is more. This is still not very popular among philosophers, who seem to be too much in love with the human subject, his psychology and idiosyncrasies.

The third reason was a lively interest in timely issues. I could not see myself working on one of the great dead philosophers, spending my life trying to rethink someone else's thoughts. It reminded me of something said by Wittgenstein when complaining about his experience as a tutor in Cambridge. He thought that it was like having his brain sucked. Well, I did not wish to be anyone's intellectual parasite; I had enough problems making sense of what I was trying to think. At the same time, I did not wish to commit myself to some theoretical investigation that would have no bearing or connection with our contemporary world. I was happy to leave the safe and comfortable garden of specialised and irrelevant scholastic philosophy in order to engage with more worldly issues. Philosophy should not talk to itself about itself, and when it does it is utterly sterile. I reasoned that, if the scientific revolution could attract a Descartes, we could definitely engage with the new information revolution, the most radical change that humanity has experienced in history for a long while. I thought this was more than sufficient to justify a robust, philosophical interest in computational and informational issues. It was a risky bet that seems to have paid back handsomely.

The last reason is the simplest: I was born to be a nerd, and what is a nerd without his computer?

It is clearly difficult and somewhat embarrassing to indicate examples from my work, and the work of others, which best illustrates the fruitful use of a computational and/or informational approach for foundational researches and/or applications.

Regarding the work of others, this volume contains contributions from some of the most influential thinkers in the field, more senior colleagues from whom I have learnt much. It would be enough to check their bibliographies to appreciate how deeply and widely the new computational and informational paradigm has influenced our ways of thinking and doing philosophy. As for my own work, this may require a longer answer. If I were to pick up one example, perhaps I would mention the use of Level of Abstractions (LoAs) in conceptual analysis. The idea is simple, powerful, and very common in computer science but its importance and fruitfulness are still not fully appreciated in philosophical circles. Let me summarise it.

The method of LoAs comes from modelling techniques developed in an area of Computer Science, known as *Formal Methods*, in which discrete mathematics is used to specify and analyse the behaviour of information systems. Before I introduce a quick summary of the method of Levels of Abstraction, an everyday example may be useful.

Suppose we wish to describe the state of a traffic light in Rome. We might decide to consider an *observable*, named *colour*, of *type* {*red*, *amber*, *green*} that corresponds to the colour indicated by the light. This option abstracts the length of time for which the particular colour has been displayed, the brightness of the light, the height of the traffic light, and so on. So the choice of type corresponds to a decision about how the phenomenon is to be regarded. To specify such a traffic light for the purpose of construction, a more appropriate type might comprise a numerical measure of wavelength. Furthermore, if we are in Oxford, the type of colour would be a little more complex, since – in addition to red, amber and green – red and amber are displayed simultaneously for part of the cycle. So, an appropriate type would be {*red*, *amber*, *green*, *red-amber*}. What we have just seen is a basic concept of *Level of Abstraction*, understood as a finite but non-empty set of observables, where an *observable* is just an *interpreted typed variable*, that is, a typed variable together with a statement of what feature of the system under consideration it stands for.

The definition of observables is only the first step in studying a system at a given LoA. The second step consists in deciding what relationships hold between the observables. This, in turn, requires the introduction of the concept of system "behaviour".

Not all values exhibited by combinations of observables in a LoA may be realised by the system being modelled. For example, if the four traffic lights in Oxford are modelled by four observables, each representing the colour of a light, the lights should not in fact all be green together (assuming they work properly). In other words, the combination in which each observable is green should not be realised in the system being modelled, although the types chosen allow it. Some technique is therefore required to describe those combinations of observable values that are actually acceptable. The most general method is simply to describe all the allowed combinations of values. Such a description is determined by a predicate, whose allowed combinations of values we call the "system behaviours". A *behaviour* of a system, at a given LoA, is defined to consist of a predicate whose free variables are observables at that LoA. The substitutions of values for observables that make the predicate true are called the *system behaviours*.

A *moderated LoA* is defined to consist of a LoA together with a behaviour at that LoA. For example, human height does not take arbitrary rational values, for it is always positive and has an upper limit of (say) nine feet. The variable h, representing height, is therefore constrained to reflect reality by defining its behaviour to consist of the predicate $0 < h < 9$, in which case any value of h in that interval is a "system" behaviour.

Since Newton and Leibniz, the behaviours of the analogue observables have typically been described by differential equations. A small change in one observable results in a small, quantified change in the overall system behaviour. Accordingly, it is the rates at which those continuous observables vary which is most conveniently described. The desired behaviour of the system then consists of the solution of the differential equations. However, this is a special case of a predicate: the predicate holds at just those values satisfying the differential equation. If a complex system is approximated by simpler systems, then the differential calculus provides a method for quantifying the approximation. The use of predicates to demarcate system behaviour is essential in any (nontrivial) analysis of discrete systems because in the latter no such continuity holds: the change of an observable by a single value may result in a radical and arbitrary change in system behaviour. Yet, complexity demands some kind of comprehension of the system in terms of simple approximations. When this is possible, the approximating behaviours are described exactly, by a predicate, at a given LoA, and it is the LoAs that vary, becoming more comprehensive and embracing more detailed behaviours, until the final LoA accounts for the desired behaviours. Thus, the formalism provided by the method of abstraction can be seen as doing for discrete systems what differential calculus has traditionally done for analogue systems.

Specifying the LoA at which one is working means clarifying, from the outset, the range of questions that (a) can be meaningfully asked and (b) are answerable in principle. In standard terminology, the input of a LoA consists of the *system* under analysis, comprising a set of *data*; its output is a *model* of the system comprising *information*. The quantity of information in a model varies with the LoA: a lower LoA, of greater resolution or finer granularity, produces a model that contains more information than a model produced at a higher, or more abstract, LoA. Thus, a given LoA provides a quantified commitment to the kind and amount of information that can be "extracted" from a system. The choice of a LoA pre-determines the type and quantity of data that can be considered and hence the information that can be contained in the model. So, knowing at which LoA the system is being analysed is indispensable, for it means knowing the scope and limits of the model being developed.

Clearly, this is a very powerful and flexible method in any context that requires rigorous conceptual analyses. In philosophy, it should be taught to undergraduates in their first year. It seems the ABC of any decent way of doing philosophy. Unfortunately, too often philosophical difficulties are generated either by a lack of attention for the specific LoA at which the investigation is being developed – as if one could ask questions in a LoA-free environment – or by the wrong choice of the LoA altogether. To give two simple examples: when people ask what a mind is, they often make the first kind mistake; and when they ask whether artificial agents might be moral agents they often make the second kind of mistake. Both could be easily avoided by paying some attention to the method of abstraction.

The proper role of computer science and information science in relation to other disciplines is probably twofold.

On the one hand, computer and information sciences are the new epistemic enablers. They play, indeed they have directly inherited, the role enjoyed by mathematics in the scientific revolution between the seventeenth and the nineteenth century. We all know that there would be no science today without the information/computational turn that took place roughly fifty years ago. This key role played by computer/information sciences will have many fundamental implications. For reasons of space, let me just highlight two of them, one positive and the other negative. Computational-informational approaches and the increasing possibility of *in silico* experiments and simulations will promote cross-disciplinarity and the development of new roles, just think of the new developments in biocomputational research, in neuroscience or in nano-technological engineering. At the same time, we will run a serious risk of disregarding those lines of research that might not easily lend themselves to a computational/informational treatment. Some mis-directed attempts to overuse computers in the Humanities, for example, are a clear sign of this limit and bias. Humanities computing and e-learning projects are well-known for attracting comparatively disproportionate funding compared to more old-fashioned projects of research. Today it is easier to obtain a grant for a database of dubious value than to hire a postdoc for philosophical research. This is where more awareness of the scope and hence of the limits of the informational revolution, in terms of applied research in the humanities, may help us to keep a healthy balance between what is doable and what needs to be done.

On the other hand, as far as philosophy is concerned, it seems that computer and information sciences can, and indeed do, play today the sort of pivotal role that physics and mathematics enjoyed in modern philosophy. If the philosophy of information is becoming our *philosophia prima*, this is also because computational and informational ideas and artefacts are today so essential for our scientific development. So, there can be a twofold relation here. We can take from computer and information sciences many interesting and valuable tools with which to identify, analyse and then seek to solve philosophical problems. And we can use philosophy to explain and make sense of issues arising in computer and information sciences and their applications. It is a symbiotic relation that is already showing excellent results.

One of the most neglected topics in late twentieth century studies of computation and information is a *philosophy of nature* in the widest sense of the word (that is, in the German sense of *Naturphilosophie* as this was used by Shelling and Hegel).[1] Partly because of the predominant debate on AI; partly because, during the twentieth century, the sciences of the

[1] The relevant entry in Wikipedia is reliable: http://en.wikipedia.org/wiki/Naturphilosophie.

physical world became increasingly mathematical and detached from the natural environment, when compared to the situation in the nineteenth century; partly for other reasons too long to explain here (the two cultures, extreme specialization, scholasticisation of the philosophical discourse, etc.), the fact is that, during the last few decades, we have witnessed a growing gap between philosophers of the artificial and philosophers of the natural. The former look exclusively (no matter whether more favourably or critically) on the development of engineered or artificial artefacts, all the way down to the emergence of whole environments. Herbert Simon is a good example. The latter are more likely to be philosophers nostalgic for some utopian state of nature, or a mythical pristine condition of Being. Heidegger is notoriously a case in question. The result has been a revival of the old tension between *techne* and *sophia*. This is unfortunate because utterly fruitless. So it will be important to overcome this divide and seek to develop a philosophy of nature that could reconcile the synthetic with the authentic. Bacon and Wiener are two important thinkers in this tradition, whose lessons should be revaluated.

In a recent series of talks and articles, I have argued that the information revolution may be described as the fourth step in the process of dislocation and reassessment of humanity's fundamental nature and role in the universe. We are not immobile, at the centre of the universe (Copernican revolution), we are not unnaturally separate and diverse from the rest of the animal kingdom (Darwinian revolution), and we are very far from being Cartesian minds entirely transparent to ourselves (Freudian revolution). We are now slowly accepting the idea that we might not be dramatically different from other informational entities and agents and smart, engineered artefacts (Turing revolution). In view of this important change in our self-understanding and of the sort of IT-mediated interactions that we will increasingly enjoy with other agents, whether biological or artificial, I suspect that the future challenge will be to develop a philosophy of the Infosphere, where this is understood as synonymous with Being/Nature, which does not privilege the natural or untouched, but treats as authentic and genuine all forms of existence and behaviour, even those based on artificial, synthetic or engineered artefacts. We need to develop a synthetic environmentalism for all entitites.