Degenerate Epistemology

Luciano Floridi

Science, especially physics, has taught us to be very cautious about our naïve certainties ("that's the way it is!"), everyday intuitions ("it must be that way!"), and common-sensical rejections ("that's impossible!"). While reading this issue of *Philosophy & Technology*, just recall that we are all travelling at about 100,000 km/h around the sun. Indeed, we are getting so used to contemporary science supporting extraordinary claims that abrasively clash with what we would consider plausible, that we might overreact, and be inclined to believe almost anything. If tomorrow some credible source tells us that unicorns have been biologically engineered in some lab, how many of us would be utterly incredulous? So when scientists come up with some incredible results, what should we believe? The problem is exacerbated by the fact that these days, experiments churn gazillions of data. The Large Hadron Collider, currently the largest highest-energy particle accelerator, pumps out approximately 15 petabytes of data per year, which require a dedicated computational grid to be refined, analysed, and put to proper use. The more data and analysis we need, the more likely it is that something might go wrong in the process. Quality standards and safety measures are serious issues in the knowledge industry too.

The problem is widely felt, not just by insecure philosophers in search of reassurance. This is why scientists require very high levels of probability when it comes to deciding whether we are witnessing an amazing discovery or just some weird glitch in our systems. Have some neutrinos really travelled faster than light (http://www.bbc.co.uk/news/science-environment-15017484)? Enter statistics.

We know that, in statistics, a result is increasingly significant the more unlikely it is that it might have occurred just by chance. To pass the test of a genuine discovery, the threshold is very high: a result must be five-sigma unlikely. As the expression indicates, the sigma notation conveys the statistical significance of an event (standard deviation) in units of "σ". If a result reaches a five-sigma level of reliability this means that obtaining it by chance would be like getting twenty times heads when

L. Floridi (✉)
Department of Philosophy, University of Hertfordshire, de Havilland Campus, Hatfield, Hertfordshire AL10 9AB, UK
e-mail: l.floridi@herts.ac.uk

tossing an allegedly fair coin. Not impossible, but so unlikely that other explanations (including a biased coin) become much more plausible. Clearly, statistical analyses and the computers that run them do not provide certainties in the same way as logic and mathematics do. They cannot. For we could ask for higher sigma values and still be dealing with likelihoods. At the roots of our philosophy of information, there is a kind of uncertainty principle: if a result R is mathematically certain then R is not empirical but rather a matter of logic (the negation of R is a contradiction), and if R is empirical then it remains in principle uncertain. Luckily, there is plenty of room between the two alternatives, and a combination of data, software, machines, and intelligence helps us to identify the negligible and opt for the best compromise. Knowledge is a balance between improbabilities. This is both reassuring and humbling. Classic epistemology loves to concentrate on what are describable, to borrow a technical term from mathematics, "degenerate cases": the single, disembodied, epistemic agent, isolated, not-embedded in any environment, rational, open-minded, well-informed, capable of having access to his doxastic content reduced to a set of beliefs, in search for certainty… This is like reducing the study of ellipses to the study of circles (degenerate ellipses with eccentricity 0), and then the study of circles to the study of points (degenerate circles with radius 0). Doable, but we are left with very little that resembles epistemic practices in the real world, where chances that we might be brains in a vat, or avatars in some futuristic software experiments are ludicrously negligible and should keep nobody awake. There is of course a price to pay for our most likely but still uncertain empirical information. This is that we cannot hold on to naïve views about some epistemologically immediate contact with things in themselves. Some forms of realism also become negligible as "degenerate". If knowledge must be Cartesianly certain, then we really cannot know the world in itself. We can only be more or less confident on the statistical significance of our data analyses about the world, whether naturally implemented in our sensory and nervous apparatus or in our computers. The aforementioned uncertainty principle holds true not just at the CERN, but at the supermarket, not only for science, but also for the ordinary knowledge we enjoy when crossing the road. It was also the view of the founding father of statistics, Karl Pearson. In his classic *The Grammar of Science* (third edition 1900, still available on Amazon) he argued for a Kantian epistemology. It makes sense. Reality provides the data that we manipulate and interpret statistically, as organisms, epistemic agents, and scientists, in order to construct the phenomenal world in which we live. Reality in itself is the source of our knowledge, yet knowledge is of the message, not of the sender. Reality as we experience it is a semantic artefact. Any talk of things in themselves is just metaphysics, of which there is a rampant resurgence these days (just google "causal powers" or the explanations of why unicorns are metaphysically impossible, despite the lab mentioned at the beginning of this article). Interestingly, the book influenced the young Einstein so deeply that he recommended it as the very first reading for the Olympia Academy. The issuing history is complex, but the story is not: Kant's conditions of possibility of knowledge became statistical conditions of significance, which in turn went hand in hand with the crucial importance of frames of reference, and it all ended up with the development of Russell's theory of types and the method of levels of abstraction in computer science, where both (types and abstractions) are needed to build systems

that can deal successfully with the data we are accumulating in our scientific experiments. Computers later caused an information revolution, and here we are today talking about the application of levels of abstraction in our philosophical understanding of what science tells us about our knowledge of the world. It takes a while to understand what it means to know that we *probably* know, although not quite certainly, and not quite so "realistically".