Calling algorithms by their name (and engineers and philosophers' novel teamwork)

Serge Abiteboul and Gilles Dowek: The age of algorithms, translated by K-Rae Nelson. Cambridge: Cambridge University Press, 2020, 166 pp, £ 14.99 PB

Simona Chiodo[1]

Optimism seems to win out over pessimism—algorithms are "a tool that makes it possible to build a world that is better, freer, and fairer" (160)—even though the authors, who aim at getting "beyond this vision of good vs. evil" (ii) know that "the choice is ours" (160), which means that the future of the relationship between algorithms and human life is constitutively hard to predict. This is the overall trajectory of *The Age of Algorithms*, written by two computer scientists offering a most accessible view on both what algorithms are (the book starts with a clearest analogy between algorithms and recipes) and how algorithms are severely changing human life. The book has twenty short chapters covering the major questions on algorithms' power for reshaping human activities: from chapters on "the end of employment", "work" and "property", to chapters on governing, community, responsibility, privacy, fairness, ecology and education, as well as on most complex issues typically addressed by philosophy and sociology of technology, such as the augmented human, on the one hand, and algorithms' intelligence and feelings, on the other hand. Moreover, the book also offers both text boxes on keywords and a kind of interlude imagining a dialogue between a robot and a human being, which makes the book even more accessible, and even written to be read also by beginners. Yet, the addition of at least essential references, which are missing, would have increased its accessibility.

In what follows, I shall focus, first, on various strengths of the book and, second, on a weakness, the former winning out over the latter, anyway. A first strength is the effort to look at algorithms also from a historical perspective, which clarifies that an algorithm is a procedure "that allows us to solve a problem without having to invent a solution each time" (6). Algorithms are remarkably old

✉ Simona Chiodo
simona.chiodo@polimi.it

[1] Politecnico di Milano, Milano, Italy

in human history, as well as common in our everyday life, even though we may be unaware that we are actually using them. Rather, what is new is the synergy between algorithms and computers. Whereas, "for a long time, algorithms and machines belonged to different cultural spheres" (20), what is new is the synergy between algorithms and a special machine, i.e. the computer as "a universal machine that can run not only several, but all symbolic algorithms" (21). This is an important point also to try to actually find a balance between nonsensical extreme optimism and pessimism. As humans, we have always been using algorithms, which are one of the most powerful human tools to solve human problems. "Our fears are the consequences of this ignorance. We fear algorithms because we see them as mysterious beings, endowed with supernatural powers, even evil intentions" (2). And, yet, we should be aware that their power has been exceedingly increased by computers, which means that, if something goes wrong, they can also lead to "the worst outcome" (2). In any case, awareness can give us a great power as well: "we must never forget that they do not, in themselves, have intention. They have been designed by human beings" (2).

A second strength of the book is making it clear that algorithms cannot do a series of things, including crucial things, not because of something related to their current state of development, but because of something constitutive:

> The problems that we can solve with an algorithm are called *computable* or *decidable*. In contrast, problems for which this is not possible are *not computable*, or *undecidable*. The idea might be disturbing, but even if all of the programmers in the world were brought together to solve one of these problems, there would be no chance of success. (36)

This kind of potentially dangerous misunderstanding can be avoided by improving computer science education, as the authors stress, which is another meaningful strength. If it is true that the synergy between algorithms and computers is literally revolutionising human life, then it is also true that historical awareness, on the one hand, and technical awareness, on the other hand, are equally literally necessary, and the particular openness of this book can help educate.

Moreover, this kind of double awareness can also help us avoid two major dangers: social unpreparedness and ethical unpreparedness. As for the former, the authors intelligently specify that the problem is not "the end of employment", "work" and "property". Human history has already experienced several ends. Rather, the problem is social abuses committed, not because of algorithms, but because we frequently use algorithms as scapegoats. Thus, instead of "considering software as a member of the community" which would allow "us to see the community as a system where entities, human or not, interact, respecting certain rules" (90), we frequently use algorithms as "a pretext for eliminating social advantages" (59), but end up having a larger pie of which too "many receive a smaller slice", if anything at all (67).

As for the latter, the use of algorithms as scapegoats is even clearer, which is a crucial issue I myself am working on. The authors' view is that ethical delegation to algorithms has a precise limit, in that "algorithms are not moral actors" (96).

On the contrary, "the responsibility for an algorithm's actions belongs to those who design and use it" (96). One of the examples analysed is quite interesting. We already happen to find "algorithm-judges on the web for small claims" (92) of buyers and sellers on eBay, for instance. Can we imagine extending the use of "algorithm-judges" to other kinds of crimes? On the one hand, we would have the following advantage: judges perfectly "rational, impartial, and incorruptible" (92), whereas human judges are constitutively characterised, for instance, by "a metabolism and we know that sentencing judges are significantly less indulgent late in the morning than at the beginning of the afternoon" (92). On the other hand, we would have the following disadvantages. First, as the authors make explicit, we would take responsibility away from the human being, and, second, as the authors do not make explicit, what we would take away from the human being is precisely what matters most, i.e. their very ability to manage responsibility itself. That is the Shakespearean question, which means becoming unprepared to face both the ordinary and the extraordinary circumstances of everyday human life. The authors end up proposing "an intermediate solution", i.e. "hybrid decision-making", which means that "the human being is 'advised' by the algorithm" (93). I should add that, if it is true that this proposal may be promising, it is also true that, paradoxically enough, the human being should not also take away their ability to make conscious mistakes, which, as in the case of a white lie, may be the fairest thing to do in exceptional circumstances. And humans are exceedingly better than algorithms at managing exceptions.

The last strength I want to stress is the quite original focus on the potentially promising relationship between algorithms and ecology, in that, for instance, "computer science has become an indispensable tool for studying climate change" (126).

As for a possible weakness of the book, I think it is the following: it is becoming more and more perspicuous that algorithms' revolutionary power requires a novel teamwork between computer scientists, on the one hand, and philosophers and sociologists of technology, on the other hand. This is a general issue characterising, not only this book, of course, but, almost anything we, including myself as a philosopher, write on algorithms' revolution. The authors seem wisely aware of some major philosophical issues challenged by algorithms: from the destiny of the human identity—"once augmented, are we still human?" (145)—to humans' possibly irreducible ontological differences—"it does appear that there are two or three little things that we perhaps do not yet share with computers: creativity, emotions, and self-awareness" (156). The aim of the book, i.e. providing beginners with essential tools, makes the lack of philosophical expertise forgivable. Yet, the book, as with several other books, can show that a novel teamwork between engineers and philosophers is an urgent need for our challenging era. When, for instance, it comes to addressing the issue of online instruction, whereas the authors claim that "online instruction is more effective with more advanced university-level students who already know how to organize their learning" (63), we know better than ever, after the coronavirus emergency, that our "university-level students" actually, and desperately, need the kind of embodied example that a human being can uniquely offer, being analogous to whom they

are starting to plan to become as adults, even though "online instruction is a cost-saving measure" (64).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.