How much can I widen my decision boundaries?

A look at bounded rationality and transaction costs in the light of machine learning and big data analysis

Remo Pareschi

Department of Bioscience and Territory, University of Molise, Pesche (IS), Italy

{remo.pareschi@unimol.it, rpareschister@gmail.com}

There is a general feeling that the trend of Big Data is forcing decision makers to think in radically new terms, and of course this also affects the concept of bounded rationality, which aims to characterize the conditions under which decision-making happens within organizations. Thus, these new directions in the implementation of practical decision-making open relevant issues for research, including the following three:

1. Does the availability of large quantities of data, over which decisions can be pondered and made, provide also a way of hypothesizing a hierarchy of bounds over which decision making can scope? For instance, the well-known characterization of Big Data in terms of Volume, Variety and Velocity defines a set of parameters whose tuning and combination identify corresponding bounds — i.e., we can assume that, by increasing empirical evidence and available information, large data volumes received within certain limits of speed and variety widen the bounds of rational decision making. By contrast, excess in speed and in variety may narrow them because of intrinsic limitations (time, memory) in the capability to deal with such high speeds and ample varieties, thus making decision makers slip into the notorious pitfalls of information overflow. Nor this issue can be light-heartedly procrastinated, because of the various pressures (of competitive, ethical, institutional nature) to make pondered decisions by taking into full account this overwhelming avalanche of data. Hence, we need to cope with the apparent paradox that more available information does not necessarily imply wider bounds for rational decision making.
2. Conversely, principles of bounded rationality appear to find their way within the requirements for Big Data management. For instance, the CAP (Consistency; Availability; Partitioning) Theorem, well-known in theoretical computer science, states that full consistency of data and their availability are incompatible within the context of massive distributed processing; in this way, it provides one such principle for information systems, which were hitherto based on a criterion of strict consistency of data and, at the same time, of full access to information.
3. In the wake of 1 and 2 above we can therefore think that the best "bounded decision makers" in a context of Big Data are given by centaur systems, which combine in such a novel environment criteria and abilities of decision making derived both from humans and from computers. (The term "centaur system" originates from chess-playing, where the performance of this type of systems, namely, chess engines guided by human players, demonstrably exceeds that of stand-alone human players as well as of stand-alone artificial players.)

The paper will illustrate through a case study one such system, describing both the aspects and the challenges related to its design and implementation, as well as the implications for decision making deriving from its use. Machine learning coupled with assessments by human experts of the sector are the key aspects that make it "centaurian". The context is given by the analysis of data on the application of labor laws (which are part of the wider realm of civil law) and the methodology can be used both to monitor the application of specific laws as well as to define the general strategy that any of the stakeholders affected by their application (eg employers, trade unions) can take in the management of related risks. Risk management strategy determines operating criteria for taking single decisions. The method can be transferred to other contexts where risks, as well as, at the opposite, benefits, can be reconstructed through the monitoring and analysis of available streams of data, which in this case are given by court judgements. Given that the outcome of the application of these laws are, in most cases, transactions, it will also highlight a new views on the notion of transaction cost through the lens of Big Data analysis.