The irrational decison. How we gave computers the power to choose for us. *Benjamin Recht.* Princeton University Press, 2026 (368 p.) isbn: 9780199607907 (ebk)

Institutional leaders make decisions with far-reaching consequences, but every one of us also has to make decisions each day, supposedly on a rational basis. Here, "rational" means that we decide on a quasi-mathematical basis, i.e. in order to optimize some benefit. This led psychologists such as Steven Pinker to posit rationality as the primary driving force for explaining human behavior and to advocate its use as a paradigm to organize society. If this is true, then AI should not only be capable of making decisions for us, but do so in a more consistent manner.

Recht, a computer scientist, narrates the origins of this rationalism, the ways people actually make decisions, and whether this rationalism is the correct way to organize society.

Of course, Descartes considered the world as organized like clockwork where everything happens according to laws prescribed by nature. But Recht claims that our rationalist thinking was enforced when we started using (digital) computers. In particular, computers made this mechanized form of rationalism possible by (1) solving optimization problems, (2) solving problems in game theory, (3) allowing wide scale randomized experiments, and (4) making predictions on a statistical basis (using machine learning).

The book starts with a historical survey of the development of digital computers during WWII and the individuals who established the theory and practice for the previously mentioned topics. Vannevar Bush was the driving force in realizing the first digital computer; Claude Shannon helped to put the logic on chips; Norbert Wiener's and John von Neumann's work on stochastic processes and prediction theory quantified uncertainty and enabled the automatic control of these processes. The subsequent chapters elaborate on these topics.

Chapter 2 spotlights George Dantzig's approach to linear programming, i.e. optimizing a linear cost function subject to a set of linear constraints. These problems admit automated solution methods, so attention turned to their algorithmic design. This was followed by dynamic programming, and as computational power improved, more sophisticated optimization algorithms using gradients or even second order approximation were applied to complex control problems. Although the world is nonlinear, Rudolf Kalman showed

that local linear approximations can be used for smooth functions in iterative processes. Optimization could also be used to improve chip design (as in the case of the Intel 386) and hence produce more powerful computers.

The next chapter is a relatively extensive discussion of game theory including (optimal) strategies for games with 1 player (just optimize the "utility of your move") or 2 players (e.g. Rock-Scissors-Paper or other zero-sum games). However, for some games (e.g. chess) though the rules are fixed, the task of determining a sequence of next moves quickly becomes computationally prohibitive because the tree of possibilities has too many branches. By reducing the number of steps that one looks ahead, the computational burden may be eased to the extent that a sufficiently powerful computer may defeat a human. Similar problems hold for checkers or the game Go, but for those games, the number of possibilities grows too fast so that some heuristics and randomization must be utilized. Only then is there a chance that the computer can win on average. Other games are also discussed: non-zero-sum games (Nash equilibrium), backgammon, dueling games, etc. Games are relatively easy if all the information is available. That is not the case for poker where also some stochastics are needed. For some of these, the computer has defeated humans (on average). Those stories are extensively covered in this chapter. But there are also games (e.g. a game as simple as matching pennies) where no strategy is possible so that winning is unpredictable and computers cannot help.

Chapter 4 considers questions such as whether it is opportune to introduce a new drug, or to launch a vaccination campaign, or to organize broad scale preventive mammography or colonoscopy testing, or whether a doctor has to advise his patient to start chemotherapy or not. Similar questions arise when one has to decide whether to admit a student, or to hire a person, or to launch an app or an advertising campaign. Criteria for such decisions may be evaluated on the basis of extensive random control trials (RCT), but there are challenges to developing good RCTs: How rigorous should the random selection be? How many people should be in the test set and how many in the control set? When is a difference in results for these two sets relevant? Several examples are discussed.

In Chapter 5 we are introduced to the world of machine learning. The now familiar AI systems had their origin in Shannon's discovery that texts in English are about 85% redundant. Using stochastic optimization and simple linear prediction under iterative improvement (learning), it is possible to predict the next letter or the next word. This requires only a large training set

from which to learn and recognize the patterns. Most of these ideas were even developed before computers were able to implement them. It also took a long time to realize that the building blocks in these digital neural nets used for pattern recognition were basically a digital version of what Frank Rosenblatt called a perceptron which was his model of a biological neuron. Indeed, this is where it all started: in studies on how to automatically recognize handwritten characters and later on how to recognize different breeds of dogs from pictures, etc. It is a matter of pattern recognition, and if the system has learned to extract the key elements from texts or pictures, then the generative AI systems just invert the process and generate text or pictures from keywords.

There are tasks that humans can do much better than computers (e.g. recognize a face) and there are tasks that are easy for computers but difficult for humans (e.g. find the parity of a bit string of a million bits). If it is difficult or impossible to formalize how humans arrive at a particular decision, then we cannot write the relevant code to automate the process. If there are precise rules for the decision-making process, then computers will eventually outperform humans.

However, humans definitely do not always decide as machines do. Chapter 6 examines the views of psychologists such as Paul Meehl on the difference between the, sometimes instinctive, decisions made by humans and the rationality of mechanical decisions of computers and whether it is sensible to organize society based on the paradigm of rationality. The Heuristics and Bias (HB) approach represented by for example Daniel Kahnemann and the Naturalistic Decision Making (NDM) school promoted by Gary Klein are compared. The former promotes the use of checklists or safety prescriptions (i.e. the rules to be followed) since humans are 'predictably irrational' (Dan Ariely) and hence need to be protected from themselves.

The concluding chapter contains Recht's personal, more pragmatic opinion about whether rationality is the proper way to organize society. Sometimes we are not able to set a unique goal, or the goal cannot be quantified. In such cases, we cannot optimize the decision-making process. There are limits to the extent of automation.

For a book on mathematical rationality, the level of mathematics presented is limited. Mathematical explanations are given via simple examples rather than using formulas. It is however an interesting read for anyone who wants to know the rationale underlying automated decision-making. It surveys the history, persons, and ideas involved, from the start of the digital

revolution to current thinking on the psychology of decision-making. Finally, the book also provides a limited and somewhat indirect introduction to the ideas behind machine learning.

Adhemar Bultheel is emeritus professor at the Department of Computer Science at KU Leuven (Belgium). He has been teaching mainly undergraduate courses in analysis, algebra, and numerical mathematics. More information can be found at his homepage https://people.cs.kuleuven.be/adhemar.bultheel/