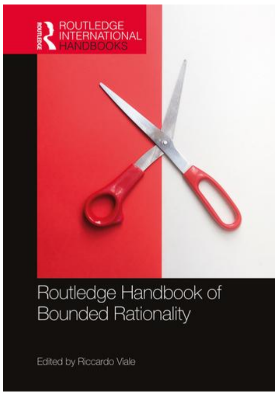

Routledge Handbook of Bounded Rationality

Riccardo Viale

Simon's legacies for mathematics educators

Publication details
https://test.routledgehandbooks.com/doi/10.4324/9781315658353-9
Laura Martignon, Kathryn Laskey, Keith Stenning
Published online on: 29 Oct 2020

8

SIMON'S LEGACIES FOR MATHEMATICS EDUCATORS

Laura Martignon, Kathryn Laskey, and Keith Stenning

Introduction

Herbert Simon is famous for having said: "The goal of science is to make the wonderful and the complex understandable and simple – but not less wonderful" (1996). The fundamental truth underlying Simon's words has an important parallel in mathematics education, where one might say that the goal is to devise ways to make wonderful, sometimes complex mathematical results understandable and simple—but not less wonderful. Explaining our ideas means decomposing them into understandable pieces and gluing them together with accessible arguments. In a proper "waterproof" explanation, these arguments must be backed by logical rigor. Logical rigor is sometimes an enemy of beauty and even, at times, of intuition. Mathematics, no less than the natural sciences, is in itself a collection of wonders. It has grown at the interface between the human mind/brain and nature, enabling the mind both to model natural phenomena and to tame uncertainty. In science in general and mathematics in particular, there is a certain risk in decomposing concepts into their components and explaining processes formally because fascination may fade through the exercise of formal decomposition and explanation. According to our version of Simon's statement, the good mathematics instructor should teach and explain mathematical facts while keeping both motivation and wonder alive in his/her students' minds. We cite a well-known mathematical example: It is a wonderful result of mathematics that there exist exactly five Platonic solids. Explaining this wonder requires a minimum of detailed formalism and going through this formalism demands special skills of a mathematics educator, if she wants to make the result clear and understandable without tarnishing its beauty.

This nugget of wisdom is just one of many that are found in Simon's writings. Simon's work on the sciences of the artificial yields innumerable bits of inspiration for the education of young students in mathematical thinking. For him, the artificial includes models and structures expressed in symbols developed by the mind! In other words, for Simon, mathematics and computer science are among the sciences of the artificial, as is the subject of how the human mind pursues these topics.

Many of the legacies left by Simon for the mathematics educator appear in his book, *The Sciences of the Artificial*, which first appeared in 1969 (Simon, 1996). Some of the chapters in that revolutionary book were specifically devoted to four topics of intense current interest in mathematics and computer science education: learning, representing information, bounded rationality, and problem solving. Our aim in this chapter is to analyze some of these legacies and discuss their relevance to education today.

Simon's predictions

Simon and his colleague Allen Newell were among the handful of founders of the field called Artificial Intelligence. In his Preface to the second edition of *The Sciences of the Artificial* (1996), Simon explained the difference between "natural" and "artificial" intelligence, contrasting the goal-oriented essence of the artificial with the evolutionary fit of natural systems to their environments. One naturally wonders how Herbert Simon, one of the fathers of artificial intelligence, would view the accomplishments of the field, were he alive today. In a 1992 interview for UBS Nobel Perspectives,[1] Simon mentioned his four predictions on the future of computers and computer programs:

1 By the mid-1960s, computers would play chess so well that they would beat the greatest chess masters.
2 Also by the mid-1960s, computers would compose aesthetically pleasing music.
3 Computers would prove at least one new theorem.
4 Most theories of the mind's cognitive processes would be expressed as computer programs.

Simon was prescient in that all his predictions have eventually come to pass, at least to some degree. In the case of chess, his timing was off by a few decades, but he was correct that computers would eventually play at the highest levels. Computers could play quite well by the 1980s, but it was not until 1997 that a computer beat World Chess Champion Garry Kasparov.[2] Simon was also early but correct regarding musical composition. In 2017, a program from Aiva Technologies became the first musical composition program to be given official status, when the French Société des Auteurs, Compositeurs et Éditeurs de musique (SACEM) conferred upon it the official status of Composer.[3] In 1976, the Four Color Theorem, a longstanding mathematical conjecture that had frustrated mathematicians since it had been proposed in 1852, became the first major theorem to be proven by computers and human minds working together. While controversial at first, because it was too complex to be checked by hand, the proof is widely accepted today. As for the fourth prediction, Herbert Simon himself saw to it that computer theories of the mind were developed, and he became a driving force behind their acceptance by the cognitive science community. It is difficult to estimate the ratio of full computational theories to less formally expressed theories of the mind, and thus to evaluate his assessment that most theories would be computational. Nevertheless, computer modeling of cognitive processes has become a vibrant, prolific, and highly successful sub-field of psychology. Simon's insight that many cognitive processes can be explained in terms of computer models was revolutionary. The question today is rather, which are those cognitive processes, which cannot be explained in terms of computer models?

One reason Simon, along with many early AI researchers, was far too optimistic about the time frame during which critical milestones would be achieved, was the contrast between how computers and humans process information. Humans are excellent at pattern recognition and analogy, still exceeding the capabilities of computers on many tasks requiring these skills. Computers are designed to perform the kind of exacting, step-by-step, logical reasoning tasks humans find especially difficult. Because computers were so good at tasks humans find so challenging, it was easy to overestimate the difficulty of making computers perform well on tasks humans find routine and even less intelligent animals can perform well.

Simon and the early AI researchers also underestimated the difficulty of developing AI systems to tackle ill-structured problems. While acknowledging that there can be no crisp, formal boundary between well- and ill-structured problems, Simon (1973) set out the characteristics of a problem that can be regarded as well-structured, i.e., one that is amenable to solution by a computer problem solver. An ill-structured problem is then a residual category, i.e., an ill-structured problem is simply one that is not well-structured. Simon argued that whether a problem is well- or ill-structured can be a matter of context, and that "any problem with a sufficiently large base of relevant knowledge can appear to be an ill-structured problem" (Simon, 1973, p. 197). He concluded: "There appears to be no reason to suppose that concepts as yet un-invented and unknown would be needed to automate the solution of ill-structured problems." Whether Simon was correct in this conclusion is still a matter of debate among AI researchers. It is unquestionable that it has taken longer than Simon anticipated for some of the more challenging problems to succumb to automation.

In 1956, von Neumann was invited to give a set of lectures comparing the ways in which computers and brains process information. He never gave the lectures, but worked on the manuscript until his death, and it was published posthumously (von Neumann, 1958). Acknowledging that he was a mathematician and not a neuroscientist, he identified a number of ways in which computers and brains operate differently:

1. Brains are analog and computers are digital.
2. Brains process in parallel and computers process serially.
3. Unlike brains, computers have the ability to use instructions stored in memory to modify control of processing.
4. Individual brain component processes are too slow to perform the serial processes at which computers excel.

Von Neumann concluded that the language and logic used by the brain must be radically different from that of computers.

Despite these fundamental differences, the project of computational modeling of cognitive phenomena has been extremely successful, and has led to many important insights. These insights have sparked advances in decision support, automation, and also education. We are here mainly concerned with Simon's seminal insights: that cognitive processes can be described by means of computer models and that the representation of information in a given problem is a key factor for finding its solution. These insights have had far-reaching consequences for mathematics and computer science education.

Ecologically rational representations

One of Simon's most famous quotes concerns the fundamental importance of representation for solving problem: "Solving a problem simply means representing it so as to make the solution transparent" (Simon, 1996, p. 132). A solution becomes transparent if it emerges from the structure in terms of which the problem has been modeled. The structure of the problem is the result of an attempt at adaptation between the problem itself and the conceptual constructions of the mind. An adaptation is successful if it is ecologically rational (cf., Gigerener et al., 1999). Ecological rationality is a fundamental characteristic of successful representations. It refers to behaviors and thought processes that are adaptive and goal-oriented in the context of the environment in which an organism is situated, thus giving survival advantage to the organism. Discovering a representation that makes a problem easily solvable gives evolutionary advantage in terms of time resources to adopters of the new representation, because they are able to solve the problem more quickly and easily than those holding on to the old representation.

Modern mathematics education treats representations of mathematical entities as a fundamental aspect of didactics in the classroom. Special attention has been devoted, for instance, to the issue of multiple representations of numerical entities and the advantages of switching between them as a means of achieving mathematical competencies for dealing with them (Dreher & Kuntze, 2014).

For many types of problems, such as fractions and probabilities, we argue that different *mathematically* equivalent representations are far from *cognitively* equivalent. That is, some representations are more adaptive and advantageous than others, because they seem to be better aligned to the cognitive systems of human problem solvers.

We begin by describing the relevance of information formats for inference, a major component of human reasoning. Both problem solving and decision-making require inference. From old knowledge, we create and acquire new knowledge by means of chains of inferences. From partial solutions, we advance to more complete solutions by means of inferences. The introduction of so-called inference machines, that is well-described systems of rules for inference, has had immense implications for science. The inferences of these inference machines are, more often than not, produced under uncertainty. Classical logic and probability provide the two most salient, standard mechanisms for inference machines of human history. We will discuss the ecological rationality of representations in the context of these two inference approaches, inverting the historical order of their inception. We begin by approaching tasks of probabilistic inference.

The ecological rationality of natural frequencies for probabilistic inference

One of the problems that makes mathematics so difficult is the already mentioned tension between intuition and rigor. In a variety of areas, mathematicians have developed elegant, abstract, logically coherent theories that encompass and generalize intuitively natural concepts. Examples include numbers, probabilities, and algebraic structures. These elegant abstract theories have a pristine beauty that attracts mathematicians, and have led to powerful innovations, but can be daunting to beginning students and to the lay public. Mathematics educators therefore face the challenge of helping students to build on their natural intuition to bridge the gap between human intuition and mathematical theory.

Natural frequency and probability

An important instance of the tension between intuition and rigor is probability theory. Our intuitive, innocent, and natural understanding of probabilities fits uneasily with the formal definition as countably additive measures on sigma-algebras of sets. Formalism and rigor, especially introduced too early, may kill intuition and with it motivation. We aim at scrutinizing the *natural Bayesian* who reasons, as in Gerd Gigerenzer's recommendation, by forming simple proportions of outcomes starting from well-defined populations of cases. The natural Bayesian goes through a sequence of stages: from multiple narratives stemming from observations that are ultimately represented as a frequency tree, to expected outcomes in future phases of updating.

The natural frequency representation is adapted to our cognitive processes in a way that makes the solution to evidential reasoning problems transparent. In other words, it is ecologically rational.

Consider a physician who must reason from evidence, such as symptoms and test results, to a hypothesis, such as whether or not a patient has a disease. To approach this problem formally, as presented in textbooks, we begin with two ingredients: (1) a *prior* probability that the disease is present; and (2) likelihoods, or probabilities that the evidence would be observed if the disease was present or absent. For our example, we imagine that the evidence is a test for the disease, and it has come out positive. We use the symbols D+ and D- to denote the presence and absence of the disease, and T+ to denote a positive test result. We use our prior probability and test likelihoods to calculate the *posterior* probability that the disease is present given the positive test result:

$$\left(D+ \mid T+\right) = \frac{P(T+ \mid D+)P\left(D+\right)}{P\left(T+ \mid D+\right)P\left(D+\right) + P\left(T+ \mid D-\right)P\left(D-\right)} \tag{7.1}$$

The formula (7.1) is called *Bayes rule* after its inventor, the eighteenth-century philosopher and minister Thomas Bayes.

Formula (7.1) gives correct answers to the evidential reasoning problem, but it requires calculations of which only a relatively few trained individuals are capable. There is an immense literature on people's difficulties with this kind of reasoning based on conditional probabilities, and some scientists have been convinced that these difficulties point to some form of human irrationality.

There is, on the other hand, a way of representing such evidential reasoning tasks that makes the solution transparent. To use this method, called *natural frequencies* (e.g., Gigerenzer & Hoffrage, 1995), the doctor imagines a population of fictitious people, say, 1,000 of them. She divides them into those who do and do not have the disease. For example, if the disease is present in only 1 percent of the population, she would partition her imaginary 1,000 patients into 10 who have the disease and 990 who do not. Of those who have the disease, let us imagine that 80 percent will test positive (). Therefore, the doctor imagines that 8 of the 10 ill patients will test positive and 2 will test negative. Now, suppose that 90 percent of those who do not have the disease will have a negative test result. In our doctor's imaginary population, this works out to 99 well patients who test positive and 891 who test negative.

This thought experiment uncovers a peculiar, and very important, characteristic of problems characterized by uncertainty: the fundamental role played by the base rate. Our doctor, like many untrained people, may be surprised to learn that, although the vast majority of ill patients have tested positive and the vast majority of well patients have tested negative, only 7 percent of the patients who test positive actually have the disease. This occurs because of the low base rate—that is, a very small proportion of our original population was ill.

Figure 8.1 compares the probability and the natural frequency representation of this evidential reasoning problem. As many an educator has discovered, students taught to apply formula (7.1) often find the result P(D | T+) = 7% surprising, sometimes even refusing to believe that the disease remains unlikely after a positive test result. By contrast, most students who work with the figure on the right immediately grasp why most patients who test positive are actually disease-free. It is readily apparent that although 107 positives is a small proportion of the 990 well patients, it is still nevertheless many more than the 8 true positives out of only 10 ill patients.

There is a wealth of literature demonstrating that people tend to neglect base rates when reasoning about this kind of problem. Even many people who have been taught Bayes rule often fall prey to this fallacy. However, performance on such tasks improves dramatically when

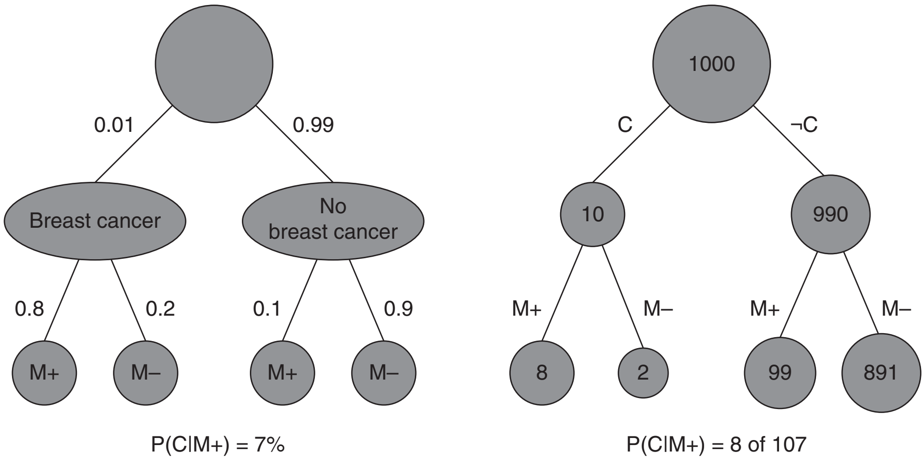

Figure 8.1 Probability and natural frequency representation of evidential reasoning

people are taught the natural frequency interpretation (Gigerenzer & Hoffrage, 1995). This improved performance has been observed in a variety of application domains (Hoffrage & Gigerenzer, 1998; Hoffrage, Lindsey, Hertwig, & Gigerenzer, 2000; Gigerenzer, 2002) and in more complex problems than reasoning with a single cue (Hoffrage, Krauss, Martignon, & Gigerenzer, 2015).

To summarize, an ecologically rational strategy is one our cognitive apparatus can perform naturally and easily, and that has adaptive value in a given environment. In other words, ecological rationality involves discovery of a representation that makes an optimal or nearly optimal solution transparent. For reasoning from evidence to hypothesis, because the natural frequency representation is ecologically rational, it allows humans to perform an otherwise difficult task quickly and easily. Because the natural frequency representation is ecologically rational, its use in educational settings facilitates understanding and improves performance.

Icon arrays and Venn diagrams

Back in the 1960s and 1970s, Venn diagrams were introduced as representations of sets in primary and secondary schools in most European countries and several other countries around the world. There were protests everywhere. In Germany, for instance, the protests both of schoolteachers and parents were so strong, that set theory was banned from primary school. This had as a consequence a reluctance to introduce probabilities earlier than in ninth grade. Also for the communication of probabilistic information in the medical and pharmaceutical domain, Venn diagrams did not enjoy acceptance.

Here again, an ecologically rational format has made its way into many sectors of communication. Otto Neurath introduced such a format during the first half of the twentieth century (he actually proposed what he called *isotypes* and arrays of such isotypes).

Icon arrays are displays of icons, which represent individuals or animals or items. The icons can exhibit special features, thus allowing for classifications. In the example in Figure 8.2a, the depicted icon array represents dogs and cats, with or without the feature "bell." One more structural way of representing the information about these house pets is

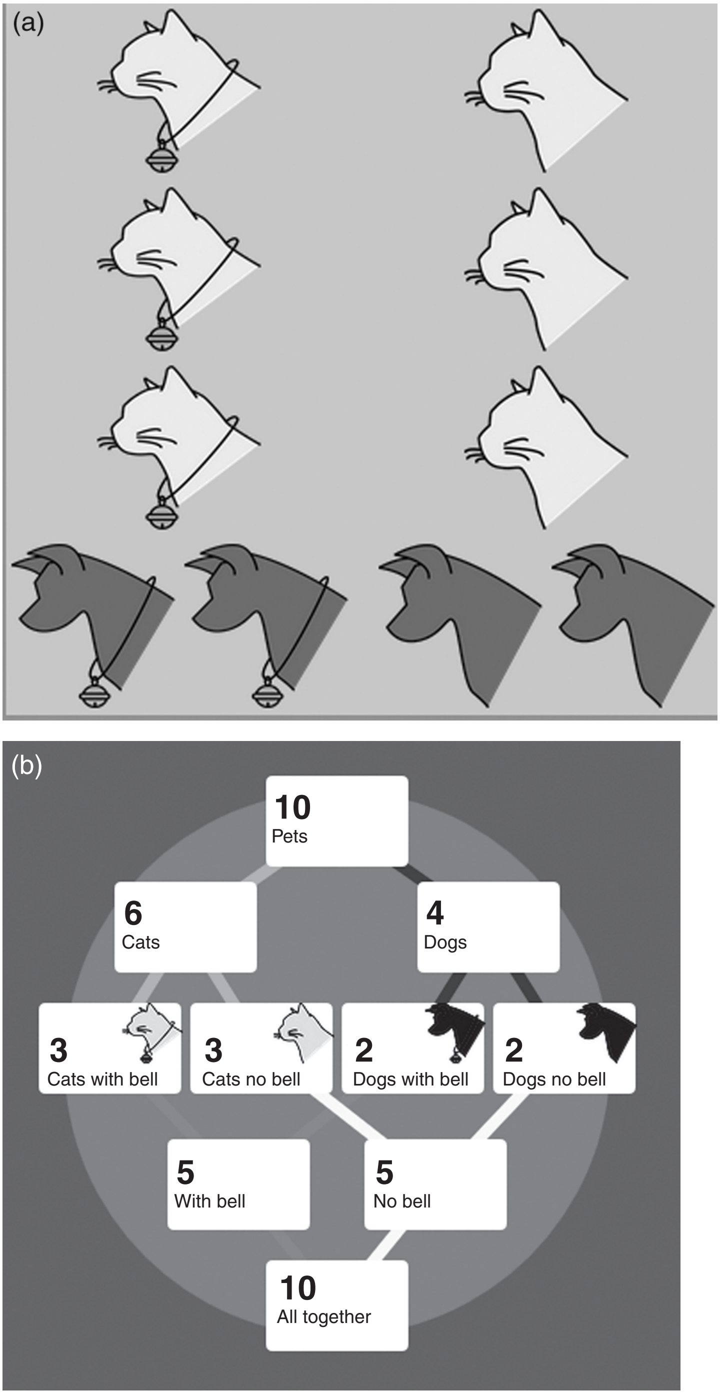

Figure 8.2 (a) An icon array representing dogs and cats, some with a bell and some without; (b) the corresponding double tree for transparent Bayesian reasoning

Source: (a) www.eeps.com/projects/wwg/ a free webpage for Risk literacy by Tim Erickson and Laura Martignon.

to organize them in a double tree: from top to bottom we first partition our house pets into cats and dogs and continue by looking in each subgroup whether the pet has a bell or not (Figure 8.2b). From bottom to top, we first look at bells, partitioning our pets into those with a bell and those without. School kids can use this representation to learn an elementary Bayesian reasoning. They learn to assess the probability—expressed in natural frequencies—that a pet with a bell is a cat, or that a dog has a bell. Thus, icon arrays combined with trees and double trees communicate information on base rates, sensitivities, specificities and predictive values in a straightforward, transparent way. They have been used in primary school with great success (Martignon & Hoffrage, 2019) coupled with corresponding natural frequencies doubletrees.

From inference based on one cue to inference based on many cues: the ecological rationality of lexicographic strategies

We commonly model situations by characterizing them in terms of a set of cues or features we can extract from them. We make inferences on these situations based on those cues. We have treated above the case of judgments or classifications based on just one cue—in that case a test—characterizing a disease. Actually, a physician usually needs more than just one test result to diagnose a disease. For instance, in the case of breast cancer, a doctor will systematically look at two cues, namely, mammography and an ultrasound test. We illustrate a possible Natural Frequency tree drawn in the *diagnostic direction*, i.e., with the end nodes corresponding to "having breast cancer" or "not having breast cancer" (Figure 8.3), in contrast with the tree of Figure 8.1, which follows *the causal direction*.

The facilitating effect of natural frequencies disappears when more cues are considered: As the number of cues grows, the size of the natural frequency tree explodes and ecological rationality is lost. The huge tree representing 6 cues, for instance, suffers from "brittleness." As it grows, the number of end nodes may become very large, while the number of cases per end node may become very small. These small numbers of cases per end node will not support reliable predictions.

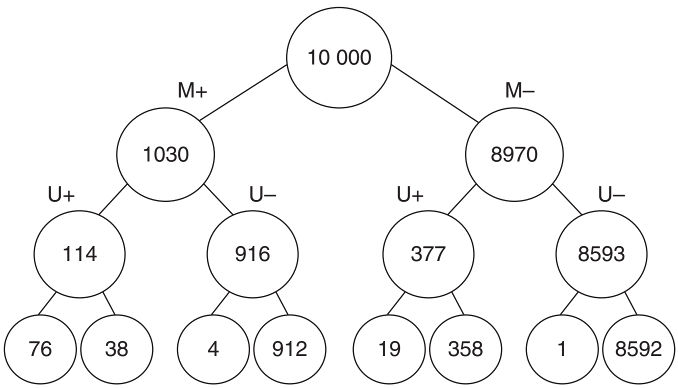

Figure 8.3 A natural frequency tree for classifying patients as having or not having breast cancer, by looking at the values of two cues, namely, a mammography and an ultrasound test

Is there an alternative ecologically rational strategy that maintains a tree structure and eliminates non-relevant information? There is, in fact, a heuristic strategy that fits the bill. Inspired by Simon, the heuristic consists of radically pruning the full tree into a minimal one based on the same set of cues. The resulting tree becomes fast-and-frugal, according to the definition of Martignon et al. (2003). The fast-and-frugal tree has a single exit at every level before the last one, where it has two. The cues are ordered by means of very simple ranking criteria. Fast-and-frugal trees are implemented step by step, reduce memory load, and can be set up and executed by the unaided mind, requiring, at most, paper and pencil.

Step-by-step procedures, that follow well-specified sequences, are a typical aspect of human behavior. We pronounce one word at a time, for instance. Time organizes our actions and thoughts in one-dimensional arrays. We are tuned for doing "first things first" and then proceeding to second and third things. This innate tendency for sequencing makes so-called lexicographic strategies for classification and decision ecologically rational. This section illustrates how sequential treatment of cues can extend and expand classifications based on one cue.

A fast-and-frugal tree does not classify optimally. Rather, to use a term coined by Simon, it *satisfices*, producing good enough solutions with reasonable cognitive effort. The predictive accuracy and the robustness of the fast-and-frugal tree have been amply demonstrated (e.g., Laskey & Martignon, 2014; Woike, Hoffrage, & Martignon, 2017; Luan, Gigerenzer, & Schooler, 2011).

Martignon et al. (2003) provided a characterization of these trees as lexicographic, supporting an ecologically rational step-by-step execution of the classification process, which can be stated as the following theorem:

> ***Theorem***: For binary cues with values 0 or 1, a fast-and-frugal tree is characterized by the existence of a unique cue profile of 0's and 1's that operates as a *splitting profile* of the tree; this means that any item with a profile lexicographically lower than the splitting profile will be classified in one of the two categories, while the rest of items will be classified in the other one.

This theorem is illustrated in Figure 8.4.

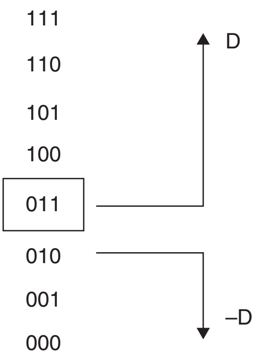

Figure 8.4 An example of a splitting profile characterizing a fast-and-frugal tree for classifying items into two complementary categories based on three cues

Source: Martignon et al. (2003).

From classification to comparison and back

The trees for classification described above belong to the class of simple heuristics inspired by Simon and analyzed for a variety of tasks by the ABC Research Group, led by Gerd Gigerenzer. Another of these heuristics tackles comparison: When two items have to be compared as to which has a higher value on a given criterion, the fast-and-frugal heuristic at hand, called Take-the-Best, proceeds as follows: It ranks the binary cues according to a simple criterion (e.g., validity) and looks at the cue profiles, comparing them lexicographically. Imagine that there were just three binary cues and the corresponding profiles were 110 and 101, as depicted in Figure 8.4. When they are compared lexicographically, 110 is larger than 101 because the first digit on the left-hand side coincides in both, but the second differs. Thus, the larger second digit corresponds to the larger number.

This simple principle is the essence of Take-the-Best. Assume, for instance, one wants to compare German cities, say, Duisburg and Ludwigsburg, as to their population. The comparison is to be based on binary cues, just as in the original example treated by Gigerenzer and Goldstein (1996). Here more than three binary cues may be necessary, if the pairs of cities to be compared can be any of all pairs in the reference of class of cities with more than 100,000 inhabitants. In fact, Gigerenzer and Goldstein considered nine binary cues, of which the first one, was simple recognition: "Do I recognize the name of the city?". If this is true for only one of the cities in the pair, then the recognized one is to be considered the larger one. If both are recognized, then one can compare the remaining "cue profiles" lexicographically. Other cues are "Is the city a state capital?," "Does the city have a soccer team playing in the National League?," or "Does it have a station where ICE trains stop?" In summary, Take-the-Best analyzes each cue, one after the other, according to the ranking by validity (i.e., its overall predictive value, or the probability of making a correct prediction if it discriminates between the two cities) and stopping the first time a cue discriminates between the items, concluding that the item with the larger value has also a larger value on the criterion.

Both the fast-and-frugal tree for classification and Take-the-Best are simple heuristics which belong in more than one of the great mathematical classes of models. On the one hand, they can be represented by trees, and, on the other hand, they are lexicographic strategies. An additional property that may be useful for computer implementation and use in applications, as well as for the comparisons in performance with models like regression (see Şimşek's Chapter 21 in this volume) is that these methods can be described in terms of linear models, which are those that "weigh and sum" cue values, in the most elementary way. In a more formal context, Martignon and Hoffrage (2002) showed that such lexicographic rules are always equivalent to weighted linear classification models with non-compensatory coefficients. A list of weights are non-compensatory if each of them is larger than the sum of all weights that follow in the list (see Figure 8.5 for an example).

On the left-hand side of Figure 8.5, we illustrate non-compensatory weights, which make a linear weighted model behave like a lexicographic strategy. On the right-hand side, we illustrate equal weights, which correspond to the tallying heuristic, which simply "counts" 1s in profiles of 1s and 0s.

In the next section, we illustrate the connection between the fast-and-frugal Take-the-Best with comparisons of natural numbers. The striking fact is that Hindu-Arabic numerals are a representation of numbers that allows the magnitude of numbers to be compared by following the Take-the-Best strategy.

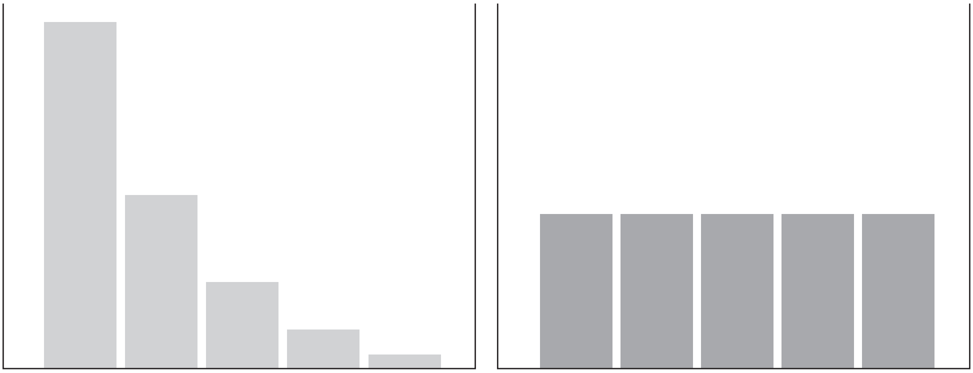

Figure 8.5 On the left side non-compensatory weights, like 2, 1, ½, ¼, 1/8. On the right side 5 weights all equal to 1.

Fast-and-frugal number comparison with the Hindu-Arabic system

Herbert Simon wrote, in Chapter 5 of *The Sciences of the Artificial*: "We all believe that arithmetic has become easier since Arabic numerals and place notation replaced Roman numerals, although I know of no theoretic treatment that explains why" (Simon, 1996).

Our aim is to use the tenets of ecological rationality to explain why arithmetic became so much easier with this new representation. We claim that "the theoretic treatment of why the decimal positional system made arithmetic calculations and comparisons so much easier than the Roman numerals" is in no small part due to the ecological rationality of the lexicographic heuristic for comparison it allows: Obviously, if a natural number has more digits than another natural number, then it is larger. What happens when they have the same number of digits? If one has to compare, say, 2056 and 2049, then one proceeds lexicographically stopping at the third place, where digits differ and declaring 2056 to be the larger one because 5 > 4. The same procedure is also valid if numbers are written in base 2, like 100000000100 and 10000000001.

Observe that these numbers written in Roman numerals are MMLVI and MMXLIX, allowing no natural, organic way for comparing them: in fact, the number with a longer representation is actually smaller, which makes little sense from the perspective of ecological rationality. More importantly, arithmetic operations become simple with place notation, being far more cumbersome with the Roman system.

For mathematics education, the connections between ecological rationality and the number system are relevant. School students learn a minimum of mathematical history. They learn about the Roman system and they learn why even Europeans finally replaced it by a system that—from the structural point of view—is far more ecologically rational. It remains a mystery why the European continent maintained the Roman numeral system for such a long time. Roman numerals were merely used to describe quantities and data, but not for comparisons nor for calculations. They mimicked, at least for numerals between 1 and 50, fingers and hands: 10 corresponded to two hands, 20 to 4 hands and 33 to 3 hands and 3 single fingers. The letter L for 50 meant "half of C," and C was the first letter of Centum, which meant 100 in Latin. The system was additive and cumulative, requiring ever more letters for representing larger numbers. Although its basic symbols may seem natural, because they depict fingers, their representation is not based on an ecologically rational structuring. Not just comparisons but also computations were cumbersome. Until the twelfth century, Italian merchants had to outsource computations on the prices of their merchandise by communicating their numbers to "abacists" who translated Roman numbers into groups of marbles and operated with these groups in a cumbersome way.

In India, and later in Persia and the Islamic countries, the positional number system had been used widely for computations since the seventh century. It was Leonardo da Pisa, also known as Fibonacci, who made the first attempts to import the Hindu-Arabic positional system to Europe, well into the twelfth century (Figure 8.6). The Church opposed this novelty and

Figure 8.6 An allegory of Arithmetic from *Margarita Philosophica*: a young woman, smiling benevolently at the young man who uses the Hindu-Arabic positional system, while his fellow at the other table uses marbles

Source: Gregor Reisch, twelfth century.

even declared the number "0," which the Arabs had called "Zifer" (from which the term "cipher" was derived), as the number of the devil. Yet Fibonacci's dream came true in the end because its ecological rationality provided such great practical utility.

The ecological rationality of defeasible logic and its role in mathematics education

In Chapter 3 of *The Sciences of the Artificial*, Simon recommends the use of multiple logics and in Chapter 5, he explains that "Multiple logics may become necessary when approaching heuristic decision making" (Simon, 1996).This is particularly true in the mathematics classroom today when instructors have to understand students' reasoning without declaring it as "wrong" simply because it does not follow the laws of classical logic. One of the responsibilities of mathematics educators today is to instruct students in identifying different forms of inferences—not just for decision making but also as tools for mathematical thinking in general. The main point we want to make here is that classical logic and probabilities are not sufficient as instruments for understanding and describing human reasoning.

Remember how, in previous sections, we treated a conditional sentence of the kind:

If she tests positive, she has the disease

We imagined a physician reasoning from evidence to hypothesis, based on knowledge of anatomical facts but also on experience. How high is the physician's confidence in this conditional? In the first part of this chapter, we have extensively discussed how the physician can treat this conditional probabilistically by estimating the probability that the patient has the disease, given that she tests positive. It is well established that physicians do not necessarily treat this implication, basing their assessment of frequencies. A fundamental treatise on physicians' decision making in inferences is David Eddy's book (Eddy, 1996), in which he presents conclusions gained from his decades of work examining his colleagues' reasoning. In fact, Eddy became known for propagating the necessity of physicians acquiring statistical literacy. One observation he made seems particularly relevant to our discussion of multiple logics: He was surprised by the fact that excellent doctors could be so bad at correctly estimating predictive values (Eddy, 1982). What they apparently had was a profound knowledge of the reasons or factors that could *enable* a positive test. Basing inferences on such "reasons" instead of frequencies is not probabilistic and it also does not follow the rules of classical logic. Nevertheless, it can be extremely successful. For example, a physician with extensive knowledge of the reasons a test could be positive without the presence of the illness in question is in a good position to order follow-up tests and to observe the patient carefully to rule out these other reasons before making a final diagnosis. Precisely here we see other logics at work, namely, defeasible logics, which take into account so-called enablers and defeaters of causal conditionals.

Psychologists have examined these different ways of reasoning and making inferences successfully. Back in 1995, Denise Cummins performed a seminal experiment. She let participants generate "defeaters" for causal conditionals, such as, "If John studies, he passes the test." She then discovered that the number of defeaters generated by a group of participants inversely predicted the confidence of another group of participants in the statements of the conditionals. In terms of heuristics, one can state that simply "tallying" defeaters inversely predicts confidence in the inference expressed by a conditional. They rely on their tallying *abnormalities* and *alternative causes*, that is "reasons" that cause conditional inferences of the *Modus Ponens* type or of the *Affirmation of the Consequent* type to be defeated on occasions. Stenning, Martignon, and Varga (2017) replicated this experiment in a within-subject design, and obtained basically the same results. They further discovered that people are ecologically rational, in that they switch from tallying defeaters or enablers to using just one of them if it is particularly strong, or sometimes two of them, if together they seem to back up their judgment strongly enough. The relevant conclusion from Cummins' experiment and our replication is that representations for inference in daily life are not just numerical or strictly logical in a classical sense but can be expressed by *narratives*, which obey defeasible logic.

In fact, strict classical logic is the instrument of adversarial communication, while defeasible logic is often the instrument of cooperative communication (Stenning, Martignon, & Varga, 2017). Adversarial communication is the mode adopted by someone who knows that peers and colleagues will be checking each step in a dialectical mode, whereas cooperative communication is the mode adopted by someone arguing to achieve progress toward a common goal. Adversarial communication can place students on the defensive and inhibit learning. Cooperative communication, on the other hand, places teachers and students in a collaboration to support learning. When a mathematics educator analyzes her students' answers, she should be aware of students' different logics before declaring that her students are committing serious mistakes. Mathematics educators are likely to be more successful if they engage their students in a cooperative pursuit of knowledge than if they act as adversaries pouncing on their students' every perceived mistake.

Conclusion

Simon devoted much of his career to promulgating the importance of what we have called ecologically rational representations and procedures. These adapt well to our cognitive apparatus and produce solutions that are good enough for the purpose. Thus, as Simon argued, solving a solution means finding a representation that makes a satisficing solution transparent. In a class attended by one of the authors, Simon argued for the necessity of satisficing with characteristic wry humor, announcing that he was insured to precisely the point that his family was indifferent as to whether he lived or died. As a brilliant pedagogical strategy, this little joke cemented in his students' mind both the impossibility of finding a genuinely optimal insurance strategy, and the necessity of *satisficing* to protect one's family.

This work has explored the role of ecologically rational representations and procedures for several common mathematical reasoning tasks: number comparison, classification, and causal inference. In each class of problems, we illustrated how ecologically rational representations and procedures make satisficing solutions transparent. As tools for education, such ecologically rational strategies support understanding, and maintain the wonder of mathematics in the minds of students.

Notes

1 The relevant excerpt can be found at this link: www.youtube.com/watch?v=ABucG05nurs.
2 See https://en.wikipedia.org/wiki/Human–computer_chess_matches.
3 See https://aibusiness.com/aiva-is-the-first-ai-to-officially-be-recognised-as-a-composer/.

References

Cummins, D. (1995) Naïve theories and causal cognition. *Memory and Cognition*, 23(5), 646–659.

Dreher, A., & Kuntze, S. (2014) Teachers' professional knowledge and noticing: The case of multiple representations in the mathematics classroom. *Educational Studies in Mathematics*, 88(1): 1–26.

Eddy, D. M. (1982) Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249–267). Cambridge: Cambridge University Press.

Eddy, D. M. (1996). *Clinical decision making: From theory to practice: a collection of essays published by the American Medical Association*. New York: Jones and Bartlett Publishers.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, 103, 650–669.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684–704.

Gigerenzer, G., Todd, P. M., & the ABC Research Group (Eds.) (1999). *Simple heuristics that make us smart*. New York: Oxford University Press.

Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *The Journal of Family Practice*, 45(3), 219–226.

Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine*, 73(5), 538–540.

Hoffrage, U., Gigerenzer, G., Krauss, S., & Martignon, L. (2002). Representation facilitates reasoning: What natural frequencies are and what they are not. *Cognition*, 84, 343–352.

Hoffrage, U., Krauss, S., Martignon, L., & Gigerenzer, G. (2015). Natural frequencies improve Bayesian reasoning in simple and complex tasks. *Frontiers in Psychology*, 6(1473), 1–14.

Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science*, 290, 2261–2262.

Laskey, K. B., & Martignon, L. (2014). Comparing fast and frugal trees and Bayesian networks for risk assessment. In K. Makar, B. de Sousa, & R. Gould (Eds.), *Sustainability in statistics education. Proceedings of the Ninth International Conference on Teaching Statistics* (*ICOTS9, July 2014), Flagstaff, Arizona, USA* . Voorburg, The Netherlands: International Statistical Institute.

Luan, S., Schooler, L. J., & Gigerenzer, G. (2011). A signal detection analysis of fast-and-frugal trees. *Psychological Review*, 118, 316–338.

Martignon, L., & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52(1): 29–71.

Martignon, L., & Hoffrage, U. (2019). Wer wagt, gewinnt? Göttingen: Hogrefe.

Martignon, L., Vitouch, O., Takezawa, M., & Forster, M. (2003). Naive and yet enlightened: from natural frequencies to fast and frugal trees. In D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment, and decision making*, (pp. 189–211). Chichester: Wiley.

Simon, H. (1973). The structure of ill-structured problems. *Artificial Intelligence*, 4, 181–201.

Simon, H. (1996). *The sciences of the artificial* (3rd edn). Cambridge, MA: MIT Press.

Stenning, K., Martignon, L., & Varga, A. (2017). Probability-free judgment: Integrating fast and frugal heuristics with a logic of interpretation. *Decision*, 4(3), 136–158.

Von Neumann, J. (1958). *The computer and the brain*. New Haven, CT: Yale University Press.

Woike, J. K., Hoffrage, U., & Martignon, L. (2017). Integrating and testing natural frequencies, naïve Bayes, and fast-and-frugal trees. *Decision*, 4(4), 234–260.