Croatian Journal of Philosophy
Vol. XXIV, No. 72, 2024
https://doi.org/10.52685/cjp.24.72.5
Received: April 3, 2024
Accepted: June 3, 2024

Human and Artificial Decision Making: A Unified View

KONSTANTINOS V. KATSIKOPOULOS
University of Southampton, Southampton, England

Machines can now match, or outperform, human performance in several reasoning and decision tasks. Some say that all that intelligence amounts to is smart computation. This is not a new thesis, dating back to Leibniz as well as Simon and Newell, but what is new is what smart means. Today it is identified with complex statistics and optimisation. Simon's meaning, however, of smart rested on bounded rationality, a unified view of human and artificial decision making. This view was fleshed out by Gigerenzer as fast-and-frugal heuristics. Interestingly, such heuristics are typically sparse, as some machine learning models are optimised to be. So, one might hope that we can make sense of artificial intelligence in human terms after all, and face the upcoming challenges with open-mindedness and courage, just like Simon, and of course Wilkes, would have done.

Keywords: Human decision making; artificial intelligence; bounded rationality; heuristics; smart computation; sparsity.

1. *Overview*

Colin Cherry was what we would today call a cognitive scientist… Or a communications engineer, or a researcher of artificial intelligence... Well, anyway, Cherry was all of the above. His "cocktail party effect" can serve to illustrate my contribution to the 2023 Kathy Wilkes Memorial Conference, also weaving in the ideas of Herb Simon. Cherry, Simon, and Wilkes align in all being open-minded, courageous researchers, who asked the tough questions of what effective human and machine communication, reasoning and decision making is, and provided answers that invite deep thought. My contribution connects to the conference contributions by Philipp Koralus (human reasoning and decision making) and Peter Millican (artificial intelligence).

First, on the cocktail party effect. Cherry (1953) run laboratory experiments where participants listened to two different messages from the same speaker and were instructed to separate them. Whether and the extent to which people can perform such tasks accurately depends on factors such as the direction from which the messages are coming, the pitch of the messages and the rate of speech. In a classic demonstration, it was found that people can detect message segments they are not actively attending to only if these segments are important to them, as when their name is spoken (Moray 1959). For a long time, human performance on cocktail party settings could not be matched by machines. But Xie et al. (2015) combined acoustic metamaterials with computational sensing to achieve competitive performance, wherein three overlapping and independent auditory sources were separated with 97% accuracy.

This machine success raised some eyebrows then, but it might have done much less so today. In the last decade, we have had the reinforcement learning algorithm Alpha Zero mastering simultaneously the games of Go, chess, and shogi at world-class level (Silver et al. 2018). And you have probably all heard enough about large language models such as Chat GPT in the past couple of years. It is definitely on the table now that machine behavior can be as intelligent as human behavior, or even more so. Furthermore, some forcefully say that intelligent behavior *only* requires smart computation.

This is not a new thesis. In the seventeenth and eighteenth centuries Gottfried Leibniz had the dream of a *Characteristica universalis*, a universal language of formal computation. Herb Simon and Allan Newell received the 1975 Turing award for their *physical symbol systems hypothesis*, stating that symbol manipulation is a necessary and sufficient condition for intelligence, be it human or artificial. What *is* new is that smart computation has recently been identified with complex statistics and optimization. But this was not Simon's meaning of "smart." The objective of this article is to present, drawing from my presentation at St Hilda's College conference at Oxford University, research that fleshes out Simon's unified view of an integral part of human and artificial behavior, decision making, as it has been developing from the seventies until today.

2. *Simon's meaning of "smart"*

Herb Simon has been the only person to date who received both the Nobel prize in 1978 and the Turing award in 1975. He received the Nobel prize for economics, the Turing award is for computer science. Economics is a social science and computer science is—in Simon's own words (1968)—a science of the artificial. Both economics and computing study behavior, of humans and computers respectively. In Simon's view, the analogy can be pushed further because he saw both humans and computers fundamentally as systems that exhibit intelligent behavior because they process information; and more specifically, as per the physical symbol systems hypothesis, they process symbols. At the highest level of abstraction, this is what being smart meant for Simon.

Now, Herb Simon was a polymath if there ever was one. He was so prolific and wrote over seven decades, that is not always easy to follow his various intellectual threads and see them fleshed out (Petracca 2021). Yes, intelligence for Simon was information processing, but is that all? Can one flesh Simon's vision out? Yes, one can and some did. Before presenting Gerd Gigerenzer's implementation, analysis, and testing of Simon's vision in the next section, at the research center he directed at the Max Planck Institute for Human Development in Berlin, let us discuss Simon's vision a little more.

Simon was gifted in formal modeling, mathematical and computational. He has produced multiple articles in areas such as statistics, decision theory and operations research. For instance, he developed fundamental methods for distinguishing spurious from genuine statistical correlations (Simon 1954) and for deriving optimal policies for stochastic dynamic programming (Simon 1956). Even though he was accused of unnecessarily, even according to some damagingly, "hardening" the social sciences, he was a dedicated scientist who understood what formal models can and cannot do for theoretical development and empirical testing (Katsikopoulos, Marewski and Hoffrage 2024).

In other words, Simon can be said to have acted respectfully to Einstein's maxim "everything should be made as simple as possible, but not more so." As such, the formal expression of smart computation Simon endorsed is considerably simpler than today's reliance on complex statistics and optimization, evident in areas such as big data analytics and statistical machine learning (Katsikopoulos and Canellas 2012). An obvious reaction to this observation is that such areas have become much more technically sophisticated since Simon's time—he passed away in 2001 and produced most of his more technical work in the fifties, sixties, seventies, and eighties—and in the most recent couple of decades. This is true of course but I do not believe that it accounts for the whole difference. For example, Simon resisted vehemently the then state-of-the-art statistical ritual of null hypothesis significance testing. And he was also pushing the envelope for developing new methods that can address stubborn problems, rather than changing problems so that the known methods can address them; hence his crusade on developing artificial intelligence. The exchange with leading applied mathematician Richard Bellman (Simon and Newell 1958; Bellman 1958), where he forcefully argued that current decision methods could not handle ill-structured problems, is particularly telling in this regard. (It would have been intriguing to hear Simon's take on today's explosion of data science and machine learning, but, alas, we do not have this privilege).

In line with his overall attitude to decision modelling, Simon (1956, 1968) voiced consistent concerns about the effectiveness of mathematical optimization outside toy problems, outside the lab, or how one might call it, "in the wild." It is a truism that something optimal according to a model of the world might be not only suboptimal, but even poor performing, in the wild. But it is a truism that modelers all too often do not heed, preferring to go about the usual business of optimizing, without even testing how benchmarks, such as non-optimizing models, perform in the wild. The fields of "soft" and "behavioral" operations research have been more sensitive than "hard" operations research, considering and acting on such points (Ackoff 1979; Rosenhead and Mingers 2001; for a historical and conceptual perspective on all these types of operations research, see Katsikopoulos 2023). But it is interesting that Simon first made such suggestions decades before (for another aligned contemporary perspective, see Kimball 1958).

What neither Simon, nor the fields of soft and behavioral operations research, did, however, is to develop *formal decision models*, computational and mathematical, that are applicable to the wild. Simon (1956) sketched the idea of *satisficing*—this word is a portmanteau of "satisfying" and "suffice"—models, which, contra optimization, do not search and settle only on the theoretically best option, but may choose another option based on criteria other than optimality, such as adaptiveness and robustness. Simon did not empirically test how far this idea can go but conjectured that it can: "The presence of uncertainty places a premium on robust adaptive procedures instead of optimization strategies that work well only when finely tuned to precisely known environments" (Simon 1968: 35). Was he right? The next section provides some answers.

3. *Gigerenzer's analysis, implementation, and testing of Simon's "smart"*

A broad and deep investigation of Simon's conjecture has been ongoing since the mid-nineties. Gerd Gigerenzer, a psychologist with a philosopher's inclination to analyze conceptually and a scientist's skill to implement and test empirically, did exactly that with Simon's concept of "smart." Whereas many have claimed to stand on Simon's shoulders—including no less the founders of the modern field of heuristics, psychologists Amos Tversky and Daniel Kahneman—scrutiny reveals that perhaps it was Gigerenzer who most closely did so. Historian Enrico Petracca has even, fittingly I believe, suggested that the work of Gerd Gigerenzer, Peter Todd and colleagues "could appear 'more Simonian than Simon'" (2021: 710). In a nutshell, Gigerenzer, Todd and the ABC research group (1999) can take credit for making clear and distinguishing three possible interpretations of Herb Simon's key idea of *bounded rationality*, as explained below.

The dominant interpretation of bounded rationality in economics (Sargent 1993) is that it is still optimization, but under constraints (e.g., cognitive, systemic). The leading interpretation of bounded rationality in psychology (Kahneman, Slovic and Tversky 1982) is consistent with the economics one, but takes the form of attributing the (supposedly regrettable) lack of (full) optimization to people's heuristics. This is the well-known *heuristics-and-biases* research program. The third interpretation of bounded rationality is more interdisciplinary, aligned with the concerns of disciplines that assess decision performance in the wild, such as human factors, operations research, and artificial intelligence (Katsikopoulos 2023). According to Gigerenzer et al. (1999), bounded rationality is not the study of what theoretically can be called lack of rationality, but the study of what practically is a real rationality that real organisms can and aim to exhibit. Bounded rationality is implemented by heuristics, some of which are of the satisficing variety originally proposed by Simon and others of which are *fast-and-frugal heuristics*, which constituted a new major research program, and according to some such as Kelman (2011), the main antipode to the heuristics-and-biases program (see also the paper by Koralus in the conference and in this collection). A main thesis of fast-and-frugal heuristics is that intelligent people, and machines, use few, informative, pieces of information, and combine those pieces in mathematically simple ways.

Fast-and-frugal heuristics are, just like most research developments, not entirely new. Early demonstrations of the concept can be found in the seventies in the work of Robyn Dawes (Dawes and Corrigan 1974) and Robin Hogarth (Einhorn and Hogarth 1975), who empirically showed that just tallying variables, without weighting them differentially as in least-squares regressions, could lead to equally, and sometimes even more, accurate predictions. Such results were, however, not typically taken that seriously. Gigerenzer's greater volume of empirical results, and of supporting theoretical analyses, with the help of some dozens of researchers at the Max Planck Institute for Human Development and in a world-wide network (Gigerenzer, Hertwig and Pachur 2012) eventually attracted more attention, and served to establish the success of fast-and-frugal heuristics across many domains in the wild (Katsikopoulos, Şimşek, Buckmann and Gigerenzer 2020). The remaining of this section samples two empirical demonstrations, both from the geopolitical sphere trying to connect to Kathy Wilkes' activism and gives a glimpse of the analytical theory.

Predicting election outcomes. Ahead of the 2016 U.S. presidential election, big data algorithms predicted a 71.4% chance of Hilary Clinton winning (Katsikopoulos et al. 2020). Furthermore, polls and prediction markets made the same prediction. Historian Allan Lichtman, on the other hand, predicted that Donald Trump would win. Lichtman (1980) *13 keys to the White House* is a tallying heuristic he derived based on domain knowledge, blending theories of politics, economics, sociology, and psychology. The keys—also called attributes or features—were fixed once and for all before the 1984 election and have been used to correctly predict all U.S. elections since. Each key is an issue that matters to voters; they are stated so that each is either true or false ahead of an election. Some of the keys are facts, others require judgment. It is of course key (pun intended) that all keys are judged and scored before the election.

Key 1: Incumbent party holds more seats in the House of Representatives after this midterm election than the previous one.
Key 2: No serious contest for incumbent-party nomination.
Key 3: Incumbent-party candidate is the sitting president.
Key 4: No significant third-party or independent campaign.
Key 5: Economy not in recession during campaign.
Key 6: Real annual per capita economic growth during the term equals or exceeds mean growth during two previous terms.
Key 7: Incumbent administration effects major changes in national policy.
Key 8: No sustained social unrest during the term.
Key 9: Incumbent administration untainted by major scandal.
Key 10: Incumbent administration suffers no major failure in foreign or military affairs.
Key 11: Incumbent administration achieves a major success in foreign or military affairs.
Key 12: Incumbent-party candidate is charismatic or national hero.
Key 13: The challenging-party candidate is not charismatic or national hero.

Lichtman proposed the following heuristic:

> *Score all keys and tally the number of false keys. If this tally is six or more, the challenger will win.*

For instance, in the 2012 election, Mitt Romney challenged Barack Obama. Lichtman counted only three keys as false (1, 6 and 12), and correctly predicted that Obama would win. In late September 2016, Lichtman found six false keys (1, 3, 4, 7, 11 and 12) and predicted, again correctly, that Trump would win (for further discussion, including some subtleties, see Katsikopoulos et al. 2020).

Unlike big data analytics, the 13-key rule is transparent. The rule contradicts campaign wisdom: All keys—except key 13—refer to the incumbent party, i.e., the party holding the White House and its candidate. That is, incumbents tend to lose, rather than challengers tend to win. The heuristic delivers a simple theory, a process-based explanation for behavior, and creates a platform for discussion, qualities important in a healthy democracy (Katsikopoulos and Canellas 2022).

Understanding and improving the operation of checkpoints. In checkpoints set up by NATO in Afghanistan between 2004 and 2009, soldiers had to classify approaching cars as a friend or a foe, and decide how to make the car stop, so that it can be respectively searched or neutralized. How did soldiers make these decisions? Did it work? Can research help do better?

The Wikileaks reports mined by Keller and Katsikopoulos (2016) referred to 1160 incidents, of which 7 were suicide attacks and 1053 civilian. Suicide attacks resulted to all car occupants and soldiers dying. The civilian incidents resulted to 204 people injured or killed. Applying standard operations research or artificial intelligence techniques is tempting but does not work because the empirical estimate of the probability of a hit (i.e., soldiers classify a car with suicide attackers as a foe) is zero. This cannot be right and would lead to extreme classifiers such as classifying all cars as civilian, which is also not right.

It seems that soldiers relied almost exclusively, for 1020 out of 1053 civilian incidents, on a heuristic that uses one attribute, *compliance*:

> *If an approaching car appears to comply with soldier instructions (e.g., slows down), then classify it as a friend and ensure peacefully that it stops and is searched. If the car does not appear to comply (e.g., speeds up), then classify it as a foe and ensure that it is neutralized, which might require shooting at it, etc.*

Can this reasonable, but to a good extent ineffective, heuristic be enhanced so that it leads to improved decision making that would have resulted to fewer than 204 civilian casualties? The authors consulted with experts, such as military personnel and teachers in military academies, about how to classify an approaching car as friend or foe and sought to combine their insights with the compliance attribute. The resulting method is more complex than a single-attribute heuristic, but still a fast-and-frugal heuristic, of the type called a *fast-and-frugal (decision) tree* (Martignon, Katsikopoulos and Woike 2008). This tree is shown in Figure 1 below (in the tree, threat cues refer to any information that makes a car seem suspicious, e.g., intelligence information).

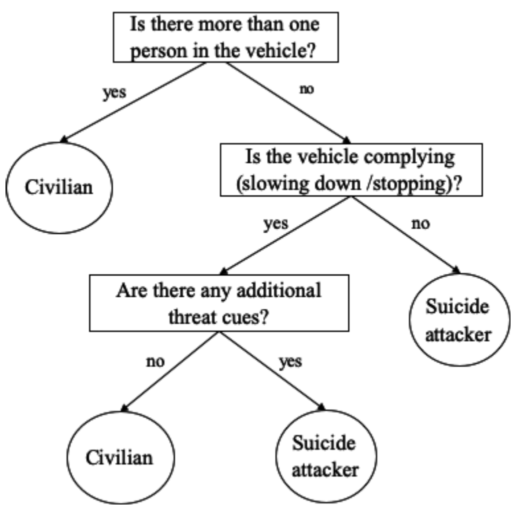

Figure 1. A fast-and-frugal decision tree for classifying cars approaching a checkpoint as friend or foe (adapted from Keller and Katsikopoulos 2016).

Decision trees are a popular type of transparent model in machine learning (Breiman et al. 1984; Bertsimas, Dunn and Mundru 2019). Fast-and-frugal trees are designed to be even more transparent. For example, the Figure 1 tree asks only a few questions, it asks those questions one at a time, and each time it asks a question it is possible that a decision is made immediately. Is the tree only transparent, or is it also accurate? Had it been applied to the *Wikileaks* dataset, it would have led to 78 casualties, that is a 60% decrease from what transpired.

Is he cherry picking? No. Katsikopoulos et al. (2020) compared the classification error of Breiman's (1984) classification and regression trees (CART) to tallying and fast-and-frugal trees in 64 classification tasks, containing 95 to 32,561 instances (median 904) and three to 1,418 cues (median 19). Across the 64 tasks, each fast-and-frugal heuristic predicted nearly as well as CART, falling behind by only half a percentage point. There is an advantage for CART in problems where the error is small, that is, in easy tasks, and an advantage for fast-and-frugal heuristic when the error is larger, that is, in more difficult tasks. Furthermore, decades of competitions among such simple heuristics and more complex optimization models such as linear regressions (including regularized versions), Bayesian networks, decision trees, random forests, and support vector machines, spread across disciplines such as psychology, economics, engineering design, statistics, and machine learning, have shown that the differences in predictive accuracy between these two model families are not that large, and that each family enjoys a region of superior performance. In other words, heuristics can be *robust*. Why? Are there explanations for these results?

Theory: The role of sparsity. Yes. For reviews, see Martignon and Hoffrage (2002), Katsikopoulos et al. (2018), and Katsikopoulos (2023: Chapter 6). A good, short answer is *sparsity*. A model is sparse if only a small proportion of its parameters are different from zero. For example, regularization techniques push regressions towards sparsity. Sparsity can make a model more predictively accurate because it does not overfit in the training set (Geman, Bienenstock and Doursat 1992; Rudin 2019).

The checkpoint fast-and-frugal tree of Figure 1 is a sparse version of full-blown CART decision trees. Tallying is a sparse version of linear regression with potentially differentially weighted attributes (Lichtenberg and Şimşek 2019). Whereas it is overall appreciated that such transparent models can be accurate as well, it should be noted that there are multiple approaches to deriving those. Bertsimas et al. (2019) and Rudin (2019) suggest that sparse models may be derived as solutions to optimization problems, while fast-and-frugal heuristics researchers may additionally generate such models by deeply observing human decision making (Gigerenzer et al. 2011; Katsikopoulos et al. 2020).

4. *Epilogue*

Cherry, Simon, and Wilkes all sought to understand human and artificial intelligence, by taking open-minded, penetrating, and courageous approaches. Humanity always faces challenges, and perhaps the artificial intelligence one will prove to be a very tough, even existential, one. May a unified view of human and artificial decision making, as the one championed by Gigerenzer and presented here, act as a resource to keep such issues at bay.[1]

References

Ackoff, R. L. 1979. "The future of operational research is past." *Journal of the Operational Research Society* 30: 93–104.

Bellman, R. 1958. "On 'Heuristic Problem Solving' by Simon and Newell." *Operations Research* 6 (3): 448–449.

Bertsimas, D., Dunn, J. and Mundru, N. 2019. "Optimal prescriptive trees." *INFORMS Journal on Optimization* 1 (2): 164–183.

Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. 1984. *Classification and Regression Trees*. London: Chapman and Hall.

Cherry, E. C. 1953. "Some experiments on the recognition of speech, with one and with two Ears." *The Journal of the Acoustical Society of America* 25 (5): 975–979.

Dawes, R. M. and Corrigan, B. 1974. "Linear models in decision making. " *Psychological Bulletin* 81 (2): 95–106.

Einhorn, H. J. and Hogarth, R. M. 1975. "Unit weighting schemes for decision making." *Organizational Behavior and Human Performance* 13 (2): 171–192.

Geman, S., Bienenstock, E. and Doursat, R. 1992. "Neural networks and the bias/variance dilemma." *Neural Computation* 4 (1): 1–58.

Gigerenzer, G., Hertwig, R. and Pachur, T. (eds.). 2011. *Heuristics*: *The foundations of adaptive behavior*. New York, NY: Oxford University Press.

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

Kahneman, D., Slovic, P. and Tversky, A. (eds.). 1982. *Judgment under uncertainty: Heuristics and biases*. Cambridge, UK: Cambridge University Press.

Katsikopoulos, K. V. 2023. *Cognitive operations: Models that open the black box and predict our decisions*. London, UK: Palgrave Macmillan.

Katsikopoulos, K. V. and Canellas, M. C. 2022. "Decoding human behavior with big data? Critical, constructive input from the decision sciences." *AI Magazine* 43 (1): 1–13.

[1] Since the time this article was written and accepted for publication, two events occurred that necessitate the following updates: First, Herb Simon is no longer the only person who has received both the Nobel prize and the Turing award as Geoff Hinton has now also received those (Nobel prize in physics in 2024 and Turing award in 2019). Second, Allan Lichtman's 13 keys to the White House has no longer predicted all U.S. elections since 1984 because the model failed to predict the outcome of the 2024 election.

Katsikopoulos, K. V., Durbach, I. N. and Stewart, T. J. 2018. "When should we use simple decision models? A synthesis of various research strands." *Omega*: *The International Journal of Management Science* 81: 17–25.

Katsikopoulos, K. V., Marewski, J. N. and Hoffrage, U. 2024. "Heuristics for metascience: Simon and Popper." In G. Gigerenzer, S. Mousavi and R. Viale (eds.). *The Herbert Simon Companion*: *Standing the Test of Time*. Northampton: Edward Elgar, 300–311.

Katsikopoulos, K. V., Şimşek, Ö. Buckmann, M. and Gigerenzer, G. 2020. *Classification in the wild: The science and art of transparent decision making*. Cambridge: MIT Press.

Keller, N. and Katsikopoulos, K. V. 2016. "On the role of psychological heuristics in operational research; and a demonstration in military stability operations." *European Journal of Operational Research* 249 (3): 1063–1073.

Kelman, M. G. 2011. *The heuristics debate*. New York, NY: Oxford University Press.

Kimball, G. E. 1958. "A critique of operations research." *Journal of the Washington Academy of Sciences* 48 (2).

Lichtenberg, J. M. and Şimşek, Ö. 2019. "Regularization in directable environments with application to Tetris." *International Conference on Machine Learning*: 3953–3962.

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

Martignon, L., Katsikopoulos, K. V. and Woike, J. K. 2008. "Categorization with limited resources: A family of simple heuristics." *Journal of Mathematical Psychology* 52 (6): 352–361.

Moray, N. 1959. "Attention in dichotic listening: Affective cues and the influence of instructions." *Quarterly Journal of Experimental Psychology* 11 (1): 56–60.

Petracca, E. 2021. "On the origins and consequences of Simon's modular approach to bounded rationality in economics." *The European Journal of the History of Economic Thought*: 1–24.

Rosenhead, J. and Mingers, J. (eds.). 2001. *Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict*. New York: John Wiley and Sons.

Rudin, C. 2019. "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead." *Nature Machine Intelligence* 1 (5): 206–215.

Sargent, T. J. 1993. *Bounded rationality in macroeconomics: The Arne Ryde memorial lectures*. New York: Oxford University Press.

Silver, D., et al. 2018. "A general reinforcement learning algorithm that masters chess, shogi, and go through self-play." *Science* 362 (6419): 1140–1144.

Simon, H. A. 1954. "Spurious correlation: A causal interpretation." *Journal of the American Statistical Association* 49(267): 467–479.

Simon, H. A. 1956. "Rational choice and the structure of the environment." *Psychological Review* 63 (2): 129–138.

Simon, H. A. 1968. *The sciences of the artificial*. Cambridge: MIT Press.

Simon, H. A. and Newell, A. 1958. "Heuristic problem solving: The next advance in operations research." *Operations Research* 6: 1–10.

Xie, Y., Tsai, T. H., Konneker, A., Popa, B. I., Brady, D. J. and Cummer, S. A. 2015. "Single-sensor multispeaker listening with acoustic metamaterials." *Proceedings of the National Academy of Sciences* 112 (34): 10595–10598.